

POST-LAUNCH CALIBRATION OF SATELLITE SENSORS

Edited by Stanley A. Morain and Amelia M. Budge

POST-LAUNCH CALIBRATION OF SATELLITE SENSORS

International Society for Photogrammetry and Remote Sensing (ISPRS) Book Series

Book Series Editor

Maxim Shoshany

Faculty of Civil and Environmental Engineering

Technion, Israel Institute of Technology

Haifa, Israel



Post-Launch Calibration of Satellite Sensors

Edited by

Stanley A. Morain & Amelia M. Budge

Earth Data Analysis Center, University of New Mexico, Albuquerque, NM, USA



A.A. BALKEMA PUBLISHERS LEIDEN / LONDON / NEW YORK / PHILADELPHIA / SINGAPORE

Coverinfo:

Symbolically, the aims and goals of this compendium are presented in its cover image, a string of mis-aligned image strips. Through radiometric and geometric calibration, the task is to bring these segments once again into a sharper and more accurate single image. The original image was prepared by personnel at Goddard Space Flight Center; the disjointed and segmented image, by Mr. Ken Christian, InDyne, Inc., Stennis Space Center, Mississippi.

Copyright © 2004 Taylor & Francis Group plc, London, UK

All rights reserved. No part of this publication or the information contained herein may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording or otherwise, without written prior permission from the publisher.

Although all care is taken to ensure the integrity and quality of this publication and the information herein, no responsibility is assumed by the publishers nor the author for any damage to property or persons as a result of operation or use of this publication and/or the information contained herein.

Published by: A.A. Balkema Publishers, a member of Taylor & Francis Group plc
www.balkema.nl and www.tandf.co.uk

ISBN 90 5809 693 9

ISSN 1572 - 3348

Table of Contents

Preface	VII
Foreword	IX

Needs, terms and procedures for calibration

How well do we understand Earth observation electro-optical sensor parameters? <i>G. Joseph</i>	1
--	---

Metrology for remote sensing radiometry <i>B.C. Johnson, S.W. Brown & J.P. Rice</i>	7
--	---

Stability and accuracy requirements for passive satellite remote sensing instrumentation for global climate change monitoring <i>R. Datla, W. Emery, G. Ohring, R. Spencer & B. Wielicki</i>	17
---	----

Calibration standards and guidelines

Validated data and removal of bias through traceability to SI units <i>N.P. Fox</i>	29
--	----

Status of standardization project ISO 19130 <i>W. Kresse</i>	43
---	----

A proposed ISO/TC 211 standards project on radiometric calibration of remote sensing data <i>L. Di</i>	53
---	----

Techniques for in-flight and on-orbit calibration

System-level approach to characterization and radiometric calibration of a space-based electro-optical sensor <i>J. Tansock, A. Thurgood & M. Larsen</i>	57
---	----

On-orbit MTF assessment of satellite cameras <i>D. Léger, F. Viallefont, P. Déliot & C. Valorge</i>	67
--	----

USGS/OSU progress with digital camera <i>in-situ</i> calibration methods <i>D.C. Merchant, A. Schenk, T. Yoon & A. Habib</i>	77
---	----

Issues and methods for in-flight and on-orbit calibration <i>K. Jacobsen</i>	83
---	----

Intersensor calibration

- In-flight intersensor radiometric calibration using vicarious approaches 93
K.J. Thome
- Validating the MISR radiometric scale for the ocean aerosol science community 103
C.J. Bruegge, W. Abdou, D.J. Diner, B.J. Gaitley, M. Helmlinger, R.A. Kahn & J.V. Martonchik

Calibration programs

- Forty years of experience with SPOT in-flight calibration 117
C. Valorge, A. Meygret, L. Lebègue, P. Henry, A. Bouillon, R. Gachet, E. Breton, D. Léger & F. Viallefont
- The Joint Agency Commercial Imagery Evaluation (JACIE) team and product characterization approach 135
V. Zanoni, M. Pagnutti, R. Ryan, D. Helder, W. Lehman, S. Roylance & G. Snyder
- Camera calibration program in the United States: past, present, and future 143
G.Y.G. Lee

Building and characterizing calibration test ranges

- In-flight characterization of spatial quality using point spread functions 149
D. Helder, T. Choi & M. Rangaswamy
- Test ranges for metric calibration and validation of high-resolution satellite imaging systems 171
G. Dial & J. Grodecki
- Locating potential vicarious calibration sites for high-spectral resolution sensors in the Negev Desert by GIS analysis 181
U. Gilead & A. Karnieli

Acronyms 189

Preface

The need for consistency in sensor calibration terminology, and the definitions for these terms, has been an issue in the Earth observing community for more than two decades. With this volume, it is gratifying to see progress in this direction; more so because ISPRS has taken a leadership role in its production. In 1998, the Committee on Earth Observing Satellites (CEOS) passed a Resolution requested by the terrain mapping subcommittee of the Working Group on Calibration and Validation (WGCV) to form a Joint WGCV/ISPRS Task Force on Radiometric and Geometric Calibration. For its part, Dr. George Joseph (then President of ISPRS Commission-I, 1996–2000) drafted a white paper outlining why confusion arises among various calibration terms used to measure spatial, radiometric, and temporal resolution of sensors. He also suggested definitions for several of these terms. His white paper is published here as the lead article.

After passing the presidency of Commission-I to Dr. Stan Morain for the 2000–2004 Congress quadrennium, a profile of perceived expertise was created by ISPRS prior to organizing the first meeting of the Joint Task Force. A founding group of some fifteen people convened at the Commission-I midterm symposium in 2002, one of the outcomes of which was to schedule an *International Workshop on Radiometric and Geometric Calibration*. Over 80 experts from seven countries met for four days in December 2003 to discuss issues and methods for post launch calibration of infrared and visible optical scanners (IVOS) onboard commercial and government satellites. Their discussions also touched upon digital aerial cameras, characterization of image quality, and design of field calibration sites and test ranges. This volume is a collection of presentations forming the technical program for that workshop.

The volume addresses only a fraction of the complexity of calibration terms and definitions. In the 2004–2008 Congress quadrennium, it is hoped that Commission-I will continue this important work through incremental improvements, and perhaps by expanding the Task Force's scope to include the full range of sensors.



John Trinder, President, ISPRS

Foreword

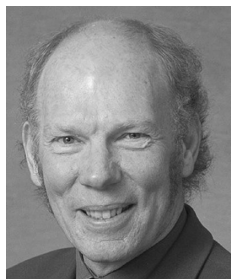
It is now generally recognised that calibration is an essential component of any sensing system; however, interest in the topic is limited to a few scientists who are directly involved. The December 2003 workshop, the CD of technical presentations, the workshop report, and this published volume are designed to review current activities in radiometric and geometric calibration, and to create discussion on this topic in context of image quality, mapping, and data analysis. The topic is a wide one involving many organizations and touching upon topics beyond just calibration. Standards are also involved, along with interoperability, traceability of units, and the presentation of data. I was pleased to see representatives from all of these communities at the workshop and I know that the discussions were constructive in preparing for the 20th ISPRS Congress in Istanbul.

Working groups coordinated by Technical Commissions carry out the technical work of ISPRS. In the 2000–2004 quadrennium there are seven technical commissions and a total of 47 working groups. Five of these working groups specifically mention *calibration* or *standards* in their terms of reference, and work with organizations such as ISO and OGC in developing standards or collaborating in developing new protocols and techniques. ISPRS also works with other international organizations, is an Associate member of CEOS, and also attends the United Nations Committee of Peaceful Uses of Outer Space. We also have memoranda of understanding with IEEE-IGARSS and SPIE.

Many different committees and groups are involved in sensor calibration and image and data quality assessment. The CEOS Working Group on Calibration and Validation (WGCV) is one. In this working group, several national space agencies and national standards laboratories are represented. ISO and OGC are also involved, as is the wider international community of photogrammetry and remote sensing, represented by ISPRS, and other organizations. It is essential that these bodies work together, not only to advance the science of calibration, but also to gain a better understanding of each others' perspectives on the topic and to avoid duplication of effort. With so many new sensors, calibration and inter-comparison is increasingly important. We should also note that new digital aerial sensors and LiDAR create new challenges.

One of the components of the December 2003 workshop leading to this volume was a set of in-depth discussions conducted by the Joint WGCV/ISPRS Task Force. This group arose from a conviction that commercial satellite data providers and space agencies should present data about sensors in a way that would ensure one's ability to intercompare sensors and products. This conviction was later formalized in a resolution of the CEOS Plenary, leading to the creation of the Task Force. The basic task could be constrained simply to listing parameters in a uniform way, but the implications are much wider.

Finally, I would like to thank Stan Morain and Amelia Budge for the effort they have invested in producing this volume. They brought together a formidable group of experts covering a wide range of topics for the workshop, then coordinated and edited the manuscripts published here. I would also like to thank Karsten Jacobsen, University of Hannover, and Manfred Schroeder, DLR, for their valuable services as consulting editors; and both NASA and USGS for their financial and scientific support for the meeting and the book.



Ian Dowman, Secretary General, ISPRS

Needs, terms and procedures for calibration

How well do we understand Earth observation electro-optical sensor parameters?

G. Joseph

Former Director, Space Applications Centre, ISRO, Ahmedabad, India

ABSTRACT: There is a strong need for consistency in terminology, definition of terms, and measurement techniques for Earth observing sensors. It appears that none of these conditions is met by either governments or industries currently involved in developing sensors, designing in-flight or on-orbit calibrations, measuring the performance of sensors, or characterizing the geometric or radiometric quality of products they produce. The paper presented by Dr. George Joseph in 1999, while serving as the President of ISPRS Commission-I, summarizes the issues surrounding these inconsistencies, and forms the basis for creating the Joint ISPRS/CEOS-WGCV Task Force on Radiometric and Geometric Calibration [editor].

1 INTRODUCTION

Extensive use of Landsat data for monitoring and managing various resources, along with detailed field and laboratory studies, have pointed out the need for Earth observation sensors of improved performance. Users are demanding data with better spatial, spectral, radiometric, and temporal resolution. Improvements in the quality of data products also are required, especially in the area of geometric (location/internal distortions) and radiometric accuracy (absolute/relative). For many applications like disaster monitoring/mitigation and agriculture, quick delivery of data is demanded.

2 SPATIAL RESOLUTION

Spatial resolution is a parameter often mentioned but little understood. Is spatial resolution the smallest object one can discern in an image? Experience shows that in Landsat TM imagery one can see roads with widths much narrower than the specified resolution of 30 m, provided there is adequate contrast with surroundings bordering the road. Sensor designers refer to this as the instantaneous geometric field-of-view (IGFOV). IGFOV can be defined as *the geometric size of the image projected by the detector on the ground through the optical system*, often called “pixel footprint.” Any optical system reduces the contrast in the image compared to the contrast of the objects imaged. This is expressed as the modulation transfer function (MTF). The question is: whether it is justifiable to

specify the sensor quality just in terms of geometric projection without any consideration of the contrast reduction it produces? How does one judge image quality of different sensors with the same IGFOV, but different MTF? Keeping this in mind, a concept of effective instantaneous field-of-view (EIFOV) was developed by NASA (1973). EIFOV is defined as *the resolution corresponding to a spatial frequency (ground resolution) for which the system MTF is 50 percent*. This concept appears to be a good measure to compare different sensors. However, if the MTF curve shapes are different, the EIFOV may not be a good indicator for intercomparison of sensor quality.

The NASA concept did not gain popularity amongst sensor designers. To add to the confusion, yet another term was added: ground sample distance (GSD). That means that data can be generated by sampling at certain specified ground distances (although GSD is sometimes used as a synonym for IGFOV). Consider the pushbroom sensors currently used in many (especially high-resolution) Earth observation systems. In the along-track direction, if radiometric and electronic performance allow, the GSD can be made smaller than the IGFOV to achieve better image quality because of reduction of smear. Does across-track performance improve just because, the data are resampled at smaller intervals compared to IGFOV? Can 5 m IGFOV sensor data, sampled at 1 m, have the same performance as a 1 m IGFOV sensor? What is the maximum allowable ratio of IGFOV over GSD?

Generally, MTF is considered as an indicator for how sharp the edges are after contrast reduction during imaging. However, MTF is also a measure of how

accurately the actual radiance from a pixel is measured, since a lower MTF indicates contribution from other pixels to the pixel under observation (and vice versa). Consider an object with 100 percent contrast (white surrounded by black objects). When the contrast in the image space is 10 percent, only 10 percent of the actual energy is measured. This could lead to problems in multispectral classification, since the radiance of a measured pixel is dependent on the nature of adjacent pixels. Thus, the same object with different surroundings will have different signatures. This concern was raised by Norwood (1974). Therefore, the question is: what is “radiometrically accurate IFOV?” Let us define this as *the resolution for which the MTF is higher than 0.95*.

Thus, the instrument designer should specify three parameters relating to resolution: (a) IGFOV; (b) radiometrically accurate IFOV (RAIFOV); and (c) MTF at IGFOV. For completeness, the designer should also give the sensor swath, i.e. the total field-of-view (TFOV). All this should be referenced to orbit height, or expressed in terms of angular measures.

3 RADIOMETRIC RESOLUTION

Radiometric resolution is the noise equivalent reflectance ($NE\Delta\rho$), or temperature ($NE\Delta T$). This can be defined as *the minimum change in reflectance (temperature) that can be detected by the sensor*. It depends on a number of parameters such as the signal-to-noise ratio (S/N), the saturation radiance setting; and the number of quantization bits. Satellite Pour l’Observation de la Terre (SPOT) has an 8-bit quantization, while the Indian Remote Sensing Satellite (IRS) Linear Imaging Self-Scanning System (LISS) cameras have 7-bit quantization. In principle, both can have the same $NE\Delta\rho$ for a specific reflectance radiance, if S/N and saturation setting are properly chosen. Current systems are being designed with 11 or more bit digitization. Such systems, unless they have a corresponding S/N, do not imply a better radiometric resolution. This is illustrated in Table 1.

Table 1. Comparison of design parameters for two sensors and the trade-offs in instrument performance.

Parameters	Sensor-1	Sensor-2
Saturation radiance (SR) (mW/cm ² /steradian/ μ m)	20	35
Quantization (bits)	7	8
Radiometric Sampling Interval (RSI)	0.16	0.14
System S/N at SR	120	150
$NE\Delta\rho$	0.17	0.23

Sensor-2 can sample at a smaller radiometric interval compared to sensor-1 and has a better S/N at the saturation radiance. However, sensor-1 gives a better performance in terms of the noise equivalent radiance at the saturation radiance. Thus, the radiometric resolution capability of an instrument in terms of quantization does not necessarily increase the accuracy with which it can measure radiance. Nevertheless, a higher number of bits increases the dynamic range, so that measurement of highly variable objects like oceans and snow can be made without changing the gain.

It is suggested that designers give not only the number of bits to which data have been quantized, but also the $NE\Delta\rho$. For example, one should state that “the data are quantized to 12 bits, (with $NE\Delta\rho \times \text{mW/cm}^2/\text{steradian}/\mu\text{m}$ for Bands 1, 2, 3...)”. Many users may not get an adequate sense of a sensor’s radiometric quality using $NE\Delta\rho$. It may be easier to compare sensors (and to get a real feel) by giving the S/N at two radiance values. In addition, S/N is a measured quantity, whereas $NE\Delta\rho$ is a derived parameter. It is suggested that for land observation sensors, the lower values of the radiance unit per micron (R_L) may be selected as 5 mW/cm²/steradian/ μ m for the VNIR region and 1 mW/cm²/steradian/ μ m for SWIR. The upper reference level (R_U) may be set as 25 mW/cm²/steradian/ μ m for VNIR and 5 mW/cm²/steradian/ μ m for SWIR. The lower value has been chosen based on average vegetation reflectance, while for ocean observation a 10 times lower value has been chosen.

Spectral bandwidth is another important sensor parameter. Each band is defined in terms of a “central” wavelength and a bandwidth. When one says that the TM band 1 is 0.45–0.52 μ m, what does this really mean? In an ideal case, the system response should be 1 between 0.45 and 0.52 μ m and 0 for wavelengths outside this range. Of course, this is not practical. There are a number of ways in which effective bandwidths are expressed. Usually, bandwidth is defined as *the 50 percent value on either side of the peak value on the lower and upper cut-off*. However, for any practical filter, there is ‘ringing’ in the pass band, making it difficult to identify the peak value. Palmer (1984)

Table 2. Comparison of the Palmer moment and full-width, half maximum (FWHM) methods.

TM Band	Moment		FWHM	
	λ_c	$\Delta\lambda$	λ_c	$\Delta\lambda$
1	486	71.5	485	66.1
2	571	88.7	569	80.6
3	660	77.1	659	68.7
4	839	134.9	841	129.1
5	1678	227	1676	216.9
7	2217	269	2272	250.2

suggested a technique called ‘moments’ to compute λ_c and $\Delta\lambda$. Table 2 taken from Palmer’s paper shows the comparison between the two methods.

While λ_c almost matches in both cases, there is about 8–10 percent difference in $\Delta\lambda$. Since the radiance received by the sensor is computed by multiplying the source radiance and the spectral bandwidth, the same instrument can give a different transfer function (radiance to count), depending on which way $\Delta\lambda$ is calculated. Therefore, there is a need to use a universally accepted methodology. Palmer’s moments method is more apt since this is independent of the ringing effect.

On another issue, what about the response below and above the 50 percent points? How do these responses affect radiometric accuracy? This is very important, especially when the spectral resolution is high, as in the case of ocean color sensors. However, there are practical limitations in the reduction of the out-of-band response that depend on the filter fabrication technology and the filter design. Here, the out-of-band contribution is referred to as the total system, which is a convolution of the response of filter, detector, and optics.

4 TEMPORAL RESOLUTION

For LANDSAT 1-3, any part of the globe (except around the poles) was imaged every 18 days. The images so taken have the same instrument view-angle for any location, which is important so that bi-directional reflectance distribution function (BRDF) differences do not influence the data. With the launch of SPOT, the term “revisit capability” was added. This novel concept is an excellent idea to image a particular place at shorter intervals than the temporal resolution of 26 days by across-track tilting of the sensor. The revisit capability should not be misconstrued to temporal resolution (also called “repeat coverage” or “repeat cycle”). The revisit of a location is carried out at the cost of not acquiring data over some other locations, and at viewing angles off-nadir. Can one ever have a coverage of the whole globe within a specified time duration, as was possible with Landsat, if the revisit capability is exercised?

It is suggested that sensor manufacturers give both repeat and revisit periods. If sensor pointing cannot be changed, then repeat and revisit periods will be same.

5 DATA COMPRESSION

Especially with high resolution imaging, the generated data rate is quite high and has to be compressed so that

the transmitted data are compatible with transmission technology. The effect of compression influences end use of the data. One should not compare the panchromatic data compression performance with that of multispectral data compression. In high resolution panchromatic imagery, one is generally trying to identify an object by shape, so a loss of data at random does not affect very much the end application, except perhaps the accuracy of stereo matching. On the other hand, when one is using multispectral data for classification, the affect can be quite different. Suppose there is 5 percent error in one band due to compression. Since the error is random, it need not be the same for a given pixel in all bands. Thus for a four band multispectral analysis, the accuracy is $(0.95)^4 = 0.81$. Thus, this will be the dominant error in classification. It may be more prudent to have a lower number of bits per pixel consistent with radiometric quality so that the compression error is less. Therefore, both the optimum number of bits per pixel and the data compression technique are important for radiometric quality of the image.

Table 3. Suggested mandatory parameters to be specified by the sensor manufacturer (for Earth observing sensors).

Parameter	Metric
1. Instantaneous Geometric Field of View (IGFOV)	IGFOV _{AL} IGFOV _{AC}
2. Radiometrically Accurate IFOV (RAIFOV)	RAIFOV _{AL} RAIFOV _{AC}
3. Ground Sample Distance (GSD)	GSD _{AL} GSD _{AC}
4. Swath (TFOV)	mmrad
5. SWR at IGFOV	SWR _{IGFOV AL} SWR _{IGFOV AC}
6. SWR at GSD	SWR _{GSD AL} SWR _{GSD AC}
7. Band to band registration	
8. Central wavelength	λ_c
9. Band width	$\Delta\lambda$
10. Percent of out of band response	
11. Saturation radiance	mw/cm ² /Strad/ μ
12. Quantization	bits
13. S/N at	R ₀ , R _L
14. Radiometric accuracy	(a) Absolute (b) Relative RMS error*
15. Data compression	
16. Polarization sensitivity	
17. Revisit period	(a) 0° Latitude (b) 40° Latitude (° off-nadir?)
18. Repeat period	

AL – Along-track, AC – Cross-track.

* The compression error depends on the scene, so one may refer to scene qualitatively as urban, farm land, etc.

Sensor designers should, among other things, specify the parameters given in Table 3. The specifications mentioned in the table are those measured at the sensor level; but the user is also concerned with product specifications like location accuracy, distortions, and post-launch calibration techniques.

In context of increasing relevance and dependence on data fusion from several sensors, the calibration community should identify a set of sensor parameters that, if standardized globally, would ensure maximal retrieval of information. The intrinsic capabilities of the sensors should be transparent so users can make their choice depending on their needs. The next step should be to formulate standard procedures for measurement of each of these parameters.

- NASA Special Publication #335 1973. Advanced scanners and imaging systems for Earth observation. *Working Group Report*. Greenbelt: National Aeronautics and Space Administration.
- Norwood, V.T. 1974. Balance between resolution and signal-to-noise ratio in scanner design for Earth resources systems. In D.A. Koso, P.N. Slater & W.L. Wolfe (eds), *Scanners and Imagery Systems for Earth Observation*. Proc. SPIE 51: 37–42.
- Palmer, J.M. 1984. Effective bandwidths for LANDSAT-4 and LANDSAT-D Multispectral Scanner and Thematic Mapper subsystems. *IEEE Trans. Geosci. & Rem. Sens.* 22(3): 336–338.

Metrology for remote sensing radiometry

B.C. Johnson, S.W. Brown & J.P. Rice

Optical Technology Division, National Institute of Standards and Technology, Maryland, USA

ABSTRACT: Metrology, the science of measurement, is discussed in terms of measurements of radiant flux and associated quantities. Radiometric measurements are naturally *remote*; that is, they are non-contact. One broad application lies in Earth science, where global measurements using satellites are spatially and temporally efficient, and sufficiently long term, to provide information on climate and global warming. Accurate radiometry requires a thorough understanding of the measurement problem, a complete description and understanding of the instruments, and mechanisms for comparing and assessing results. Here, an introduction to useful terminology and review some of the National Institute of Standards and Technology (NIST) activities are given in support of remote sensing.

1 INTRODUCTION

Measurements of radiant flux from the Earth, Sun, Moon, or planets using optical instruments deployed remotely are used to determine a wide range of physical parameters. For the Earth-Sun system, a critical long-term goal is to acquire an accurate and detailed understanding of processes affecting global climate change, in particular the role of human activities. From radiometric data, it is possible to study and quantify the Earth's temperature distribution, energy budget, properties and dynamics of the atmosphere and oceans, land use, carbon cycle, and other systems that relate to climate. The accuracy of these data depends on the radiometric calibration and characterization of the instruments involved. Recognizing the stringent requirements for radiometric stability and accuracy, and the difficulty of radiometric measurements, verification activities involving experts in metrology are needed.

This paper is a brief description of radiometric metrology with emphasis on remote sensing for the spectral interval from the ultraviolet (UV) to the thermal infrared (TIR). First, some background is provided on metrology, followed by terminology and definitions. These are followed by examples relating to characterization, calibration, traceability, and validation.

2 NATIONAL METROLOGY INSTITUTES

The International System of Units (SI) is the metric system of measurement for scientific, technical, and

commercial endeavors. Adopted internationally, the SI defines base and derived units, and an associated system of physical quantities (BIPM 1998). National metrology institutes (NMIs), with an eye towards public safety, science, technology, and the global marketplace, realize these quantities, perform research to improve their experimental and theoretical foundations, establish methods to disseminate their values to users, participate in international comparisons and related activities, and carry out specialized activities such as measurement assurance programs. In summary, NMIs represent the expertise in metrology. International interactions among these organizations (such as through the Treaty of the Meter) help ensure the robustness of global measurement science. This is necessary for international activities in science such as climate change monitoring.

The NMI for the USA is the National Institute of Standards and Technology (NIST), a non-regulatory agency in the Department of Commerce's Technology Administration. The measurement and standards laboratories within NIST interact with users through various established mechanisms, such as: 1) calibration services; 2) standard reference materials; 3) reference information; 4) databases; 5) special publications; 6) training courses; 7) workshops and conferences; 8) cooperative research and development programs; and 9) special tests that are typically an extension of an established calibration service. In addition, collaborative efforts between various USA federal agencies and NIST are undertaken in areas of national interest or broad scope, such as defense, health, or global climate change. Similar NMI functions and activities exist in other countries.

Several agencies involved with remote sensing for environmental applications have started or expanded radiometric collaborative efforts with the Optical Technology Division (OTD) in the Physics Laboratory at NIST over the past decades. The advantage to these programs includes independent validation of their methods by direct measurement and access to metrology programs that are permanent. NIST benefits because the often difficult measurement problems faced by users lead to incorporation of improved technologies, which are then available for all to use, as well as increased sophistication in the NIST expertise.

The spectral interval discussed in this paper is from about 0.3 μm to 20 μm . In terms of radiant flux from the Earth, it covers much of the reflected solar and a portion of the emitted thermal radiation. This spectral region overlaps with the measurement programs in the OTD. The needs of the remote sensing community in other regions of the electromagnetic spectrum are also served at NIST, but not discussed here. For example, the NIST Synchrotron Ultraviolet Radiation Facility (SURF) is used to calibrate sensors for solar physics (Woods et al. 1993). In the microwave region, NIST is developing a calibration/validation program with the National Polar-orbiting Operational Environmental Satellite System (NPOESS).

3 METROLOGY OF RADIOMETRY

A brief introduction to this topic is presented here. For additional descriptions of the NIST activities in remote sensing, see [Rice & Johnson \(2001\)](#) and [Brown & Johnson \(2003\)](#).

3.1 General terminology

It is necessary to begin with terminology. A frequently used phrase is “NIST-traceable.” But what does this mean? The definition recommended by NIST and other NMIs is from the *International Vocabulary of Basic and General Terms in Metrology*, known as the VIM (ISO 1993a). Traceability is defined as the *property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties*. It is NIST policy that for claims of traceability of measurement results, or values of standards, the provider is responsible for documenting support of the claim, and the user is responsible for assessing the validity of the provider’s claim. These *verification steps* may not involve NIST; but, if they do, the procedures can be selected from existing tools or made specific (e.g. a joint radiometric calibration program).

Proper assessment and statement of uncertainties is required for the measurements used to establish claims of traceability. Evaluation and expression of uncertainty is often an area of confusion, and it is not unusual to have incomplete information on this topic stated by the provider, or reported by the user. Fortunately, again one can turn to international documents, in this case the *Guide to the Expression of Uncertainty in Measurement*, known as the GUM (ISO 1993b). Other useful documents are available from NCSL International (NCSL 1997 & 1995) and NIST (Taylor & Kuyatt 1994).

The VIM, GUM, and related documents provide clear definitions and recommendations of usage for common metrological and statistical terms typically encountered when describing the results of measurements; examples include error, uncertainty, accuracy, and repeatability. For example, the term “combined standard uncertainty” refers to the combination of the uncertainty components with consideration of their interdependence, or correlation. Uncertainty components can be distinguished as Type A or Type B. Type A uncertainties are evaluated using statistical methods and Type B uncertainties are evaluated using models or other external information. The term “standard uncertainty” refers to an estimated standard deviation. Type B components of standard uncertainty are evaluated to be equivalent to one standard deviation. The expanded uncertainty is the product of the combined standard uncertainty and the coverage factor k , where the value of k is chosen based on the desired level of confidence (typically $k = 2$). In reporting the uncertainty for a measurement, the components of standard uncertainty should be listed and their designation stated (A or B). It should be clear if the final result corresponds to a combined standard or expanded uncertainty, and, if the latter, the value of the coverage factor.

For more information on terminology, including how to acquire the reference documents, see the NIST Technology Services Division website (<http://ts.nist.gov/>).

3.2 Radiometric quantities

3.2.1 Radiant energy quantities

Three radiant energy quantities are discussed that are significant to remote sensing radiometry, but other quantities exist. For a more complete discussion, see [Hengstberger \(1989\)](#), [DeCusatis \(1998\)](#), or [McCluney \(1994\)](#). The distinguishing features for the quantities are with respect to the geometry of the measurement situation. The first, radiant flux or radiant power, is simply the *total amount of power in a defined optical beam, measured in watts*. When measuring optical power with a simple radiometer consisting of an aperture in front of a detector, the optical beam must be

well defined, such as from a laser, and it must underfill the aperture.

More generally, the optical beam of radiant flux at the detector is constrained by the geometry of the source and the radiometer, and this leads to the second and third radiant energy quantities – irradiance and radiance. Irradiance is *radiant flux per unit area*, with the reference area located with respect to the radiometer. For the simple radiometer described above, let radiant flux, such as from a distant point source, overfill the aperture. If the spatial distribution of the irradiance is uniform across the aperture (for this case it should be), then the irradiance from the source at the location of the aperture is the radiant flux measured by the detector divided by the area of the aperture. An example from remote sensing is determination of the solar irradiance at the top of the Earth's atmosphere, e.g. Lawrence et al. 2003 & Frohlich et al. 1995.

Radiance is *radiant flux per unit area per unit solid angle*. Fundamentally, radiance is a differential optical beam concept, and the direction of the beam is part of the definition. For radiance, the area of interest is the projected source area in the viewing direction. The solid angle (unit steradian) is an example of a SI dimensionless derived unit. Typically in radiometry, a pair of apertures separated by some distance defines the solid angle. If both apertures are filled by the source, radiance is measured, since both the cross-sectional area and angular extent of the beam have been defined. An example from remote sensing is the radiance of the Earth in the TIR, from which the thermodynamic temperature can be determined using Planck's Law, if the emittance is known, e.g. Moeller et al. 2003, Minnett et al. 2001.

3.2.2 Material properties

Quantities describing the optical properties of materials are also of interest to remote sensing radiometry. Examples are reflectance, transmittance, and emittance. Quantification of material properties is necessary to realize of radiometric quantities. In addition, prior to final sensor assembly, it is often important to determine the optical performance of critical components such as filters, mirrors, beam-splitters, or lenses.

The elementary reflectance quantity is the bi-directional reflectance distribution function (BRDF), which is defined as *the ratio of radiance leaving a target to the irradiance incident on the target* (Barnes et al. 1998). Samples subjected to BRDF studies are typically “diffuse,” that is, flux incident from one direction is scattered into the entire hemisphere. The reflectance of mirrors, which are highly specular, does not have to be measured for all exitant angles in the full hemisphere at each incident angle, although in some critical applications this is necessary. In general, reflectance depends on polarization and wavelength.

The transmittance of a sample is the ratio of the incident to transmitted flux. Scattering within the sample and sensitivity to temperature, beam divergence, humidity, and air pressure must be considered in addition to the factors mentioned above, especially for optical elements coated with thin film multilayers.

Finally, the emittance of a sample is *the ratio of its radiance to the radiance emitted from an ideal black-body at the same thermodynamic temperature of the sample*.

3.3 Methods of traceability

The realization of values for radiant quantities is usually based on either electrical and dimensional metrology, or on temperature metrology. The former utilizes electrical substitution radiometry (ESR) and the latter, blackbody physics.

3.3.1 ESR

The basic principle of ESR is that an optical watt is the same as an electrical watt (Hengstberger 1989). Optical power is compared to electrical power in an absorbing receiver by direct substitution using a shutter and external electrical heating of the receiver. Since the scale is realized on an optical detector, this is often termed detector-based radiometry.

One ESR method utilizes intensity-stabilized lasers and cryogenic receivers to determine values of radiant flux very accurately. A scale of radiant flux responsivity is then realized on transportable, stable radiometers by direct substitution; the units are amps per watt (Gentile et al. 1996). Here, the radiometers are as described in §3.2.1 – an aperture in front of a detector. However, use of a single detector element has several disadvantages when one recognizes the primary objective is to realize values for other radiant energy quantities. Instead, for the spectral range corresponding to silicon photodiodes, multiple photodiodes are arranged in a light-trapping configuration. This is termed a “trap” detector.

The disadvantages of the single detector are related to the “like-to-like” rule in radiometry: to avoid introducing bias in the measurement result from systematic effects, the measurement conditions in the traceability “chain of comparisons” should be as similar as possible. However, this is not always achievable. For example, for general detector spectral responsivity realizations, radiometers are calibrated with laser beams and then used as the reference standards in a system where the source is the exit slit of a lamp-illuminated monochromator (Larason et al. 1998). Hence the radiometer is used with sources that differ in terms of their coherence, polarization, degree of monochromaticity, $f/\#$, flux level, and spatial uniformity. Compared to a single photodiode, the trap detector is less sensitive to these source parameters (Eppeldauer & Lynch 2000).

3.3.2 *Blackbody physics*

The basic principle here is Planck's Radiation Law, which predicts the spectral radiance from an ideal blackbody in terms of its thermodynamic temperature and the measurement wavelength. Since the scale is realized with an optical source, this is often termed source-based radiometry. In "fixed-point" blackbodies, the temperature is set by surrounding the blackbody cavity by a pure molten metal that is undergoing a freeze. With the temperature at a constant value during this phase transition, the spectral radiance is calculable. Values for the relevant freezing temperatures of the metals are given in The International Temperature Scale of 1990 (ITS-90) (Preston-Thomas 1990).

Fixed-point blackbodies are expensive, tedious to operate, and exhibit substantial temperature gaps in their coverage. Instead, variable temperature blackbody standards are constructed with various materials and instrumented with calibrated contact thermometers, such as platinum resistance thermometers (PRTs). Thus the traceability of spectral radiance values is established using temperature standards that are calibrated according to ITS-90. Of course, this procedure will not work at higher temperatures where contact thermometers cannot function.

If a cavity is black (unity emittance) and its temperature is uniform, the uncertainty of the radiance is determined only by the uncertainty of the temperature and associated fundamental constants. Real blackbodies are only an approximation of this ideal, and therefore they must be characterized to determine the emittance, spatial uniformity, stability, and so forth. For large-area blackbodies typically used in the TIR to calibrate Earth-observing sensors, it is difficult to achieve unity emittance. If the background environment is not cooled to low temperatures, the reflection of the background radiance from the blackbody must be considered.

3.4 *Scale realization at NIST*

The previous section gave examples of detector-based radiometry for realizing radiant-flux and spectral-flux responsivity, and source-based radiometry for realizing spectral radiance. Using these concepts, various other methods at NIST are used to realize spectral irradiance and radiance. For both irradiance and radiance, limiting apertures play a critical role. At NIST, these areas are determined using optical metrology in a dedicated facility (Fowler & Litorja 2003).

3.4.1 *Radiance*

The Low Background Infrared (LBIR) facility at NIST has a vacuum chamber with 20 K cryoshrouds. The radiance temperature of blackbodies is determined using a cryogenic ESR (Carter et al. 2003). Radiant flux from the aperture of the blackbody overfills the aperture on the radiometer. The area of the apertures

and the distance between them is measured. The radiometer is designed to absorb equally at all wavelengths, so the measured flux is proportional to the blackbody exitance (Stefan-Boltzmann Law). Consideration of the geometric factors allows determination of the radiance temperature. For small blackbody apertures, diffraction effects must be considered. From the LBIR measurements, it is possible to assess the accuracy of the user's source-based radiance scale that is based on the temperature standards.

In the Medium Background Infrared (MBIR) and Ambient Background Infrared (ABIR) facilities at NIST, large area (10 cm diameter apertures) blackbodies serve as standard sources of spectral radiance. The cryoshrouds in the MBIR vacuum chamber are operated at 80 K, or room temperature (ambient). With 80 K cryoshrouds, the MBIR cryogenic blackbody operates over the temperature range 200 K to 350 K (Fowler et al. 1998). In the ABIR facility, which is a standard laboratory environment, water- and oil-bath based blackbodies cover the temperature range 288 K to 460 K. The blackbody temperatures are traceable to ITS-90 by use of calibrated PRTs.

In the Facility for Automated Spectroradiometric Calibrations (FASCAL), a gold-point blackbody (1337.33 K) is the source standard for spectral radiance (Walker et al. 1987). A prism-grating spectroradiometer is used to measure spectral radiance ratios of the gold-point blackbody at 654.6 nm to higher temperature sources (tungsten filament lamps, or high temperature graphite blackbodies). The spectral radiance values of lamp-illuminated integrating sphere sources are then determined by comparison to the high temperature blackbody. Calibrated integrating sphere sources are used as spectral radiance standards in the remote sensing community.

The Spectral Irradiance and Radiance Responsivity Calibrations with Uniform Sources (SIRCUS) facility is detector-based and designed for determination of absolute spectral irradiance or radiance responsivity of complete radiometers (Brown et al. 2000 & Eppeldauer et al. 2000). The output of tunable lasers is input into an integrating sphere. The irradiance of the laser-illuminated sphere is determined using a standard radiometer such as the aperture/trap detector radiometer described in §3.3.1. As in the LBIR facility, the sphere is fitted with a known aperture, and the distance between the source and detector apertures is measured. This allows determination of the radiance of the laser-illuminated sphere.

Integrating spheres with large exit aperture diameters that overfill the field-of-view of the radiometer under test are used for radiance responsivity, and smaller ones that underfill the field-of-view of the radiometer under test are used for irradiance responsivity. Currently, the laser systems in SIRCUS are continuously tunable from 350 nm to about 5 μm , with

discrete output at CO and CO₂ laser wavelengths (5 μm to 7 μm and 10 μm to 11 μm, respectively). Several methods are under development for operation to 20 μm. For the longer wavelengths, the reference radiometers utilize Indium Gallium Arsenide (InGaAs) photodiodes, or an electrically substituted bolometer. In addition, the interior of the sphere is coated with diffuse gold, instead of the polytetrafluoroethylene (PTFE)-based material used in the visible.

3.4.2 Irradiance

NIST has recently modified the procedure for realizing spectral irradiance. In the past, spectral irradiance values were traceable to the gold-point blackbody spectral radiance at 654.6 nm using multiple comparison steps. Spectral radiance values were assigned to a small integrating sphere source as described in §3.4.1. Then, the prism-grating spectroradiometer was configured with irradiance fore-optics (a small integrating sphere with a circular input aperture) and, using the integrating sphere source as a standard, the spectral irradiance responsivity of the spectroradiometer was determined. Using the calibrated spectroradiometer, spectral irradiance values were assigned to 1000 W quartz-tungsten standard irradiance lamps, type FEL (not an acronym).

In 2000, a new NIST irradiance scale was introduced (Yoon et al. 2000). It is detector-based, not source-based. The spectral irradiance values of the NIST FEL working standard lamps are determined by comparison to a high temperature (3000 K) blackbody using a spectroradiometer. The temperature of the blackbody is determined using filter radiometers calibrated for absolute spectral irradiance responsivity by separate determination of the aperture area and the filter/detector spectral flux responsivity. The great benefit is a factor of between 2 and 10 times reduction in uncertainty for the assigned spectral irradiance values, with the greatest improvement in the short-wave infrared (SWIR).

It is possible to realize spectral radiance using the FEL spectral irradiance standards to illuminate a diffuse reflectance standard. If the BRDF of this target is known (see §3.2.2), then the spectral radiance is calculable (Mueller & Austin 2003). Therefore, there are two methods to establish traceability to NIST spectral radiance standards – via a calibrated integrating sphere source, or by use of an irradiance and diffuse reflectance standard. In an ideal world, one would utilize both methods and then intercompare the results routinely. In the practical world, the choice between the two methods should involve discussion of issues such as size limitations, operating environment, planned frequency of calibration schedule, methods in place to monitor the stability in spectral radiance, and sensitivity to uniformity, scattered flux, and so forth.

4 CHARACTERIZATION

The authors have described how traceability of values can be established for detector responsivity, irradiance, and radiance. Although not described here, determination of reflectance, transmittance, or emittance values also utilize sources and radiometers, e.g. Barnes et al. 1998 & Hanssen 2001. One can see how, as measurements progress through the “chain of comparisons” required for traceability, it is possible to violate the “like-to-like” rule. For proper assessment of uncertainties, the sources and radiometers must be characterized.

4.1 Measurement equation

A measurement equation is a mathematical expression that describes the role and effect of all of the influencing parameters, the wavelength, polarization, direction, beam size, and flux level of the radiant quantity (Nicodemus et al. 1976). The responsivity of a radiometer depends on, and the specifications of a source involve, these parameters. Time is also an essential variable, as it relates to both the measurement precision (short term stability) and drift or degradation (long term stability). Temperature of the environment, intervening medium, or instrument components can be a significant influencing parameter.

One usually starts by assuming that the parameters can be treated independently, which allows for separate characterization measurements for each parameter. This assumption should either be confirmed, or the experiments designed to minimize possible systematic effects. An example of testing would be to perform the measurements verifying that linearity of the radiometer output with flux level is independent of integration time. An example of reducing sensitivity to systematic effects would be to fill the entire entrance pupil of the radiometer with monochromatic flux during spectral responsivity characterizations, taking care to match the $f/\#$ of the beam to final measurement configuration. For cases where correlations must be considered, see Gardner (2000) & Gardner (2003).

4.2 Spectral responsivities

Determination of radiometer spectral responsivity is critical for the correct interpretation of measurements of natural sources. For filter radiometers or spectroradiometers at a particular wavelength setting, both the in-band and out-of-band spectral responsivities are of interest. The measurement equation with all of the other influencing parameters neglected (or at some reference values) is:

$$i = \int L(\lambda) R(\lambda) d\lambda \quad (1)$$

where i is the measured output current, $L(\lambda)$ is the source spectral radiance, λ is the wavelength, and $R(\lambda)$ is the absolute radiance responsivity at wavelength λ . Note that Equation 1 could have been written in terms of source irradiance and absolute irradiance responsivity.

One way of determining the function $R(\lambda)$ is in two steps – by measuring the relative spectral responsivity and then performing a calibration measurement using a source standard. The relative spectral responsivity values are typically determined using dispersive or interferometric systems that are illuminated with continuum sources. The determination of the relative spectral responsivity is sometimes done at the component level, with the relative spectral responsivity for the entire sensor calculated as the product of the individual values for transmittance, reflectance, and related measures. This can introduce bias from systematic effects such as interreflections and scatter that are present in the complete system, but not for the individual components. Modeling of such effects is possible and provides useful insight, but such estimations of bias may not have the required uncertainty.

A second way to determine $R(\lambda)$ is to use the method employed in the SIRCUS facility. The absolute responsivity is determined in one step. Compared to methods that use filtering of continuum sources, it is straightforward to achieve the correct beam geometry, flux level, and bandwidth. The output of SIRCUS is continuously tunable and has a wide dynamic range in flux. In the TIR, the laser can be chopped as a means to eliminate thermal background radiation.

4.2.1 In-band example

Often it is convenient to refer to a measurement channel of a filter radiometer in terms of its measurement wavelength, λ_m , and bandwidth, $\Delta\lambda$:

$$\lambda_m = \frac{\int \lambda R(\lambda) d\lambda}{\int R(\lambda) d\lambda} \quad (2)$$

$$(\Delta\lambda)^2 = 2.354 \frac{\int (\lambda - \lambda_m)^2 R(\lambda) d\lambda}{\int R(\lambda) d\lambda} \quad (3)$$

In general, the limits of integration in Equations 1–3 correspond to the entire region of finite detector responsivity. The in-band profile is defined by restricting the region around λ_m , with the location of the upper and lower wavelength defined in various ways, e.g. where the responsivity is a certain percentage of the maximum responsivity.

Measurements to determine the in-band profile can be incorrect through improper assignment or understanding of the wavelength scale, inadequate spectral

resolution or bandwidth, or mistreatment of interreflections, scattered light, and thermal effects. Although the example of a filter radiometer is used for Equations 2 and 3, similar issues exist in spectroradiometers.

The in-band example concerns the Robotic Lunar Observatory (ROLO), which is a pair of Ritchey-Chrétien telescopes, each fitted with a rotating filter wheel and an array detector (Anderson et al. 1999). Relative spectral responsivities were estimated from the transmittance traces provided by the filter vendor. However, the $f/\#$ for these measurements was not the same as for the illumination conditions in the ROLO telescopes. For interference filters, the wavelength scale depends on $f/\#$ because of the sensitivity to the angle of incidence, and errors of this type impact the ROLO atmospheric correction algorithm as well as the values determined for the lunar radiances.

The solution involved construction of a collimator-source assembly with beam properties designed to simulate lunar illumination (Stone et al. 2003). A portable version of SIRCUS, “Traveling SIRCUS,” was deployed to ROLO outside Flagstaff, Arizona. Using a tunable Ti:sapphire laser configured for intracavity doubling that was fiber-optic coupled to the collimator, the relative in-band profiles for eight channels were determined. Shifts of up to 2.5 nm were observed.

4.2.2 Out-of-band example

As seen from space in the visible and near-infrared (VNIR), the total radiance when viewing the sunlit oceans is a function of the atmospheric contribution, reflection at the ocean surface, and flux scattered from the oceans. The latter component, the water-leaving radiance, depends on the optical properties of the oceans, which are a function of bio-physical parameters, e.g. chlorophyll concentration. The water-leaving radiance is a small component of the total radiance, and to achieve the desired accuracies for ocean products continuous *in-situ* measurements from optical sensors such as the in-water marine optical buoy (MOBY) are used to revise the satellite calibration coefficients (Gordon 1998 & Barnes et al. 2000).

Typically, lamp-based sources are used for the pre-flight satellite sensor and the *in-situ* radiometer calibration. Lamp-based sources produce an overabundance of red and near-infrared flux compared to natural ocean scenes. This is another example of unavoidable violation of the “like-to-like” rule, and results in bias if the sensors have significant spectral out-of-band response that is not properly accounted.

In the MOBY instrument, two charged-couple device (CCD) spectrographs are used to cover the interval from 360 nm to 940 nm. In the overlap region of the two spectrographs (560 nm to 640 nm), the derived radiant values did not agree. However, a simplified version of Equation 1 was in use, which

assumed that the sensor output at a measurement wavelength was proportional to the product of spectral responsivity and spectral radiance at that wavelength. As Equation 1 demonstrates, the relative effect of spectral out-of-band response on the total output depends on the spectral shape of the source. Since the sources measured are quite different, the effect can be significant. It is most apparent in spectral regions where the system responsivity is small, e.g. where a dichroic beamsplitter is limiting the flux, as the fraction of measured signal from the spectral out-of-band effect is significant. In MOBY, a dichroic beamsplitter directs the incoming flux to the spectrographs, and spectral out-of-band (or “stray light”) was suspected as the cause of the discrepancy in the region of overlap.

The size and deployment schedule of MOBY made it impossible to perform characterizations at NIST, so Traveling SIRCUS was deployed to the MOBY Operations Facility in Honolulu, Hawaii. A dye laser and several discrete lasers were included in addition to the Ti:sapphire laser. By fine-tuning the Traveling SIRCUS wavelength, the relative in-band profile at selected regions on each CCD for both spectrographs was determined.

The features observed in the images of monochromatic illumination were identified – they arose from the image of the slit, haze and diffuse reflections, and interreflections of higher-order diffraction. Using the laser characterization data along with the lamp-based calibration data, an iterative solution that results in “stray-light corrected” values for the MOBY spectral responsivities, as well as the measured up-welling and water-leaving radiances was implemented. As anticipated, the magnitude of the discrepancy in the region of overlap was reduced with application of the stray-light correction. In the spectral region between 410 nm and 550 nm, the values for water-leaving radiance reported by MOBY were revised by several percent, with concomitant improvements in the uncertainties for water-leaving radiance (Brown et al. 2003a).

5 CALIBRATION & VALIDATION

The results of instrument characterization measurements are methods to describe sensor performance for all operating conditions, e.g. a system of equations and associated parameters, lookup data tables, results of thermal models, and uncertainty budgets. Calibration of a sensor involves measurement of a known source, with the roles reversed for calibration of a source. As a result, radiometric quantities are known in terms of measured quantities. In the case of a *primary* device, such as an ESR or blackbody, a calibration measurement is not performed. In either case, the uncertainty

is based on the measurement equation and the characterization experiments.

Validation is the process of verifying radiometric performance by using an independent method. With values traceable to SI quantities, as is recommended for remote sensing (CCPR 1995), the results of inter-comparison measurements should be consistent within the combined uncertainty because all values are traceable to the same system of measurement. Validation of an instrument can be performed by comparison of detector- and source-based calibration approaches. Primary instruments are intercompared, or used in specialized experiments that determine fundamental constants radiometrically. Here are mentioned results of validation programs with the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the Departments of Energy and Defense (DOE and DOD).

5.1 VNIR

The Earth Observing System (EOS) Project at NASA established a calibration/validation program with NIST in the mid-1990s. One part of the program involves validation of the spectral radiance of sources using NIST-designed, built, and deployed transfer radiometers. The sources are used to assign preflight calibration coefficients for EOS sensors, and they consist of lamp-illuminated integrating sphere sources (VNIR) or blackbodies (TIR).

The EOS/NIST Visible Transfer Radiometer (VXR) (Johnson et al. 2003) and the Shortwave Infrared Transfer Radiometer (SWIXR) (Brown et al. 1998) are filter- and grating-based instruments, respectively, that cover the spectral interval from 412 nm to 2.5 μm . They have been used in conjunction with transfer radiometers from other institutions to independently determine spectral radiance values of radiance standards. Because of resource constraints, it is generally not possible to perform validation tests during the calibration of flight hardware. Therefore, the measurement protocol is for the hosting institution, immediately prior to the comparison measurements, to assign spectral radiance values to the source using the same methods of establishing traceability that were used for the calibration of the EOS sensor. The VNIR results to date confirm that the uncertainties in spectral radiance assigned by the manufacturer (± 3 percent, $k = 1$) are valid.

As for radiometers, their calibration can be validated by using different methods and comparing the results. For example, the VXR was calibrated using SIRCUS and FASCAL and the results agreed within the combined uncertainties (Johnson et al. 2003). In the NOAA/MOBY project, a portable version of the optical system used in the buoy was characterized and

calibrated using the SIRCUS facility at NIST, and also used to measure a lamp-illuminated radiance source that was calibrated on FASCAL. The stray-light corrected system responsivities derived from the source agreed with those from SIRCUS to within 2 percent (Brown et al. 2003b).

5.2 TIR

The NIST/EOS Thermal-infrared Transfer Radiometer (TXR) is a two-channel portable filter radiometer that operates in vacuum or ambient conditions (Rice & Johnson 1998). The detectors, filters, and reflective optics are built into a liquid-nitrogen cryostat with a ZnSe window. The two TXR channels are at 5 μm (photovoltaic InSb detector) and 10 μm (photovoltaic mercury cadmium telluride detector). A variable-temperature, vacuum compatible blackbody that can rotate in front of the TXR window under computer control is used to monitor the stability of the TXR radiance responsivity during deployments.

The motivation for the TXR was as described above for the VNIR spectral regions: to validate the spectral radiance scale of standard sources using a portable, NIST-calibrated radiometer. There are differences worth noting, however. For the VNIR, the radiance values are traceable to NIST standards using a chain of measurement steps that involve radiometric artifacts – typically spectral irradiance standard lamps and diffuse reflectance standards. For the blackbodies used in thermal/vacuum chambers to calibrate infrared sensors, the spectral radiance values are traceable to NIST temperature standards. As pointed out in §3.3.2, there are uncertainties associated with the emittance of the blackbody source, and these are often difficult to evaluate. Finally, in the TIR spectral region, all “warm” structures ($> 100\text{ K}$) are an additional source of radiant flux. Reflection of this background flux from a non-ideal blackbody results in a calibration that is dependent on the temperature of the surrounding environment.

Calibration of the TXR has been performed using multiple approaches. Initially, the NIST water-bath blackbody (Fowler 1995) was used (Rice & Johnson 1998). More recently, the TXR was calibrated in the MBIR chamber using its cryogenic blackbody. The blackbody was characterized and the emittance determined using the TXR with the MBIR thermal background at 300 K and 80 K. At the present time, the TXR uncertainty is $< 0.1\text{ K}$ at 300 K ($k = 2$). In the future, the TXR will be calibrated on SIRCUS and an ESR based on high- T_c superconducting temperature sensors will be used to determine the total irradiance from the MBIR blackbody, thus validating the calibration of the TXR and the blackbody using independent methods.

The TXR has been used on multiple occasions to verify TIR spectral radiance values. In July 1999 and August 2001, it was deployed to the remote sensing radiometric calibration facility at Los Alamos National Laboratory (LANL) in support of DOE programs. The LANL system has two chambers connected by a gate valve – one with a cold shroud at $\approx 80\text{ K}$ for the sensor, and one at room temperature for the standard blackbody, two cold blackbody sources, and a collimator for coupling radiant flux to the sensor. With the TXR located in the cold chamber at the location normally used by a remote sensor, the measurements validated the LANL blackbody-collimator system. The 1999 results at 5 μm agreed in spectral radiance to within 0.16 percent, or 50 mK at 300 K ($k = 2$) (Rice et al. 2003). At the time of the 1999 deployment, the 10 μm channel had offsets of up to 0.5 K with cryocycling. Measurements of a check-source were used to correct for this effect, and the resulting agreement with LANL was within 0.1 K (Rice et al. 2000).

In May 2001 the TXR was used in an EOS sea-surface temperature (SST) intercomparison at the University of Miami (UM) (Rice et al. 2004). Ship-based measurements of SST are used to validate satellite results using different types of infrared radiometers and field-deployable blackbody sources. Members of the international Committee on Earth Observing Satellites (CEOS) helped organize and support the 2001 comparison, which included ship-based measurements (Barton et al. 2004), and deployment of the TXR and the NIST water-bath blackbody source as part of the laboratory comparison. Five blackbody sources were measured using the TXR – the NIST source plus four sources used for field calibration of SST radiometers. The comparison was performed in a University of Miami (UM) laboratory at ambient temperature. All blackbodies agreed to within 0.1 degrees C near ambient; measurements at other temperatures provided information on the emittance of the SST sources (Rice et al. 2004).

In July 2001 the TXR was deployed to the Geostationary Operational Environmental Satellite (GOES) Imager radiometric calibration chamber at ITT in Ft. Wayne, Indiana. The chamber is not equipped with cryoshrouds. The GOES Imager is calibrated from measurements of the Earth Calibration Source (ECT) and a cold blackbody source. The ITT procedure involves correcting for temperature gradients in the ECT, which were believed to be driven by the thermal background of the chamber (Farthing 1993).

During the GOES TXR deployment, measurements were made at a range of temperatures of the ECT and the TXR check-standard blackbody, as well as the cold blackbody source. An analysis procedure was developed that enabled parameterization of the results in terms of a non-unity emittance and a temperature

gradient in the ECT. The results are in qualitative agreement with the existing GOES model, thus providing valuable confirmation to the project.

More recently, in the fall of 2003 the TXR was used to measure a blackbody source at Santa Barbara Remote Sensing (SBRS) in Goleta, California in support of DOD research. Plans are in place for TXR measurements of a second blackbody source at SBRS in early 2004. This source was used to calibrate the MODerate Resolution Imaging Spectroradiometer (MODIS) sensors on the EOS satellites Terra and Aqua. Plans call for use of this blackbody for the calibration of the Visible/Infrared Imager/Radiometer Suite (VIIRS) sensor that will be flown as part of NPOESS.

6 CONCLUSIONS

Metrology and radiometry are established fields and substantial reference material is available regarding standards, terminology, and protocols. International organizations and NMIs represent a valuable resource – they maintain and disseminate the SI, validate their own results through international comparisons, and engage with industry and the research community in specialized measurement assurance programs in broad efforts such as remote sensing.

Characterization and calibration of sources and radiometers, and validation of measurement results are all essential for reliable products such as that required for climate quality data. The expanded use of tunable laser facilities such as SIRCUS at commercial facilities appears possible and would benefit the remote sensing community. Calibration of radiometric artifacts should be in terms of the SI to produce long-term data sets on a common measurement system. Validation is the process of verifying a result with an independent approach, and it should be performed whenever a technically feasible method with appropriate uncertainties is available. Examples include using a common instrument with multiples of its radiometric counterpart, e.g. the TXR and several blackbody sources, using source- and detector-based calibration methods with the same instrument, and investigating ways to validate model results.

ACKNOWLEDGMENTS

The authors thank all of the members of the OTD who have participated in these programs and helped to ensure their success. Also, we thank all members of other agencies and the remote sensing community who have either funded these efforts or participated in the intercomparisons. Without this spirit of cooperation, programs such as those described here would not be possible.

REFERENCES

- Anderson, J.M., Becker, K.J., Kieffer, H.H. & Dodd, D.N. 1999. Real-time control of the robotic lunar observatory telescope. *Astron. Soc. Pac.* 111:737–749.
- Barnes, P.Y., Early, E.A. & Parr, A.C. 1998. *Spectral Reflectance. NIST Special Publication 250-48*. Gaithersburg: U.S. Dept. of Commerce.
- Barnes, R.A., Eplee, R.E., Robinson, W.D., Schmidt, G.M., Patt, F.S., Bailey, S.W., Wang, M. & McClain, C.R. 2000. The calibration of SeaWiFS on orbit. In W.L. Barnes (ed.) *Earth observing systems*. Proc. SPIE 4135: 281–293.
- Barton, I.J., Minnett, P.J., Maillet, K.A., Donlon, C.J., Hook, S.J., Jessup, A.T. & Nightingale, T.J. 2004. The Miami2001 infrared radiometer calibration and intercomparison: Part II. Ship-board results. *J. Atmos. & Oceanic Tech.* 21(2): 268–283.
- BIPM 1998. *The International System of Units (SI)* 7th ed. Sèvres: Bureau International des Poids et Mesures.
- Brown, S.W., Johnson, B.C. & Yoon, H.W. 1998. Description of a portable spectroradiometer to validate EOS radiance scales in the shortwave infrared. *The Earth Observer* 10(3): 43–48.
- Brown, S.W., Eppeldauer, G.P. & Lykke, K.R. 2000. NIST facility for spectral irradiance and radiance responsivity calibrations with uniform sources. *Metrologia* 37(5): 579–582.
- Brown, S.W. & Johnson, B.C. 2003. Advances in radiometry for ocean color. In W.L. Barnes (ed.), *Earth observing systems VIII*. Proc. SPIE 5151: 441–453.
- Brown, S.W., Johnson, B.C., Flora, S.J., Feinholz, M.E., Yarbrough, M.A., Barnes, R.A., Kim, Y.S., Lykke, K.R. & Clark, D.K. 2003a. Stray-light correction of the Marine Optical Buoy. In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Rev. 4, Vol. VI: Special Topics in Ocean Optics Protocols and Appendices. NASA/TM–2003-211621/Rev4-Vol. VI*: 87–124. Greenbelt: National Aeronautics and Space Administration.
- Brown, S.W., Johnson, B.C., Feinholz, M.E., Yarbrough, M.A., Flora, S.J., Lykke, K.R. & Clark, D.K. 2003b. Stray-light correction algorithm for spectrographs. *Metrologia* 40(1): S81–S84.
- Carter, A.C., Jung, T.M., Smith, A., Lorentz, S.R. & Datla, R. 2003. Improved broadband blackbody calibrations at NIST for low-background infrared applications. *Metrologia* 40(1): S1–S4.
- CCPR 1995. *Recommendation P1 (1994) of the Comité Consultatif de Photométrie et Radiométrie (CCPR) submitted to the Comité International des Poids et Mesures (CIPM)*. 13th CCPR Meeting Report. Sèvres: Bureau International des poids et Mesures.
- DeCusatis, C. (ed.) 1998. *Handbook of Applied Photometry*. New York: American Institute of Physics Press.
- Eppeldauer, G.P., Brown, S.W., Larason, T.C., Racz, M. & Lykke, K.R. 2000. Realization of a spectral radiance responsivity scale with a laser-based source and Si radiance meters. *Metrologia* 37(5): 531–534.
- Eppeldauer, G.P. & Lynch, D.C. 2000. Opto-mechanical and electronic design of a tunnel-trap Si radiometer. *J. Res. Natl. Inst. Stand. & Tech.* 105(6): 813–828.
- Farthing, W.H. 1993. *Imager S/N 03 Channel 2 Non-linearity*. Swales Aerospace Memorandum.

- Fowler, J. 1995. A third generation water bath based blackbody source. *J. Res. Natl. Inst. Stand. & Tech.* 100(5): 591–599.
- Fowler, J., Johnson, B.C., Rice, J.P. & Lorentz, S.R. 1998. The new cryogenic vacuum chamber and blackbody source for infrared calibrations at the NIST's FASCAL facility. *Metrologia* 35(4): 323–327.
- Fowler, J. & Litorja, M. 2003. Geometric area measurements of circular apertures for radiometry at NIST. *Metrologia* 40(1): S9–S12.
- Frohlich, C. et al. 1995. VIRGO: experiment for helioseismology and solar irradiance monitoring. *Solar Phys.* 162(1–2): 101–128.
- Gardner, J.L. 2000. Correlated colour temperature – uncertainty and estimation. *Metrologia* 37(5): 381–384.
- Gardner, J.L. 2003. Correlations in primary standards. *Metrologia* 40(1): S167–S171.
- Gentile, T.R., Houston, J.M. & Cromer, C.L. 1996. Realization of a scale of absolute spectral response using the National Institute of Standards and Technology high-accuracy cryogenic radiometer. *Appl. Optics* 35(22): 4392–4403.
- Gordon, H.R. 1998. In-orbit calibration strategy for ocean color sensors. *Rem. Sens. Env.* 63(3): 265–278.
- Hanssen, L.M. 2001. Integrating sphere method for absolute transmittance, reflectance, and absorbance of specular samples. *Appl. Optics* 40(19): 3196–3204.
- Hengstberger, F. (ed.) 1989. *Absolute Radiometry*. New York: Academic Press.
- ISO 1993a. *International Vocabulary of Basic and General Terms in Metrology* 2nd ed. Geneva: International Organization for Standardization.
- ISO 1993b. *Guide to the Expression of Uncertainty in Measurement*.
- Johnson, B.C., Brown, S.W., Eppeldauer, G.P. & Lykke, K.R. 2003. System-level calibration of a transfer radiometer used to validate EOS radiance scales. *Int. J. Rem. Sens.* 24(2): 339–356.
- Larason, T.C., Bruce, S.S. & Parr, A.C. 1998. *Spectroradiometric Detector Measurements: Part I – Ultraviolet Detectors and Part II – Visible to Near-infrared Detectors*. NIST Special Publication 250-41. Gaithersburg: U.S. Dept. of Commerce.
- Lawrence, G.M., Kopp, G., Rottman, G., Harder, J., Woods, T. & Loui, H. 2003. Calibration of the total irradiance monitor. *Metrologia* 40(1): S78–S80.
- McCluney, R. 1994. *Introduction to Radiometry and Photometry*. Boston: Artech House.
- Minnett, P.J., Knuteson, R.O., Best, F.A., Osborne, B.J., Hanafin, J.A. & Brown, O.B. 2001. The marine-atmospheric emitted radiance interferometer: A high-accuracy, seagoing infrared spectroradiometer. *J. Atmos. & Oceanic Tech.* 18(6): 994–1013.
- Moeller, C.C., Knuteson, R.O., Tobin, D., Revercomb, H.E. & Menzel, W.P. 2003. Assessment of Aqua MODIS and AIRS TIR band LIB radiances using ER-2 based observations during TX-2002. In W.L. Barnes (ed.), *Earth observing systems VIII*. Proc. SPIE 5151: 355–366.
- Mueller, J. & Austin, R. 2003. Characterization of oceanographic and atmospheric radiometers. In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Rev. 4, Vol. II: Instrument Specifications, Characterizations, and Calibration*. NASA/TM – 2003-211621/Rev4-Vol.II: 17–33. Greenbelt: National Aeronautics and Space Administration.
- NCSL 1995. *Handbook for the Interpretation and Application of ANSI/NCSL Z540-1-1994*. Boulder: National Conference of Standards Laboratories (NCSL) International.
- NCSL 1997. *U.S. Guide to the Expression of Uncertainty in Measurement*. ANSI/NCSL Z540-2-1997. Boulder: National Conference of Standards Laboratories (NCSL) International.
- Nicodemus, F.E., Kostkowski, H.J. & Hattenburg, A.T. 1976. Introduction. In *Self-Study Manual on Optical Radiation Measurements: Part I – Concepts, Chapters 1 to 3*. NBS Tech. Note 910-1: 1–9. Gaithersburg: U.S. Dept. of Commerce. Available electronically from <http://physics.nist.gov>.
- Preston-Thomas, H. 1990. The international temperature scale of 1990. *Metrologia* 27(1): 3–10 & 107.
- Rice, J.P. & Johnson, B.C. 1998. The NIST EOS thermal-infrared transfer radiometer. *Metrologia* 35(4): 505–509.
- Rice, J.P., Bender, S.C. & Atkins, W.H. 2000. Thermal-infrared scale verifications at 10 micrometers using the NIST EOS TXR. In W.L. Barnes (ed.), *Earth observing systems V*. Proc. SPIE 4135: 96–107.
- Rice, J.P. & Johnson, B.C. 2001. NIST activities in support of space-based radiometric remote sensing. In C. Londono (ed.), *Harnessing light: optical science and metrology at NIST*. Proc. SPIE 4450: 108–126.
- Rice, J.P., Bender, S.C., Atkins, W.H. & Lovas, F.J. 2003. Deployment test of the NIST EOS thermal-infrared transfer radiometer. *Int. J. Rem. Sens.* 24(2): 367–388.
- Rice, J.P., Butler, J.J., Johnson, B.C., Minnett, P.J., Maillet, K.A., Nightingale, T.J., Hook, S.J., Abtahi, A., Donlon, C.J. & Barton, I.J. 2004. The Miami 2001 infrared radiometer calibration and intercomparison: Part I. Laboratory characterization of blackbody sources. *J. Atmos. & Oceanic Tech.* 21(2): 258–267.
- Stone, T.C., Kieffer, H.H. & Becker, K.J. 2003. Modeling the radiance of the Moon for on-orbit calibration. In W.L. Barnes (ed.), *Earth observing systems VIII*. Proc. SPIE 5151: 463–470.
- Taylor, B.N. & Kuyatt, C.E. 1994. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST Tech. Note 1297. Gaithersburg: U.S. Dept. of Commerce.
- Walker, J.H., Saunders, R.D. & Hattenburg, A.T. 1987. *Spectral Radiance Calibrations*. NBS Special Publication 250-1. Gaithersburg: U.S. Dept. of Commerce.
- Woods, T.N., Ucker, G.J. & Rottman, G.J. 1993. Solar stellar irradiance comparison I: 2. instrument calibration. *J. Geophys. Res. (Atmospheres)* 98(D6): 10679–10694.
- Yoon, H.W., Gibson, C.E. & Barnes, P.Y. 2002. Realisation of the National Institute of Standards and Technology detector based spectral irradiance scale. *Appl. Opt.* 41(28): 5879–5890.

Stability and accuracy requirements for passive satellite remote sensing instrumentation for global climate change monitoring

R. Datla

NIST, Gaithersburg, Maryland, USA

W. Emery

University of Colorado, Boulder, Colorado, USA

G. Ohring

NOAA-NESDIS (consultant), Camp Springs, Maryland, USA

R. Spencer

NASA Marshall Space Flight Center, Huntsville, Alabama, USA

B. Wielicki

NASA Langley Research Center, Langley, Virginia, USA

ABSTRACT: In the remote sensing community, stability and accuracy are two critical terms that need to be used in a consistent and clear way for their use to be of maximum benefit. Their definitions in combination with the vocabulary of the International Organization for Standardization (ISO) Guide on uncertainty analysis have been applied to long time series measurements using satellite instruments for climate monitoring. Based on the definitions of these terms, requirements have been developed for satellite instrumentation at a calibration workshop in November 2002 organized by National Institute of Standards (NIST), National Oceanic and Atmospheric Administration (NOAA), National Polar-orbiting Operational Environmental Satellite System – Integrated Program Office (NPOESS-IPO) and the National Aeronautics and Space Administration (NASA). The background analyses that lead to some of these requirements are discussed.

1 INTRODUCTION

To assess the impact of anthropogenic effects on global climate and to distinguish them quantitatively from natural effects is a daunting scientific task. To meet this challenge, satellite instruments must be capable of observing atmospheric temperature trends as small as 0.1 degrees C/decade, ozone changes as small as 1 percent/decade, and variations in the sun's output as small as 0.1 percent/decade.

The importance of understanding and predicting climate variation and change has increased significantly in the last decade. In 2001, the White House requested the National Academy of Sciences (NAS) National Research Council (NRC) (NRC 2001) to review the uncertainties in climate change science. One of the three key recommendations from the NRC's report is to "ensure the existence of a long-term monitoring system that provides a more definitive

observational foundation to evaluate decadal- to century-scale changes, including observations of key state variables and more comprehensive regional measurements". To accelerate research and reduce uncertainties in climate change science, the Federal government created the Climate Change Research Initiative (CCRI) in 2001.

To develop recommendations for improving the calibration of satellite instruments to meet the challenge of measuring global climate change, the National Institute of Standards and Technology (NIST), National Polar-orbiting Operational Environmental Satellite System-Integrated Program Office (NPOESS-IPO), National Oceanic and Atmospheric Administration (NOAA), and the National Aeronautics and Space Administration (NASA) organized a workshop in November 2002 (NIST 2004). Some 75 scientists, including researchers who develop and analyze long-term data sets from satellites, experts in satellite

instrument calibration, and physicists working on state-of-the-art calibration sources and standards, participated.

The final report: (a) defines the required absolute accuracies and long-term stabilities of global climate data sets; (b) translates the data set accuracies and stabilities to required satellite instrument accuracies and stabilities; and, (c) evaluates the ability of current observing systems to meet these requirements.

The focus of the report is on passive satellite sensors that make observations in spectral bands ranging from the ultraviolet to the microwave. The climate change variables of interest include: (a) solar variables (Earth radiation budget, and clouds, total solar irradiance, spectral solar irradiance, outgoing long-wave radiation, net incoming solar radiation, cloudiness); (b) atmospheric variables (temperature, water-vapor, ozone, aerosols, precipitation, and carbon dioxide); and, (c) surface variables (vegetation, snow cover, sea ice, sea surface temperature, and ocean color).

This list is not exhaustive. The variables were selected using the following criteria: (a) importance to decadal scale climate change, (b) availability or potential availability of satellite-based climate data records, and (c) measurability from passive satellite sensors.

There has been a number of previous reports that have also discussed accuracy and stability measurement requirements for long-term climate data sets, e.g. Hansen et al. 1993, Jacobowitz 1997 & NPOESS 2001, and for calibration issues Guenther et al. 1997, NRC 2000a, NRC 2000b. The next section discusses the terminology of accuracy and stability as an extension of the vocabulary of the Guide to the Expression of Uncertainty in Measurement (ISO 1993), or simply referred to here as the ISO Guide. The section following it describes the requirements of accuracy and stability for climate change variables given in the Workshop report. The final section discusses the analysis that led to requirements, giving examples of those variables such as solar irradiance, total column water-vapor, total column ozone and sea surface temperature.

2 ACCURACY AND STABILITY: EXTENSION OF CURRENT VOCABULARY OF ISO GUIDE

Measuring small changes over extended time periods necessarily involves the concepts of accuracy and stability of time series. Accuracy is defined as the *closeness of the agreement between the result of the measurement and the true value of the measurand*. Specifically the quantity to be measured is called the measurand and truth is the true value of the measurand. So the measurands are the various climate variables. This same definition is also given in the International

Vocabulary of Basic and General Terms in Metrology, commonly abbreviated as VIM (ISO 1993). In the notes on the definition of the term accuracy, the ISO guide observes it as a qualitative concept. As such, Taylor & Kuyatt (1994) commented "Because 'accuracy' is a qualitative concept, one should not use it quantitatively, that is associate numbers with it: numbers should be associated with measures of uncertainty instead." However, the community of experts at NOAA have already established use of the term accuracy as defined above, but used in a quantitative fashion (see for example Hansen et al. 1993). It has been widely used in scientific as well as policy making and contractual documents for various satellite missions of NOAA, NPOESS and NASA.

Rather than change this widely accepted and used terminology, its commonly accepted definition was re-emphasized and used as an extension of the ISO Guide. So, the term accuracy may be thought of as *closeness to the truth* and is measured by the bias or systematic error of the data; that is, *the difference between the short-term average measured value of a variable and the truth*. The short-term average value is the *average of a sufficient number of successive measurements of the variable under identical conditions such that the random error is negligible relative to the systematic error*. The ISO Guide elaborates the meaning of the systematic error as that which would *result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus the value of the measurand*. Elaborating further, "repeatability conditions include the same measurement procedure, the same observer, the same measuring instrument used under the same conditions, with repetition over a short period of time." The systematic error is called bias in the case of a measuring instrument. The ISO Guide takes the position that since the value of the measurand cannot be completely known, the systematic error and its causes cannot be completely known. However, the measurement goal is to arrive at a true value of the measurand. Therefore, one should follow the procedure of the ISO Guide to determine the systematic error. For example, "if a device is tested through a comparison with a known reference standard and the uncertainties associated with the standard and the comparison procedure can be assumed to be negligible relative to the required uncertainty of the test, the comparison may be viewed as determining the error of the device (ISO 1993)." Methods to establish the true value of a variable (the measurand) should be consistent with the internationally adopted methods and standards, thus establishing the *Système International d'Unités* (SI Units) traceability (BIPM 1998, NIST 1995). The resolution adopted by the 20th Conference Generale des Poids et Mesures (CGPM 1995) states: "that those responsible for studies of Earth resources, the

environment, and related issues ensure that measurements made within their programs are in terms of well-characterized SI units so that they are reliable in the long term, be comparable world-wide and be linked to other areas of science and technology through the world's measurement system established and maintained under the Convention du Metre".

The term "stability" may be thought of as the extent to which the accuracy remains constant with time. Stability is measured by the maximum excursion of the short-term average measured value of a variable under essentially identical conditions over a decade. The smaller the maximum excursion, the greater the stability of the data set. Climatology experts for monitoring climate change variables arbitrarily chose the decadal unit of time for ascertaining stability. Stability could be considered as an extension of the ISO Guide terminology of "reproducibility." It is *the closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement* (ISO 1993). In defining stability, the only changed condition of measurement is time. Reproducibility assessed quantitatively over a decade of time determines stability. It is *the mean excursion over a decade of the short-term, e.g. yearly average of the measured value of the measurand*. The unchanged conditions of measurements include principles of measurement, method of measurement, observer, measuring instrument, reference standard, and conditions of use.

For this report, the spatial scale of interest is generally global averages. This is not to say that regional climate change is not important. On the contrary, all climate changes are regional, e.g. desertification, monsoonal changes, ocean color (coral death), and snow/ice cover (retreating snowlines and decreasing sea ice cover/receding glaciers). Since trends in globally averaged data will generally be smaller than those of regional averages, meeting global average requirements can often be more demanding than meeting regional climate monitoring requirements because of the smaller fluctuation requirements.

It should be pointed out that achieving the instrument measurement requirements does not guarantee determining desired long-term trends. Superimposed on these trends is climatic noise, or short-term climate variations, that may mask the signal to be detected, or reduce one's confidence in the derived trend.

Well-validated stability and accuracy statements are critical to the success of a satellite sensor. Achieving the required stability and accuracy on-orbit has been plagued by a myriad of problems, from degradation of plaques to unforeseen stray-light effects. An accuracy claim supported by an uncertainty statement with traceability to SI units as maintained by NIST, and validated by comparisons with other instruments, ideally provides the necessary information to establish

the absolute value of a climate variable for absolutely calibrating climate change models. It also facilitates the search for the unforeseen or for new physics, i.e. the discrepancy between model and measured atmospheric transmission, and provides traceability to measurements made by past and future generations and by completely different instruments. A reliable accuracy statement also provides a route to make comparisons between two different sensors when satellite overlap is not possible due to premature failure or policy decision.

Stability allows one to look for the small drift in climate variables, at levels often well below the accuracy of the instrument. Only by having a true understanding of the accuracy of the instrument can one attribute these drifts to the climate variable of interest, and not to instrument drift, or changes in other climate variables such as solar irradiance. The absence of such an understanding requires that satellite sensors measuring a common measurand have sufficient time overlap that their measurements can be compared to assess on-orbit accuracy and to provide continuity to establish that the measured drifts are related to the expected small decadal drifts in climate change variables. Certainly, reasonable scenarios are possible where instrument drift between two sensors of unknown accuracy would give the appearance of change in a climate variable. Some of these scenarios could be eliminated with satellite time overlap and a proper understanding of an instrument's absolute uncertainty. The ability to attribute these drifts to changes in climate variables requires continual truthing by other measurements. Satellite time overlap itself is not sufficient to ensure the continuity of the measurements, as often the exact orbital paths of the satellites will differ and additional corrections will need to be performed. A sophisticated understanding of the measurement accuracy will ensure that corrections are properly applied and that comparisons are possible.

3 OUTCOME OF THE NOVEMBER 2002 WORKSHOP

3.1 *Climate variables*

The required accuracies and stabilities of the climate variable data sets were established with consideration of changes in important climate signals based on current understanding and models of long-term climate change. Such signals include the following: (a) climate changes or expected trends predicted by models; (b) significant changes in climate forcing or feedback variables, e.g. radiative effects comparable to that of increasing greenhouse gases; and, (c) trends similar to those observed in past decades (NIST 2004).

The first step in the process is specifying the anticipated signal in terms of expected change per decade.

Table 1a. Required accuracies and stabilities for solar irradiance, Earth radiation budget, and cloud variables. Column labeled “signal” indicates the type of climate signal used to determine the measurement requirements.

Variables	Signal	Accuracy	Stability (per decade)
Solar irradiance	Forcing	1.5 W/m ²	0.3 W/m ²
Surface albedo	Forcing	0.01	0.002
Downward longwave flux: surface	Feedback	1 W/m ²	0.2 W/m ²
Downward shortwave radiation: surface	Feedback	1 W/m ²	0.3 W/m ²
Net solar radiation: top-of-atmosphere	Feedback	1 W/m ²	0.3 W/m ²
Outgoing longwave radiation: top-of-atmosphere	Feedback	1 W/m ²	0.2 W/m ²
Cloud base height	Feedback	0.5 km	0.1 km
Cloud cover (fraction of sky covered)	Feedback	0.01	0.003
Cloud particle size distribution	Feedback	TBD*	TBD*
Cloud effective particle size	Forcing: water	Water: 10%	Water: 2%
	Feedback: ice	Ice: 20%	Ice: 4%
Cloud ice water path	Feedback	25%	5%
Cloud liquid water path	Feedback	0.025 mm	0.005 mm
Cloud optical thickness	Feedback	10%	2%
Cloud top height	Feedback	150 m	30 m
Cloud top pressure	Feedback	15 hPa	3 hPa
Cloud top temperature	Feedback	1 K/cloud emissivity	0.2 K/cloud emissivity
Spectrally resolved thermal radiance	Forcing/climate change	0.1 K	0.04 K

* to be determined.

The second step is determining the accuracies and stabilities needed in the data set to permit detection of the signal. Excellent absolute accuracy in the measurement of the climate variable is important for understanding climate processes and changes. However, it is not necessary for determining long-term changes or trends as long as the data set has the required stability. For most satellite instruments, stability appears to be less difficult to achieve than accuracy. Difficulties arise because of the many known and unknown systematic uncertainties that must be accounted for in calibrating the instrument on the ground to establish its absolute accuracy. Further challenging the effort is the need to maintain or adjust the calibration and continually monitor it on-orbit. Stability, on the other hand, is the measure of repeatability and reproducibility of the metrological characteristics of the instrument with time. Thus, a key attribute for climate data sets is long-term stability. The required stability is some fraction of the change in the atmospheric variable, assumed to be one part in five (1/5) in this report. If one cannot achieve the above stability, e.g. if one can only achieve a stability of 0.5 of the signal, there would be an increased uncertainty in the determination of the decadal rate of change.

The factor 1/5, or 20 percent, is arbitrary. It should be reevaluated periodically. If the climate change

signal is one unit per decade, a 20 percent stability would imply an uncertainty range of 0.8 to 1.2, or the ratio of minimum to maximum excursion is a factor of 1.5 in the estimate of the signal. One basis for choosing such a factor is related to the uncertainty in climate model predictions of climate change. Thirty-five climate model simulations yield a total range of 1.4 K to 5.8 K, or factor of about 4 in the change in global temperature by 2100 (IPCC 2001). Thus, a stability of 20 percent should lead to a considerable narrowing of the possible climate model simulations of change. Achieving the stability requirement does not guarantee determining these long-term trends. Superimposed on these trends is climatic noise (short-term climate variations) that may mask the signal being detected or reduce confidence in the derived trend.

Although excellent absolute accuracy is not critical for trend detection, it is crucial for understanding climate processes and changes. Continuous efforts should be undertaken to improve the accuracy of satellite instruments. Accuracy is essential also when satellite overlap is not possible for comparing measurement across generations, for proving an absolute scale, for modeling, or for related requirements.

Tables 1a–1c summarize required accuracies and stabilities of the data sets for solar irradiance, Earth

Table 1b. Required accuracies and stabilities for atmospheric variables. Column labeled signal indicates the type of climate signal used to determine the measurement requirements.

Variables	Signal	Accuracy	Stability (per decade)
<i>Temperature</i>			
Troposphere	Climate change	0.5 K	0.04 K
Stratosphere	Climate change	0.5 K	0.08 K
Water vapor	Climate change	5%	0.26%
<i>Ozone</i>			
Total column	Expected trend	3%	0.2%
Stratosphere	Expected trend	5%	0.6%
Troposphere	Expected trend	10%	1.0%
<i>Aerosols</i>			
Optical depth (troposphere/stratosphere)	Forcing	0.01/0.01	0.005/0.005
Single scatter albedo (troposphere)	Forcing	0.03	0.015
Effective radius (troposphere/stratosphere)	Forcing	Greater of 0.1 or 10%/0.1	Greater of 0.05 or 5% /0.05
Precipitation		0.125 mm/hr	0.003 mm/hr
Carbon dioxide	Forcing/Sources – sinks		

Table 1c. Required accuracies and stabilities for surface variables. Column labeled “signal” indicates the type of climate signal used to determine the measurement requirements.

Variables	Signal	Accuracy	Stability (per decade)
Ocean color		5%	1%
Sea surface temperature	Climate change	0.1 K	0.04 K
Sea ice area	Forcing	5%	4%
Snow cover	Forcing	5%	4%
Vegetation	Past trend	3%	1%

radiation budget and cloud variables, atmospheric variables, and surface variables. The table also indicates which one of the above climate signals (climate changes, climate forcings, climate feedbacks, or trends similar to recent trends) forms the basis for the requirement.

3.2 Satellite instruments

The requirements for data sets must be translated into the required accuracies and stabilities of satellite measurements. In some cases, e.g. solar irradiance and top-of-atmosphere (TOA) Earth radiation budget, there is a one-to-one correspondence. For other climate variables, this translation is more complex; and for a few of the variables, additional studies are needed to match data set accuracies and stabilities with satellite accuracies and stabilities.

Because of the difficulties in achieving necessary accuracies, e.g. exo-atmospheric total solar irradiance (Quinn & Frohlich 1999), a key attribute for satellite instruments is long-term stability. This may be achieved by either having an extremely stable instrument, or by monitoring the instrument's stability by various methods while it is on-orbit. An ideal external calibration source is one that is nearly constant in time and able to be viewed from different orbit configurations. In addition to other attributes, if there is scientific evidence regarding the degree of stability of such a source, and it is believed to be at an acceptable level for long-term climate studies, then the stability of the satellite sensor can be assessed independent of other reference standards. With such monitoring, instrument readings can be corrected for drift. However, this brings up a measurement challenge for establishing the degree of stability of the external reference source. Obviously the methods and instruments testing the stability of those sources must have stability requirements far more stringent than given here. One method that has been successfully implemented for the reflected solar spectral interval is lunar observations by the sensor while on-orbit. An example is the ocean color satellite Sea-viewing Wide Field-of-view Sensor (SeaWiFS), which used lunar observations to correct for degradation in the near-infrared channels. The required lunar data are being supplied by a dedicated ground-based facility (Anderson et al. 1999).

Since satellites and their instruments are short-term (NPOESS satellites and instruments have design lives of about 7 years), satellite programs launch replacement satellites to continue the observations. Thus, the

Table 2a. Required accuracies and stabilities of satellite instruments to meet requirements of Table 1a. The instrument column indicates the type of instrument used to make the measurement.

Variables	Instrument	Accuracy	Stability (per decade)
Solar irradiance	Radiometer	1.5 W/m ²	0.3 W/m ²
Surface albedo	Vis radiometer	5%	1%
Downward long-wave flux: surface	IR spectrometer and Vis/IR radiometer	See tropospheric temperature, water vapor, cloud base height, and cloud cover	See tropospheric temperature, water vapor, cloud base height, and cloud cover
Downward short-wave radiation: surface	Broad band solar and Vis/IR radiometer	See net solar radiation: TOA, cloud particle effective size, cloud optical depth, cloud top height, and water vapor	See net solar radiation: TOA, cloud particle effective size, cloud optical depth, cloud top height, and water vapor
Net solar radiation: top-of-atmosphere	Broad band solar	1 W/m ²	0.3 W/m ²
Outgoing long-wave radiation: top-of-atmosphere	Broad band IR	1 W/m ²	0.2 W/m ²
Cloud base height	Vis/IR radiometer	1 K	0.2 K
Cloud cover (fraction of sky covered)	Vis/IR radiometer	TBD*	TBD*
Cloud particle size distribution	Vis/IR radiometer	TBD*	TBD*
Cloud effective particle size	Vis/IR radiometer	3.7 μm: water 5%, ice 10%, 1.6 μm: water 2.5%, ice 5%	3.7 μm: water 1%, ice 2%, 1.6 μm: water 0.5%, ice 1%
Cloud ice water path	Vis/IR radiometer	TBD*	TBD*
Cloud liquid water path	Microwave and Vis/IR radiometer	Microwave: 0.3 K Vis/IR: see cloud optical thickness and cloud top height	Microwave: 0.1 K Vis/IR: see cloud optical thickness and cloud top height
Cloud optical thickness	Vis radiometer	5%	1%
Cloud top height	IR radiometer	1 K	0.2 K
Cloud top pressure	IR radiometer	1 K	0.2 K
Cloud top temperature	IR radiometer	1 K	0.2 K
Spectrally resolved thermal radiance	IR spectroradiometer	0.1 K	0.04 K

* to be determined.

long-term data record for any climate variable will consist of contributions from a series of satellite instruments, some using different measurement techniques. To assess the reproducibility of the measurement results, to assist in understanding the differences that arise even with instruments of similar design, and to create a seamless data record, it is essential that the satellites be launched on a schedule that includes an overlap interval of the previous and the new instrument. Acquiring multiple independent space-based measurements of key climate variables would also help insure maintenance of stability in the event of a single instrument failure.

Tables 2a–2c summarize the required accuracies and stabilities of satellite instruments for solar irradiance, Earth radiation budget and cloud variables, atmospheric variables, and surface variables. The tables also indicate types of satellite instruments used for the measurements.

4 ANALYSIS OF REQUIREMENTS

4.1 *Solar irradiance, Earth radiation budget and clouds*

Overall, the variables in this section are linked in their role in the energetics of the climate system. The Sun is the dominant source of energy for Earth's climate. Once thought to be a steady, constant energy source (hence the term “solar constant”) to express the amount of solar radiation reaching the Earth, it varies significantly on a decadal time scale. Accurate measurements of solar irradiance are key to defining climate radiative forcing, and its accuracy requirements are specified in that context. Changes in surface albedo can represent both changes in climate forcings due to human caused land-cover change, and climate feedbacks due to changes in ecosystems and in snow and ice cover resulting from climate changes. Cloud

Table 2b. Required accuracies and stabilities of satellite instruments to meet requirements of Table 1b. The instrument column indicates the type of instrument used to make the measurement.

Variables	Instrument	Accuracy	Stability (per decade)
<i>Temperature</i>			
Troposphere	MW or IR radiometer	0.5 K	0.04 K
Stratosphere	MW or IR radiometer	1 K	0.08 K
Water vapor	MW radiometer	1.0 K	0.08 K
	IR radiometer	1.0 K	0.03 K
<i>Ozone</i>			
Total column	UV/VIS spectrometer	2% (λ independent), 1% (λ dependent)	0.2%
Stratosphere	UV/VIS spectrometer	3%	0.6%
Troposphere	UV/VIS spectrometer	3%	0.1%
Aerosols	VIS polarimeter	Radiometric: 3% Polarimetric: 0.5%	Radiometric: 1.5% Polarimetric: 0.25%
Precipitation	MW radiometer	1.25 K	0.03 K
Carbon dioxide	IR radiometer	3%	Forcing: 1% Sources/sinks: 0.25%

Table 2c. Required accuracies and stabilities of satellite instruments to meet requirements of Table 1c. The instrument column indicates the type of instrument used to make the measurement.

Variables	Instrument	Accuracy	Stability (per decade)
Ocean color	VIS radiometer	5%	1%
Sea surface temperature	IR radiometer	0.1 K	0.01 K
	MW radiometer	0.03 K	0.01 K
Sea ice area	VIS radiometer	12%	10%
Snow cover	VIS radiometer	12%	10%
Vegetation	VIS radiometer	2%	0.80%

feedback remains the largest single factor in the current large uncertainty in climate sensitivity (IPCC 2001). Cloud properties are critical to understanding and defining the role of clouds as feedback mechanisms in the climate system. Earth's radiation budget is critical to the climate system, and is a key diagnostic for a wide range of climate forcings (aerosol), feedbacks (clouds, ice/snow), and climate responses (heat transport). Accuracies for clouds and radiation budget are defined at levels sufficient to be at, or above, estimates of unforced natural climate variability in current climate models. These accuracies also must be sufficient to observe directly decadal changes in clouds and radiation budget that would constrain potential cloud feedback mechanisms in climate models.

The largest time and space scales will drive accuracy and stability requirements. For solar irradiance and surface albedo, climate radiative forcing drives the requirements. For clouds and radiation budget, climate feedbacks drive the requirements. Recent studies of the last two decades of cloudiness (International Satellite Cloud Climatology Project [ISCCP]) and

radiation budget data (Earth Radiation Budget Experiment [ERBE], Scanner for Radiation Budget [ScaRaB], and Clouds and Earth's Radiant Energy System [CERES]) have indicated significant interannual to decadal variability in the tropics between 20S and 20N latitudes. This variability is not shown in current climate model simulations and is representative of changes that are critical to assess accurately from observations, and to be able to predict from climate models. A climate observing system that cannot rigorously observe such changes with high confidence is very unlikely to be able to constrain and verify cloud feedbacks within climate prediction models.

Accuracy requirements can also be determined by considering the amount of climate change likely over the next few decades. For example, many climate change models use a one percent/year increase in carbon dioxide to simulate a nominal doubling of CO_2 in 70 years. This doubling is a radiative forcing of the climate system of about 4 Wm^{-2} , or about 0.6 Wm^{-2} per decade. A change in global average cloud fraction sufficient to offset this radiative forcing would be

about 0.015 if all other cloud properties remained fixed. This would be a cloud feedback so strong that climate change due to greenhouse forcing would become negligible. It is suggested that a minimum signal-to-noise ratio (SNR) of at least 5 is needed to detect such change, leading to a requirement for stability per decade in global cloud cover of 0.003. This would be sufficient to detect a cloud feedback. This approach essentially follows that used by Hansen et al. (1993) in a workshop report that summarized accuracies required for long-term monitoring of global climate forcings and feedbacks. The accuracy requirements in this section are, in general, similar to those in Hansen et al. where the same climate variable was evaluated. As in that report, the November 2002 workshop concluded that the appropriate scaling for climate requirements are the radiative flux changes that can potentially alter climate: either forcing or feedback.

The NPOESS project convened a workshop to assess climate measurement requirements for the Integrated Operational Requirements Document (IORD) variables (Jacobowitz 1997). The report influenced the IORD to add or change stability requirements, but had little effect on other IORD requirements, which were focused on instantaneous observations and, often, high spatial resolution. Climate spatial scales run from 50 km through global, and climate time scales run from a few weeks to centuries for current global change concerns. The requirements in this report and in Hansen et al. (1993) for clouds and radiation budget are often more stringent than in the NPOESS climate workshop. The NPOESS workshop does not appear to have used a consistent definition of the radiative forcings and feedbacks. Many of its threshold stability values would not be able to detect the decadal changes expected for forcings and feedbacks. Following Hansen et al. (1993), the November 2002 workshop addresses the requirements in a consistent radiative forcing or feedback metric (NIST 2004). It also assumes that the forcing or feedback must be detected accurately enough to assess decadal change at the level of 20 percent of the anticipated greenhouse gas forcings per decade. If four forcing and/or feedback mechanisms are found to be significant at this level, and the data verify that a future climate model predicts them to this accuracy, then the uncertainty in future predictions by the climate model is composed of four likely independent errors, each of which is 20 percent of the base greenhouse forcing. One might anticipate in this scenario that the uncertainty in future predictions would be: $20 \text{ percent} \times \text{square root}(4) = 40 \text{ percent}$. This would be a dramatic improvement over the current factor of 4 or larger uncertainty. But it also suggests that the stricter stability requirements in the November 2002 workshop and in Hansen et al. (1993) are to be thought of as thresholds or minimum values, not as desired

objectives. The objectives should be set even tighter by a factor of 2 to 4 (10 percent to 5 percent of the greenhouse forcing).

NIST (2004) does not discuss spatial, angular, and time sampling requirements in depth, because the focus was on calibration. But climate data record (CDR) accuracy includes these issues as well. For an observing system with fixed sun-synchronous orbits such as NPOESS, angular and time sampling biases are primarily a function of the orbit. Time sampling for many of the cloud and radiation variables can be augmented by incorporating the geostationary satellite data sets (imager and sounder), especially where they can be routinely intercalibrated with climate instruments to provide consistent data. Angle sampling errors are being reduced markedly through the efforts of the new multi-angle Polarization and Directionality of the Earth's Reflectances (POLDER), Multiangle Imaging SpectroRadiometer (MISR), and CERES observations. Spatial sampling errors become significant for instruments that only view nadir, such as the new active LiDAR and Radar systems. Nadir viewing primarily limits their climate-monitoring role to zonal and global means, but in some cases they can be sufficiently accurate for 1000 km scale annual mean regional values.

Regional climate change signals will be larger than zonal or global climate signals, but internal climate system noise will also be larger on these regional scales. The tradeoff of the internal climate noise versus signal has yet to be clearly defined for all of the variables in this report. There should be a continuing effort to estimate climate noise for each variable at a range of time and space scales. This information can then be used to refine the observing system requirements. There is little justification to measure more than a factor of 2 more accurately than the background climate noise. In the current analysis, climate model noise estimates have been used to help set requirements for several climate variables.

4.1.1 *Solar irradiance as an example*

IORD-II requirements were reviewed and were endorsed for both total irradiance and spectral irradiance accuracy and stability (NPOESS 2001). The threshold for absolute accuracy of total irradiance is 1.5 Wm^{-2} (0.1 percent), and for stability 0.02 percent/decade over many decades. As for many instruments, the stability of the active cavity radiometers greatly exceeds the absolute accuracy. At least a one-year overlap of observations is needed to remove instrument differences in absolute calibration. A 0.02 percent/decade stability requirement is sufficient to detect a 0.3 Wm^{-2} change in solar irradiance over a decade. This stability will constrain solar radiative forcing of the Earth's climate to within $(0.3)(0.25)(0.7) = 0.05 \text{ Wm}^{-2}$ per decade. The factor of 0.25

converts solar constant to the global average insolation over the Earth's surface, while the factor of (0.7) is the approximate fraction of energy absorbed by the Earth. This stability requirement will also allow rigorous tests of decadal to century time scale variability in solar output as the length of the data record grows. The system would be capable of detecting 0.5 W m^{-2} per century change in solar forcing. Even this subtle change would be a significant fraction of anticipated greenhouse gas forcing over the next century.

Spectral irradiance requirements are in general about a factor of 10 less stringent, but details vary with wavelength as indicated in the IORD-II. Spectral irradiance measurements are crucial for properly specifying the way that solar radiative energy enters the climate system. Absorption, scattering, and reflection (in the atmosphere, at the surface, and in the mixed layer of the ocean) all depend on wavelength. Solar radiation at different wavelengths has different variability. As an example, UV radiation that is deposited in the stratosphere, and that influences ozone, varies by one to two orders of magnitude more than visible radiation that reaches the Earth's surface. IR radiation varies least. Hence solar radiation at different wavelengths is deposited in different ways, depending on geography and altitude. Measurements of total irradiance alone provide no information about the spectral content of the irradiance variability, so a physical understanding of the processes by which climate responds to solar forcing requires measurement of spectral irradiance. Spectral irradiance observations are also important for verifying solar physics models. On the whole, the accuracy and stability requirements of this important variable, solar irradiance, map directly into radiometer requirements.

4.2 Atmospheric variables

The expected decadal changes in a variety of atmospheric variables were used to determine accuracy and stability requirements. This usually involved using the expected response to global warming estimated from general circulation model experiments. As in the previous section, one assumes that an SNR of at least 5 is required to detect these changes reliably from an instrument stability standpoint. The instrument accuracy, as has been discussed above, is less of an issue. As long as overlapping satellite records can be constructed to determine the calibration offsets between instruments, absolute accuracy requirements can be relaxed to what is expected (and indeed already achievable) from a variety of sensor technologies in the coming decade. This is not to minimize the importance of understanding the sources of absolute accuracy errors, since some of these sources could affect the stability required for climate monitoring. For many passive microwave or infrared technologies, instrument

absolute accuracies of 0.5 degrees C can meet requirements, as long as these accuracy numbers are dominated by a systematic bias that can be accounted for during satellite overlap periods.

4.2.1 Total column water-vapor as an example

Again, the accuracy (bias) associated with measured humidity is less important than the long-term stability of that measurement. A somewhat arbitrary assumption of a 5 percent accuracy requirement, which for deep layer averages, or vertically integrated water vapor, is being achieved already from the Special Sensor Microwave Imager (SSM/I). This is considerably more stringent than that listed in IORD II (20–25 percent), primarily because of large uncertainties in the retrieval of humidity in *shallow layers* to meet NPOESS weather forecasting requirements.

Assuming that constant relative humidity is maintained during global warming, at least in the lower troposphere, then an absolute humidity (or total water vapor) increase of about 1.3 percent/decade would be expected (global average) for a warming of +0.20 degrees C/decade. Again, very substantial regional deviations from this average value would be expected. Utilizing a factor of (1/5) leads to a 0.26 percent/decade long-term stability requirement. This is substantially more stringent than the 2 percent threshold stability requirement in the NPOESS IORD II (2001). Again, this is the requirement to observe the global moistening of the atmosphere associated with global warming – regional changes could be much larger and would have a much less stringent stability requirement.

Weak water vapor absorption lines in the infrared (on the wings of the $6.7 \mu\text{m}$ band or in the water-vapor continuum from $11 \mu\text{m}$ to $12 \mu\text{m}$) or in the microwave (around the 22 GHz water vapor line) are used to observe emission from the lower atmosphere. Discrimination of water vapor in the lower troposphere is dependent on the relative contrast between surface emission and atmospheric emission. In the infrared, both ocean and land surfaces have emissivities near 1.0, creating a low sensitivity to lower tropospheric water vapor. In the microwave near 22 GHz, ocean emissivity ranges from 0.5–0.6 but land emissivity is near 1.0. Because of this, there is good contrast in the microwave and a greater sensitivity to changes in lower tropospheric water vapor over the ocean versus infrared techniques. In the microwave, the water vapor weighting function, i.e. change in transmittance with change in logarithm of pressure, is stable and the radiance is related linearly to brightness temperature. In contrast, in the infrared, the weighting function is more highly variable (and is a function of the water vapor profile), and radiance is a non-linear function of temperature (about T^8 near $6.7 \mu\text{m}$).

Both microwave and infrared water vapor measurements operate at frequencies where the expected increase in vapor accompanying, say, a 1 K warming leads to a larger instrument response than 1 K, i.e. from a 2 K increase at microwave frequencies to 0 K to 4 K decreases at infrared wavelengths, depending upon the channel frequency. Thus, the signal magnitude of increased humidity might be expected to be larger than the expected global warming signal by a factor of 2 to 4. Unfortunately, since water vapor is not a uniformly mixed gas like oxygen (for microwave temperature) or carbon dioxide (for infrared temperature), there are significant data interpretation problems when trying to retrieve water vapor in the atmosphere from passive measurements.

In the microwave, total column vapor can be measured near the 22 GHz water vapor line, while tropospheric profiles of vapor can be retrieved with several frequencies near the 183 GHz water vapor line. Using a 2:1 instrument response factor just described, one can double the temperature and stability requirements, i.e. 1.0 degrees C absolute accuracy and 0.08 degrees C/decade, for microwave water vapor measurements.

In the infrared, the response of individual channels varies widely, but one can assume an average response factor of around 2 to 4. For the global warming case in which relative humidity remains approximately constant, the global average brightness temperature also remains approximately constant. This is because the radiative impact of the warmer temperature profile offsets the effect of increased specific humidity in the free troposphere. The approximate simple relation between an infrared channel brightness temperature and upper tropospheric humidity is:

$$a + bT_b = \ln(UTH P_{ref}) \quad (1)$$

where T_b = brightness temperature, $UTH P_{ref}$ = upper tropospheric humidity at P_{ref} , a reference pressure level, and $b = -0.115$.

Although UTH depends on both water vapor mixing ratio and atmospheric temperature, observations indicate that the main variations are due to water vapor. This equation indicates that to detect a 0.3 percent change in water vapor requires a stability of 0.03 K in brightness temperature.

4.3 Surface variables

Surface variables include land vegetation, snow cover, sea ice, ocean color, and temperature. One problem with defining requirements for satellite measurements of the Earth's surface is the wide range of surface types covered. A fundamental concern is the need for both accurate pre-launch calibration and post-launch validation of all satellite instruments. By their very nature satellite measurements do not

sense directly the parameter of interest, and it is only through these calibration and validation efforts that one can develop methods to estimate the desired parameters from satellite data. These concerns apply both to present and future satellite measurements.

4.3.1 Sea surface temperature as an example

Climate models predict air temperature increases of about 0.2 K/decade due to greenhouse warming. Sea surface temperature (SST) can be expected to increase at about the same rate. To measure this change requires a data set stability of 20 percent of 0.2 K/decade or 0.04 K/decade. Accuracy of 0.1 K is considered adequate.

Ocean buoys measure the SST at 1 m to 2 m below the surface representative of the "bulk SST" that were also measured by ship buckets prior to the 1950s, and ship injection SSTs since then. Satellite measured SSTs are sensitive to the topmost "skin" layer of the ocean, but they are generally corrected to bulk SSTs. The Surface Panel recommends that the satellite SST program include an *in-situ* program of calibration/validation measurements that combine both skin and bulk SSTs.

The required SST stability and accuracy are 0.04 K/decade and 0.1 K. Sea surface temperatures are generally measured at infrared atmospheric window wavelengths. The relevant equation is:

$$SST = T + 2.5(T_1 - T_2) \quad (2)$$

where, T_1 and T_2 are IR brightness temperatures at two IR window wavelengths.

Error analysis of this equation assuming that T_1 and T_2 have the same absolute errors leads to a stability requirement of about 0.01 K for each window wavelength brightness temperature. Required accuracy for the measurements is 0.1 K. For sensors with additional channels, such as the Visible/Infrared Imager/Radiometer Suite (VIIRS) and MODIS, other SST algorithms may be more effective. The proposed SST algorithm for VIIRS is a "dual split window" that uses a brightness temperature difference at the shorter 4 micron channels together with the longer 11 micron channel difference to give a more stable SST estimate. But the above error analysis should hold for any split window type of SST measurement.

Microwave observations at 6.9 GHz can also be used to measure SST. A 1 K change in SST causes about a 0.33 K change in observed brightness temperature. The reduction in sensitivity is due to the low microwave ocean emissivity of about 0.5 and wind roughening effects. Thus, to maintain a stability of 0.04 K in SST requires about a 0.01 K microwave instrument stability. Required accuracy is 0.03 K.

These values ignore the influence of sensor pointing angle on SST accuracy for passive microwave

sensors. Both the passive microwave and thermal infrared sensors will require an *in-situ* calibration/validation program to insure that these requirements are met. This *in-situ* program must include both skin and bulk measurements of SST and should be continuous.

5 CONCLUSIONS

A more detailed discussion of the definitions of stability and accuracy as an extension of the vocabulary and usage of the ISO guidelines on uncertainty determination in measurements is presented. The stability and accuracy requirements for long time series data sets for climate change variables have been developed. Explanations of how the requirements have been arrived at for various climate change variables with a few examples taken from the full report are presented.

REFERENCES

- Anderson, J.M., Becker, K.J., Kieffer, H.H. & Dodd, D.N. 1999. Real-time control of the robotic lunar observatory telescope, *Astron. Soc Pacific* 111: 737–749.
- BIPM 1998. *The International System of Units (SI)* (7th ed). Paris: Bureau International des Poids et Mesures.
- CGPM 1995. Resolutions adopted by the 20th Conference Generale des Poids et Mesures. Paris: Bureau International des Poids et Mesures (BIPM).
- Guenther, B., Butler, J. & Ardanuy, P. 1997. Workshop on strategies for calibration and validation of global change measurements. *NASA Reference Publication* 1397.
- Hansen, J., Rossow, W. & Fung, I. (eds) 1993: *Long-term monitoring of global climate. forcings and feedbacks; Wkshp. Proc., 3–4 February 1992*. New York: Goddard Inst. for Space Studies.
- IPCC 2001. *Intergovernmental panel on climate change: third assessment report – climate change 2001*. World Meteorological Organization, UNEP. Cambridge: Univ. of Cambridge Press.
- ISO 1993. *International vocabulary of basic and general terms in metrology* (2nd ed). Geneva: International Organization for Standardization.
- Jacobowitz, H. (ed.) 1997. Climate measurement requirements for the National Polar-orbiting Operational Environmental Satellite System (NPOESS). *Wkshp. Report*. Boulder: Univ. Consort. Atmos. Res.
- NIST 1995. Guide for the use of the International System of Units (SI). *Publication 811*. Washington DC: U.S. Government Printing Office.
- NIST 2004. Satellite instrument calibration for measuring global climate change. *NISTIR 7047*. Washington DC: U.S. Government Printing Office.
- NPOESS 2001. Integrated Operational Requirements Document II (IORD II) – <http://npoeplib ipo.noaa.gov/>
- NRC 2000a. Improving atmospheric temperature monitoring capabilities. *Letter report from the panel on reconciling temperature observations*, Climate Research Committee. Washington DC: The National Academies.
- NRC 2000b. Issues in the integration of research and operational satellite systems for climate research: part II. *Implementation*. Committee on Earth Studies, Space Studies Board. Washington DC: National Academy Press.
- NRC 2001. *Climate change science: an analysis of some key questions*. Commission on Geosciences, Environment and Resources (CGER). Washington DC: National Academy Press.
- Quinn, T. & Fohlich, C. 1999. Accurate radiometers should measure the output of the sun. *Nature* 401(6756): 841.
- Taylor, B.N. & Kuyatt, C.E. 1994. Guidelines for evaluating and expressing the uncertainty of NIST measurements results. *NIST Tech. Note 1297*. Gaithersburg: U.S. Dept. of Commerce.

Calibration standards and guidelines

Validated data and removal of bias through traceability to SI units

N.P. Fox

National Physical Laboratory, Teddington, Middlesex, UK

ABSTRACT: The paper describes recent advances in establishing primary radiometric scales and transfer standards, and their application to improving accuracy and traceability to Standard International (SI) units of Earth Observation measurements. It further discusses how to ensure that all steps in generating a data product, data collection, processing and validation, have associated with them an analysis of their uncertainty and assessment of any known biases so that end users can compare and combine reliably all levels of data product that may be derived from them. Finally, the paper concludes with options for how future accuracy and traceability can be improved significantly.

1 INTRODUCTION

1.1 Background

In recent years there has been an increasing demand for improved accuracy and reliability of Earth observation (EO) data. This is stimulated not only by the desire to understand better the workings of Earth and the causes and impacts of climate change, but also because improvements in models allow us to discriminate better between data. A further driver is the increasing reliance on the combination of data from different sensors and sources (satellites, aircraft and *in-situ*) to establish more sophisticated data products. These will be important for “operational services” of the future, as envisaged in initiatives such as Global Monitoring of Environment and Security (GMES) of the European Union (EU) and European Space Agency (ESA). This combination of data from different sources can only be carried out if each data set has a clear and reliable statement of uncertainties.

However, when only relative spatial maps are required, only radiometric resolution (signal-to-noise) and uniformity of response of the sensor are critical. In these applications, absolute accuracy is of lower importance. However, as soon as temporal information is required, combination of data from more than one source (satellite or ground) or where the data may form a baseline for a future long-term study (as in most climate change research), accuracy and a clear knowledge of its associated uncertainty, is essential. This in turn requires “traceability” of measurement to the internationally agreed system of units (SI). The term traceability in this context has a clear definition requiring

a full analysis and description of uncertainties and an unbroken chain of calibrations back to an internationally recognized primary standard, usually as maintained by a National Metrology Institute (NMI) such as NPL or NIST. This requirement has been formally recognized by the Committee for Earth Observation Satellites (CEOS) with the adoption at CEOS Plenary 14 in November 2000 of the following resolutions:

1. All Earth observation (EO) measurement systems should be traceable to SI units for all appropriate measurands.
2. Pre-launch calibration should be performed using equipment and techniques that can be demonstrably traceable to, and consistent with, the SI system of units; and traceability should be maintained throughout the lifetime of the mission.

These resolutions follow closely those adopted by the 20th General Conference of Poids et Mesures (CGPM 1995) (the international body responsible for SI system of units). This meeting concluded that: *those responsible for studies of Earth resources, the environment, human well-being and related issues ensure that measurements made within their programmes are in terms of well-characterized SI units so that they are reliable in the long-term, are comparable world-wide and are linked to other areas of science and technology through the world's measurement system established and maintained under the Convention du Mètre.*

This paper describes how recent improvements in the development of traceability routes from the primary standard can help meet the needs of user applications. It is limited to passive optical radiometric sensors, but

the principles translate to other EO technology disciplines. The paper will start from the establishment of the primary scale before describing the development and use of specialized transfer standards tailored to the calibration of a satellite/airborne sensor or a ground based field-spectrometer.

1.2 Terminology

To ensure that the emerging commercial EO user community understands the quality statements associated with both low and high level data products, it is essential that the underpinning metrological terminology be clearly and consistently defined. At present, relatively common terms are frequently misinterpreted and mis-used when applied to metrology, causing lack of clarity and understanding. Table 1 contains formal, internationally agreed upon definitions (ISO 1993), and it is appropriate that they be used when writing and reviewing documents and articles for EO work.

Table 1. Definition of key metrological terms as agreed by the international community, ISO, IEC, and SI.

SI units	The coherent system of units adopted and recommended by the General Conference of Weights and Measures (CGPM).
Accuracy	Closeness of the agreement between the result of a measurement and a true value of the measurand.
Precision	No metrological definition except to state that it should never be used in the context of “accuracy” and, because of possible confusion, its use should normally be avoided in metrological applications.
Repeatability	Closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.
Reproducibility	Closeness of the agreement between the results of the measurements of the same measurand carried out under changed conditions of measurement.
Uncertainty	Parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.
Error	Result of a measurement minus a true value of the measurand.
Stability	Ability of a measuring instrument to maintain constant its metrological characteristics with time.
Traceability	Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually through an unbroken chain of comparisons all having stated uncertainties.

2 PRIMARY SCALES

2.1 Introduction

Figure 1 is a schematic representation of the inter-linking and interdependence of the primary optical radiometric quantities as a result of innovation and research over the last 20 years. This figure illustrates the key role of filter radiometry, which provides the interface between the primary monochromatic scales and those (polychromatic) of the real world.

2.2 Cryogenic radiometry and spectral responsivity

The primary standard of Figure 1 could in principle be a detector or source. However, the majority of NMIs are currently using a special form of electrical substitution radiometer (ESR) known as a cryogenic radiometer (Fox 1996) as the primary standard. An ESR is essentially a detector that compares the heating effect of optical radiation with that due to electrical power. Electrical substitution has been used in radiometry for more than 100 years, but it is only since the 1970s when Quinn & Martin (1985) successfully operated an instrument at liquid helium temperatures that uncertainties of a few tenths of a percent have been possible in disseminated scales (Fox 1996) (the uncertainty in the primary scale being <0.005 percent).

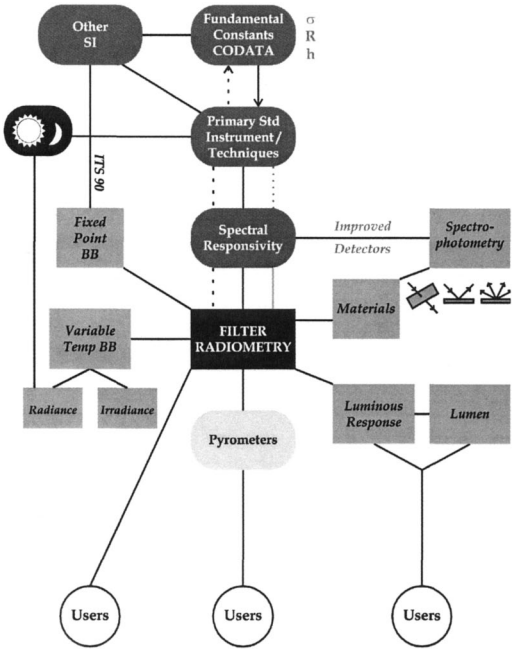


Figure 1. A schematic representation showing the inter-linking and interdependency of primary radiometric quantities and their route to the user.

The cryogenic radiometer described by Martin et al. (1985), and in a modified form by Fox et al. (1996), can be used to calibrate the spectral responsivity of a detector by determining the radiant power in a beam of monochromatic radiation, which is then made to fall on the detector. The process can be repeated at different wavelengths until sufficient spectral data have been obtained for interpolation of a continuous scale. The most commonly used source for this purpose is an intensity stabilized laser (Geist et al. 1977, Fox 1996). Such a source is highly stable (power variations of <0.001 percent over many hours can be achieved), monochromatic of known wavelength, and highly collimated, making it easy to manipulate. However, to achieve continuous spectral coverage requires a large suite of lasers, which are only available in one or two NMIs, e.g. the National Laser Radiometry Facility (NLRF) of NPL (Fox 2001) and Spectral Irradiance and Radiance Responsivity Calibrations with Uniform Sources (SIRCUS) of NIST (Brown et al. 2000).

Fortunately, solid-state photodiodes exist with relatively smooth variation in their spectral response, allowing continuous spectral responsivity scales to be established by interpolation with relatively few direct calibrations against the cryogenic radiometer. This process is further simplified through specially packaged photodiodes “trap detectors” as a transfer standard. Trap detectors were initially conceived as a means of independently establishing primary spectral responsivity scales (Zalewski & Duda 1983) and consist of three photodiodes arranged so that the reflection of radiation from each photodiode surface is incident onto a subsequent detector as shown in Figure 2. In this way, incident radiation undergoes five internal reflections before exiting the device, making it insensitive to changes in surface reflectance. Since photodiodes can be chosen that have a near unity internal quantum efficiency (conversion efficiency of absorbed photon into detected electron), they also have spectral responsivities that can be modeled easily, making them highly suitable for interpolating between calibration points (Fox 1991).

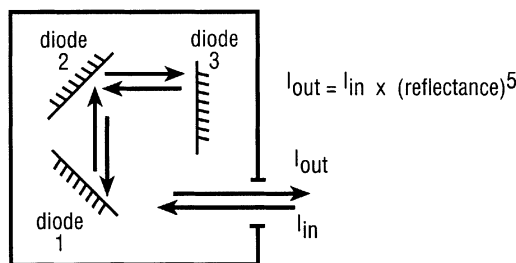


Figure 2. Schematic of the trap device. I_{in} and I_{out} refer to the total radiation entering and exiting the device, respectively.

2.3 Spectral irradiance

To determine the spectral radiance or irradiance from a broad band source, e.g. an incandescent lamp or reflected sunlight from the Earth, a detector with a quasi-monochromatic spectral response has to be used, i.e. a filter radiometer. A filter radiometer can be a simple interference-filter-detector combination or a grating spectrometer-diode-array imaging camera, as used in many remote sensing instruments. It is simply an instrument capable of defining a known narrow spectral response band out of a broad continuum source.

It is the development of methods to accurately calibrate the filter radiometer's spectral response that has led to improvements in accuracy of primary spectral emission scales established by NMIs. In particular, use of tuneable lasers makes possible the calibration of a filter radiometer so that it can measure the spectral radiance of a source with an uncertainty of <0.05 percent (in any spectral region – UV, VIS, NIR) (Anderson et al. 1992, Fox et al. 1991). Since most sources have only slowly varying spectral emissions, Anderson & Fox (1991) showed that it is necessary only to have a few filter radiometers to define the general shape, allowing interpolation between calibration points. In practice, because of the need to calibrate a wide range of different sources, including those with line emission, NMIs generally calibrate an intermediate high temperature blackbody (operating temperature around 3000 K) to establish a spectrally continuous scale (Johnson et al. 1993, White et al. 1996, Sperfeld et al. 1996, Yoon et al. 2002). The procedure for establishing such scales is relatively simple in concept, with the most important measurement being determination of the radiant temperature of the blackbody using a radiometrically calibrated filter radiometer. In principle, only a single spectral-band filter radiometer is needed to determine the absolute spectral irradiance or radiance from the blackbody and thus, through Planck's law, thermodynamic temperature. In practice however, for the most accurate measurements a number of filter radiometers spanning the spectral range of interest is generally used to take account of any non-Planckian characteristic of the blackbody cavity.

After determining its thermodynamic temperature, the blackbody is then viewed by a monochromator spectroradiometer to give a short-term calibration of its spectral response before it is used to view a transfer standard lamp, in effect transferring the calibration from the blackbody. A schematic representation of the NPL Spectral Radiance and Irradiance Primary Scales (SRIPS) facility is shown in Figure 3. NPL currently uses a blackbody operating at temperatures up to 3500 K to establish its scales down to 200 nm (Woolliams 2003). These new “detector-based” scales have much lower uncertainty than previous realizations,

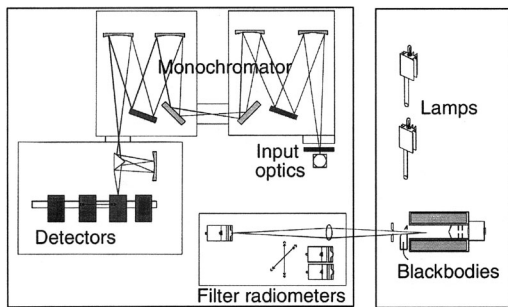


Figure 3. Schematic representation of the NPL primary spectral irradiance facility.

Table 2. Uncertainty budget for the NPL spectral irradiance scale.

Uncertainty component	Individual uncertainty (K)	Effective uncertainty, 3050 K, blackbody		
		300 nm (%)	550 nm (%)	2000 nm (%)
<i>Impact on temperature</i>				
Blackbody uniformity	0.327	0.17	0.09	0.03
Blackbody stability	0.258	0.13	0.07	0.02
FR absolute responsivity	0.115	0.06	0.03	0.01
Lens transmission	0.103	0.05	0.03	0.01
Blackbody emissivity	0.089	0.05	0.03	0.01
SRIPS repeatability		0.26	0.08	0.28
Lamp alignment		0.06	0.06	0.06
Total scale uncertainty (1σ)		0.37	0.17	0.30

which will allow significant improvements in the accuracy of derived quantities such as radiance.

The last international comparison of spectral irradiance was carried out in 1990 and generally demonstrated agreement between participants within 2 percent in the visible region of the spectrum (Walker et al. 1991). A new comparison piloted by NPL is nearing completion, and it is hoped that the introduction of the new detector-based scales described above (and their reduced uncertainties) will lead to closer agreement.

Table 2 presents a summary of the uncertainty components of the new NPL spectral irradiance scale for a few specific wavelengths (many of the components have a spectral sensitivity). Uncertainty sources with an

effect less than 0.03 percent include blackbody temperature, blackbody absorption, blackbody-integrating sphere geometry, SRIPS linearity, and monochromator wavelength error. Elements unique to blackbody temperature are: size of source effect; mathematical approximations; geometric factor; electronics; and filter radiometer relative shape. This table is included as an example of the detail that needs to be considered and presented when describing a measurement, and should be encouraged within the EO community. Because of the added difficulty in using a blackbody, it is rare that customer lamps are calibrated directly. Instead, a hierarchical calibration chain is generally established so as to minimize burn hours, and thus drift, on any particular lamp. Such a process is illustrated in the left and middle panels of Figure 4.

2.4 Diffuse reflectance

The primary scale of diffuse reflectance in most NMIs is that of total hemispherical reflectance. It is this quantity that is generally compared internationally. For this reason, most of the world's NMIs who have established such a primary scale have done so by developing methods that collect all the radiation reflected from a reference target scattered over a hemisphere using a well characterized integrating sphere. Some of these NMIs have also developed independent facilities for measuring angular resolved diffuse reflectance, or radiance factor, using some form of goniometer operating either at one or more fixed angles or continuously. NPL is unique in being the only NMI that has established its primary hemispherical diffuse reflectance scale by using a goniometer. In this approach, the total hemispherical reflectance is determined through modeling and summing together of the angularly resolved reflectances (Chunnill et al. 2003).

In terms of the EO community, the most important quantity is that of angularly resolved reflectance and more specifically bidirectional reflectance distribution function (BRDF). Recently, a comparison was organized between NPL, NIST and PTB to compare primary realizations of gonio-diffuse reflectance. The results (Fig. 5) show agreement to better than ± 0.5 percent between all laboratories for most of the angles and most wavelengths. Figure 5 shows the results for 400 nm, one of the more difficult spectral regions.

3 APPLICATION TO EO MEASUREMENTS

3.1 Introduction

Although spectral irradiance has some direct applications in the EO sciences, e.g. the measurement of input solar irradiance, spectral radiance is a far more

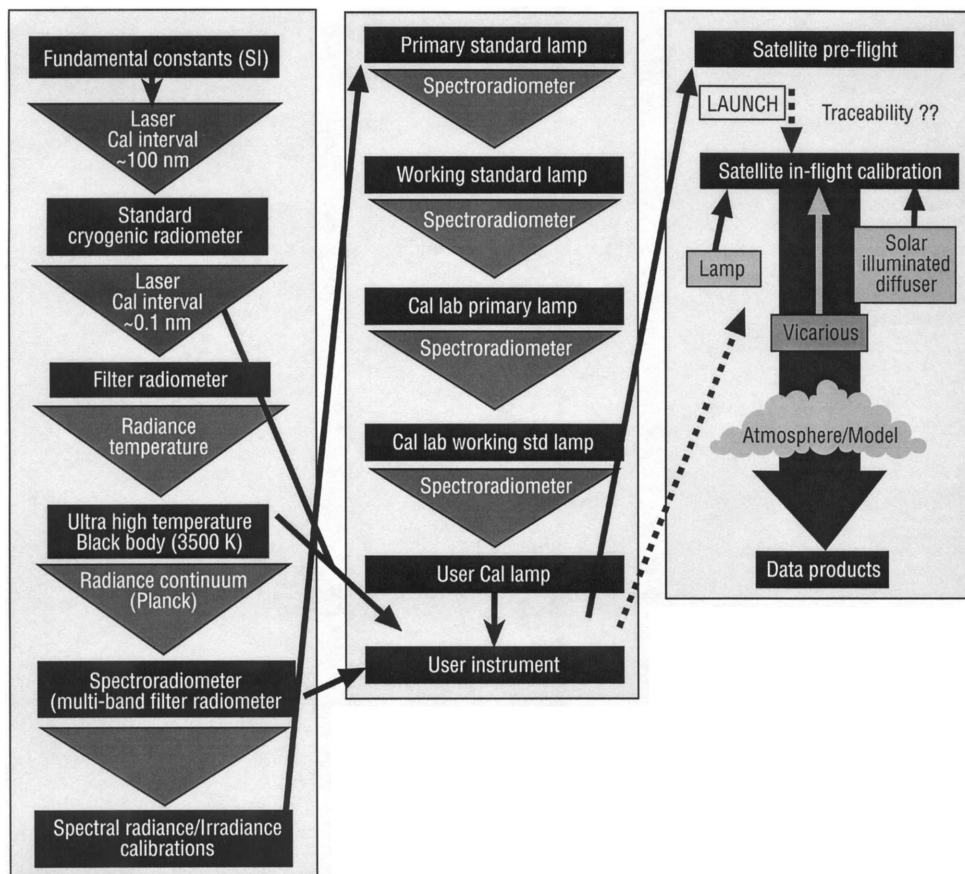


Figure 4. Schematic representation of traceability for the EO community. The left and middle panels describe the typical procedure from the primary standard to the user. The right panel illustrates the difficulty in obtaining traceability for a post-launch instrument. Solid lines connecting boxes show good traceability routes; the dashed a “best efforts.”

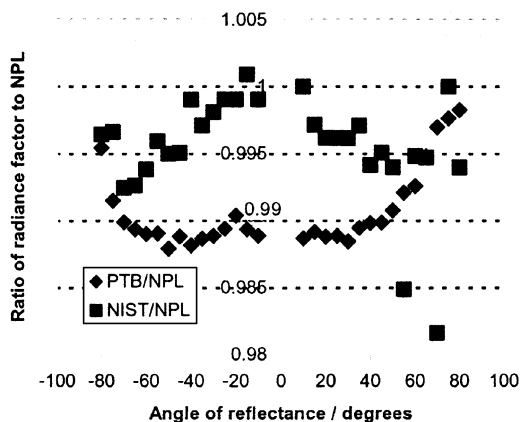


Figure 5. Comparison of radiance factor at 400 nm as a function of angle between NIST, PTB, and NPL.

important parameter. This is not surprising because most Earth viewing instrumentation measures spatially resolved Earth reflected sunlight, or emitted thermal radiation, which are radiances. Of course, in the case of solar reflected radiation, the real biogeophysical parameter of interest is reflectance; but it is determined through radiometric measurement. Although, in principle, satellite and aircraft based measurements could be processed directly in terms of reflectance (by assuming or measuring input solar spectral irradiance), the need to correct measurements for transmission through the atmosphere makes it more convenient to use radiance units. However, by contrast, measurements made *in-situ*, to validate data products of the satellite or aircraft instrument, or provide vicarious calibration, are often made in reflectance terms and then converted to radiances.

Spectral reflectance of the ground is usually determined by reference to a standard of known reflectance,

for example a white diffusing panel made of spectralon. The measuring instrument (spectroradiometer/filter radiometer) alternately views the solar illuminated panel or solar illuminated surface at the same viewing angle, and the reflectance is then calculated. If the spectral BRDF characteristics of the diffusing panel are known, the BRDF of the surface can be determined. This is usually converted to radiance by using a known value for solar spectral irradiance and correcting for atmospheric loss. There are many subtleties in this process and a good review of the subject, at least in terms of desert sites, based on the efforts of the University of Arizona can be found in Duingard & Slater (1999).

Care needs to be taken when using these reflectance based methods because, although appearing to be a relatively simple technology, differences in the level of reflectance from the reference diffuser (~100 percent) and the target surface (~30 percent) can result in errors due to differences in interreflections between the spectroradiometer and the respective measured surface. Therefore the *in-situ* measuring community ought to make measurements directly in spectral radiance units for which these interreflection errors are not present. This would result in little change to the measurement instruments and procedures being used, except to calibrate the spectroradiometer for spectral radiance. This could be carried out in a similar manner to that of other instruments described below.

3.2 Spectral radiance

Spectral radiance is a primary quantity and its direct route of dissemination in the visible spectral region from NMIs is often through a calibrated tungsten strip lamp. These lamps have an emitting ribbon of a few

millimeters width. In the thermal infrared, the transfer standard usually takes the form of a small aperture blackbody linked to SI through the International Temperature Scale (ITS-90), usually using a platinum resistance thermometer. Neither of these transfer standards is very suitable for calibrating a large aperture or field-of-view instrument. However, it is not difficult to see how a thermal infrared (TIR) radiance traceability route can be established by simply increasing the size of the aperture of the blackbody and modeling/measuring the change in emissivity as compared to a small area blackbody. There are difficulties with this method if care is not taken at each of the steps involved, e.g. if the contact thermometer is not measuring the emitting surface temperature, or if the emissivity model is not truly representative of the blackbody. But, in principle, a link to ITS-90 can be established through a fairly simple modification of the normal transfer standard.

This is not the case for the UV/VIS/NIR region of the spectrum. In this region, an alternative method has been widely adopted. A spectral irradiance standard lamp is obtained (with a traceable calibration) and positioned at a measured distance from a diffusing panel of known angular diffuse reflectance (radiance) factor, or BRDF. If this is a good Lambertian diffuser such as spectralon, the spectral radiance reflected from the panel is defined by the output of the spectral irradiance standard lamp, and the reflectance characteristics of the diffuser (Fig. 6). The diffusing panel is then viewed directly by an instrument needing calibration, or used to calibrate an intermediate spectroradiometer, which is then used to calibrate another radiance source, e.g. a large lamp-illuminated hemisphere or sphere (Fig. 7).

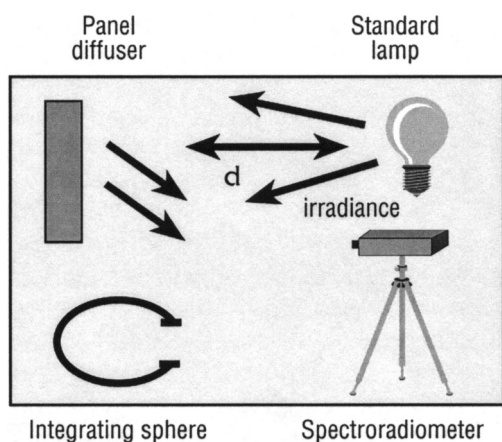


Figure 6. Establishing spectral radiance scale using an irradiance lamp and diffusing panel.

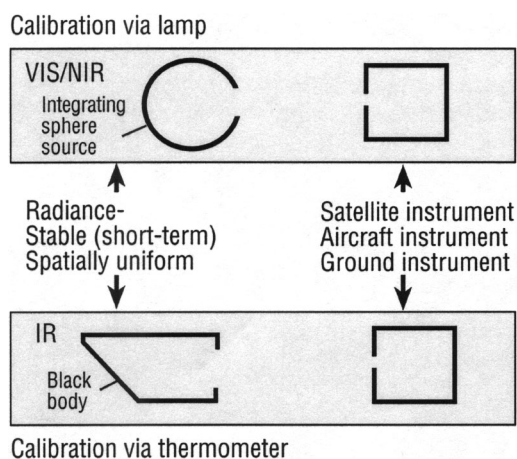


Figure 7. Spectral radiance calibration of an EO instrument using a large aperture source.

Until recently this has been the only route available for a calibration traceable to the SI. Even smaller lamp integrating sphere sources used as transfer standards from NMIs, were calibrated in this way. As described in §3.1, large errors can occur through interreflection effects when using a highly reflecting diffusing panel. For the method shown in Figure 7, a spectral irradiance lamp is generally placed very close (~ 0.5 m) to the highly reflecting panel, so interreflection problems can be large when illuminated as a large bright lambertian source. New transfer standards are now available that overcome this problem.

3.3 Transfer standards

One of the principle difficulties facing users of spectral irradiance is the transfer standard itself. The most frequently used source is a 1 kW tungsten halogen lamp of the FEL type. Such sources can operate at color temperatures of ~ 3200 K and thus give reasonably high levels of UV/blue irradiance. In the Key Comparison of spectral irradiance (see above) more than 60 such lamps have been used, each being measured under highly controlled conditions. NPL was the organizer for this comparison and thus has been able to investigate the performance of this large set of lamps, which have undergone various forms of transportation (hand carried, air-freight, and road) in their journeys between participants across the globe. The results indicate that the spectral irradiance of approximately one in three lamps changed significantly (>1 percent) on transportation. This result highlights the need to improve the performance of transfer standards for source related quantities.

Anticipating this result, NPL (Fox et al. 1998, Windsor et al. 2000, Woolliams et al. 2003) and others (Sperling & Bentlage 1998) have developed new irradiance sources based on coupling a tungsten halogen lamp with a group of stable filter radiometers. In such sources, one filter radiometer can be used as a reference to control the output of the lamp by varying the current supplied to it from the power supply in a feedback control circuit. Other filter radiometers, with peak spectral response spread across the spectrum of interest, can be used to monitor spectral changes. In this way, drifts and changes in the output of the lamp can either be corrected or at least monitored, to allow the user to assess their significance for their application.

Similarly, this technology can be applied to lamp illuminated integrating spheres used as radiance sources. NPL has constructed transfer standard absolute radiance sources (TSARS) specifically for application in the EO sector (Woolliams et al. 2002). In addition to detector stabilization, these sources also have been optimized for spatial uniformity of output radiance, and non-uniformities of $< \pm 0.2$ percent have been achieved over emitting areas of 80 mm

diameter. Sources such as TSARS allow a wide range of EO instrumentation to be calibrated for spectral radiance, either in the lab or in the field, without the need to use diffusing panels. Since the TSARS are self-monitoring, they can alert the user to any problem of drift in the calibration caused by aging or damage.

3.4 Shortening the calibration chain

The previous section illustrated a number of improvements in transfer standards that improve the accuracy of optical radiometric measurements. However, even greater accuracy, and in some cases flexibility, can be achieved by slightly more radical approaches. These approaches are highlighted in the schematic shown in Figure 4, and represent a shortening of the calibration chain, in effect reducing the number of steps to the primary standard.

Calibrated filter radiometers can be used directly as the transfer standard to generate absolute radiance, instead of using an incandescent lamp and diffuser (panel or sphere). Filter radiometers measure the radiance of an intermediate source, which might still be a lamp and diffuser, but in this case the source only has to be lambertian and stable in the short-term, since its radiance is being measured directly by the filter radiometer (Fig. 8). Fox (1995) showed that for the highest accuracy, the spectral shape of the response of the remote sensing instrument can be matched by the filter radiometers, in principle achieving accuracies of < 0.1 percent. In practice, other sources of error are likely to limit the accuracy and not the primary calibration and traceability to SI units. This approach has been widely used in the United States for preflight calibration, at least as a means of comparing different calibration facilities led by NIST (Johnson et al. 2004).

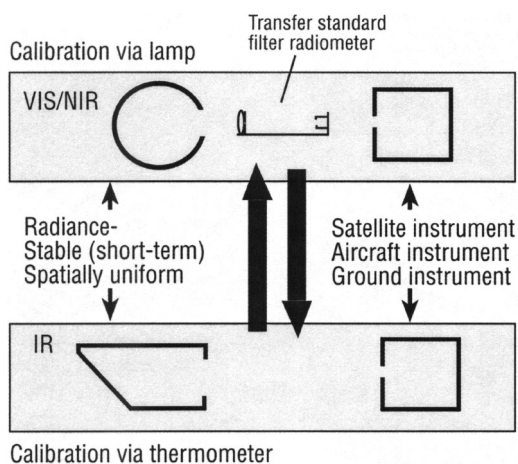


Figure 8. The use of filter radiometers in the determination of spectral radiance of the calibration source.

The technique could be applied equally for *in-situ* measurements with the filter radiometer being used directly to view the ground target.

The technique is not limited to the UV/VIS/NIR region and also has many advantages for the TIR, allowing spectral radiances of large aperture blackbodies to be measured directly rather than estimated from models and contact thermometry (Theocharous et al. 1998, Rice & Johnson 1998).

Multispectral and hyperspectral space instruments provide an additional challenge for this approach because they need greater spectral resolution. However, tuneable filter radiometers, or absolute spectral radiometers, are being developed which are expected to meet this challenge (Fox 1996, Fox et al. 1998).

An alternative approach is to calibrate instruments directly against the primary standard blackbody source. NPL recently performed such a calibration, using its ultra high temperature blackbody (UHTBB) with an optical imager called CHRIS (Compact High Resolution Imaging Spectrometer) built by SIRA Electro Optics Ltd (currently onboard the ESA Proba platform). Calibration required the construction of a class 100 clean environment in the NPL calibration laboratory and the blackbody was then used to calibrate CHRIS in terms of spectral irradiance and radiance, the former for an in-flight calibration monitor.

Instruments can also be calibrated using a laser illuminated integrating sphere, which can provide a source of spectrally tunable, stable, spatially uniform sources of radiation. This can easily be tailored to meet the specific needs of any EO sensor. NIST has used such sources in a number of measurements (Johnson et al. 2004) and NPL recently used such a source in a calibration of GERB 2 (Geostationary Earth Radiation Budget) due for flight onboard the Eumetsat satellite Meteosat Second Generation (MSG).

3.5 Traceability to SI in-flight

As shown in Figure 4 (right hand panel), the dominant issue for the EO community is establishing traceability to SI in-flight, and consequently determining accuracy and uncertainty of the primary data products: radiance or reflectance. Current use of onboard calibration systems limit accuracy to at best 3 percent because of the potential for degradation between preflight calibration and in-flight operation. They also are prone to degrade in-flight. Vicarious based methods have until now been limited to similar uncertainty levels due to difficulties in establishing traceability to SI. The Moon, for example, is highly stable and offers the prospect of an excellent calibration target if its absolute radiance can be determined with sufficient accuracy (Keiffer & Wildey 1996). Similarly deserts or lakes (for the TIR band) may offer large enough areas of spatially uniform terrain

to fill the pixel of the sensor (this is significantly easier with the high spatial resolution imagers). Together with improved transfer standards this offers the prospect of improved calibration accuracy. However, in this case, the poor long-term instability of the best-characterized sites, means that their calibration coefficients need to be regularly updated and ideally tied to the time of overpass. By contrast, other sites, e.g. Northern Africa, have been shown to be highly stable in the long-term, but due to inaccessibility are much more difficult to characterize on the ground.

4 TRACEABILITY

4.1 Traceability to SI and the Mutual Recognition Arrangement (MRA)

In October 1999, Directors of the NMIs of 38 countries, signatories of the convention of the meter, agreed to a new arrangement under which calibration and measurements certificates issued in one country are accepted in another without the need for further bilateral agreements. This arrangement is the Mutual Recognition Arrangement (MRA) (www.bipm.org). It operates through comparisons organized by the consultative committees, e.g. Consultative Committee of Photometry and Radiometry (CCPR) of the Comité International des Poids et Mesures (CIPM).

Most important are the “Key Comparisons” of the most basic quantities associated with each SI base unit (generally <10 for each SI unit) and involve a subset of NMIs that have a proven historical record of research activity in the technical area, and that also span the globe. The quantities chosen for these Key Comparisons are selected to evaluate the principle methodologies within a technology area so as to underpin the broader range of measurement services provided by NMIs. In context of radiometric measurements for the EO community, they are: spectral irradiance, radiance, responsivity, transmittance and diffuse reflectance. It is worth noting that spectral radiance is not one of these quantities, as in principle it is derivable from irradiance and reflectance, and uses the same principle techniques as irradiance. However, although not a Key Comparison, there is a “Supplementary Comparison” of spectral radiance being carried out under the auspices of the CCPR. Since this comparison is in terms of tungsten strip lamps, i.e. emitting areas of a few millimeters, it is of little practical use to the EO community.

The results of these Key Comparisons establish the degree of equivalence between participating NMIs for the specific quantities being compared. This is generally followed by a series of regional comparisons of the same quantity to link all the other NMIs into the

system. The results of all the comparisons, together with the degree of equivalence between each laboratory, are then entered into a database and available to all via the internet. This then underpins the process whereby all calibrations performed within any country can be compared to those in any other, providing that appropriate checks are made on the quality procedures of any secondary calibration laboratory. In addition to the results of comparisons, the database also contains an appendix that provides peer reviewed information (including uncertainties) for a wide range of standard measurement services. This appendix is structured so that each NMI entry is presented against a predetermined consistent set of measurement service categories to allow users to readily compare information.

A long-term goal for the EO community might be to establish a similar database of EO specific quantities populated by information from the calibration teams of the world about their measurement capabilities, but with the quality of the data underwritten by formal comparisons/audits organized by an independent authority, e.g. an NMI. Such information then could be reliably used to compare results made with different instruments or at different calibration sites, allowing the user to remove calibration offsets and biases easily without having to debate, which is the "correct" answer.

4.2 *Traceability to SI in the EO community*

Successful strengthening of traceability for EO data sets requires field personnel to have a clear understanding of its benefit. The principle benefit is to improve the likelihood of data products derived from a remote sensing instrument being a quantitative description of the biogeophysical parameter that it is seeking to measure, and that such measurements are invariant with time and are robust enough for regulatory, policy or commercial decisions.

While pre-launch activities help evaluate the extent to which the instrument meets specifications, it is in the post-launch environment that the issue of traceability to SI units becomes critical. This is particularly so for post-launch calibration of satellite sensors in the visible and near infrared where there are many examples of pre-launch calibration coefficients needing revision due to changes in the sensor caused by storage and launch into orbit. Frequent revisions need to be made throughout the life of the mission due to degradation of the instrument while in the space environment. Measurements by Los (1998) of NDVI (Normalized Difference Vegetation Index) for a nominally stable desert site demonstrate the difficulty of ensuring long-term stability in-flight, and consistency between instruments, even when they are of the same design. In the Los example, when there is

no onboard calibration system, results can be used as a vicarious ground calibration to apply correction factors to normalize results to a common baseline.

How such calibration updates should best be determined is a fairly hot debate between the advocates of onboard calibrators and those of vicarious techniques. The latter offers a range of options for determining correction factors usually based on solar reflected radiation from a "stable" natural target such as desert, snowfield, or the Moon. In most cases, there is little to choose between the techniques. All are adequate for monitoring change, but when determining radiometric accuracy in terms of SI units, all are relatively poor, with accuracies ranging from 3 to 10 percent. If such uncertainties are to be improved, then the following issues need to be addressed: (a) what is the attainable accuracy of radiation measurements in the visible and near-infrared with onboard calibrators?; (b) what is attainable through vicarious techniques?; (c) are the results from onboard and vicarious techniques in agreement?; and, (d) what is the accuracy for the user?

The solution to all these issues requires a common measurement system against which results can be compared. This is best achieved by ensuring that all measurements are made traceable to SI and have reliable estimates of their uncertainty. Without this, it is impossible to compare meaningfully the results and performance of different research groups, techniques, instruments, or calibration sites. It is important that traceable calibration includes not only preflight activities, but also all the equipment associated with ground calibrations and validations throughout the life of a measurement mission (which may exceed any one satellite's lifetime). These include the field spectroradiometers, both portable and site-based, aircraft spectroradiometers, sun photometers, and ocean buoys. These support instruments need to have associated with them well documented, auditable traceability chains demonstrating the uncertainty of the measurements made with the same rigour of the satellite mission. Each instrument should have been calibrated by an organization using a procedure that has in some way been through an auditable process to assure traceability to SI. This audit could take a number of forms, but is likely to be either a direct bilateral comparison, or participation in some more formal intercomparison or round-robin exercise. Traceability cannot be demonstrated sufficiently if workers rely solely upon a manufacturer's data sheet. Even a calibration certificate would require clear demonstrable evidence that the calibration has independent verification of its uncertainty statement.

It is, of course, important to emphasize that even if full traceability for these primary measurement quantities, e.g. radiance/reflectance is achieved, there are many additional sources of uncertainty that will effect

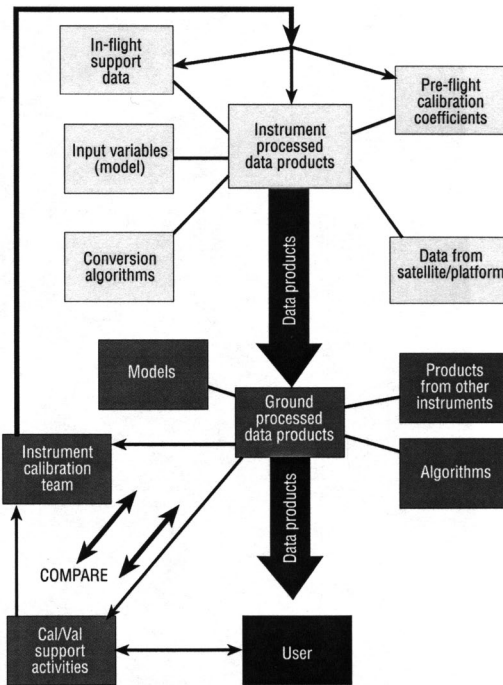


Figure 9. Schematic representation of the production process for a data product from a typical satellite sensor.

the final accuracy of the users' higher level data products due to various processing steps. Figure 9 illustrates how validation activities (comparing the processed data products with real variables) may also result in feedback to the process leading to further correction of the calibration coefficients. Such feedback mechanisms make it difficult to evaluate in an independent manner the real uncertainty of the final released data product. It is therefore essential that any strategy to improve the quality assurance of the EO product, includes all the steps in the production chain: data collection; correction; and processing, each having clear traceability for all input or measured variables. Data processing algorithms also require validation through an independent route, through appropriate testing and/or comparison to understand any associated biases.

Obviously, simple improvements in the traceability to SI of primary input variable, such as spectral radiances or reflectances, will not improve immediately the EO data product. Without demonstrable traceability in these primary measurands, however, it is hard to see how improvements in accuracy can be made in the processing chain, since any incremental gain is likely to be masked by large uncertainties of the input parameters.

One option for improving accuracy and traceability of EO data significantly would be to implement a dedicated calibration mission such as proposed in TRUTHS (Traceable Radiometry Underpinning Terrestrial- and Helio- Studies) (Fox et al. 2003a, 2003b). This proposal was for a mission to establish a set of calibration reference targets, Earth deserts, Sun, and Moon to transfer calibrations of radiance or irradiance to other in-flight sensors. Broadly, this is similar to current best practice for vicarious calibrations. However, in TRUTHS, the calibration coefficients of these targets would be updated regularly through observation and calibration by instruments onboard a small satellite. TRUTHS instruments would be calibrated using a novel onboard procedure that mimics those performed on the ground by NMIs when establishing primary scales. This procedure includes the flight of a primary standard, so traceability to SI can be established fully and regularly in-flight with very high radiometric accuracy (<0.5 percent for spectral radiance) avoiding any problems due to drift either pre- or post-launch.

While TRUTHS offers complete solution, the principles and techniques it proposes can be used independently. For example, the in-flight calibration system can be incorporated onto any Earth viewing imager.

Similarly, it is perhaps timely to consider establishing a global network of calibration test sites to acquire benchmark data sets. Global Instrumented And Networked Test Sites (GIANTS) has been proposed by Teillet et al. (2001). Use of "standard" ground reference calibration test sites as a means of cross-calibration and validation of satellite sensors is well established. In many cases, dedicated campaigns have been organized using teams supported by the respective instrument. In some cases, particularly atmospheric chemistry applications, use has been made of existing ground networks of validation equipment, much of which is automated. In the case of land imagers, some test sites have become recognized "standards," e.g. White Sands National Monument and Railroad Valley Playa in the western USA, and La Crau in Southern France. These and other sites have been well characterized and shown to be relatively homogeneous spatially, and are relatively stable temporally (at least in the short-term). However, significant differences have been observed by different sensor teams when using the same target area for vicarious calibration activities because of biases originating from subtle differences in the methodologies used, instrumentation, and calibration traceability. Such biases also can occur for networked sites, although these can be reduced by using common instrumentation and standard methodologies. Each site requires a common set of automated instrumentation, including sun photometers, standard

meteorological parameters, video images of the site in real time, down-welling solar irradiance, and surface spectral reflectance/radiance. All instruments should be automated and should transmit data independently. Continuous year round availability of a single calibration site is difficult to achieve and highly susceptible to local weather conditions, e.g. snow. However, having a global network of essentially interoperable test sites overcomes this limitation.

In context of the TRUTHS mission, data from the ground will correlate with absolute information from the satellite such that other satellite sensors need only be stable in the short-term, which is easier to achieve than absolute calibration. However, in the near-term, in the absence of such a satellite, calibration updates would need to be carried out by ground support teams or aircraft overpass.

6 CONCLUSIONS

Recent developments in the techniques used to establish primary radiometric scales, in particular the calibration of filter radiometers, makes it possible to improve significantly the accuracy and traceability of EO instrumentation. These improvements and techniques also make it possible to carry out detailed comparisons to understand and remove biases that may exist between calibration methods employed by different calibration teams, both in instrument calibration and *in-situ* calibration/validation activities.

The paper also has highlighted the importance of traceability for all steps in the production of an EO data product, from data collection to data processing. It recommends that this process is best carried out through some form of independent quality assurance (QA) assessment that likely will involve comparisons or audits and that NMIs are well placed provide. It further suggests that the results of such QA assessments could be stored in a publicly accessible database to aid users when trying to compare and combine data sets from different sources.

In looking to the future, options were briefly described that offer the possibility for improving significantly the accuracy and traceability of many EO data products. The adoption of one or more of these options is likely to be essential to ensure that the accuracy of future EO products fully meets the needs of the user community. This is particularly true for those studying climate change, but it is also true for customers of future commercial operational services.

REFERENCES

Anderson, V.E. & Fox, N.P. 1991. A new detector based spectral emission scale. *Metrologia* 28(3): 135–140.

- Anderson, V.E., Fox, N.P. & Nettleton, D.H. 1992. A novel, highly stable, monochromatic and tunable optical radiation source and its application to high accuracy spectrophotometry. *Appl. Opt.* 31(4): 536–545.
- Brown, S.W., Eppeldauer, G.P. & Lykke, K.R. 2000. NIST facility for spectral irradiance and radiance responsivity calibrations with uniform sources. *Metrologia* 37(5): 579–583.
- CGPM 1995. *Resolutions adopted by the 20th Conference Generale des Poids et Mesures*. Paris: Bureau International des Poids et Mesures.
- Chunnillall, C.J., Clarke, F.J.J. & Shaw, M.J. 2003. Diffuse reflectance scales at NPL. In A. Springsteen & M. Pointer (eds), *Fourth Oxford conference on spectroscopy*. Proc. SPIE (4826): 129–136.
- Dinguirard, M. & Slater, P.N. 1999. Calibration of space-multispectral imaging sensors: a review. *Rem. Sens. Env.* 68(3): 194–205.
- Fox, N.P. 1991. Trap detectors and their properties. *Metrologia* 28(4): 197–202.
- Fox, N.P., Martin, J.E. & Nettleton, D.H. 1991. Absolute spectral radiometric determination of the thermodynamic temperatures of the melting/freezing points of gold, silver and aluminium. *Metrologia* 28(6): 357–374.
- Fox, N.P. 1995. Improved radiometric calibration techniques for multi-spectral remote sensing satellite instrumentation. In E.T. Engman, G. Guyot & C.M. Marino (eds), *Remote sensing for agriculture, forestry, and natural resources*. Proc. SPIE 2585: 257–266.
- Fox, N.P. 1996. Radiometry with cryogenic radiometers and semiconductor photodiodes. *Metrologia* 32(6): 535–534.
- Fox, N.P., Haycocks, P.R., Martin, J.E. & Ul-haq, I. 1996. A mechanically cooled portable cryogenic radiometer. *Metrologia* 32(6): 581–584.
- Fox, N.P., Chunnillall, C.J. & White, M.G. 1998. Detector based transfer standards for improved accuracy in spectral irradiance and radiance measurements. *Metrologia* 35(4): 555–561.
- Fox, N.P. 2001. Developments in optical radiation measurement at NPL: part I. In C. Londono (ed.), *Harnessing light: optical science and metrology*. Proc SPIE 4450: 140–154.
- Fox, N.P., Aiken, J., Barnett, J.J., Briottet, X., Carvell, R., Frohlich, C., Groom, S.B., Hagolle, O., Haigh, J.D., Kieffer, H.H., Lean, J., Pollock, D.B., Quinn, T., Sandford, M.C.W., Schaepman, M., Shine, K.P., Schmutz, W.K., Teillet, P.M., Thome, K.J., Verstraete, M.M. & Zalewski, E. 2003a. Traceable Radiometry Underpinning Terrestrial and Helio- Studies (TRUTHS). In H. Fujisada, J.B. Lurie, M.L. Aten & K. Weber (eds), *Sensors, systems, and next-generation satellites VI*. Proc. SPIE 4881: 395–406.
- Fox, N.P., Aiken, J., Barnett, J.J., Briottet, X., Carvell, R., Frohlich, C., Groom, S.B., Hagolle, O., Haigh, J.D., Kieffer, H.H., Lean, J., Pollock, D.B., Quinn, T., Sandford, M.C.W., Schaepman, M., Shine, K.P., Schmutz, W.K., Teillet, P.M., Thome, K.J., Verstraete, M.M., Zalewski, E. 2003b. Traceable Radiometry Underpinning Terrestrial- and Helio-Studies (TRUTHS). *Adv. Space Res.* 32(8): 2253–2261.
- Geist, J., Lind, M.A., Schaefer, A.R. & Zalewski, E.F. 1977. *Natl. Bur. Stand. U.S. Tech. Note 954*.
- ISO 1993. *International vocabulary of basic and general terms in metrology* (2nd ed). Geneva: International Organization for Standardization.

- Johnson, B.C., Cromer, C.L., Saunders, R.D., Eppeldauer, G., Fowler, J., Sapritsky, V.I. & Dezsi, G. 1993. Cryogenic solar absolute radiometer. *Metrologia* 30(4): 309–315.
- Johnson, B.C., Brown, S.W. & Rice, J.P. 2004. Metrology for remote sensing radiometry. In S.A. Morain & A.M. Budge (eds), *Post-launch Calibration of Satellite Sensors*. Rotterdam: Balkema.
- Kieffer, H.H. & Wildey, R.L. 1996. Establishing the Moon as a spectral radiance standard. *J. Atm. & Oceanic. Tech.* 13(2): 360–375.
- Los, S.O. 1998. Estimation of the ratio of sensor degradation between NOAA-AVHRR channels 1 and 2 from monthly NDVI composites. *IEEE Trans. Geosci. & Rem. Sens.* 36: 206–213.
- Martin, J.E., Fox, N.P. & Key, P.J. 1985. A cryogenic radiometer for absolute radiometric measurements. *Metrologia* 21(4): 147–155.
- Quinn, T.J. & Martin, J.E. 1985. A radiometric determination of the Stefan-Boltzmann constant and thermodynamic temperatures between -40°C and $+100^{\circ}\text{C}$. *Phil. Trans. Roy. Soc. London* 316: 85–181.
- Rice, J.P. & Johnson, B.C. 1998. The NIST EOS thermal infrared transfer radiometer. *Metrologia* 35(4): 505–509.
- Sperfelf, P., Raatz, K.H., Nawo, B., Möller, W. & Metzendorf, J. 1996. Spectral-Irradiance scale based on radiometric blackbody temperature measurements. *Metrologia* 32(6): 435–439.
- Sperling, A. & Bentlage, V. 1998. A stabilised transfer-standard system for spectral irradiance. *Metrologia* 35(4): 437–440.
- Teillet, P.M., Thome, K.J., Fox, N. & Morissette, J.T. 2001. Earth observation sensor calibration using a global instrumented and automated network of test sites (GIANTS). In T. Zhang, B. Bhanu, & N. Shu (eds), *Image extraction, segmentation, and recognition*. Proc. of SPIE 4550: 246–254.
- Theocharous, E., Fox, N.P., Sapritsky, V.I., Mekhontsev, S.N. & Morozova, S.P. 1998. Absolute measurements of black-body emitted radiance. *Metrologia* 35(4): 549–554.
- Walker, J.H., Saunders, R.D., Jackson, J.K. & Mielenz, K.D. 1991. Results of a CCPR intercomparison of spectral irradiance measurements by national laboratories. *J. Res. Natl. Inst. Stand. Tech.* 96(6): 647–668.
- White, M.G., Fox, N.P., Ralph, V.E. & Harrison, N.J. 1996. The characterisation of a high temperature blackbody as a basis for the NPL spectral irradiance scale. *Metrologia* 32(6): 431–434.
- Windsor, S.A., Harrison, N.J. & Fox, N.P. 2000. The NPL detector stabilized irradiance source. *Metrologia* 37(5): 473–476.
- Woolliams, E.R., Hunt, T.M., Harrison, N.J., Windsor, S.A. & Fox, N.P. 2002. Improved transfer standards sources for calibration of field spectrometers used for Earth observation applications. In H. Fujisada, J.B. Lurie, M.L. Aten & K. Weber (eds), *Sensors, systems, and next-generation satellites VI*. Proc SPIE 4881: 386–394.
- Woolliams, E.R. 2003. *Development and evaluation of a high temperature blackbody source for the realisation of NPL's primary spectral irradiance scale*. PhD Thesis, University of Manchester.
- Yoon, H.W., Gibson, C.E. & Barnes, P.Y. 2002. Realisation of the National Institute of Standards and Technology detector based spectral irradiance scale. *Appl. Opt.* 41(28): 5879–5890.
- Zalewski, E.F. & Duda, C.R. 1983. *Appl. Opt.* 22(18): 2867–2873.

Status of standardization project ISO 19130

W. Kresse

University of Applied Sciences, Neubrandenburg, Germany

ABSTRACT: Worldwide official standards are developed by the International Organization for Standardisation (ISO) and two other similar organizations. The ISO Technical Committee 211 (ISO/TC 211) is responsible for geographic information including imagery data. In 2001, ISO project 19130 “Sensor and data models for imagery and gridded data” was approved. The project team has developed an International Standard for auxiliary data that relate imagery type data to the Earth. The focus of the first phase of standardization projects by ISO/TC 211 was geographic vector data. In the second phase, the project dealt primarily with imagery type data. Two related new standardization projects are ISO 19101-2 “Reference model – imagery” and ISO 19115-2 “Metadata – Part 2: Metadata for imagery and gridded data.”

1 INTRODUCTION

After decades of rapid innovation, remote sensing and geographic information systems have reached maturity. Today many excellent or exotic ideas do not survive because commercial success has gained the first priority. Photogrammetry has moved from its expert's corner to wider acceptance. Standardization has become a top issue.

2 ROLE OF ISO

ISO is one of several international standards-developing organizations. Others include the International Electrotechnical Commission (IEC) and the International Telecommunication Union (ITU). The Bureau International des Poids et Mesures (BIPM) sometimes is considered an international standards organization whose primary focus is on standard international (SI) units.

ISO was founded in 1947. The letters “ISO” are derived from the Greek word “isos” meaning “equal.” The full name is “International Organization for Standardization.” IEC and ITU have their specific topics. The ISO is responsible for all other fields of standardization that are not covered by the IEC and the ITU. IEC and ITU follow standardization procedures that are comparable to those of the ISO.

The ISO and the IEC have formed the Joint Technical Committee-1 (JTC-1) that is dedicated to all questions of computer science. Many well-known

programming languages and imagery formats have become standards developed by this committee.

ISO, as well as IEC and ITU, are official organizations. They are composed of the national standardization organizations and of external liaison members. The national standardization organizations represent individual nations. They have the right to vote. The external liaison members are technical international organizations that provide expertise to the development process and that have the right to take part in the discussions during the meetings. In addition there are internal liaison members that are other ISO Technical Committees serving in neighboring fields of work.

The ISO is subdivided into Technical Committees consisting of Working Groups and in some cases, Subcommittees.

The development of ISO standards follows a formal procedure that takes up to five years. It has recently been shortened from seven years to keep pace with accelerated technical developments. After the publication of an International Standard, it is reviewed at least every five years. Technical committees such as the ISO/TC 211 are standing committees. It means that they care for their standards on a long-term basis that is enabled by agreements with the national standardization organizations around a guaranteed annual budget.

Standards documents can be purchased via the Internet. Details are on the homepages of the organizations. In some cases, standards can be downloaded from the Internet at no cost, or they may be found in other publications.

3 ROLE OF ISO/TC 211 "GEOGRAPHIC INFORMATION/GEOMATICS"

ISO/TC 211 (Geographic information/ Geomatics) is one of about 190 Technical Committees of ISO. According to its name, it develops standards in the field of geographic information. According to the ISO, geographic information is a specialization of computer science. The title "Geographic information" is meant in a fairly wide sense. This is best reflected in the spectrum of the presently more than 20 external liaison members. They include organizations like the Committee on Earth Observation Satellites (CEOS), International Society for Photogrammetry and Remote Sensing (ISPRS), International Association of Geodesy (IAG), Open GIS Consortium (OGC), International Hydrographic Organization (IHO), Food and agriculture Organization of the United Nations (FAO), Digital Geographic Information Working Group (DGIWG), and regional initiatives.

As of 2003, the work of ISO/TC 211 is subdivided into five Working Groups and about 40 project teams. Each project team results in one International Standard, or in some other ISO-deliverable. The standards range in number from ISO 19101 to ISO 19140. Some project teams already have finished their work and their standards have been published.

From the ISO perspective, photogrammetry and remote sensing are a subtopic of geographic information. Therefore, it is advisable to allocate future standards for remote sensing to ISO/TC 211. However, relevant standards may have been developed already by other ISO technical committees, e.g. photography, intelligent transport systems, or the ISO/IEC JTC-1. for hypertext markup language (HTML), virtual reality markup language (VRML), or basic image interchange format (TIFF).

4 TYPE OF PROJECTS IN ISO/TC 211

Standardization projects may be categorized in one of at least four types: infrastructure standards, basic standards, imagery standards, and implementation standards. Those categories are not official, but they help to understand the overall structure.

Infrastructure standards serve as a basis for the others and include the reference model (ISO 19101), terminology (ISO 19104), testing of conformance to the standards (ISO 19105), and the creation of profiles (ISO 19106). Those standards have all been completed.

Basic standards cover the classical world of GIS. They include the application model (ISO 19109), spatial schema for vector-data processing (ISO 19107), methodology for feature cataloguing (ISO 19110), the coordinate reference model (ISO 19111), metadata

(ISO 19115), portrayal (ISO 19117), encoding (ISO 19118), and others. Those standards have been completed also. They are profiled for specific needs and are implemented in various GIS environments. Their completion made the ISO/TC 211 a powerful player in the GIS-world.

Imagery standards are under development. The concept includes imagery, gridded, and coverage data. Imagery data have pixels that represent a physical value such as light or temperature. The resulting images may be pictures or temperature maps. Gridded data do not necessarily represent a physical value such as population data. Coverages have been standardized in ISO 19123. A coverage is a widely used concept in geographic information systems. It is well comparable with the concept of a "field" in physics.

Implementation standards are not considered the primary goal of ISO/TC 211, but they are necessary to bridge the gap between the original abstract standards of ISO/TC 211 and users. In addition, many requirements from specific user communities need a home for finalizing their standardization process. Examples for implementation standards are the Web Map Server interface (ISO 19128), and the Geography Markup Language (GML, ISO 19136). It is still questioned whether the implementation standards are well-positioned within the ISO/TC 211 suite of standards (Kresse & Fadaie 2003).

5 RELEVANT PROJECTS FOR PHOTOGRAMMETRY AND REMOTE SENSING

Since the ISO/TC 211 has focused its work on imagery, interest among ISO experts in photogrammetry and remote sensing has grown. Both sciences are a major data source for GIS. Though technical details of individual sensors or systems are not a topic for ISO-standards, both sciences were considered mature enough to standardize all important data interfaces between sensor and GIS.

Imagery has been split into a number of standardization projects. Work began almost three years ago when the first two projects were approved: "Imagery, gridded, and coverage data framework" (ISO 19129) and "Sensor and data models for imagery and gridded data" (ISO 19130). During 2003 two other projects were approved: "Reference model – imagery" (ISO 19101-2 and "Metadata – Part 2: Metadata for imagery and gridded data" (ISO 19115-2). Further standardization projects on encoding for imagery and calibration/validation of sensors are planned (Di 2004).

Work on "Coverage geometry and functions" (ISO 19123) was the earliest to start. Though not completely finished, the document has matured. Because of its very generic approach towards the definition of

images, ISO 19123 has influenced the creation of all other standards for imagery.

6 FRAMEWORK (ISO 19129)

The goal of “Imagery, gridded data, and coverage framework” (ISO 19129) is to create a roadmap towards individual imagery standardization projects within the ISO/TC 211. ISO 19129 contains a decision support list for all existing standards of the ISO 19100 family. Each entrance is a work item necessary to qualify the existing ISO 19100 set of standards for imagery purposes. In some cases minor or even no amendments were necessary for an extended validity

Table 1. Distribution of imagery topics to the existing standards of ISO/TC 211 according to ISO 19129.

ISO 19100 Standards	Imagery and gridded data elements	Solutions
19101		ISO 19101-2
19104		
19105	TBD	Annex to ISO
19107	Surfaces required	Make appropriate references. e.g. to GM_Tin
19109	Apply rules to imagery, gridded data, and coverages	Create template imagery application schemas in ISO 19129
19110	No additional feature elements necessary for imagery and gridded data	No action
19111	Need for Earth Centered Inertial coordinate reference system (ECI) and Earth Centered Rotating coordinate reference system (ECR)	Apply 19111 to ECI and ECR in 19130
19115	FGDC IHO: S61 DGIWG Quality elements (ISO 19113, ISO 19114)	ISO 19115-2
19117	Look-Up Table, color maps Color-systems: Commission International de l'Eclairage(CIE) Munsell etc. Scaling and rotation (resampling) Styling: ISO 19136 (GML) W3C SMIL (Synchronized Multimedia Integration Language)	Present: discussion in ISO 19129; Future: technical amendment to ISO 19117

to include imagery. In some cases new ISO standards are required. The ISO 19101-2 and the ISO 19115-2 are examples for such cases. Table 1 is a shortened version of the decision support list.

In addition, some imagery topics of major importance and minor size are a part of the ISO 19129 itself. The related sections of the document deal with the structure of gridded data, with n-dimensional grids, with the shape of grid cells, with radiometric processing, and with data compression.

7 REFERENCE MODEL (ISO 19101-2)

ISO 19101-2 provides a reference model for imagery based on concepts developed in computer science. The core model is standardized as ISO/IEC 10746 “Information technology – Open distributed processing – Reference model.” It describes five points of view that enable a comprehensive understanding of a computer system before, and during, its development.

Three of the points of view are relevant for imagery standardization: (a) *computational* describes the functional decomposition of a system; (b) *information* describes the semantics of a system; and (c) *engineering* describes for example, the interaction between sensor and platform. To draw a complete picture, the remaining two points of view (enterprise and technology) are mentioned, though they have a very limited meaning to imagery standardization.

Standards of the ISO 19100 family follow a model-driven approach. As a consequence, the core elements of every standard are mandatory models written in a conceptual schema language, which is the Unified Modeling Language (UML) in the ISO 19100 family. The models of all standards are bound to a comprehensive UML model that governs the whole ISO/TC 211 series. For imagery standards this top level model is provided by the ISO 19101-2.

The top level model distinguishes between two different schemes for classifying imagery information, operations on imagery, and refinement of imagery information.

The model provides six classes of *operation*: (a) image sensing, (b) image processing, (c) image analysis, (d) image interpretation, (e) image synthesis, and (f) image presentation. It also provides five classes of *image refinement*: (a) data assimilation, (b) data fusion, (c) knowledge from imagery, (d) identifying features from imagery, and (e) decisions based on geographic imagery (Percivall 2003).

8 SENSOR AND DATA MODELS (ISO 19130)

The primary goal of “Sensor and data models for imagery and gridded data” (ISO 19130) is the provision

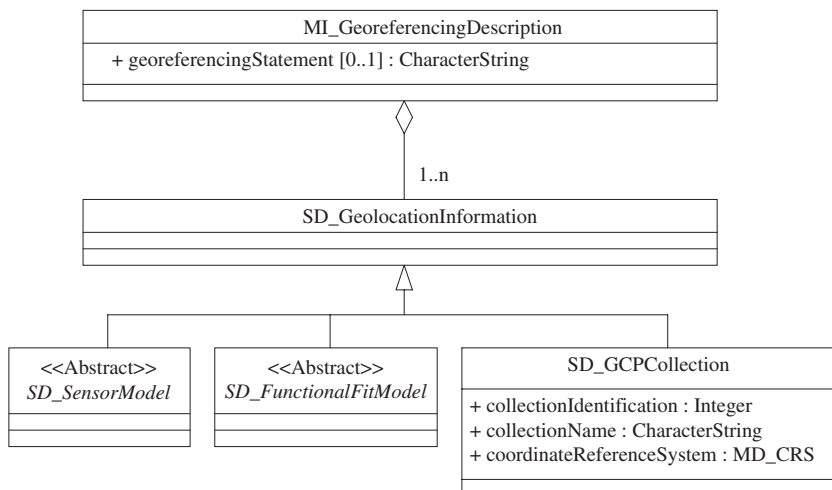


Figure 1. Sensor model, functional fit model, and GCP collection.

of parameters necessary to georeference imagery. The standard includes sensor models that are traditionally applied in the domain of photogrammetry, and the functional fit models that are widely used in remote sensing. Ground control points (GCPs) may be used on their own as a georeference or in conjunction with sensor models or functional fit models. The GCPs reside in a category called “sensor and data models” (SD), or more specifically, *SD_GCPCollection* (Fig. 1). ISO 19130 addresses all sensors that are common in the domain of ISO/TC 211. However, in the version of the standard that has been elevated to the committee draft (CD)-stage, only linear sensors, frame cameras, paper and film scanners, and a “virtual sensor” remained. Light Detection and Ranging (LiDAR), Synthetic Aperture Radar (SAR) and hydrographic sonar were deleted because of the lack of technical contributions. By convention within TC 211, names of UML classes include a two-letter prefix that identifies the standard and the UML package in which the class is defined. In the following figures, SD stands for “sensor and data model” and MI stands for “meta-data for imagery”.

8.1 Sensor models

A sensor model describes the physical reality of a sensor in mathematical terms as far as necessary for a rigorous 3-D transformation of points with a controllable error budget.

The sensor model consists of the sensor type, the location model, the platform, and the process (Fig. 2). Sensor components are associated with the sensor model. A specific sensor such as the scan linear array

or the frame camera consists of components that are taken from the pool of components defined in the class *SD_Sensor-Components*. The sensor optic with its calibrated focal length is an example of a component; film with fiducial marks is another. The georeference of the sensor is defined in the location model. It contains the exterior orientation parameters, the orbit parameters, and others.

In remote sensing, sensors are often referenced to their platform. The parameters involved are defined in the class *SD_Platform*. This class includes stationary and dynamic platforms, and distinguishes between spaceborne, airborne, and landborne platforms. The geometric relation between a sensor and its platform is called “the mounting of the sensor.”

The virtual sensor shown in Figure 3 is a collection of hardware-describing parameters that are independent of any type of sensor. These parameters are grouped in the simple case for constant and “unlimited” translation and rotation; and in the linear case, for a more detailed description such as start and stop distance. The virtual sensor enables the application of standards to new sensors even though they might not fall into one of the other categories of sensor types.

Coordinate systems (CSs) are an essential part of a sensor model. They are attached to physical units such as the sensor, the platform, and the Earth. Different applications use different transformation pipelines between the CSs. For instance, photogrammetry usually sets a transformation between the sensor (camera) CS and the projected (ground) CS.

A complete sensor model takes all physical steps into account, beginning with the data taken at the focal plane of an imaging system to the creation of a

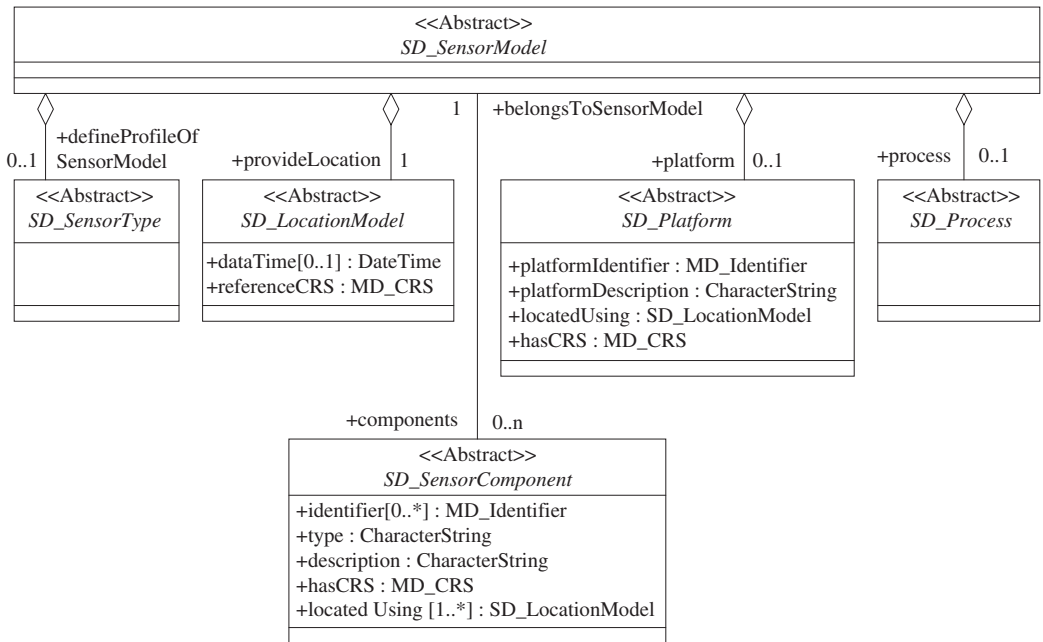


Figure 2. Sensor model with sensor type, location model, platform, process, and sensor components.

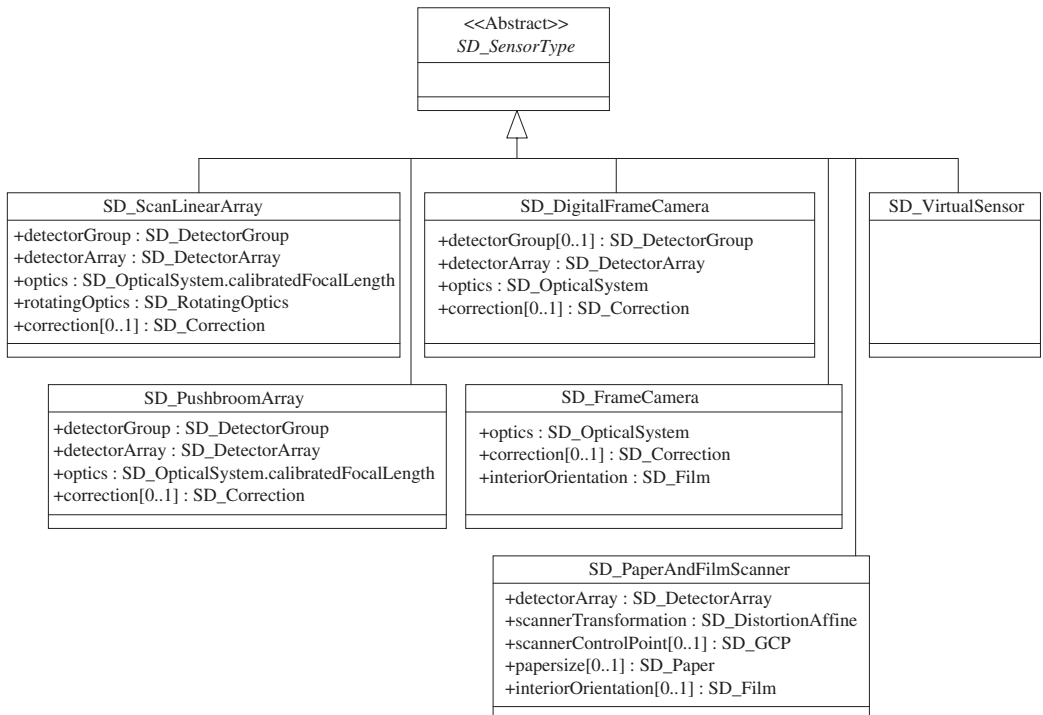


Figure 3. Sensor types described in ISO 19130.

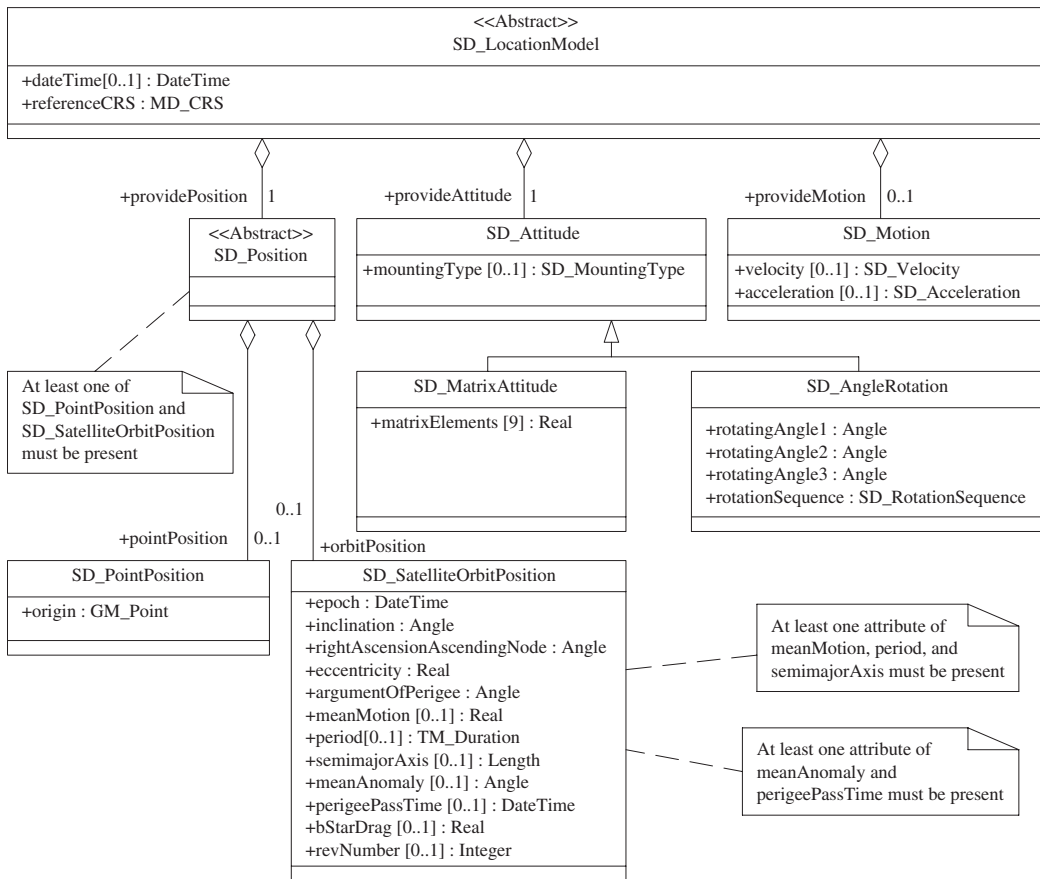


Figure 4. Location model with position, attitude, and motion.

georectified data set. The process requires a sequence of transformations that are related to coordinate systems (CSs). The ISO 19130 defines the sensor coordinate system (sensor CS), the platform CS, the orbit CS, the ECR CS, the ECT CS, the projected CS, and some specialized CSs such as the stereo model CS.

The majority of the coordinate systems in ISO 19130 are defined as coordinate reference systems within the frame provided by the ISO 19111 “Spatial referencing by coordinates.” The ECI CS cannot be derived yet from the ISO 19111.

One of the major contributors to the development of ISO 19130 is the sensor model language (SensorML). It is an extensible mark-up language (XML) based implementation of sensor models that has been developed independently of ISO. As the models of ISO 19130 and SensorML are similar, both developments are merged presently to enable SensorML to function as an implementation of the abstract standard ISO

19130 (Botts 2002). Figures 2–5 show the sensor, sensor type, location, and functional fit models.

8.2 Functional fit models

A functional fit model provides relationships between an image and a ground coordinate system. The simplest functional relation is an affine transformation with six parameters. Often used ground coordinates are the state plane coordinates of a projected coordinate system.

ISO 19130 defines four types of functional relations: (a) *polynomial fit*, (b) *ratios of polynomials*, (c) *universal real time model*, and (d) *grid interpolation*. The *polynomial fit* model has two-dimensional polynomials to describe the relation. Though the grade of the polynomials may have any number, it is hardly ever greater than three. The *ratios of polynomials* model has ratios of two-dimensional polynomials,

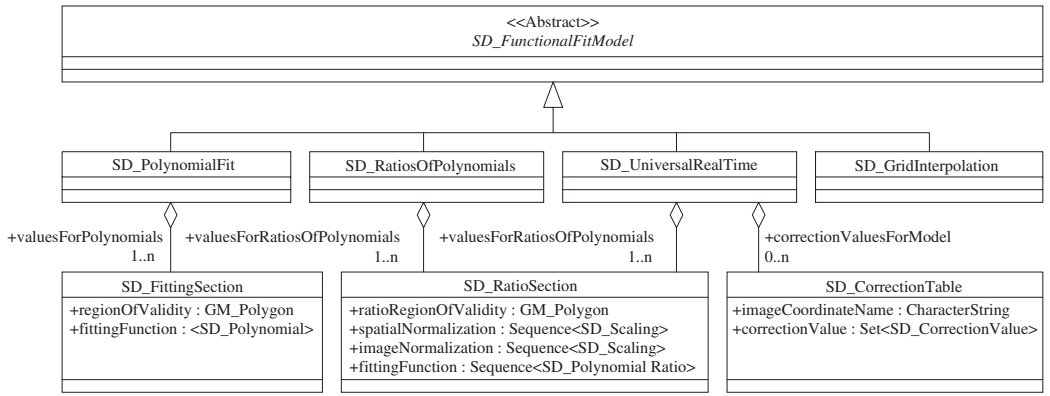


Figure 5. UML model of the functional fit model.

while the *universal real time model* has piecewise ratios of polynomials. The pattern of these pieces is mostly rectangular. Finally, the universal real time model is often called a rapid positioning model as it allows precise online positioning calculations even with slower computers than today.

The *grid interpolation* model provides image-to-ground relations at the vertices of a grid pattern. Values within the grid have to be interpolated.

The basic equations for the *polynomial fit* model and the *ratios of polynomials* in the functional fit model are:

Polynomial fit model:

$$p = \sum_{k=0}^{r_1} \sum_{l=0}^{r_2} \sum_{m=0}^{r_3} a_{klm} x_n^k y_n^l z_n^m \quad (1)$$

Ratios of polynomials model:

$$i_n = [p_1(x_n, y_n, z_n) / q_1(x_n, y_n, z_n)] \quad (2)$$

$$j_n = [p_2(x_n, y_n, z_n) / q_2(x_n, y_n, z_n)] \quad (3)$$

with

$$p_i = \sum_{k=0}^{r_1} \sum_{l=0}^{r_2} \sum_{m=0}^{r_3} a_{klm} x_n^k y_n^l z_n^m \quad (4)$$

$$q_i = \sum_{k=0}^{r_1} \sum_{l=0}^{r_2} \sum_{m=0}^{r_3} b_{klm} x_n^k y_n^l z_n^m \quad (5)$$

8.3 Timeframe for ISO 19130

Development of an ISO standard follows a well defined timetable and must not take longer than five years. The ISO 19130 project was approved by the

members of ISO/TC 211 in 2001. By this approval, the project changed its status from “New Work Item Proposal” (NWIP) to “Working Draft” (WD). For more than two years the project team has tried to find consensus on the working draft document. The long time was mainly caused by insufficient expert contributions for the subtopics LiDAR, SAR/InSAR, and hydrographic sonar. In addition, the ideas of SensorML delivered a major input to the structure of the ISO 19130 document. In early 2004, the project team reached consensus about the future ISO standard 19130.

The working draft has changed its status to “Committee Draft” (CD). A committee draft is a public document within the Technical Committee. The CD will be reviewed by members of the other project teams and, in particular, by representatives of liaison organizations. Within the CD status, a standard shall become mature. People who need to have their views included as part of the new standard should contact their national standardization organizations or one of the liaison members to gain access to the document and to place their comments.

Once the CD stage is completed, the document is forwarded to the central secretariat of ISO in Geneva. Subsequent drafts are called “Draft International Standard” (DIS) and “Final Draft International Standard” (FDIS). Usually no major technical changes are applied to the document during these phases. Finally the ISO central secretariat publishes the document as an “International Standard” (IS).

9 METADATA FOR IMAGERY (ISO 19115-2)

The already published International Standard “Metadata” (ISO 19115) provides guidelines for describing geographic information and services. It contains the complete listing of all metadata elements

of the ISO 19100 family as far as the published standards are concerned. The requirements for imagery and gridded data were not fully defined at the time the standards were drafted.

“Metadata – Part 2: Extensions for imagery and gridded data” (ISO 19115-2) is presently under development. It will include all metadata elements that are addressed in the ISO 19100 standards for imagery and gridded data. The listings will not be created from scratch. Rather, existing metadata catalogues will be used. Prominent examples of such catalogues are the Content Standards for Digital Geospatial Metadata (FGDC 1998), the IHO:S61, DGIWG, and the Deutsches Institut für Normung (DIN 1995).

Metadata listings of ISO 19115-2 also will contain the elements required to georeference imagery according to ISO 19130. The aforementioned standard mainly explains how these elements must be applied to establish unambiguously the transformation pipeline between the sensed image data and coordinates on the ground, or in some other user environments.

10 ENCODING FOR IMAGERY

The standardization project “Encoding for imagery” has not yet started. The need for a standard is obvious. The most successful former development is the GeoTIFF format that combined a TIFF image with header information about the georeference of the data (GeoTIFF 2003). The technical and semantic limitations of GeoTIFF are known. However, it is difficult to establish a better or more comprehensive solution.

Two aspects of the problem are the great number of existing and already standardized image data formats, and awareness about the long-term validity of decisions to be made. Presently, the existing image data standards are reviewed. In particular, the work of the ISO/IEC JTC-1/Subcommittee 24 “Computer graphics and image processing” and the *de-facto* standards need special consideration. Important examples of these standards are TIFF and BIFF.

Once the prerequisites have become clearer, the new ISO standards certainly will be built upon the combination of existing standards or *de-facto* standards for image data and an environment described in XML.

11 COVERAGES

Coverages represent a concept to handle spatial and temporal data and their attributes. Coverages may have two or three spatial dimensions and zero or one temporal dimension. Space and time are subdivided into a pattern of vertices. Each position is the anchor

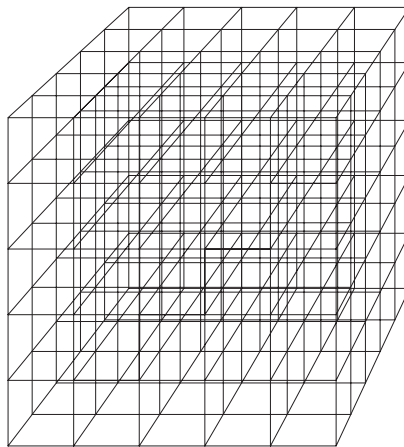


Figure 6. Three-dimensional coverage.

point for an associated record of attributes. Space and time are called the spatiotemporal domain. The attributes form an attribute range.

A 2-D grid with a soil-type attribute at each grid point is a simple example of a coverage. A 3-D grid (Fig. 6) with a temperature attribute at each grid point may represent an ocean-body. A digital image with gray value information at each pixel may be viewed as a coverage also. Coverages represent the most general concept to describe digital imagery.

The general conceptual approach and the almost completed development of the “Coverage geometry and functions” (ISO 19123) make the coverage standard a basic component for all other imagery and for gridded data standards of the ISO/TC 211 family. As stated in ISO 19123, *coverages support mapping from a spatiotemporal domain to feature attribute values where feature attribute types are common to all geographic positions within the spatiotemporal domain. A spatiotemporal domain consists of a collection of direct positions in a coordinate space that may be defined in terms of up to three spatial dimensions as well as a temporal dimension. Examples of coverages include rasters, triangulated irregular networks, point coverages, and polygon coverages. Coverages are the prevailing data structures in a number of application areas, such as remote sensing, meteorology, and mapping of bathymetry, elevation, soil, and vegetation.*

12 CONCLUSIONS

ISO 19130 standardizes the models and the associated parameters for georeferencing photogrammetric and remote sensing imagery. The development of this

standard is almost completed. Other ISO standards for imagery and gridded data, such as reference model and metadata, are under development.

REFERENCES

- Botts, M. 2002. Sensor model language (SensorML) for in-situ and remote sensors. *Open GIS Consortium (OGC)*. Document 02-026, Huntsville: Univ. of Alabama, Huntsville.
- Di, L. 2004. A proposed ISO/TC 211 standard project on radiometric calibration and validation of remote sensing data. In S.A. Morain & A.M. Budge (eds), *Post-launch Calibration of Satellite Sensors*. Rotterdam: Balkema.
- DIN 1995. Photogrammetry and Remote Sensing – Part 1: General terms and specific terms of photogrammetric data acquisition. *Deut. Inst. für Normung (DIN)* 18716-1. Berlin: Beuth.
- FGDC 1998. Content standard for digital geospatial metadata. FGDC-STD-001-1998 *Federal Geographic Data Committee*: Washington D.C. 78 pp.
- GeoTIFF 2003. *GeoTIFF-homepage*.
<http://www.remotesensing.org/geotiff/geotiff.html>
- Kresse, W. & Fadaie, K. 2003. *ISO Standards for Geographic Information*. Springer: New York, Heidelberg.
- Percivall, G. 2003. Geographic imagery reference model, preparatory work. *Presentation at the ISO/TC211 meetings, Berlin*.

A proposed ISO/TC 211 standards project on radiometric calibration of remote sensing data

L. Di

*Laboratory for Advanced Information Technology and Standards, George Mason University,
Fairfax, Virginia, USA*

ABSTRACT: ISO/TC 211 is the technical committee responsible for setting international standards on geographic information. Its Working Group 6 (WG6) is responsible for developing imagery-related standards. In recent years, ISO TC 211 has worked intensively on setting standards for geographic imagery data, including those from remote sensing. The ISO 19130 project is setting a standard for information that will allow determination of the *geographic* location of sensor data, which is provided with its location given in a coordinate system relative to the sensor. ISO does not have standards to describe the *radiometric* properties of remote sensing data. Consequently, ISO/TC 211 WG6 decided to initiate preparatory work for setting an ISO standard on radiometric calibration and validation of remotely sensed data under the ISO 19130 project team. The decision was reported to the ISO/TC 211 Plenary in May 2003. This paper discusses the status of preparatory work, current thoughts on the scope and content of the planned standard, and the work plan. It also discusses the plan to cooperate with relevant international stakeholders for developing such an important international standard.

1 INTRODUCTION

Remotely sensed data are a major contributor to geospatial data. Large amounts of remotely sensed data have been collected by government agencies and the private sector. For example, the NASA EOS program generates more than two terabytes of data per day (King et al. 2003, McDonald & Di 2003). Most of these data are in the form of imagery that needs to be processed further to derive useful information. To facilitate further processing, useful metadata have to be available.

Remote sensing data are very useful in many aspects of socio-economic activities, ranging from military operations to natural disaster mitigation. A single type of data can be used in many diverse applications. On the other hand, current application developments call for integrated analysis of multiple types of remotely acquired data from multiple sources (Asrar & Dozier 1994). Required data for any given project may be collected and archived by multiple international space agencies and data providers. To do so requires international interoperability and sharing. To achieve interoperability, international standards on remote sensing data have to be developed.

ISO/TC 211 is a technical committee within the International Organization for Standardization (ISO) responsible for setting standards on geographic data

and information. Since remotely acquired data are a major form of geographic data, ISO/TC 211 has worked on setting standards in the past few years through its Working Group 6 (WG6): "Imagery." After several exploratory projects on imagery and gridded data, ISO/TC 211 began developing its first standard on imagery and remote sensing: ISO 19130 "Sensor and data models for imagery and gridded data" in March 2001 (Di et al. 2002). Currently, the standard is near the end of the development stage, and a committee draft (CD) is to be released for public review in early 2004.

Remote sensing data have two major sets of properties: radiometric and geometric. Standardizing the descriptions of these two sets is one of the keys for data interoperability and sharing, especially data at low levels of processing. Standardization of the *geometric* properties of remote sensing data is being covered by ISO 19130. However, there is no standard or standards project in ISO/TC 211 to facilitate description of *radiometric* properties of sensor data and products. Therefore, the ISO 19130 project team and WG6 feel it is necessary to create a radiometric standard for remote sensing data, and work toward this goal was initiated in May 2003. This paper summarizes the preparatory work for the planned ISO project.

2 ISO 19130 AND ITS RELATIONSHIP TO THE PLANNED NEW PROJECT

The ISO 19130 standard specifies a sensor model for describing the physical and geometrical properties of each kind of sensor producing imagery of type two or higher dimensions (Di et al. 2002). It also defines a conceptual data model that specifies the minimum content requirement for each kind of sensor and the relationship among the components of the raw data that were recorded in an instrument-based coordinate system. The purpose is to make it possible to geolocate and analyze the data. Figure 1 shows the top level Unified Modeling Language (UML) model from the ISO 19130 Committee Draft (CD). In the UML model, all classes starting with an SD prefix are defined in ISO 19130, while classes that start with MD or MI are defined in ISO 19115, or ISO 19115-2, respectively. The prefix VC indicates that they will be part of the UML model for validation and calibration. This VC prefix is tentative and may be different in the actual standard.

This top level UML model shown in Figure 1 defines a georeferenceable data set that consists of sensor measurements and metadata. Any georeferenceable data sets that comply with the ISO 19130 standard have made both components available to data users. Sensor measurements represent the matrix of values measured by the sensors. MD_Metadata has been defined by ISO 19115: “Geographic Information-Metadata,” which is

an International Standard. However, the metadata defined in ISO 19115 are general geospatial metadata (ISO 2003). Some of the specific metadata needed by remote sensing and imagery are lacking. Consequently, ISO 19115 is being extended, with FGDC remote sensing metadata as a basis (Di et al. 2000, FGDC 2002, Di 2003). These extensions are ISO 19115-2: “Geographic Information – Metadata – Part 2: Extensions for imagery and gridded data.” One of the classes defined in ISO 19115-2 is MI_GeolocationDescription. SD_Geolocation-Information is the top level geolocation class defined in ISO 19130. It specifies three types of geolocation information; namely, sensor models with platform information, ground control points (GCPs), and functional fitting. Because more than one kind of geolocation information may be supplied with the data, the aggregation of SD_GeolocationInformation into MI_Geo-referencingDescription has multiple cardinalities (1..n). Metadata for quality description has been defined in ISO 19113 “Geographic information – Quality principles,” ISO 19114 “Geographic information – Quality evaluation procedures,” and the ISO 19138 “Geographic information – Data quality measures” standard under development. However, additional quality description metadata are needed for imagery data. The only component that has not been covered by ISO/TC 211 is radiometry. It is likely that ISO 19115-2 will develop a general description of the radiometric properties of imagery and gridded data. However, it was decided that VC_Radiometry

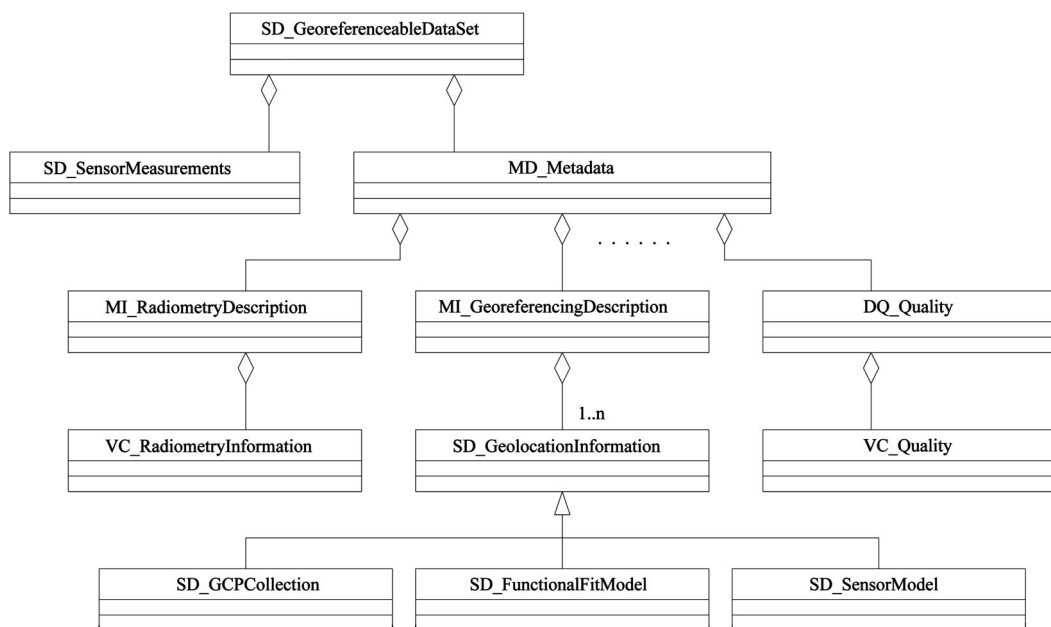


Figure 1. The top level UML model for a georeferenceable data set as shown in ISO 19130 CD 1.

Information and VC_Quality and their subclasses should be defined in a separate International Standard.

3 A PROPOSED ISO PROJECT ON RADIOMETRY

Large quantities of geographic data are being produced and distributed by government mapping and remote sensing agencies and commercial data vendors. Many of these data originate as measurements made by sensors such as frame cameras, whiskbroom scanners, pushbroom scanners, synthetic aperture radar, and lasers. Without a common standard, information required by the user to understand the radiometric characteristics of data from different providers may be in different forms, requiring them to understand and handle data from each data source separately. Existence of a common standard will allow data providers to package their calibration and validation information in a way that will make it accessible to the widest possible number of users, promoting interoperability of data between application systems and facilitating data exchange.

In May 2003, ISO/TC 211 WG6 met in Thun, Switzerland to discuss future standards development for imagery and gridded data. It was concluded that it was necessary to develop an ISO standard on the radiometric properties of remote sensing data. WG6 assigned the preparatory work for setting the new ISO standard to the ISO 19130 project team. Since then, some preparatory work has taken place. The project team discussed the matter in several project team meetings. The planned project also was discussed with other relevant international organizations, including the Committee on Earth Observation Satellites (CEOS), the International Society for Photogrammetry and Remote Sensing (ISPRS), and the IEEE Geosciences and Remote Sensing Society (IGARSS). Based on the results of those communications and discussions, a new work item proposal has been drafted and circulated for comments.

The tentative title for the new ISO standard is *Geographic Information-Radiometric calibration and validation of remotely sensed data*. The purpose of this International Standard will be to define standards for radiometric calibration and validation of the results. An essential aspect of deriving geographical information from sensor measurements is the conversion of raw data values into physical variables that can be interpreted in context of geographical information. For this conversion to be done, the instrument must be calibrated: its radiometric properties and its response to input signals must be known and available to data users. The utility of derived geographic information depends upon the quality of input data and of

the calibration information. Other standards in the ISO 19100 series define the principles for data quality and the measures for them. This standard will define the precise information that must be provided for different kinds of measuring systems as this process is carried out.

The standard will cover two groups of remote sensing data: radiometrically uncalibrated data, e.g. data still in sensor readout units such as digital numbers (DNs), and radiometrically calibrated non-thematic data, e.g. radiance images.

For uncalibrated data, the standard will define all necessary parameters required to derive physical values quantitatively from raw sensor measurements for each type of sensor. It also will specify the information needed to evaluate the quality of measurements, such as the response of the sensor system to input signals, and signal-to-noise ratio.

For calibrated, non-thematic data, the standard defines mechanisms for reporting the methods used for radiometric calibration/correction, the accuracy and precision of the calibration, and the methods for determining the accuracy and precision.

Development of this standard will complete the process of defining the content required to use and evaluate geographical information obtained through sensor measurements.

4 THE PROPOSED CONTENT OF THE STANDARD

Current thoughts on the content of the standard envision two parts. The first part will define the terms and definitions for parameters and measurements required for radiometric correction of raw remote sensing data acquired by each type of sensor. The second part will define the methods for reporting radiometric correction and validation procedures applied to radiometrically corrected data, and the radiometric fidelity of such data.

For this standard, remote sensors are divided into two groups, based on the way they work. Active sensors emit signals and measure those returned by objects being sensed. Examples of such sensors include LiDAR and Radar. Passive sensors measure signals emitted or reflected by objects being measured. The sensors themselves do not emit the signals. Examples of such sensors include frame cameras and linear scanning sensors.

Based on the wavelength at which the sensor works, each group of sensors will be further classified into five types: visible/near infrared (optical), thermal infrared, microwave, radio wave, sound wave. It is very common for a remote sensing instrument to have multiple bands working at different wavelengths. In this standard, each band (or group of bands

if they work in the same wavelength region) will be considered as a separate sensor because the parameters to describe the radiometric properties of the data from each band are different.

5 THE DEVELOPMENT PLAN

An ISO standard typically has the following development stages: Proposal, Working Draft (WD), Committee Draft (CD), Draft International Standard (DIS), Final Draft International Standard (FDIS), and International Standard (IS). The project team is responsible for developing the working draft and submitting it to the ISO/TC as a CD. Once the standard is in the CD stage, it will be reviewed by the public, and an editing committee will edit the CD based on comments from member nations and liaison organizations.

As a result of preparatory work, a new work item proposal has been drafted for comment by stakeholders. The proposal was briefed at the WG6 meeting in Berlin in October 2003, and also has been available on the WG6 Forum for comments. Many valuable comments have been received from members of WG6. Based on consensus between WG6 and the U.S. national body for ISO/TC 211, the proposal will be submitted to ISO/TC 211 through the body as a U.S. national contribution. The proposal must first be approved by the International Committee for Information Technology Standards/Technical Committee L1 INCITS/L1, which is responsible for setting U.S. national standards on geographic information and acts as the U.S. national body to ISO/TC 211. It is expected that the proposal will be submitted to INCITS/L1 for review and vote in mid 2004. If L1 approves the proposal, it will be submitted to ISO/TC 211 for vote by all principal members of the TC. If TC 211 approves, it is likely that a project team will be formed before the end of 2004, consisting of a project chair, an editor, and experts nominated by national and organizational members of ISO/TC 211 to write the working draft. Based on new ISO rules on standards development, the project team has 12 months to draft the standard before moving the draft into editing stage. Under the above development plan, the standard should reach the CD stage by June 2005.

6 CONCLUSIONS

Standard setting is a consensus-building process. A successful standard should be developed by as many

stakeholders as possible. Because of the importance of this planned standard to remote sensing, the international remote sensing community should participate fully in its development. The International Society for Photogrammetry and Remote Sensing (ISPRS), the Committee on Earth Observation Satellites (CEOS), and the IEEE Geoscience and Remote Sensing Society (IGARSS) have expressed interest in participating in developing a standard. The International workshop on Radiometric and Geometric Calibration organized by the ISPRS/CEOS joint task force in December 2003 was a very important venue for reporting the best international practices on radiometric and geometric calibration. It is expected that documents from the workshop will serve as one of the information sources for drafting international standards. It is also expected that the above-mentioned international organizations will nominate experts to the project team.

REFERENCES

- Asrar, G. & Dozier, J. 1994. *EOS: Science strategy for the Earth Observing System*. New York: American Institute of Physics Press.
- Di, L., Schlesinger, B. & Kobler, B. 2000. U.S. FGDC content standard for digital geospatial metadata: extensions for remote sensing metadata. In *Int. Arch. Photogr. & Rem. Sens. (ISPRS)*, Vol. XXXIII. Amsterdam: Int. Soc. Photogr. & Rem. Sens. (CD-ROM).
- Di, L., Kresse, W. & Kobler, B. 2002. The progress and current status of ISO standard development project 19130-sensor and data models for imagery and gridded data. In J. Chen & J. Jiang (eds), *Int. Arch. Photogr., Rem. Sens. & Spat. Info. Sci.* XXXIV(2): 107–110.
- Di, L. 2003. The development of remote sensing-related standards at FGDC, OGC, and ISO/TC 211. In *IEEE Int. Geosci. & Rem. Sens. Symp.*; Toulouse, 21–25 July 2003.
- FGDC 2002. *Content standard for digital geospatial metadata: extensions for remote sensing metadata, FGDC-STD-012-2002*. Reston: Federal Geographic Data Committee.
- ISO 2003. *ISO 19115 Geographic Information – Metadata. ISO 19115:2003*. Geneva: International Organization for Standardization.
- King, M.D., Closs, J., Spangler, S. & Greenstone, R., (eds), 2003. *EOS Data Products Handbook*. NP-2003-4-544-GSFC. Washington DC: National Aeronautics and Space Administration.

Techniques for in-flight and on-orbit calibration

System-level approach to characterization and radiometric calibration of a space-based electro-optical sensor

J. Tansock, A. Thurgood & M. Larsen

Space Dynamics Laboratory, Utah State University, Logan, Utah, USA

ABSTRACT: State of the art electro-optical sensors being designed for today's space-based applications require more complete characterization and thorough calibration to permit increasingly smaller calibration uncertainties needed for many applications. This is especially true in the remote sensing community to assess global climate change. The Space Dynamics Laboratory (SDL) at Utah State University has been a pioneer in calibrating IR electro-optical sensors for over 30 years. Based on SDL's experiences, this paper describes a system-level calibration approach that addresses each phase of calibration, from planning to on-orbit operations. This approach leads to a thorough understanding of sensor performance, provides a complete calibration, and minimizes risk of not meeting measurement requirements. The overall calibration philosophy is presented along with example calibration parameters. Considerations for estimating uncertainty and lessons learned are also discussed.

1 CALIBRATION PHILOSOPHY

A thorough system-level calibration should address calibration throughout a sensor's lifetime, from the design phase through on-orbit operations. To accomplish this goal, an efficient, cost effective approach to calibration should be developed early in the sensor's design stage. Through this approach, a thorough understanding of sensor performance is achieved, a complete calibration is obtained, and risks of not meeting measurement requirements are reduced.

1.1 Complete calibration

A complete sensor calibration provides thorough understanding of sensor operation and performance. It verifies that the sensor meets mission instrument requirements, verifies a sensor's readiness for flight, quantifies radiometric and goniometric performance, provides traceability to appropriate standards, and quantifies measurement uncertainties.

Five responsivity domains should be addressed in a complete calibration of radiometric, spectral, spatial, temporal, and polarization parameters (Wyatt 1978). The radiometric responsivity domain includes calibration parameters that describe the radiometric response of a sensor and includes radiance and/or irradiance calibration coefficients, response linearity, array detector-to-detector response uniformity, nominal and outlying detector identification, and radiometric calibration of internal calibration sources.

The spectral domain describes the spectral responsivity of a sensor. This responsivity can be made relative by applying a relative normalization factor, or scaling to appropriate engineering units.

The spatial responsivity domain includes calibration parameters that quantify spatial and/or angular figures of merit of a sensor. These parameters include point or line responses, effective field of view, modulation transfer function, field of regard, optical distortion, and optical scatter.

The temporal response domain includes repeatability for a specified time period, i.e. short, medium, and/or long, and amplitude response as a function of optical input temporal frequency.

The polarization domain characterizes the sensor's response as a function of polarized light input. This domain can be quantified with Mueller matrix coefficients and Stokes parameters (Peterson et al. 2000a, b). These parameters can then be used to calculate polarization properties such as S and P vectors, degree of polarization, and angle of polarization.

The goal of calibration is to characterize each of these responsivity domains independently. However, various dependencies between domains can complicate calibration, making it difficult to quantify each domain independently. An example of this is spectral and spatial dependence, which can be caused by out-of-band filter scattering known as the Stierwalt effect (Fequa et al. 2003). Sensors must therefore be designed with calibration taken into consideration to minimize the impact of these dependencies.

1.2 Calibration parameters

An electro-optical sensor responds to optical or radiometric input energy. A sensor-specific calibration equation, which incorporates certain calibration parameters, is then used to convert the sensor output to engineering units, like counts or volts, or to equivalent SI units. Calibration parameters that are not included in the calibration equation, but are required to fully calibrate or characterize the instrument, are incorporated into the radiometric model. This allows the calibration parameters to be grouped into two categories: parameters associated with the calibration equation, and parameters that comprise the radiometric model.

1.2.1 Calibration equation

The form of the calibration equation depends on sensor design and specific measurement application. For example, a sensor designed to measure the degree and angle of polarization will have a different calibration equation than a sensor designed to measure extended source radiance. A generic radiance (extended source) calibration equation for an imaging array-based infrared radiometer is given by:

$$L_{M,k,t} = \left(\frac{1}{\mathfrak{R}_L \Delta \mathfrak{R}_{L,t}} \right) r_{k,t} \\ = \left(\frac{1}{\mathfrak{R}_L \Delta \mathfrak{R}_{L,t}} \right) \left(\frac{B_k G_I [F_{NL,k}(r_{T,k,t} - r_{O,k,t}) - r_{B,k,t}]}{F_{FF,k} \Delta F_{FF,k,t}} \right) \quad (1)$$

where $L_{M,k,t}$ = measured radiance ($\text{W}/\text{cm}^2 \text{sr}$), \mathfrak{R}_L = array-average peak-radiance responsivity (counts per $\text{W}/\text{cm}^2 \text{sr}$), $\Delta \mathfrak{R}_{L,t}$ = array-average peak-radiance responsivity change vs. time (unitless), $r_{k,t}$ = corrected detector response (counts), B_k = bad detector mask function (unitless), G_I = gain or integration normalization factor (unitless), $F_{NL,k}()$ = nonlinearity correction function (counts), $r_{T,k,t}$ = raw detector response (counts), $r_{O,k,t}$ = detector offset correction (counts), $r_{B,k,t}$ = linearity corrected response due to telescope thermal emission (counts), $F_{FF,k}$ = flat-fielding or non-uniformity coefficient (unitless), $\Delta F_{FF,k,t}$ = flat-fielding or non-uniformity coefficient change vs. time (unitless), k = detector index (parameter is unique to each detector), and t = time (parameter varies as a function of time).

1.2.2 Radiometric model

As previously discussed, the radiometric model quantifies all parameters that are not included in the calibration equation but are required to fully calibrate or characterize the instrument. Table 1 lists possible radiometric model parameters that may be required for calibrating an imaging array-based infrared radiometer.

Table 1. Radiometric model parameters for an imaging array-based infrared radiometer.

Radiometric model parameters
Noise equivalent radiance
Saturation equivalent radiance
Noise equivalent irradiance
Noise equivalent temperature
Saturation equivalent irradiance
Illuminated short-term repeatability vs. radiance
Saturation response
Relative spectral responsivity
Effective field of view
Point response function
Modulation transfer function
Near angle scatter
Array coalignment
ICU response settling time
ICU response repeatability
ICU response uniformity
Dark offset repeatability
Dark offset drift rate
Point source repeatability (short, medium, and long)
Extended source repeatability (short, medium, and long)
Angular measurement accuracy and precision
Polarization sensitivity
Uncertainty
Off-axis rejection
Optical distortion

1.3 Considerations for estimating uncertainty

Measurement results cannot be compared with reference standards unless they are reported in standard units and include an estimate of uncertainty. When possible, it is recommended that calibration results be expressed in the International System of Units (ISO) (Pollock et al. 2003). Guidelines for estimating uncertainty have been established (ISO 1992, Wyatt et al. 1998). The following guidelines apply to a system-level approach to calibration:

1. Establish an uncertainty budget early in the calibration planning process that can be used to track uncertainty estimates against instrument requirements throughout the planning process. Direct efforts to reduce the largest uncertainties to give the smallest possible overall uncertainty.
2. When possible, report results in SI units and make use of established guidelines for estimating uncertainty. Recognize the need for other programs to use your results.
3. Do not underestimate the effort required to quantify a comprehensive estimate of uncertainty. It is not uncommon to spend more time estimating uncertainty than deriving calibration results.
4. Be realistic and honest when estimating uncertainty. No one ultimately benefits by providing an overly optimistic level of uncertainty.

5. Establish multiple traceability paths to physical standards. This will verify or help quantify uncertainty estimates.
6. Provide sufficient detail when documenting uncertainty estimates. This provides confidence that all sources of uncertainty were considered and allows the reader to make an independent assessment, if new uncertainty information becomes available at a later date.

2 PHASES OF CALIBRATION

Several phases of calibration are required to ensure a thorough calibration and complete understanding of sensor performance throughout the sensor's lifetime. These include calibration planning during sensor design, ground calibration measurements, and on-orbit calibration measurements.

2.1 Calibration planning during sensor design

Calibration planning should begin early in the sensor design phase to ensure that the design can accommodate the methods used to calibrate the instrument. The calibration approach should be optimized during this phase to meet the program's performance requirements, and to minimize schedule and cost risk by understanding what is required to perform a successful calibration.

Figure 1 is a block diagram of the calibration planning process. Mission requirements flow down to instrument or sensor requirements. These requirements then flow down to the sensor design effort. From a calibration perspective, they are also used to formulate calibration equations and to identify the needed radiometric model parameters. This information then flows down to detailed calibration planning, which is essentially a trade study among cost, schedule, measurement uncertainty, required ground support equipment (GSE), e.g. hardware and software, and risk. This process should also allow interaction and feedback between calibration planning and sensor design.

2.2 Ground measurements

Measurements made during ground calibration are used to verify proper instrument operation, to quantify calibration equation and radiometric model parameters, and to estimate measurement uncertainties.

For many sensor applications, detailed calibration is most efficiently performed during ground calibration because on-orbit time dedicated to calibration is limited and expensive. In addition, on-orbit calibration will not provide a sufficient number of calibration sources with the required radiometric traceability, spectral content, and/or flux levels.

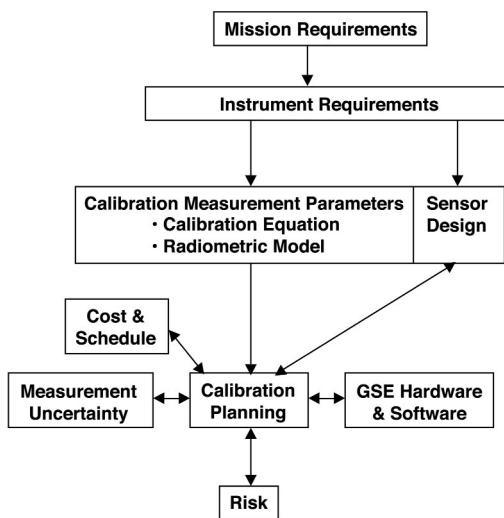


Figure 1. Block diagram of the calibration planning process.

Ground measurements are grouped into four categories: subsystem and/or component level, engineering calibration, sensor-level calibration, and optical bench or satellite-level integration and test.

2.2.1 Subsystem and component measurements

Subsystem and component measurements should be made prior to system assembly to predict sensor performance, evaluate sensor design, and assist with other planning activities such as calibration and mission planning. Obtaining subsystem and component measurements will also minimize schedule risk by identifying problems at the lowest level of assembly.

Component-level measurements should not be merged to predict final end-to-end sensor level calibration parameters because it is difficult to predict interactions between subsystems (Hansen et al. 2003, Lorentz & Rice 2003). These interactions may produce calibration uncertainties that are unknown until they are quantified at the sensor or system level.

2.2.2 Engineering calibration

An engineering calibration is performed prior to the sensor-level calibration to characterize the sensor sufficiently to optimize the physical adjustment of the integrated measurement system. The sensor may not be in its final flight configuration during this phase, but information can still be obtained to verify the combined operation of the sensor, GSE, test configurations, and calibration procedures. In addition, preliminary data can be obtained that can be used to evaluate sensor performance. Performing these tests several weeks prior to the final calibration allows time to make

appropriate procedural or hardware updates based on results, and/or lessons learned.

2.2.3 Sensor-level ground calibration

An end-to-end sensor-level calibration accounts for interactions between subsystems and components, and should be performed at the highest possible level of assembly. For many space-based systems, it is practical to perform this calibration at the sensor level, and then perform cross checks after the sensor has been integrated into the system or satellite.

To bracket the instrument operational bounds, it is important to perform calibration measurements under expected operational conditions for the intended application. For many space-based sensors, this will involve operating the sensor in a thermal vacuum chamber to simulate the appropriate thermal environment. For infrared applications where room temperature or stray thermal emission may affect the radiometric purity of calibration measurements, cold radiometric baffles, should be considered in the design to reduce these errors to levels consistent with measurement requirements.

A complete sensor-level calibration prior to launch will minimize problems during integration and promote mission success during on-orbit operations.

2.2.4 Optical bench or satellite-level integration and test

For platforms that contain a suite of sensors, it is recommended that the calibration approach contains cross checks after integration for each sensor installed on the

satellite or optical bench. Emphasis should be placed on verifying flight readiness, and/or identifying if sensor calibrations have changed compared to the baseline established during sensor-level ground calibration.

2.2.5 Ground calibration and data collection facilities

Sufficient time and resources must be allocated during the calibration planning phase to develop and acquire the facilities needed to perform ground calibration measurements. In addition, calibration hardware and software should be tested and characterized prior to the actual calibration to minimize schedule delays, and to prevent a possible degradation in the calibration due to equipment issues.

If possible, calibration measurements should be integrated into a single facility. This approach minimizes calibration time by reducing or preventing repeated cycle times associated with pump, cool-down, warm-up, and configuration times. This is particularly important for time-constrained projects where calibration is on the critical path.

Figure 2 shows an example of the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) calibration facility (Tansock et al. 2003), which was designed and built to measure individual parameters of the calibration equation and radiometric model. This facility consisted of a test chamber interfaced with a collimator, and provided all the required radiometric calibration testing in a single test setup. SABER was mounted and operated in the test chamber, which also provided blackbody and knife

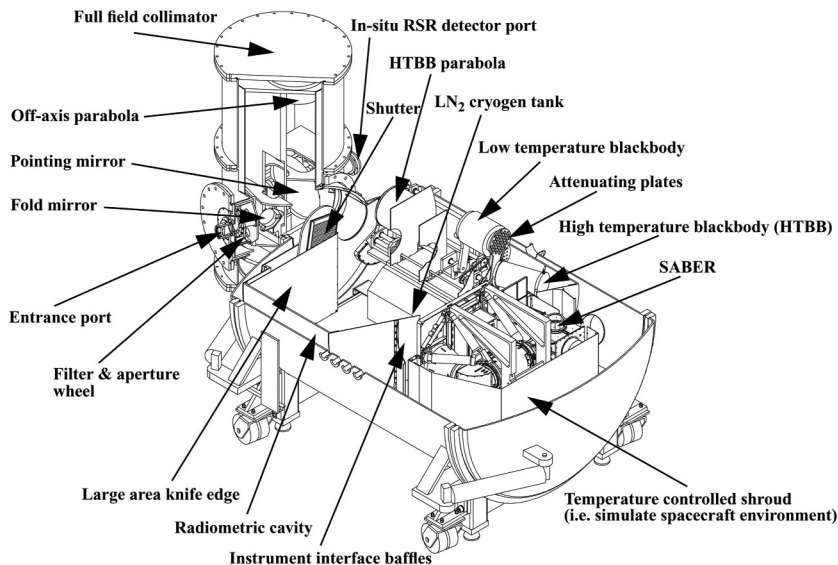


Figure 2. SABER calibration facility.

edge calibration sources. The collimator provided source configurations for spectral, spatial, linearity, and temporal calibrations.

The data collection system required for ground calibration should include a GSE computer to control and view the status of the sensor, a GSE computer to control and view the status of the GSE, a data collection computer to initiate and execute data collection and to control and monitor status, and a quick look analysis station. Figure 3 shows a data collection environment for an infrared sensor calibration performed at SDL. This picture shows an example of an efficient data collection environment that is obtained by locating equipment, tables, chairs, and computers strategically within a relatively small working area located next to the sensor under test.

Data collection should be implemented to ensure that calibration measurements are complete, performed efficiently, and are readily repeatable. This can be accomplished by developing and using detailed and mature data collection procedures that describe preparation activities, step-by-step test procedures, time and data volume requirements, and associated as-run documentation, i.e. script files, data collection log sheets, notes, and similar information.

Data collection should be automated when practical. Scripting languages or other high-level application programs may be used for this automation, including Java (Jscript), Visual BasicV(Bscript), Perl and Tool Command Language/Tool Kit (TCL/TK). The high-level software code used for automation should be developed and maintained by a calibration engineer. Also, automation scripting or software code can be used to document as-run data collection, making the measurements repeatable.

As part of automation, ensure that all environmental conditions and measurement configurations for the GSE and sensor under test are logged at an appropriate rate and can be cross-referenced easily to specific calibration files. In addition to monitoring vacuum,

pressure, and temperatures, contamination should also be monitored, and the levels at which corrective action is required should be identified.

Quick look analysis should be performed in near real-time to verify the quality of the calibration data. The first step of data quality verification is performed during data collection by viewing the real-time sensor output monitor. The sensor output monitor should have sufficient flexibility to show sensor response in a way that can be used to verify that the sensor is responding as expected for a given calibration source and setting.

The second part of data quality verification is performed in parallel with data collection and shortly after a specific test. Data are processed through specified data management steps. Preliminary analysis is then performed on these data by a calibration analysis engineer to ensure data quality for the intended application. To avoid discovering degraded or insufficient data quality after calibration testing, data quality verification should be completed before the end of calibration testing.

For many sensor development projects, calibration is performed just before delivery of the sensor and it is on the critical schedule path. For this reason, there is usually pressure to complete calibration in the shortest time possible. Since it is impractical to maintain a large calibration crew throughout the duration of a project, sensor engineers, and/or technicians, who have familiarity with the sensor under test can assist with data collection, if sufficiently detailed procedures are developed in advance. This will facilitate the data collection schedule by making shift work possible.

2.3 On-orbit calibration

On-orbit calibration is the last phase of calibration. Its goal is to maintain calibration throughout a sensor's operational lifetime by quantifying calibration uncertainty and updating calibration coefficients, if necessary, to meet measurement requirements. A thorough, complete, and sufficiently bounded ground calibration will minimize the satellite operational time required for on-orbit calibration.

A sensor contains internal calibration sources that are common to all phases of calibration, and are used in trending the response of a sensor. These sources are thoroughly characterized and calibrated during sensor-level ground calibration to establish baselines. Assuming these sources are stable and repeatable, they can be used to quantify uncertainty and even update specific calibration coefficients, if necessary.

On-orbit calibration is comprised of two phases: early on-orbit and normal operations. A comprehensive system check-out is performed during early on-orbit operations to ensure the system is ready for normal operations. A series of on-orbit calibration experiments should also be conducted during this period to assess

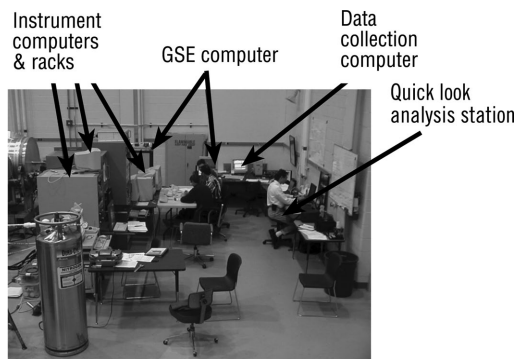


Figure 3. Example calibration data collection environment.

calibration readiness for normal operations. During normal operations, the same calibration experiments conducted during early operations should be performed periodically.

To maintain calibration, it is desirable to obtain reference calibration source measurements throughout the operational phase of a sensor. The frequency of these measurements will depend on requirements and the stability of the sensor. A common practice is to bracket science experiments with calibration measurements. Internal calibration sources provide a convenient means for this purpose because they require less experiment time than a calibration measurement that requires satellite attitude adjustments. They are, however, prone to errors. Often they do not use the entire optical path, and they are sometimes variable themselves.

On-orbit sources provide a means for calibration external to the sensor that can be used in addition to the internal calibration sources to track and trend the response of a sensor. For many applications, these types of measurements require special satellite maneuvers. As a result, they are performed less frequently than measurements using internal calibration sources.

In recent years, efforts have been made to develop and characterize on-orbit sources. Each measurement community tends to concentrate on specific sources that meet their needs for specific processing techniques, and/or measurement requirements. There are a number of on-orbit sources that may be considered for a specific application.

Stars are used for both goniometric and radiometric calibrations. In the visible and short wave IR, there is a number of available star catalogs, including Air Force Geophysics Laboratory (AFGL), Infrared Astronomical Satellite (IRAS), Midcourse space Experiment (MSX), and 2 Micron All-sky Survey (2MASS). For the mid and long wavelength regions, a number of stable standard stars have been identified including α Boo, α Lyra, α Tau, α Cma, β Gem, and β peg (Cohen et al. 1999, Russell et al. 2003).

Other celestial objects such as planets, asteroids, and the moon may be candidate calibration sources depending on specific measurement requirements. The moon is a stable source in the visible and near IR and has a radiance that is generally in the dynamic range of downlooking sensors in these bands (Kieffer & Anderson 1998). Little radiance information is available for the moon in the mid and long-wavelength infrared; however, the moon is a good candidate for off-axis scatter characterization in these bands.

In addition to celestial objects, space-based Earth-observing geological remote sensing applications tend to use other vicarious techniques for their calibrations. These include using large panels on the surface of the Earth with known emissivity for radiometric calibrations and known geometry for spatial calibrations.

In addition, geodesic point targets are used for goniometric calibrations.

3 CONCLUSIONS

State of the art electro-optical sensors being designed for today's space-based applications require more complete characterization and thorough calibration to allow for the increasingly smaller calibration uncertainties needed for many applications.

This paper discussed SDL's philosophy for calibrating electro-optical sensors. It presented a system-level calibration approach that addressed each phase of calibration including planning, ground measurements, and on-orbit operations. Through this approach, a thorough understanding of sensor performance is achieved, a complete calibration is obtained, and risks of not meeting measurement requirements are reduced.

REFERENCES

- Cohen, M., Walker, R., Carter, B., Hammersley, P., Kidger, M. & Noguchi, K. 1999. Spectral irradiance calibration in the infrared. X. A self-consistent radiometric all-sky network of absolutely calibrated stellar spectra. *Astron. J.* 117: 1864–1889.
- Fequa, P., Barrie J. & Presser N. 2003. Observation and characterization of the Stierwalt effect in dielectric filters with model coating defects. In B.F. Andersen, G.F. Fulop & M. Strojnik (eds), *Infrared technology and applications XXVII*. Proc SPIE 4820: 878–884.
- Hansen, S., Peterson, J., Esplin, R. & Tansock, J. 2003. Component level prediction versus system level measurement of SABER relative spectral response. *Int. J. Rem. Sens.* 24(2): 389–402.
- ISO 1992. Guide to the expression of uncertainty in measurement. 1992. ISO, IEC, OIML, BIPM, ISO/TAG 4/WG 3. Geneva: International Organization for Standardization.
- Kieffer, H.H. & Anderson, J.M. 1998. Use of the moon for spacecraft calibration over 350 to 2500 nm. In H. Fujisada (ed.), *Sensors, systems, and next-generation satellites II*. Proc SPIE 3498: 325–336.
- Lorentz, S. & Rice, J. 2003. System level vs. piece parts calibration: NIST traceability – when do you have it and what does it mean? *Proc. CALCON; Logan, Utah, 15–18 September 2003*. Utah State Univ.: Space Dynamics Laboratory. CD ROM.
- Peterson, J., Jensen, G. & Kristl, J. 2000a. Imaging polarimetry capabilities and measurement uncertainties in remote sensing applications. In D.B. Chenault, M.J. Duggin, W.G. Egan & D.H. Goldstein (eds), *Polarization analysis, measurements, and remote sensing III*. Proc SPIE 4133: 221–232.
- Peterson, J., Jensen, G., Kristl, J. & Shaw, J. 2000b. Polarimetric imaging using a continuously spinning polarizer element. In D.B. Chenault, M.J. Duggin, W.G. Egan & D.H. Goldstein (eds), *Polarization analysis, measurements, and remote sensing III*. Proc SPIE 4133: 292–300.

- Pollock, D.B., Murdock, T.L., Datla, R.U. & Thompson, A. 2003. Data uncertainty traced to SI units. *Int. J. Rem. Sens.* 24(2): 225–235.
- Russell, R., Rossano, G., Masuz, A., Lynch, D., Chatelain, M., Venturini, C., Prater, T., Kim, D., Mazuk, S. & Ostrander, M. 2003. The Aerospace spectral energy distribution stellar radiometric calibration database: approach, application to SBIRS programs, sample data, plans, and status. *Proc. CALCON; Logan Utah, 15–18 September 2003*. Utah State Univ.: Space Dynamics Laboratory. CD ROM.
- Tansock, J., Hansen, S., Paskett, K., Shumway, A., Peterson, J., Stauder, J., Gordley, L.L., Wang, Y., Melbert, M., Russell, J.M. III & Mlynczak, M.G. 2003. SABER ground calibration. *Int. J. Rem. Sens.* 24(2): 403–420.
- Wyatt, C. 1978. *Radiometric calibration theory and methods*. New York: Academic Press.
- Wyatt, C., Privalsky, V. & Datla, R. 1998. *Recommended practice: symbols, terms, units and uncertainty analysis for radiometric sensor calibration*. NIST Handbook 152. Washington, DC: U.S. Government Printing Office.

On-orbit MTF assessment of satellite cameras

D. Léger, F. Viallefont & P. Déliot

ONERA, DOTA/QDO, France

C. Valorge

CNES, DSO/OT/QTIS, France

ABSTRACT: The modulation transfer function (MTF) is a technique for characterizing the spatial resolution of optical sensors. Therefore, the MTF of satellite cameras is one of the image quality parameters assessed during on-orbit operations. Since launching the first Satellite Pour l'Observation de la Terre (SPOT 1) in 1986, several methods of on-orbit MTF assessment have been studied and implemented. The main methods are the step edge method and the point source method. Trials have been made also with the bi-resolution method and periodic target method. A neural network method is in progress. All these methods give an absolute MTF assessment. Other methods give the relative MTF in the field-of-view. The paper presents the principles of the various methods and their implementation, and comments about their advantages and drawbacks. Characteristics of SPOT cameras are given for illustration. Vibrations during the launch and transition from air to vacuum may slightly defocus the cameras and degrade their MTF. Therefore, SPOT cameras are refocused before measuring their MTF. The paper also presents the defocusing assessment method of SPOT cameras and other methods in progress for future satellites.

1 INTRODUCTION

The spatial resolution of satellite-borne cameras is usually described by the modulation transfer function (MTF), i.e. the Fourier transform (FT) of the impulse response (response at a point source or point spread function (PSF)). MTF results from the cumulative effects of instrument optics (diffraction, aberrations, focusing error), integration on a photosensitive surface, charge diffusion along the array, and image motion induced by the motion of the satellite during imaging.

This important parameter for image quality has to be checked in orbit to be sure that launch vibrations, transition from air to vacuum, or thermal state have not spoiled the sharpness of the images. Some MTF losses may be compensated thanks to a refocusing mechanism. Therefore, the amount of SPOT camera defocus is assessed and the cameras refocused before measuring their MTF. Moreover, an accurate knowledge of MTF is useful to adjust deconvolution filters used in the SPOT 5 Très Haute Résolution (THR) mode.

The first section describes the SPOT focusing procedure and alternative methods for other cameras. The second section concerns the relative MTF measurements, i.e. the comparison of one instrument with another and, for one instrument, the MTF quality in

the field-of-view. The third section addresses absolute MTF measurements, presents results, and describes methods in progress.

2 CAMERA REFOCUSING

The SPOT cameras, Haute Résolution Visible (HRV) aboard SPOT 1-3, Haute Résolution Visible et InfraRouge (HRVIR) aboard SPOT 4 and Haute Résolution Géométrique (HRG) aboard SPOT 5, are equipped with a refocusing mechanism. On the ground, this mechanism is positioned on the best step measured in a vacuum. Nevertheless, possible shift during launch may upset this focusing so that the position of each mechanism has to be checked and, if necessary, altered.

The SPOT refocusing method is based on the SPOT satellites having two identical cameras. Future satellites without this feature will need other methods.

2.1 Two-camera method

The SPOT refocusing method (Begni et al. 1984, Meygret & Léger 1996) uses two cameras simultaneously viewing the same target. Several scene pairs

having complex details, such as urban areas, are selected from the refocusing image sequence. The focus of one camera is left fixed and is used as a reference, while the focusing mechanism of the other camera is varied around its initial position. For each scene pair, the ratio of the FTs of corresponding images is calculated. The best focus is reached while maximizing this spectrum ratio, which varies as a near-parabolic function around its vertex.

Column-wise and row-wise spectrum ratios are computed to characterize camera astigmatism. Moreover, the method is performed in several areas of the field-of-view: each array for SPOT 1-4; the center and two edges for SPOT 5. This focus measurement is performed in the panchromatic mode, i.e. the sharpest mode.

The first version of the method was to effect a large defocusing and to fit merely a parabola on measurements by a least-square optimization. Therefore, at the beginning of the SPOT program, the refocusing operation was done only if necessary. This was a careful decision based on risk of mechanism failure. After a technological experiment of the method on SPOT 1 in 1994 (Meygret & Léger 1996), and the improvement of the method explained below, refocusing operations have become nominal since SPOT 4.

To improve accuracy and to limit the mechanism moving, a defocusing model replaced the parabola (Steel 1956). The Steel defocus model is an adaptation of the geometrical defocus model to approach results obtained with the diffraction theory. As before, the model is fitted on measurements by a least-square optimization. The vertex of this curve gives the best focus position. To have a cautious protocol, the mechanism was moved only a few steps around its base position, p_0 (± 8 steps, or about $10\text{ }\mu\text{m}$, for SPOT 5). Therefore, the focusing mechanism was tested without reaching large defocusing values. If the vertex of the curve is not found between $p_0 - 8$ and $p_0 + 8$, the vertex position is extrapolated from the measurements, but the resulting best focus determination can be inaccurate. Therefore, the mechanism is operated again with a sufficient walk to overpass the best focus (Figs 1–2).

The global best focus is a weighted average of the best foci for the different field areas and both measurement directions (rows and columns). Once its best focus is determined, the first camera is focused and can be used as a reference for the second one. Table 1 gives an example of best focus position in the field for the HRG cameras of SPOT 5.

2.2 Methods in progress

Other methods are currently studied for future satellites having only one camera.

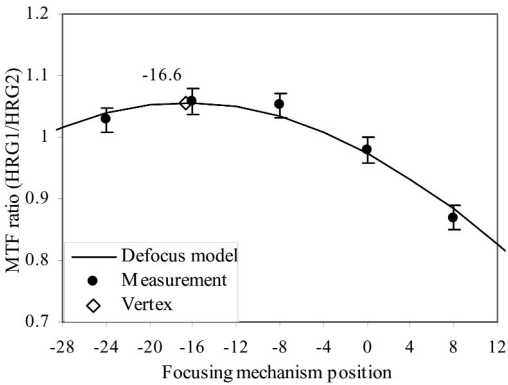


Figure 1. SPOT 5 HRG 1 refocusing (field center – rows).

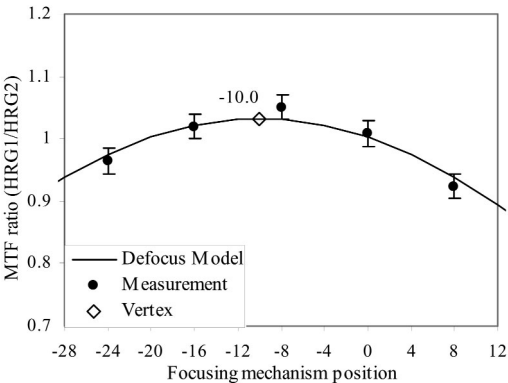


Figure 2. SPOT 5 HRG 1 refocusing (field center – columns).

Table 1. Best focus position with respect to initial position, p_0 .

Field area	HRG 1			HRG 2		
	Left	Center	Right	Left	Center	Right
Rows	-12.8	-16.6	-13.0	0.7	-8.7	-14.6
Columns	-5.9	-10.0	-8.7	2.6	-5.3	-7.1
Mean	-9.4	-13.3	-10.8	1.7	-7.0	-10.8
Astigmatism	-6.8	-6.6	-4.3	-1.9	-3.4	-7.5

2.2.1 Neural network

The method for evaluating the defocus with an artificial neural network (ANN) is a univariate method: it needs no reference imagery. However, this method still acquires images taken with different steps of the focusing mechanism. The first step consists of training the neural network with information coming from images having positive focusing steps. The second step uses images having negative focusing steps

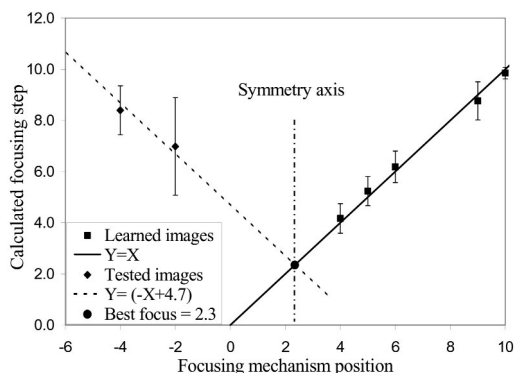


Figure 3. Best focusing step obtained with a neural network method using an assumption of symmetry.

(y axis on Fig. 3). The ANN gives results corresponding to the positive focusing mechanism position it has learned (x axis on Fig. 3). Assuming symmetry about defocusing, it is then possible to find the best focus position, as shown in Figure 3. From a typical image of SPOT (6000 by 6000 pixels), several hundred subsets can be extracted to train the neural network. Thus, only a few images are needed at any given focusing mechanism position.

The characteristic input vector describing the images must provide relevant information linked with the output, i.e. the current focusing step, and knowledge of the landscape. These elements, because of their image frequency properties, are quite similar to those described in §4.7.1 concerning the absolute MTF measurement. The training set of images also must be representative of the whole set of images in order to train the network efficiently. In that case, the ANN will have a property called the “generalization capability”: it will give a proper interpolated output for any image. Because simulated images are not used, the generalization process must be limited to selected landscapes. This is done by selecting the characteristic vector in a hypersphere centered on the mean vector, and with a radius of twice the standard deviation value.

The accuracy of this method is improved by treating row and column defocus independently to take into astigmatism into account. Good results have been achieved on SPOT 1 images, giving the best focus position at ± 0.6 step (Fig. 3). A further stage will be to test an operational implementation of this method.

2.2.2 Fitting model on step edge

The edge method is a widely used way to assess in-flight MTF (see §4.3). For high spatial resolutions, artificial targets are available to ensure good radiometric quality of the edge. This enables one to make accurate measurements. The idea is to use this accuracy to

estimate the defocus of the camera. This is done using a transfer function parametric model. The transfer function is written as the product of the well-focused camera transfer function by the Steel defocus function.

An ideal edge described by two parameters is built. It is convolved with the line spread function (LSF) computed by inverse FT of the three parameter transfer function model. The quadratic difference between this edge and the edge obtained from the image is computed and minimized thanks to having five parameters. The first trial based on simulated images exhibits a very accurate MTF estimation, but the physical meaning of the parameters is not ensured. This problem can be solved by using a widely defocused image of the edge. In this case, the behaviors of the focused and defocused term of the model are different enough to separate the effects and to solve properly the minimization problem. First results are very promising and studies are continuing, particularly for future focus assessment of the Pleiades camera.

3 RELATIVE MTF MEASUREMENT

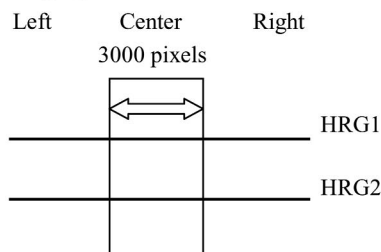
The spectrum comparison tool used for refocusing also can compute the MTF ratio of the two cameras near the frequency $0.3 f_s$ where f_s is the sampling frequency. One can compare directly one instrument with another when both instruments view the same landscape. As for the focusing measurements, the comparison can be done versus the field area: each array for SPOT 1-4, the center and the two edges for SPOT 5. To compare the MTF in the field of one instrument, e.g. SPOT 5, one needs image pairs where the same landscape is viewed by the field center of one instrument, say, HRG 1, and a field edge of the other instrument, say, left field of HRG 2 (Fig. 4). Thus, it is possible to obtain successively the MTF ratio for both HRG center fields and for HRG 2 left field against HRG 1 center field. It is easy to deduce the ratio of MTF HRG 2 left field to MTF HRG 2 field center and, in the same way, all other ratios. The method is used for both directions (rows and columns) and for every spectral band. The accuracy of such a comparison of image frequency contents is estimated by the measurement dispersion, which is 2 percent (see error bars in Fig. 1).

HRG MTF relative measurements showed that both SPOT 5 cameras have equivalent MTFs and have a uniform quality in the field-of-view.

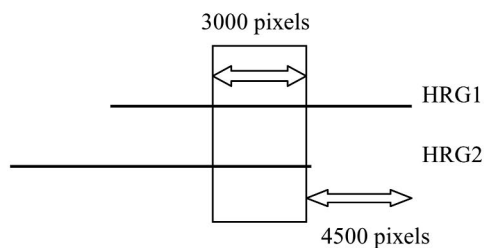
4 ABSOLUTE MTF MEASUREMENT

In spite of its accuracy, methods based on *comparison* only give relative MTF measurements. Since the launch of SPOT 1, several methods have been studied

First imaging mode



Second imaging mode



Third imaging mode

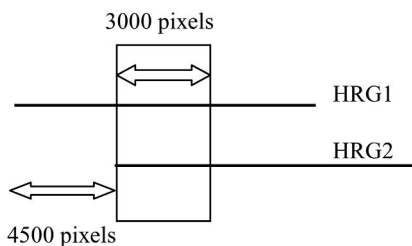


Figure 4. Successive modes for relative MTF measurement.

and brought into operation to evaluate *absolute* MTF. The on-orbit commissioning phase of SPOT 5 has given the opportunity to perform most of them.

4.1 Visual assessment

The first method used to evaluate the SPOT MTF is a photointerpretative method. Before launch, a reference image catalog simulating SPOT panchromatic images was prepared with a set of various parameterized MTFs. During flight, corresponding SPOT images are acquired. Then the SPOT images are ordered into the reference catalog by visual comparison.

The catalog method was used for the commissioning phases of SPOT 1-3. It was very important for SPOT 1 because it was the only method of MTF evaluation. Its use led to a conclusion that the MTF specification was satisfied for both HRV cameras, and the refocusing operations were not necessary (Léger et al.



Figure 5. Spotlight located on a grassy area in ONERA Fauga-Mauzac center near Toulouse.

1986). Some years later, a technological refocusing experiment confirmed that the SPOT 1 defocusing was too small to be distinguishable visually (Meygret & Léger 1996). Catalog images from both SPOT 2-3 were compared visually to those from SPOT 1. Ultimately the catalog was considered obsolete and the visual method was abandoned. Nevertheless, a visual examination of first on-orbit images gave a subjective assessment of the sharpness of image quality.

4.2 Point source methods

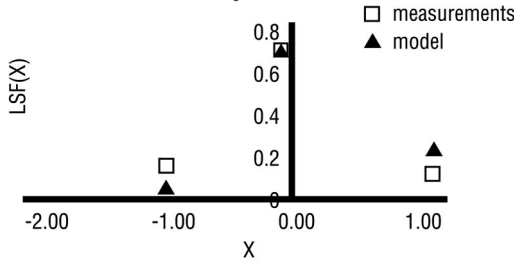
The MTF can be obtained by making the FT of the PSF of the system. It is obtained by imaging a point source lying on the ground. In a general case, the ground sampling distance (GSD) is too large to get the MTF directly by computing a FT of the point image; in other words, the PSF is undersampled. This difficulty can be overcome in two ways: by using one point image and fitting it to a PSF parametric model (§4.2.1); or using a combination of several images with different sub-positions on the sampling grid to rebuild an over-sampled PSF from which the MTF can be computed without undesirable aliasing effect (§4.2.2).

As a point source, a spotlight aimed at the satellite can be used. The spotlight is located in the middle of a grassy field, i.e. in a quasi-uniform area (Figs 5, 7). The spotlight can be oriented in site and azimuth by means of two rotating plates. The satellite position when looking at the spotlight location can be calculated from orbitography predictions. The spotlight Xenon lamp power is 3 kW and the beam divergence is about 6 degrees.

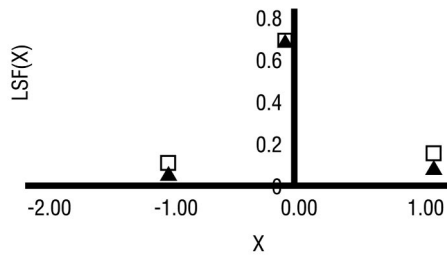
4.2.1 Single point source method

The cheapest way to overcome the aliasing problem is to use a model. The method consists of using one

Case 1: Inaccurate matching



Case 2: Accurate matching



Case 3: Accurate matching

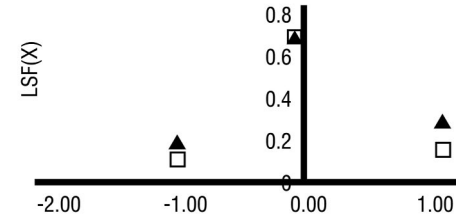


Figure 6. Matching measurement and model.

spotlight image and fitting it to a PSF parametric model to compute two parameters: one for the MTF shape; the other, for the sub-position on the sampling grid. To simplify the problem, the process is limited to one dimension. Therefore, the method consists of fitting a LSF model to the observed LSF (Léger et al. 1994).

For SPOT, an analytic MTF approach based on theoretical and experimental criteria was modeled in the principal directions. For the SPOT panchromatic band, the model is:

$$MTF(f_x, 0) = \exp(-\lambda_x f_x) \text{sinc}(\pi f_x) \quad (1)$$

$$MTF(0, f_y) = \exp(-\lambda_y f_y) \text{sinc}^2(\pi f_y) \quad (2)$$

where f_x and f_y are normalized frequencies. Similar models have been obtained for spectral bands. The 1D-MTF model allows computing oversampled LSFs.

To use the model, it is necessary to normalize the samples of the spotlight image by the sum of the

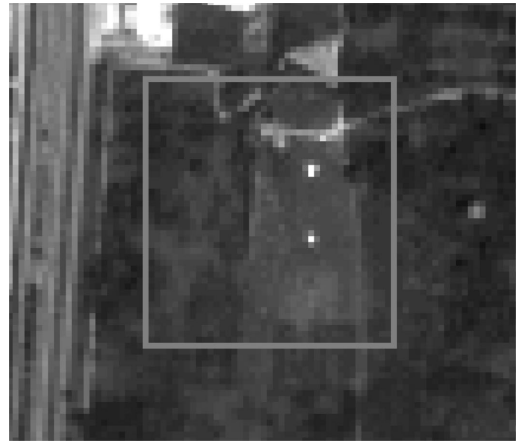


Figure 7. SPOT 5 THR image of Fauga-Mauzac center with two spotlight experiment.

samples. The observed LSFs are then obtained by summation over rows or columns. The last step consists of searching the best matches between the modeled LSFs and the observed LSFs (Fig. 6).

4.2.2 Multiple point sources

The second way to overcome undersampling uses a combination of several images with different subpositions on the sampling grid to rebuild an oversampled PSF from which the MTF can be computed without undesirable aliasing effect. This method was used for Landsat TM (Schowengerdt et al. 1985, Rauchmiller & Schowengerdt 1988) with an array of 16 passive point sources. The implementation of an array of spotlights has been studied (Robinet et al. 1991), but the method is too cumbersome. A way of simplification is to rebuild a LSF rather than the PSF. The needed number of sources decreases from 16 to 4. Furthermore, if one can consider that the MTF is negligible beyond $f = 1$, it is sufficient to have 2 point sources. Indeed, a trial with two spotlights was executed during the commissioning phase of SPOT 5 (Fig. 7).

Another way to obtain multiple point images is to acquire several images over time. The single point method allows obtaining the position on the pixel grid. With a sufficient number of acquisitions, one can have two images with a 0.5 pixel shift in one direction. Therefore, one can obtain a twice-oversampled LSF in this direction and calculate the 1-D MTF by a discrete FT.

For the panchromatic band of the SPOT 5 HRG cameras, there is a peculiar case where the THR mode directly gives a sufficiently sampled image that is equivalent to a two-point image with a classical mode.

4.3 Step edge method

The step edge (or knife edge) method is well known in optics to obtain an MTF measurement. This method uses a natural or artificial target with sharp transitions between dark and bright uniform areas. The principle is that the derivative of Heaviside step function is a Dirac. Therefore, the derivative of the step image is the camera LSF.

Again, with a sampled image it is necessary to overcome undersampling. The edge method has been used widely to estimate the resolution of satellite cameras (Anuta 1984, Malaret 1985). The answer to undersampling is to use a transition slightly inclined with respect to the row-wise (or column-wise) direction. Therefore, one can interleave the successive rows (or columns) in order to get an oversampled response to the Heaviside function, and then the 1-D MTF (Fig. 8).

Actually, rather than the derivative method, one can calculate the MTF as the ratio of the FT of the observed step to the FT of the ideal Heaviside step.

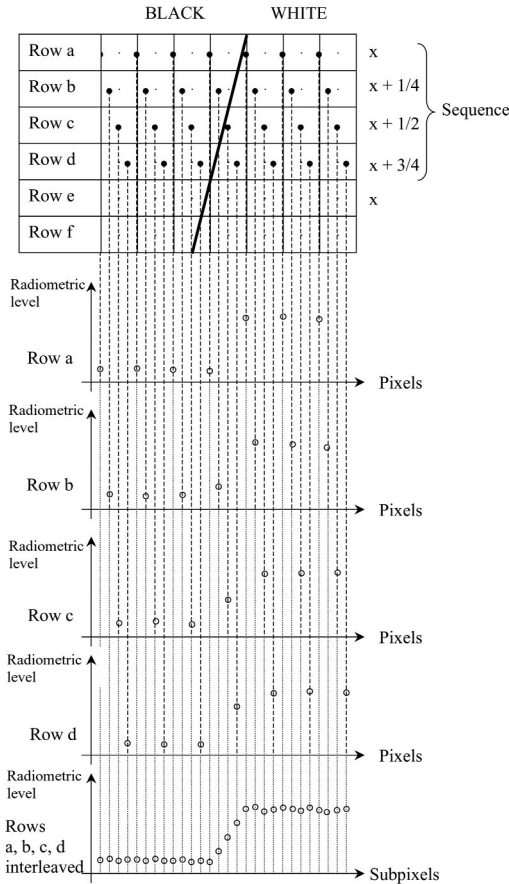


Figure 8. Principle of 1-D oversampling edge response.

4.3.1 Natural edge

The edge method was used during the commissioning phase of SPOT 4 (Kubik et al. 1998). Some edges were extracted from images of the Maricopa site, near Phoenix. This site includes a number of agricultural field edges having good contrast. However, the method requires uniform areas. Such natural areas may be difficult to select. Another difficulty is the eventual presence of tracks between the fields. After 2-D and 1-D visualization, only very few were kept and used for the MTF assessment. Therefore, the method was not used with SPOT 5 HRG because there were other methods: spotlight or artificial target. However, the method was still used with SPOT 5 Haute Résolution Stéréoscopique (HRS) for lack of other methods: the spotlight was used for HRG and the artificial target was too small. As for SPOT 4, images of the Maricopa site have been acquired as shown in Figure 9.

4.3.2 Artificial edge

The Office National d'Etudes et de Recherches Aérospatiales (ONERA) and the Centre National d'Etudes Spatiales (CNES) have laid out an artificial edge target at Salon-de-Provence in the south of France. The target looks like a checkerboard (Figs 10, 11).



Figure 9. Maricopa site seen by SPOT 5 HRS.

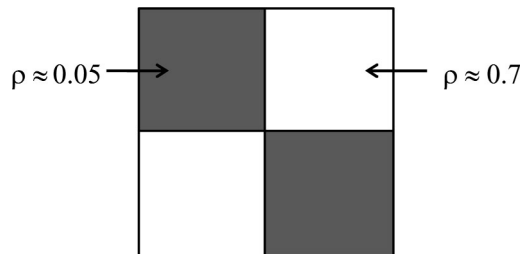


Figure 10. Checkerboard target.

This shape allows performing row-wise and column-wise measurements and distinguishing between upward and downward steps. This last point is useful in case of lag effects. The overall size of the target is $60\text{ m} \times 60\text{ m}$. The target is a tar pad painted with road paints. The white diffuse reflectance is about 0.70 and the dark one is 0.05. Due to its dimension, it can be used only for the SPOT 5 panchromatic mode.

(Fig. 12) shows an MTF curve measured with the edge target. The curve is straight obtained by Fourier calculations, only followed by a scale change to compensate for the effect of edge inclination. The method could be improved by fitting an MTF model to the raw curve.

4.4 Bi-resolution

Another method to assess in-flight MTF relies on image couples of the same landscape acquired in the same spectral band with two spatial resolutions. The higher resolution image stands for the landscape so that the



Figure 11. SPOT 5 THR image of Salon-de-Provence target.

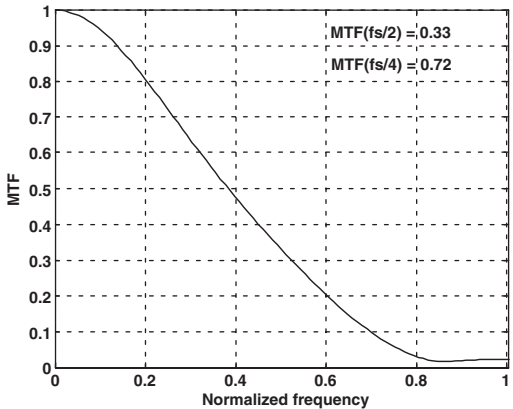


Figure 12. SPOT 5 HRG1 MTF (edge method).

ratio of the image spectra gives the lower resolution sensor MTF (Schowengerdt et al. 1985).

For most satellites, the only way to obtain an adequate image couple is to image the chosen landscape with an airborne camera. This has been attempted for the HRVIR sensor on SPOT 4. Use of this method is much easier if both sensors are on the same satellite. This is the case for the VêGéTation (VGT) and HRVIR couple aboard SPOT 4 for the B1 and B2 spectral bands.

The method is sensitive to aliasing. Work has been done to overcome this problem (Viallefont-Robinet 2003). A theoretical expression of the image spectra ratio has been proposed and checked. This ratio is not far from the quadratic summation of the transfer function and the aliased transfer function. This relation is taken as an equation to be solved. One way to solve it is to take a parametric model for the transfer function and to minimize the difference between the image spectra ratio and the quadratic summation of the transfer function, and the aliased transfer function. The parameter value obtained allows one to compute an in-flight MTF of the sensor. This is illustrated, in Figures 13 and 14.

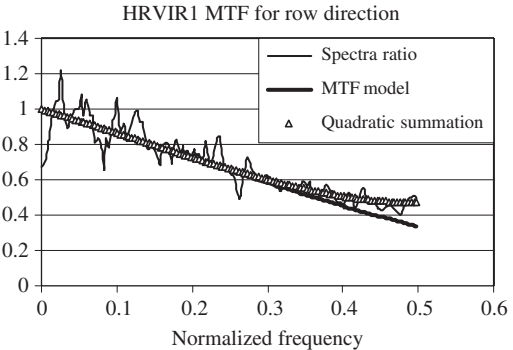


Figure 13. HRVIR 1 B2M row-wise MTF.

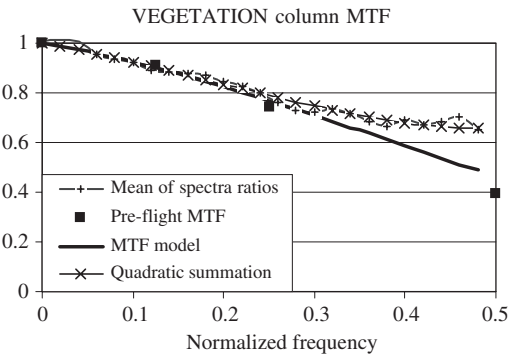


Figure 14. VGT B2 column-wise MTF.

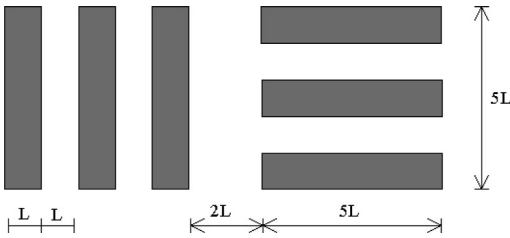


Figure 15. Standard three-bar target.

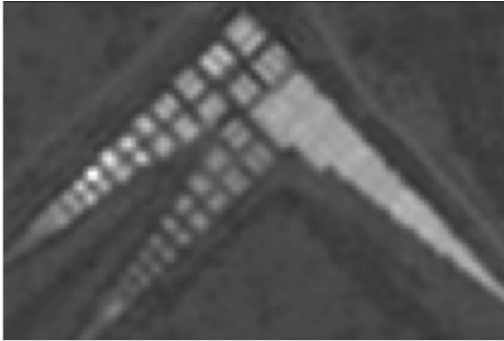


Figure 16. THR SPOT 5 image of Fort Huachuca target.

4.5 Periodic targets

Periodic targets are other strategy for evaluating MTF. Such targets were studied at Stennis Space Center (SSC) (McKellip et al. 1997). A possible use of these targets also is studied at ONERA and CNES. An advantage of periodic targets is that they can be examined visually. The commissioning phase of SPOT 5 gave an opportunity to acquire such targets.

4.5.1 Three-bar target

The three-bar target, also known as United States Air Force (USAF) target, was designed and standardized at the beginning of the 1950s. It is used widely to evaluate spatial resolution. Each target consists of a series of three black and white bar pairs (Fig. 15). The progression factor of the series is based upon a root of two, usually $2^{1/6}$. The target was designed at the time of analog imagery. Its use can be difficult with digital images due to sampling effect.

During the commissioning phase of SPOT 5, an opportunity arose to acquire imagery of the three-bar target located near the Fort Huachuca Air Force base in Arizona. One can see on Figure 16 that no target is resolved, but some frequency dependent information is perceptible in the first targets. This is not surprising. The period of the biggest target is 3 m whereas the Nyquist frequency of THR mode in the 45 degree

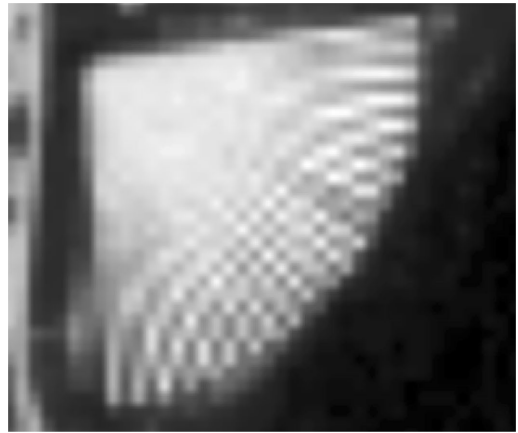


Figure 17. SPOT 5 HRG 2 HM image of radial SSC target.

direction is $1/3.5 \text{ m}^{-1}$. The perceptible structure in the target corresponds to aliasing.

4.5.2 Radial target

A radial target was laid out at SSC. Unlike a three-bar target, a radial target accommodates different flight tracks. Radial targets provide a continuous estimate of the MTF. Recent studies have shown that a radial target must give a better MTF estimation than the three-bar target. To confirm this, an attempt was made to evaluate MTF with an image of the SSC target. The calculations were not easy because the target was not in good maintenance condition. Nevertheless, some results were obtained and compared to other methods (discussed in § 4.6).

Figure 17 shows the SSC radial target image in the Haute Résolution Monomode (HM) mode. In this mode, the GSD is 5 m. It is beyond the target range. One can only see aliasing. Figure 18 shows a radial target image in the THR mode. The THR 2.5 m mode is created by combining two 5 m images (called HMA and HMB) each having 5 m GSD. Now, the Nyquist frequency is in the target range, which allows MTF calculations (Fig. 19).

4.6 Comparison of results

Three methods were used to measure the in-flight MTF of SPOT 5 HRG cameras. The HRG 1 results are shown in Table 2. In-flight results also are compared to ground measurement and to instrument design specifications. Results from the different methods are close. May be the radial target result is a little too high. This method needs dark and white references that are not easy to measure on the SSC target.

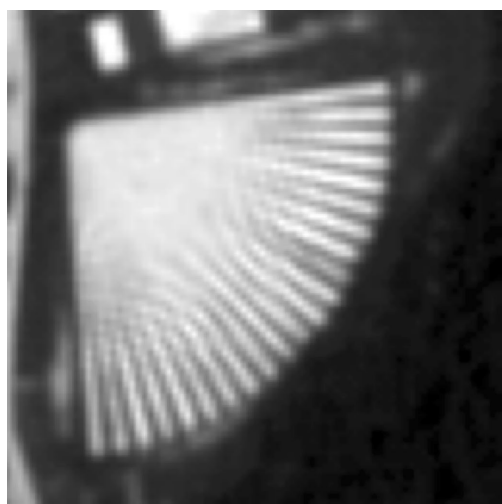


Figure 18. SPOT 5 HRG 2 THR image of radial SSC target.

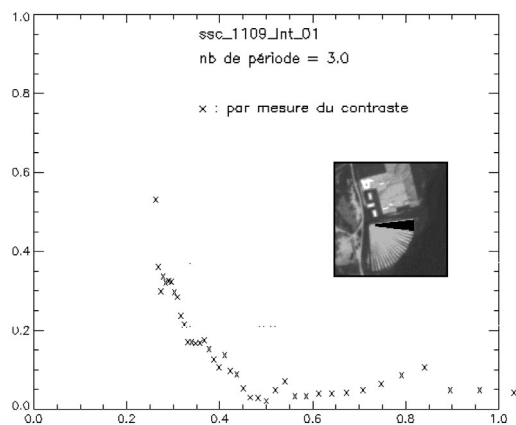


Figure 19. MTF results from THR image of radial SSC target.

Table 2. Comparison of SPOT 5 HRG 1 MTF measurements.

Direction	Rows	Columns	Diagonal
Spotlight	0.35	0.32	0.15
Step edge	0.33	0.30	
Radial target		0.38	0.18
Ground	0.31	0.36	
Specification	0.25	0.23	

The spotlight allows measurements in the 45 degree direction, while the edge method is limited to rows and columns. The radial target orientation does not allow a row-wise measurement in the SPOT case.

4.7 Methods in progress

There are two new MTF calibration methods being investigated. These are based on artificial neural networks and least squares.

4.7.1 Neural network

Artificial neural networks can be used to evaluate the MTF. The principle is first to use ANN to learn the MTF of simulated, or perfectly known images, then to use it to assess MTF parameters of unknown images.

To estimate any parameter (output) with a neural network, it is necessary to find a characteristic vector correlated to the problem. This characteristic vector (input) is the entrance of the neural network and is computed for a set of images. The neural network learns how to establish a correspondence between the inputs and the outputs for known examples. In the end, the neural network is able to estimate any output from any image through its learning process. The characteristic vector includes general information relative to the image derived from the first four moments of the image. Image frequency properties are added using the FT. A transformation on the FT on an image was built to define a vector that has a wide variation with the parameters of the MTF model. This characteristic vector is supplemented by components describing the landscape, which in turn is described by the image variogram. The variogram of an image is a measurement of the mean variation of the gray level (squared) of two points separated by a number of pixels (Raffy & Ramstein 1989). Variation can be modeled with a 3-parameter function named "Landscape structure parameters." An interesting property of the model is that two landscape parameters are invariant for the MTF (Delvit et al. 2002). The method has been studied and validated using simulated images. The next stage is to apply it to actual in-flight images.

4.7.2 Least square method

An alternative technique supposes that there is a higher resolution image of the same landscape viewed by the satellite camera being evaluated. The method consists of searching the PSF that minimizes the difference between the lowest resolution image and its simulation obtained from the highest resolution image by convolution with the PSF, then under-sampling. The PSF is calculated from a two-parameter MTF model. Two other parameters are used for a PSF phase shift. At last, two coefficients allow matching the image radiometric levels. In all, there are six parameters.

Simulations have yielded good results. The next step is to evaluate the method with actual images, e.g. SPOT 5 images to compare with classical methods.

5 CONCLUSIONS

Successive commissioning phases of SPOT satellites have resulted in using several MTF measurement methods. Three methods were compared with SPOT 5: the step edge, the point source, and the radial target.

The step edge is well suited to high-resolution sensors because the method needs an artificial edge target to give the best results. The GSD limit with the Salon-de-Provence target is 4–5 m depending on the angle between the edge and the satellite track. Target orientation is a constraint of the method, and only two measurement directions are obtained. The method is rather precise. Furthermore, it can be improved by fitting a model. A drawback is that target maintenance is expensive. It must be repainted every five years.

The spotlight method can be used with lower resolution cameras, typically with a GSD up to 30 m. The advantage is that there is no orientation constraint. The drawback is that the method is more difficult to operate and needs a team on the ground when the satellite is imaging the spotlight.

The bi-resolution method is attractive if there are several cameras with different GSDs aboard the satellite.

The radial target allows a visual assessment of the resolution in addition to the calculated MTF. Unlike the bar target, it is insensitive to track orientation. As for the step target, the main drawback is its cost.

All these methods, with the exception of bi-resolution, present the same constraint: the need to image a specific test site. That is why new methods should be explored. Preliminary results with neural network and least square methods are promising.

ACKNOWLEDGMENTS

Since the launch of SPOT 1, MTF measurements have been obtained with the collaboration of colleagues at ONERA and CNES. The authors are grateful to: G. Begni, E. Breton, X. Briottet, M. Dinguirard, J. Duffaut, O. Hagolle, P. Henry, Ph. Kubik, C. Latry, M.C. Laubies, L. Lebègue, M. Leroy, F. Masson, A. Meygret, J. Perbos, and Ph. Rolland.

The authors would like also to thank the ONERA team in Salon-de-Provence.

REFERENCES

- Anuta, P.E., Bartolucci, L.A., Dean, M.E., Lozano, D.F., Malaret, E., McGillem, C.D., Valdes, J.A. & Valenzuela, C.R. 1984. Landsat-4 MSS and Thematic Mapper data quality and information content analysis. *IEEE Trans. Geosci. Rem. Sens.* 22(3): 222–235.
- Begni, G., Léger, D. & Dinguirard, M. 1984. An in-flight refocusing method for the SPOT HRV cameras. *Photogr. Eng. & Rem. Sens.* 50(12): 1697–1705.
- Delvit, J.M., Léger, D., Roque, S. & Valorge, C. 2002. Modulation Transfer Function measurement using non specific views. In S.B. Serpico (ed.) *Image and signal processing for remote sensing VIII*. Proc. SPIE 4885: 34–45.
- Kubik, P., Breton, E., Meygret, A., Cabrieres, B., Hazane, P. & Léger, D. 1998. SPOT 4 HRVIR first in-flight image quality results. In H. Fujisada (ed.), *Sensors, systems, and next-generation satellites II*. Proc. SPIE, 3498: 376–389.
- Léger, D., Leroy, M. & Perbos, J. 1986. SPOT MTF performance evaluation. In D. Léger, M. Leroy & J. Prebos (eds), *Earth remote sensing using the Landsat Thematic Mapper and SPOT sensor systems*. Proc. SPIE 660: 93–97.
- Léger, D., Duffaut, J. & Robinet, F. 1994. MTF measurement using spotlight. *IGARSS'94, Pasadena, August 1994*.
- Malaret, E., Bartolucci, L.A., Lozano, D.F., Anuta, P.E. & McGillem, C.D. 1985. Landsat-4 and Landsat-5 Thematic Mapper data quality analysis. *Photogr. Eng. & Rem. Sens.* 51(9): 1407–1416.
- McKellip, R. 1997. Resolution verification targets for airborne and spaceborne imaging systems at the Stennis Space Center. In W.R. Watkins & D. Clement (eds), *Targets and backgrounds: characterization and representation III*. Proc. SPIE 3062: 199–205.
- Meygret, A. & Léger, D. 1996. In-flight refocusing of the SPOT 1 HRV cameras. In A.E. Iverson (ed.), *Algorithms for multispectral and hyperspectral imagery II*. Proc. SPIE 2758: 298–307.
- Raffy, M. & Ramstein, G. 1989. Analysis of the structure of radiometric remotely sensed images. *Int. J. Rem. Sens.* 10(6): 1049–1073.
- Rauchmiller, R.F. & Schowengerdt, R.A. 1988. Measurement of the Landsat Thematic Mapper modulation transfer function using an array of point sources. *Opt. Eng.* 27(4): 334–343.
- Robinet, F., Léger, D., Cerbelaud, H. & Lafont, S. 1991. Obtaining the MTF of a CCD imaging system using an array of point sources: evaluation of performances. In *IGARSS'91, Helsinki, June 1991*: 1357–1361.
- Schowengerdt, R.A., Archwamety, C. & Wrigley, R.C. 1985. LANDSAT Thematic Mapper image-derived MTF. *Photogr. Eng. & Rem. Sens.* 51(9): 1395–1406.
- Steel, W.H. 1956. The defocused image of sinusoidal gratings. *Optica Acta* 3(2): 65–73.
- Viallefont-Robinet, F. 2003. Removal of aliasing effect on MTF measurement using bi-resolution images. In R. Meynart, S.P. Neeck, H. Shimoda, J.B. Lurie & M.L. Aten (eds), *Sensors, systems, and next-generation satellites VII*. Proc. SPIE 5234: 468–479.
- Amata, P.E., Bartolucci, L.A., Dean, M.E., Lozano, D.F., Malaret, E., McGillem, C.D., Valdes, J.A. & Valenzuela, C.R. 1984. Landsat-4 MSS and Thematic Mapper data

USGS/OSU progress with digital camera *in-situ* calibration methods

D.C. Merchant, A. Schenk & T. Yoon

The Ohio State University, Department of Civil, Environmental Engineering and Geodetic Science, Columbus, Ohio, USA

A. Habib

Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, Canada

ABSTRACT: The United States Geological Survey (USGS), in cooperation with The Ohio State University (OSU) and Calgary University (CU), have developed procedures and software for calibrating metric quality aerial cameras; both film- and charged-couple device (CCD) based sensors. With the advent of the global positioning system (GPS), and its use for positioning the exposure station of an aerial camera, it became necessary to establish accurately camera interior orientation under operational circumstances (*in-situ* calibration) to assure agreement of the photogrammetric procedure results with positional data provided by GPS. Disagreement between photogrammetric resection methods based on laboratory calibration and GPS results were consistently as large as one part in 1000 of the flight height. This paper describes the development of the airborne method of camera calibration, software development, and some results of accuracy improvements when using the *in-situ* method of camera calibration.

1 INTRODUCTION

With the advent of global positioning system (GPS) technology and other airborne sensors, it became necessary to revise the traditional concept of camera calibration. Influences on the camera and its spatial relationships to other sensors require that data used in calibration be collected under conditions closely approximating those expected in application of the photogrammetric system and its associated sensors (the *in-situ* method).

This paper presents a revised concept of calibration of aerial photogrammetric systems. Justification for this revised approach is presented by comparing photogrammetric results to corresponding GPS results. In terms of positional information these results should be in agreement if the full spatial accuracy of GPS is to be exploited.

The US Geological Survey (USGS) sponsored development of appropriate software programs for use with the aerial method of calibration. Two programs are described for image measurement and calibration purposes. These programs are specialized to work in connection with use of aerial imagery, controlled by GPS, and taken over a suitably targeted control range. Either conventional film based or digital cameras may be treated for calibration.

2 BACKGROUND

Rapid acceptance of digital cameras range across the photographic industry from low-end 35 mm, hand-held cameras to special purpose mapping cameras. At the low end, digital cameras offer advantages over traditional film-based cameras for measurement applications. Digital advantages in range of light sensitivity (speed) over a wider spectral response, and immediate access to imagery, are well known. For measurement applications, digital cameras provide a direct means of recovery of the internal orientation, since its basic character provides a known and stable relationship between the lens and the captured image. As a consequence, there is a growing interest in using digital cameras for measurement applications, generating a need for an efficient and relatively inexpensive means of camera calibration. This need has recently been met by the USGS facility for camera calibration located in South Dakota.

For cameras intended for aerial surveying and mapping applications, both digital and film-based, the introduction of added sensors such as GPS and inertial navigation system (INS) require that careful consideration be given to a systems approach to calibration to assure the added sensors are fully exploited metrically.

In recognition of this need, the United States Geological Survey (USGS) awarded a grant to provide and demonstrate a means for camera system calibration for aerial mapping cameras. During the contract a Nikon D1X camera equipped with wide angle and narrow angle lenses was flown over a specially developed camera calibration range. Trimble GPS receivers controlled the imagery in the air. In addition, film-based photography was collected by a Zeiss LMK 15/23 camera, also controlled by GPS receivers. Programs were developed for image measurement and subsequent camera calibration using Visual C++ software.

During this development, a number of interested organizations participated along with USGS. Software development was a cooperative program between The Ohio State University and the University of Calgary with a grant from USGS. The Madison Test and Calibration Range was developed and maintained by the Ohio Department of Transportation / Aerial Engineering Office. Flight testing of the digital cameras was accomplished by Topo Photo, Inc. of Columbus, Ohio.

2.1 *In-situ approach to calibration*

In a benchmark publication by Eisenhart (1962), a rationale for calibrating measurement systems is set forth. His work provides clear guidance for designing and applying a calibration program for airborne sensors associated with the aerial mapping industry. To summarize Eisenhart's concept of calibration: (a) establish first the "Measurement System Specifications;" and, (b) exercise the measurement system as specified and compare results to a standard of higher accuracy until sufficient information is available to achieve a "*State of Statistical Control*."

System specifications describe all aspects of the system including hardware, software, environment of the system, and operational procedures necessary to achieve the final measurement. Procedures may include specified ranges within which the system will operate. In the aerial case, a range in altitude can be specified that represents an important practical aspect of the concept.

By accepting the concept of measurement system calibration, it is clear that certain aspects of camera calibration, as traditionally practiced, need to be reconsidered. For aerial cameras, independent of the added sensors, the primary difference in measurement accuracy for application is due to temperature differences between a laboratory and an *in-situ* method of data collection. This difference is most evident for an open-port windowed aircraft, most typical of the aerial industry. For the closed-port case, the addition of the window adds an additional optical component to the system not conveniently treated during laboratory calibration.

Experimental data collected under operational circumstances tend to support the need for adopting an *in-situ* approach to calibrating aerial cameras.

For the aerial case, the choice of camera platform, as with digital cameras, offers a wide range of possibilities with the technically best being the most expensive. The choices discussed below range from a minimum cost, single-engine, open-port aircraft to a multi-engine, windowed-port aircraft.

2.1.1 *Single-engine open-port*

In the United States one popular, relatively inexpensive aircraft is the Cessna 207. The aircraft represents an adequately stable platform for carrying the camera, mount, and supporting equipment. However, when considering use of airborne GPS for survey control, one must assume that a significant disturbance is generated in the volume of air below the aircraft through which the camera must function. The engine exhaust may be diverted; however, cooling air for the conventional reciprocating engine can cause a rise in temperature from external ambient to cowl exit of 66°C. In addition, as with all open-port systems, the influences of temperature differences between cabin and external air will have an influence on the metric characteristics of the camera.

2.1.2 *Multi-engine open-port*

Probably the most widely used aircraft for collecting photography of photogrammetric quality is a light twin, open-port platform. With this aircraft, no significant disturbance to the volume of air beneath the aircraft is expected. However, the influence of the temperature difference between cabin and outside air can be extreme. This difference causes a change in relationship between the optical and image collecting components of the aircraft, which usually is seen as a centering error that can be represented as a corresponding change in the camera constant. Accordingly, recording temperatures within the cabin and at the camera lens may become a means for accounting for the open-port installation errors.

2.1.3 *Multi-engine windowed-port*

The ideal, but most expensive aerial platform includes a windowed-port. The window consists of high quality glass, as specified by military and other standards. Clearly, in operation, the window becomes part of the optical system and must be included as part of the photographic system during calibration. When operating without cabin pressurization, the influences of temperature differences are mitigated. When pressurized, the differences between cabin and external pressure generates a stress/strain relationship on the window, producing an image deformation that requires additional mathematical modeling during the calibration process (a component of the system specification).

3 USGS/OSU PROJECT

The USGS project is an initial step toward camera system calibration on a national basis. Software was developed for the image measurement process and for the subsequent computation of interior orientation, the primary components of camera calibration. Subsequently, a series of flights was conducted over the Madison range to verify the systems approach to camera calibration for both digital and film-based cameras.

3.1 Software development

Programming was accomplished in Visual C++ language.

3.1.1 Image measurement program

The first program, termed "Image Measure" (IM) software, was designed for measuring image coordinates and producing files for subsequent introduction to the calibration program. Observation of target images is facilitated by computing a single photo resection after the first four targets have been identified and measured manually. At that point, the program indicates the residuals of the fit to control, and selects only those targets that appear within the current photograph. This is followed by automatic movement of the measurement mark to the first of the imaged targets in the selected set of targets. When the coordinates of the first point have been determined, the observer can record them rapidly before the software automatically directs the observer to the next target in the set. Auto-location is accurate to several pixels for a system of low distortion. After all the reduced target list images have been brought forward for fine pointing, the observer saves results and moves on to the next photograph.

For processing of film-based images, the film is first scanned and the fiducials are measured. This is followed by a 2-D transformation into a fiducial coordinate centered system. All subsequent image measurements on this frame are transformed accordingly, resulting in photo coordinates in a fiducial system, but corrected for film deformation. Digital image coordinates are measured directly from the photo file, then transformed by a rigid-body transformation to the photo center, resulting in a conventional photo image coordinate system.

Additional input files provide the GPS coordinates of the antenna phase center, the survey coordinates of the targets, and the first approximations to parameters of both interior and exterior orientation along with associated variance/covariance files for weight constraint purposes.

A right-handed coordinate system and right-handed rotations are assumed in all cases. When all images on a given photo are measured and transformed to photo

coordinates, a final single photo resection is computed, resulting in the angles relating the photo coordinate system to the ground system of coordinates. This transformation of coordinate systems proceeds from the object space to the image space. Given the rotation matrix expressing the relationship of ground to camera coordinate systems, its inverse is used to transform the photo parallel offsets, and GPS phase center to camera entrance node into corresponding components in the ground control system. The exposure station then is computed by addition of the transformed spatial offsets to the phase center coordinates of the antenna.

For any given photo, final processing applies atmospheric refraction correction using the Saastamoinen model (1972). The final step applies the transformed spatial offsets, and antenna phase center to camera node directly to the GPS coordinates for any given exposure. Results of this program are data files containing refined photo coordinates of targets (lens distortions remain) and exterior orientation.

3.1.2 Camera calibration program

The calibration program titled "Bundle Adjustment with Self Calibration" (BASC) is designed to use files produced by the image measurement program (PIC). Additional files used by the program include a description of the camera including first approximations to the interior orientation, target survey coordinates, and variance/covariance information for all parameters describing interior and exterior orientation, image measurements, and target coordinates.

The mathematical model used is the SMAC model as defined by the USGS, a model that represents focal length correction, symmetrical and decentering distortion, and location of the principal point.

In accord with this SMAC model, radial distortion is expressed as: (δx , δy)

$$\delta x = (x - x_p) (K_0 + K_1 r^2 + K_2 r^4 + K_3 r^6 \dots) \quad (1)$$

$$\delta y = (y - y_p) (K_0 + K_1 r^2 + K_2 r^4 + K_3 r^6 \dots) \quad (2)$$

Where: x_p , y_p are photo coordinates of the principal point

$$r^2 = (x - x_p)^2 + (y - y_p)^2 \quad (3)$$

K coefficients representing radial, symmetrical distortion.

The distortion due to decentering of the compound objective is expressed as: (Δx , Δy)

$$\Delta x = (1 + P_3 r^2) (P_1 (r^2 + 2x^2) + 2P_2 xy) \quad (4)$$

$$\Delta y = (1 + P_3 r^2) (2P_1 xy + P_2 (r^2 + 2y^2)) \quad (5)$$

Where: P coefficients represent decentering distortion.
The corrected photo coordinates are then:

$$x_c = x + \delta x + \Delta x \quad (6)$$

$$y_c = y + \delta y + \Delta y \quad (7)$$

Note that the K_0 represents a scalar term for photo coordinates. Accordingly, it accounts for small differences in the chosen value of focal length. This permits use of an arbitrary but close approximation when using the nominal focal length associated with the lens design in the computations.

3.2 Flight test verification

Flight testing of both digital- and film-based cameras was conducted concurrently with the development of the software programs. This assured that all elements of the calibration process could be identified and treated accordingly during development of the programs. It also verified that digital cameras, even with narrow fields of view, can be accommodated by the *in-situ* approach to aerial camera calibration. In addition, these flight tests demonstrated the contrast in results between a laboratory and *in-situ* form of calibration. These differences further justify the need for a systems approach to aerial camera calibration.

3.2.1 Madison test and calibration range

The Madison range currently consists of about 100 targets located within a 1.6 km by 2.6 km region, 50 km west of Columbus, Ohio. Target coordinates were measured by GPS methods with elevations augmented by spirit leveling. Adjustment results indicate that the internal accuracy of the network is better than 2 cm on each axis and includes the base station, MAD1. The base station is located 5 km from the range center at the Madison County Airport. The range was constructed and is maintained by the Office of Aerial Engineering of the Ohio Department of Transportation (ODOT). The Range is centered at latitude 39° 56' 25" N. and longitude 83° 31' 28" W (Fig. 1).

Targets are painted on existing asphalt roads and are centered on magnetic PK nails. Targets are 2.4 m in diameter with a 0.80 m flat white center. The targets were designed to provide optimum images for automatic pointing and recording of image coordinates for film-based cameras flying at about 1200 m above the field.

To image sufficient targets on a single photo when using conventional digital cameras, targets were densely distributed in the vicinity of the intersections of US40, Potee Road, and Markley Road. In the vicinity

of the intersection, targets were separated by 10 m. As targets radiated from this intersection, the intervals were sequentially increased by the cube root of 2. In this way, relatively narrow field cameras flying at low altitudes can acquire sufficient target images for calibration purposes. The target distribution for use by low-flying or narrow field camera systems is shown in Figure 2.

3.2.2 Resection comparisons to GPS exposure station

To demonstrate the improvements offered by an *in-situ* approach to camera calibration when compared to a



Figure 1. The Madison test and calibration range.



Figure 2. High density range for digital cameras.

conventional laboratory calibration, two single photo resection computations were computed. The first used results from a laboratory calibration; the second computed results of an *in-situ* calibration. The exposure station coordinates for each case were compared to the station coordinates derived from GPS.

3.2.3 Flight tests of the Zeiss LMK/15-23 camera In open-port aircraft

The standard film-based mapping camera was flown in a Patenavia light twin aircraft equipped with an open port. The comparison of results obtained from single photo resections based on laboratory and *in-situ* calibrations, compared to exposure station coordinates provided by GPS, are presented in Figure 3. The differences in resected elevations [Z] from those provided by GPS are seen clearly in Figure 3a. These

differences are produced by what may be termed a *centering error*, an error that may be corrected by choice of an appropriate value for calibrated focal length. It is also interesting to note that both calibrations produce nearly the same horizontal components of root mean square error (RMSE) and bias. Figure 3b provides comparisons for the same aircraft under the same circumstances but at a higher altitude above ground. The resected results at 3070 m are the same as those for 1260 m except for the magnitude of the centering error, due in part to the differences in scale of the imagery.

3.2.4 Flight tests of the Wild RC30 camera in a windowed-port aircraft

To demonstrate the influence of a port window on results of resection, a Cessna Citation aircraft was flown

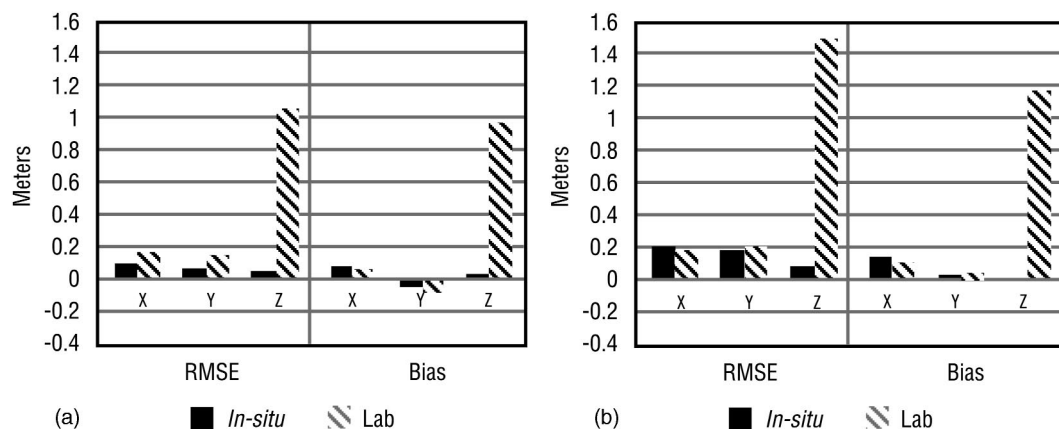


Figure 3. Single photo resection comparisons for an open-port, twin engine aircraft: (a) LMK 15/23 camera at 1260 m AGL for 7 photos; (b) LMK 15/23 camera at 3070 m AGL for 9 photos.

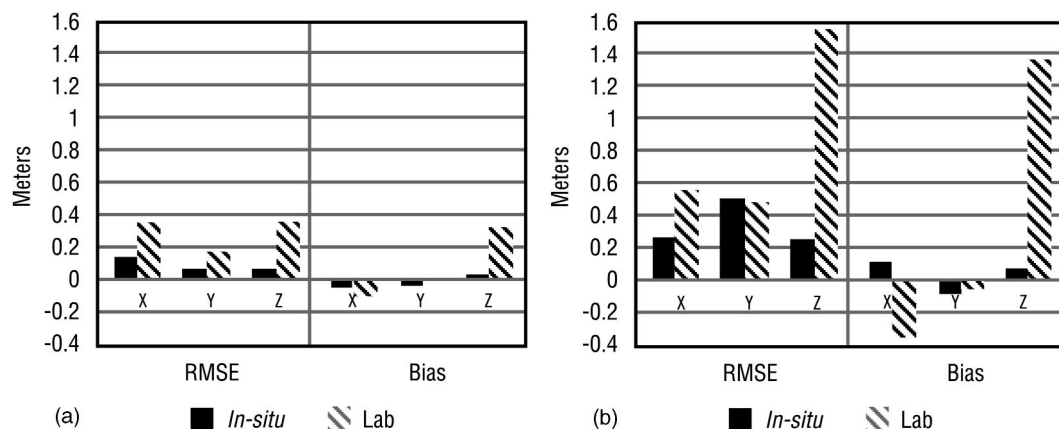


Figure 4. Single photo resection comparisons for a windowed-port, multi-engine aircraft with Wild RC30 15/23 camera: (a) unpressurized camera at 1316 m AGL for 12 photos; (b) pressurized camera at 5817 m AGL for 5 photos.

over the Madison Range at two different altitudes. The lower altitude did not use cabin pressurization, the higher did.

Normally, at the lower altitude of 1316 m AGL, no pressurization is used. Figure 4 indicates the character of the window's influences. When no pressurization of the cabin is used, the window tends to introduce a moderate centering bias and a small increase in RMSE (Fig. 4a). This is probably due to the inability of the mathematical model used for calibration to represent the deformations of imagery introduced by the window.

The influences of cabin pressurization for the higher altitude flight is indicated in Figure 4b. Of particular interest in the pressurized case is the inability of the calibration model to account for image deformations for both the *in-situ* and laboratory procedures. This is implied by the large bias errors in elevation, and for horizontal components when compared to the GPS result.

4 CONCLUSIONS

Recognizing significant differences between laboratory and *in-situ* methods of calibration, and preparing appropriate software to conduct an *in-situ* calibration has taken the USGS closer to having a calibration

procedure that can accommodate added airborne sensors such as GPS effectively.

The Eisenhart concept of "measurement system calibration" provides guidelines that can be adapted well to calibration of aerial cameras and its supporting equipment and procedures. The results of an *in-situ* calibration represent one element in the "measurement system specification". It still remains to establish a "state of statistical control" through an ongoing process of testing the measurement system by comparing it to a standard of higher accuracy, such as provided by a calibration and test range.

This work was undertaken through support given by the USGS. In addition, recognition and thanks are given to the Aerial Engineering Office of the Ohio DOT for range preparation and film based camera flight tests, and to Topo Photo Inc. for conducting flight testing of the digital cameras.

REFERENCES

- Eisenhart, C. 1962. Realistic evaluation of the precision and accuracy of instrument calibration systems. *J. Res. Natl. Bur. Stand. - C. Eng. & Instr.* 67C(2): 161-187.
- Saastamoinen, J. 1972. Refraction. *Photogr. Eng.* 38(8): 799-810.

Issues and methods for in-flight and on-orbit calibration

K. Jacobsen

University of Hannover, Hannover, Germany

ABSTRACT: The result of a geometric laboratory and an in-flight calibration of analogue or digital cameras may be not the same, because of different environmental conditions and mechanical stress during transport. The mathematical model for handling images has to respect instant geometry. Not all components of laboratory calibration will change. Distortion values usually are stable over longer time. On the other hand, focal length and principal points are not very stable. By self-calibration with additional parameters, the real geometry of a sensor can be determined and the results can be used partially as a pre-correction for other images. Individual additional parameters must be able to solve special geometric problems, so the formulas used have to be fitted to the special geometric problems of the sensor. The residuals at ground control points and in the images must be analyzed for unsolved geometric problems. Not all parameters can be determined by self-calibration. In particular, the focal length only can be improved if control points with quite different elevations in the same test area are available; or, if the position of the projection centers is known. Using a larger number of additional parameters usually results in high correlation causing misleading results. For this reason additional unknowns have to be checked for their significance and for exceeding a correlation tolerance limit.

1 INTRODUCTION

A correct determination of the 2-D or 3-D geometry of imaged objects requires reconstructing the imaging rays. This includes exterior and interior orientation. The exterior orientation describes the location of the projection center and the attitude of the bundle of rays, while the geometry of the bundle of rays will be reconstructed by the measured image position and the interior orientation. There is one projection center for perspective images like those taken by an analogue frame camera or digital cameras with a charged-couple device (CCD) array. Digital cameras based on a linear CCD, or panoramic cameras, have the same exterior orientation only for each image line. Corresponding to this, the interior orientation is restricted to this line. For these sensors, reconstruction of the image geometry requires mixing the interior and exterior orientation. Still more difficult are systems using just an imaging detector (like Landsat). Here the interior orientation is reduced to the relationships of the moving parts.

The simple mathematical model is based on the collinearity condition. Object point, projection center, and image point are located on a straight line and the image is generated in a perfect plane or line. Of course this does not respect the real physical situation: refraction has to be considered, the imaging surface is not a

perfect plane, the CCD-lines are not totally straight lines, and the imaging rays do not transit the optics without change of direction. The expression “imaging ray” is a simplification because of the wave-structure of light; nevertheless it is used for geometric reconstruction. The field-of-view is a function of the focal length. Usually it is determined by laboratory calibration, but it changes with environmental conditions. Hence, the calibrated focal length is only an approximation.

2 AERIAL APPLICATIONS

2.1 *Perspective cameras*

A schematic for a perspective camera is shown in [Figure 1](#). The difference between the entrance nodal point, which is identical to the projection center in the object space, and the exit nodal point (the projection center in the image space), will only be respected in the case of direct observations of the projection centers as, for example for combined block adjustment with projection center coordinates, or by direct sensor orientation. Usually a simplified model with the positive location of the image is used, as shown in [Figure 2](#). The geometric condition of the original image above the projection center is identical to the positive location

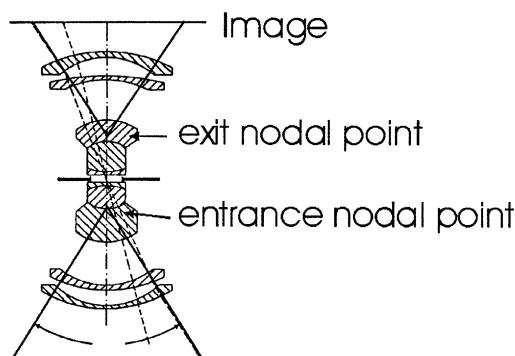


Figure 1. Optical system.

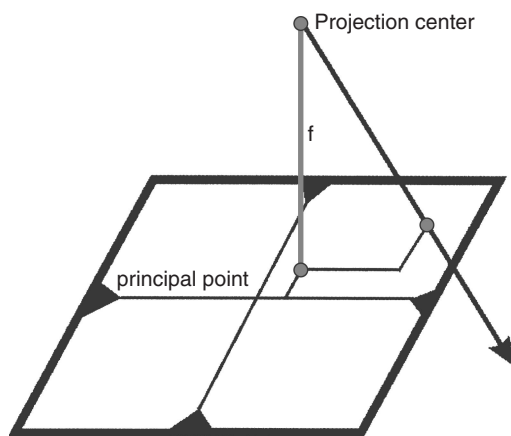


Figure 2. Simplified model.

below the projection center. In this simplified model there is a projection centre with the distance of the focal length, f , from the image plane and a principal point which has the closest distance from the image plane to the projection center. Like the focal length, the location of the principal point will be determined by laboratory calibration, but it also may change slightly over time or by thermal influences. Analogue film cameras describe the location of the principal point indirectly by means of the fiducial marks. CCD-array cameras do not need fiducial marks; the CCD-array is fixed in the camera, so the location of the principal point is known from laboratory calibration using rows and columns.

Analogue film is not as stable as it should be. It changes dimensions during processing. Investigations by Koch (1987) have shown a scale reduction of up to 0.4 percent or 92 microns across a 230 mm film format. The dimensional change is different in the x- and y-directions up to 0.3 percent or 69 microns have been detected. In addition, there may be also an angular

affinity up to 0.15 percent or 35 microns. The change in film dimensions is not the same for all photos on a film roll; a variation of 0.1 percent or 23 microns is usual. But the change of the film dimension is not a problem; it will be compensated by the standard affinity transformation to the calibrated fiducial mark coordinates. Only a deformation of the fiducial mark location will cause an affinity deformation of all images taken by a given camera.

Over-determination of image orientation by bundle block adjustment permits self-calibration using additional parameters. Systematic discrepancies of image coordinates against the mathematical model used can be detected and addressed using additional unknowns. There are different sets of formulas for additional parameters of perspective images. Gotthardt (1975) has developed a set of parameters that is able to eliminate systematic image errors at the position of the Gruber points (3×3 positions). This set of formulas was used by Ebner (1976) and is known under his name. Grün (1979) extended this set of formulas to 5×5 image positions (Gruber points and the points on half the distances). The formulas used by Ebner are optimal if the image points are just located in the Gruber points. The set extended by Grün often causes problems by having too many parameters that are strongly correlated. The author developed the set of parameters shown in Table 1.

This set of parameters is a combination of physical values extended by other general parameters. Parameters 1 and 2 describe an angular affinity and an affinity deformation. Parameters 9, 10, and 11 compensate radial symmetric distortions, while 7 and 8 characterize tangential distortion (Fig. 3). The other parameters are similar to a Fourier series in polar coordinates and are able to fit remaining systematic effects. If the image positions are well distributed, as is usually the case, this set of parameters shows little correlation one to another, and has the additional advantage of providing physical explanations. A tangential distortion can be caused by off-center location of individual lenses. This problem is not normally checked by standard laboratory calibrations, but it often exists.

Very often, especially for close range applications, a set of polynomial parameters is used to characterize radial symmetric distortion like $K1 \cdot r^3$, $K2 \cdot r^5$, $K3 \cdot r^7$. This set has the disadvantage of strong correlation and the higher degree values are only effective in the extreme image corner. With a set of 11 aerial images, the values $K1$ and $K2$ have been adjusted, leading to a correlation between both of 0.91 which is above the critical value. The same data set processed with parameters 9, 10 and 11 in Table 1 lead to a maximal correlation of 0.32. All additional parameters of radial symmetry are highly significant with student test values larger than 4.5. Because of the strong correlation, the parameter $K2$ was not accepted, so only $K1$, the

Table 1. Additional parameters for perspective images used in the Bundle adjustment, University of Hannover program (BLUH).

Parameter	x' equation	y' equation
1	$x' = x - y \cdot P1$	$y' = y - x \cdot P1$
2	$x' = x - x \cdot P2$	$y' = y + y \cdot P2$
3	$x' = x - x \cdot \cos 2b \cdot P3$	$y' = y - y \cdot \cos 2b \cdot P3$
4	$x' = x - x \cdot \sin 2b \cdot P4$	$y' = y - y \cdot \sin 2b \cdot P4$
5	$x' = x - x \cdot \cos b \cdot P5$	$y' = y - y \cdot \cos b \cdot P5$
6	$x' = x - x \cdot \sin b \cdot P6$	$y' = y - y \cdot \sin b \cdot P6$
7	$x' = x + y \cdot r \cdot \cos b \cdot P7$	$y' = y - x \cdot r \cdot \cos b \cdot P7$
8	$x' = x + y \cdot r \cdot \sin b \cdot P8$	$y' = y - x \cdot r \cdot \sin b \cdot P8$
9*	$x' = x - x \cdot (r^2 - 16384) \cdot P9$	$y' = y - y \cdot (r^2 - 16384) \cdot P9$
10*	$x' = x - x \cdot \sin(r \cdot 0.049087) \cdot P10$	$y' = y - y \cdot \sin(r \cdot 0.049087) \cdot P10$
11*	$x' = x - x \cdot \sin(r \cdot 0.098174) \cdot P11$	$y' = y - y \cdot \sin(r \cdot 0.098174) \cdot P11$
12	$x' = x - x \cdot \sin 4b \cdot P12$	$y' = y - y \cdot \sin 4b \cdot P12$

* $r^2 = x^2 + y^2$ arctan $b = y/x$. Constant values in parameters 9–11 are scaled to the actual image size. The listed constant values are valid for maximal radial distance in the image of $162.6 \text{ mm} = 115.0 \cdot 2^{1/2}$.

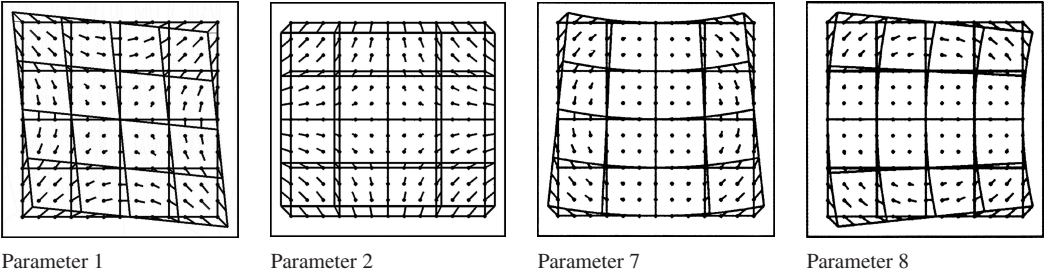


Figure 3. Effect of some additional parameters used in program BLUH.

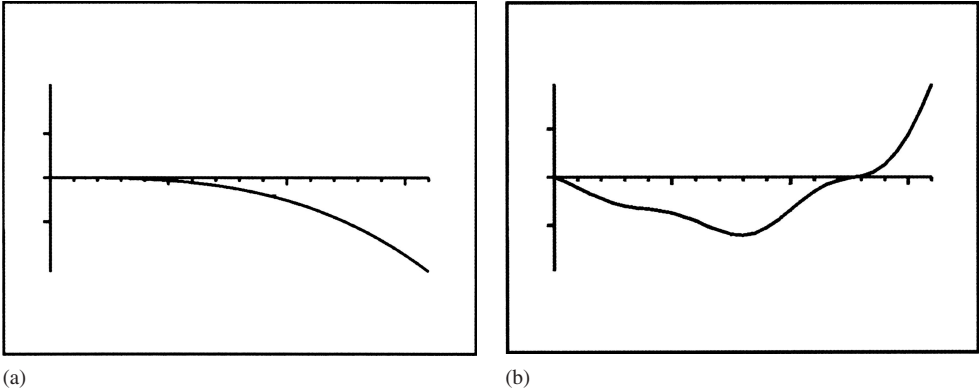


Figure 4. Radial symmetry lens distortion: (a) polynomial K; (b) parameters 9–11.

radial distortion shown in Figure 4a, could be characterized. With parameters 9, 10, and 11, the more detailed situation shown Figure 4b was derived.

Empirical tests showed the negative influence of highly correlated and unjustified additional parameters

(Jacobsen 1980, Jacobsen 1982a). The unjustified parameters should be excluded from the adjustment, as well as the very strongly correlated ones. In program system BLUH the additional parameters are analysed by a student test. The parameters with values

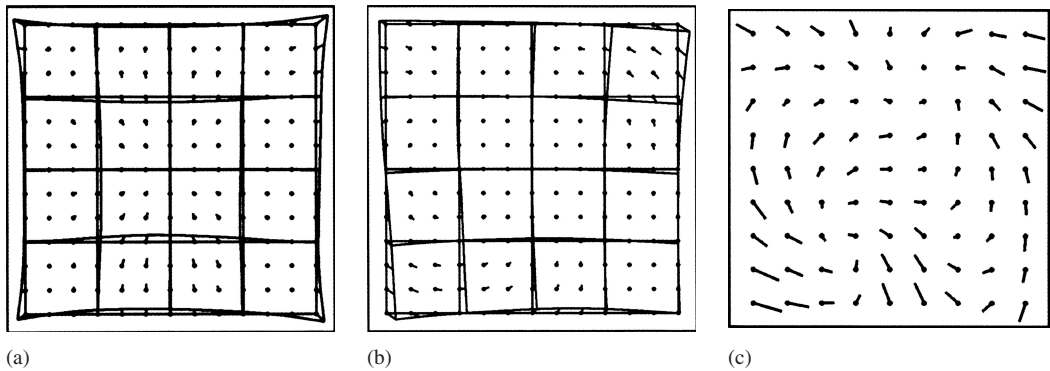


Figure 5. Systematic image errors: (a) OEEPE test block, company 2, averaged image coordinate residuals (adjustments without self-calibration); (b) OEEPE test block, company 2, showing tangential distortion (b) (the largest vector is 20 μm); (c) OEEPE test block, company 1 (largest vector 7 μm).

below $T = 1.0$ are excluded, as well as parameters causing correlation coefficients exceeding 0.85 (Jacobsen 1982b).

Using additional parameters to determine image coordinates characterizes systematic image errors. Figure 5 shows these “systematic image errors” determined by self-calibration using the set of additional parameters given in Table 1 for images taken with aerial cameras. The bundle block adjustment used the test field at Frederikstad, Norway with the data of the OEEPE test block (Heipke et al. 2001). The middle data set (151 photos, 4203 photo points) has been handled also without self-calibration. The resulting image coordinate residuals have been averaged in 9×9 image sub-areas (Fig. 5a). The trend of the discrepancies agrees very well with the corresponding systematic image errors (Fig. 5b). Such a good fit requires a higher number of observations. In general, the analysis of residuals can be used to determine remaining systematic effects that cannot be handled by using additional parameters.

Systematic image errors determined by additional parameters can be used for pre-correcting image coordinates. This is only valid if the systematic image errors do not change. Image geometry of analog aerial photos is not only dependent upon the camera itself, but also on the film magazine and pressure plate. Moreover, systematic image errors change from mission to mission. Only the radial and tangential lens distortion, and the affinity deformation, are more stable. Radial and tangential deformation can be determined with only a few images and with only a few control points. The affinity deformation requires control points or crossing flight strips.

Under usual conditions of an aerial mission, the focal length cannot be determined by bundle block adjustment with self-calibration. Geometric information depends on the Z-variation of the control points.

The standard deviation of well defined ground point heights determined from wide angle photos taken from a flying height of 1000 m is $SZ = 48$ mm under the optimal condition of an x-parallax accuracy of 5 μm. For a vertical height difference between control points:

$$SDZ = SZ \cdot \sqrt{2} = 68 \text{ mm} \quad (1)$$

To achieve a standard deviation of the focal length of 15 μm, a height difference between vertical control points of $153 \text{ mm} / 15 \mu\text{m} \times 68 \text{ mm} = 680 \text{ m}$ is required. Of course, with 9 control points in both height levels, this would be reduced by the factor of 3, or a height difference of 330 m at a flying height of 1000 m. But even under such conditions there are configuration problems and a standard deviation of 15 μm in the focal length is not sufficient. A different situation exists if the projection center coordinates are available from relative kinematic GPS-positioning.

With this, the focal length can be determined, but it will be totally correlated to constant Z-errors of the GPS-positioning which cannot be avoided. Only in a combined adjustment with images taken from at least 2 different height levels, can the constant GPS-errors be separated from the focal length, as shown in Table 2.

With the data set of the referenced OEEPE test block, standard deviations of the focal length of $Sf = 3$ to 4 μm are noted. Even here there is a limitation because the focal length changes with air temperature and pressure (Meier 1978). Using the determined focal length for other missions with an image scale outside the range used for in-flight calibration, is limited, but it is a better estimation of real conditions than a focal length derived from laboratory calibration. Camera calibration under flight conditions requires correct data handling. The effect of Earth curvature and the net projection should be accounted. Proper data handling requires an adjustment in an

Table 2. Correction of focal length and principal point by bundle block adjustment using GPS values of projection centers in 2 different height levels (program BLUH).

	GPS shift			Standard deviation		
	X	Y	Z	SX	SY	SZ
GPS values (set 1)	.001	.003	.086	.007	.007	.004
GPS values (set 2)	-.016	.054	.257	.014	.014	.006
Change of focal length			-.037			
GPS shift (absolute)			-.095			
Corresponding shift of principal point	.001	.000				

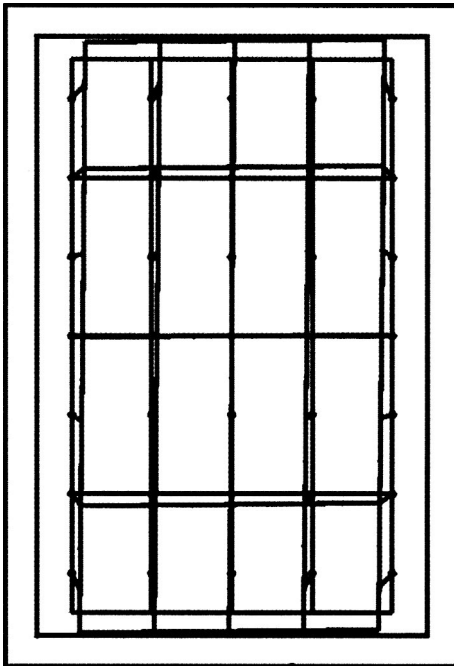


Figure 6. Systematic image errors of Z/I DMC, largest vector = 11 μ m.

orthogonal coordinate system, like a tangential system, to the Earth ellipsoid. Also, refraction must be considered. Most published formulas for the refraction correction are valid up to a flying height of about 14 km. For space application, polynomial formulas are leading to totally wrong results. Deformation of the radiance by atmospheric refraction can be corrected by using image coordinates.

The situation for digital perspective cameras is less complex. CCD-arrays have the advantage of a fixed surface. Usually they are perfectly flat and have no

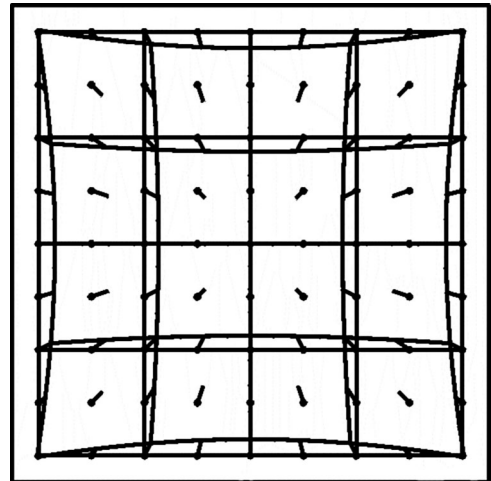


Figure 7. Systematic image error of Rollei Q16, $f = 41$ mm, largest vector = 270 μ m.

temporal deformations. In addition, the positions inside the CCD are nearly perfect. Existing distortions are dominated by the influence of the optics. Experiences with the Z/I Imaging digital mapping camera (DMC) have shown high internal stability of this system (Fig. 6). The synthetic image generated from the DMC is based on a configuration of 4 individual CCD-cameras, each of which is calibrated, and the calibration values are respected. The synthetic image should be free from systematic image errors. A bundle block adjustment over a test field with some crossing flight lines showed only a small affinity deformation. A check of the CCDs confirmed that the elements were not exactly square. This affinity deformation was taken into account, leading to images free of systematic image errors.

Other digital cameras equipped with only one CCD-array showed a dominating radial symmetric characteristic (Fig. 7). The lens systems used are off-the-shelf

and not made specifically for photogrammetric purposes. No special care has been taken for distortion-free geometry, and radial symmetric distortions in the range of $200\text{ }\mu\text{m}$ are usual. The distortion of the Rollei Q16, 41 mm-optics, shown in Figures 7 and 8, has been determined from 12 overlapping images. The radial symmetric additional parameters 9, 10, and 11 have student test values between 302 and 67 (size of the parameters in relation to their own standard deviation). The significance level always starts at the value 3. The correlation between these parameters does not exceed 0.10, indicating the significance of the results.

Cameras with unstable optics, like the thermal camera ThermScan with a variable focus, may show a tangential distortion like that shown in Figure 9. A variable focus does not allow calibration, even if it is set to infinity. After changing the focus and resetting it to infinity, the geometric properties will not be the same. The same problem exists if the optics are exchanged. If they are mounted again, the geometric

properties will not be the same. Of course, if the optics are fixed, they can be calibrated, and may be stable until there is a change in the optical system. This is not the case for zoom lenses. The inner optical system usually is not fixed very well, so the camera geometry may change as a result of even mild vibration.

2.2 Other systems

Not all perspective aerial cameras are panoramic cameras and CCD-line scanners. Self-calibration by additional parameters of panoramic cameras is not a calibration. With the geometry having been determined, the dynamic effect of the scan can be computed (Fig. 10). It is dependent not only on the scan speed and the aircraft speed, but also upon angular effects of roll,

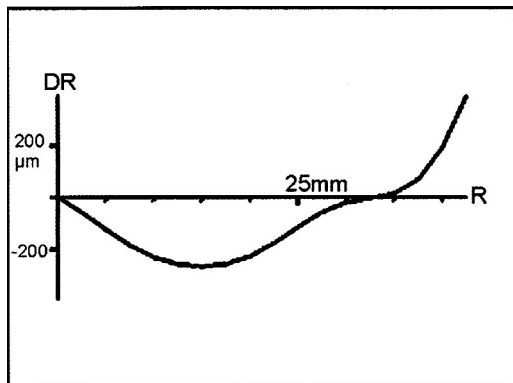


Figure 8. Radial symmetric distortion of Rollei Q16, $f = 41\text{ mm}$.

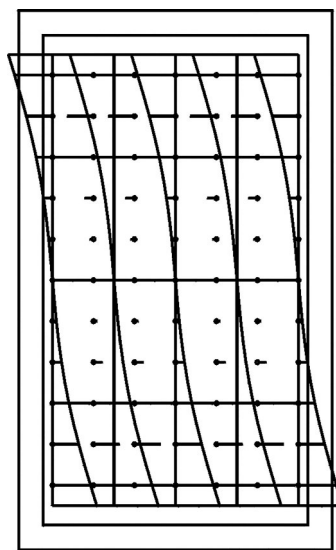
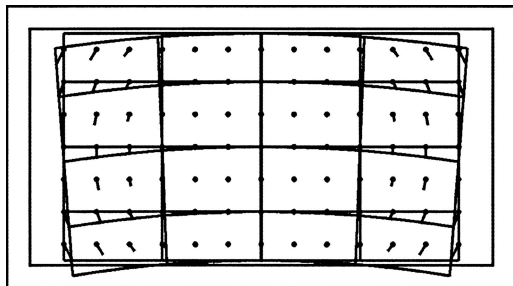
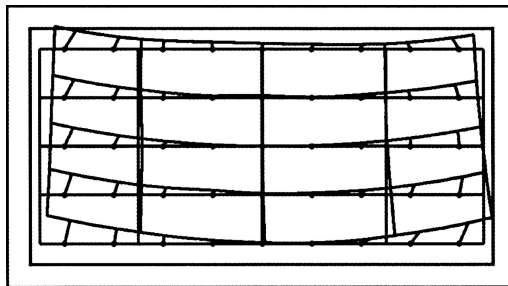


Figure 10. Systematic discrepancies of a panoramic camera.



(a)



(b)

Figure 9. Systematic image errors of the ThermScan camera: (a) wide angle optics, largest vector 1.5 pixels; (b) normal angle optics, largest vector 1.7 pixels.

pitch and yaw. The linearity of the image line usually cannot be calibrated under flight conditions but in any case and is negligible compared to the dynamic effects. In Figure 11, the laboratory calibrations of two high-resolution stereo camera (HRSC) CCD-lines are shown. These include the non-linearity of the CCD-line and the rotation against the reference direction. The general trend can be determined also under flight conditions, if crossing flight lines are available; or, there is a higher number of control points. For the HRSC, like the corresponding Leica airborne digital sensor (ADS40), the relation of the inertial measurement unit (IMU) to the CCD-line cameras is

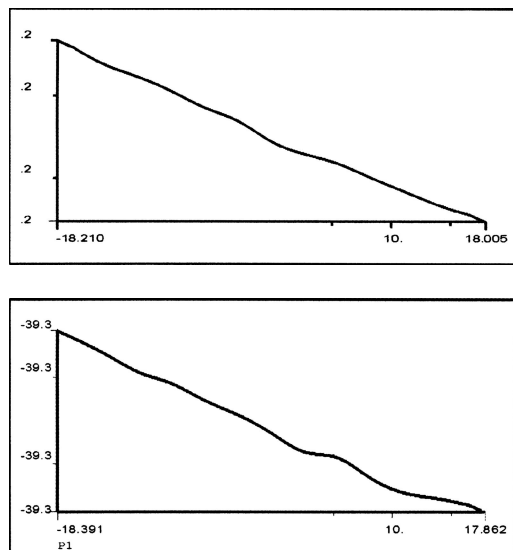


Figure 11. HRSC CCD-line linearity.

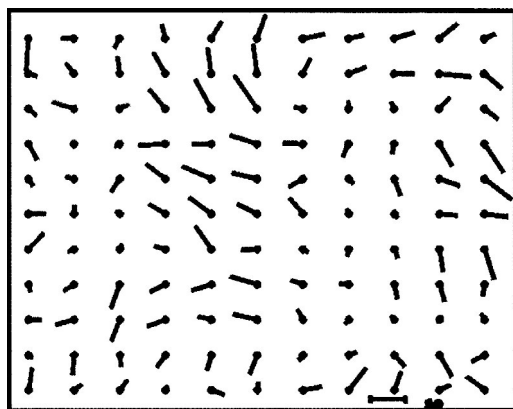


Figure 12. TK350 systematic film deformation determined with 1050 reseau points, mean square deformation = $10 \mu\text{m}$.

more important. This boresight misalignment has to be determined under flight conditions.

3 SPACE APPLICATIONS

By theory, there are no differences between aerial and space applications. The difference lies in the sensors used and the restriction of the orbit. Because of the large areas covered by space images, a mathematically correct ground coordinate system is required. Earth curvature also requires consideration.

Only Russia is still acquiring photos from space. For these photos, conditions similar to those for aerial images exist. Some images, like the TK350, have a reseau platen. This reduces the problems with film stability and the imaging plane.

The systematic film deformation of a TK350 photo determined by the reseau can be seen in Figure 12. This is an individual film correction not requiring calibration. The systematic image errors in Figure 13, however, indicate that the camera has geometric problems. Of course the results of a single determination of image geometry has to be verified by independent data sets.

The determined geometry of aerial images computed by self-calibration with additional parameters is based mainly on the over-determination of a bundle block adjustment. If crossing flight directions are available, ground control points (GCPs) have little influence, but it is different for space images where there are single images or image pairs. The geometric characteristics of single images can be determined

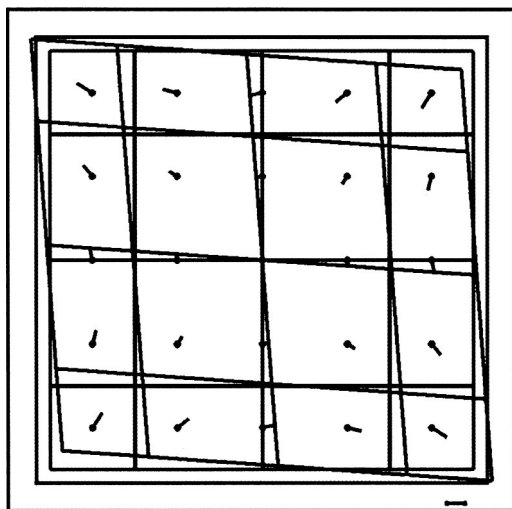


Figure 13. TK350 systematic imager errors of central part, largest vector $280 \mu\text{m}$.

only by control points. For overlapping image pairs, the influence of the GCPs is dominating. So, for geometric analysis of space images, a higher number of accurate and well distributed GCPs is required. Points located on corners cannot be identified as well as symmetric points. Corner points always are shifted from the bright image parts to the dark parts. For symmetric points, such a shift is unimportant because it is equal in all directions.

Only small satellites are equipped with CCD-array cameras, like the group constructed by Surrey Satellite Technologies (SSTL), e.g. the UoSAT-12 and the BILSAT-1. Calibration of these systems is controlled by the quality of the optics (the radial symmetric and the tangential distortion), as well as the focal length and the location of the principal point. In addition, the CCD-array may have a slightly different dimension in both directions. The very narrow field-of-view does not really require determining the principle point since the bundle of rays will not be deformed by a small offset from the center. Also the focal length is not so important because the rays are close to parallel. The radial and tangential lens distortion can be determined without problems with a limited number of control points and overlapping images.

Most space sensors used for mapping purposes are CCD-line sensors. Areas are mapped via the movement of the satellite along-track and/or by rotating the system.

The geometric relation of the Indian Remote Sensing satellite (IRS-1C) Pan-camera has been investigated using the original sub-images for an area near Hannover where there is ample ground control (Jacobsen 1998). The IRS-1C Pan-camera has three CCD-lines. The parts imaged at the same instant on the ground are shown in Figure 14a, b. The time delay of CCD2 can be respected by matching the slightly overlapping

sub-images. With a special group of additional parameters (Table 3), the relation of the roughly merged sub-images were determined by bundle orientation using the Hannover program BLASPO for processing satellite line scanner images. The geometric relation can be seen in Figure 14e. Obviously the CCD-lines are not exactly aligned. This calibration information has to be used for matching sub-images to a more or less error free synthetic image.

This calibration is possible just by using one scene, based on control points. The configuration derived from of three subscenes taken from different directions has the advantage that one sub-scene stabilizes the others (Fig. 14d), connected by tie points. So by theory, calibration with additional parameters is also possible with a minimum of control points.

The described calibration has to be made for all space systems based on the configuration of CCD-lines. It belongs to the calibration phase, which for very high resolution space systems takes between one and two months.

Bundle adjustments of different original space images have been made using the BLASPO program. In no case was a deformation of the imaging line detected; otherwise, an additional calibration of the

Table 3. Special additional parameters for IRS-1C calibration.

Parameter	
$X = X + P11 * (X - 14)$	if $x > 14$
$X = X + P12 * (X + 14)$	if $x < -14$
$Y = Y + P13 * (X - 14)$	if $x > 14$
$Y = Y + P14 * (X + 14)$	if $x < -14$
$X = X + P15 * \sin(X * 0.11) * \sin(Y * 0.03)$	

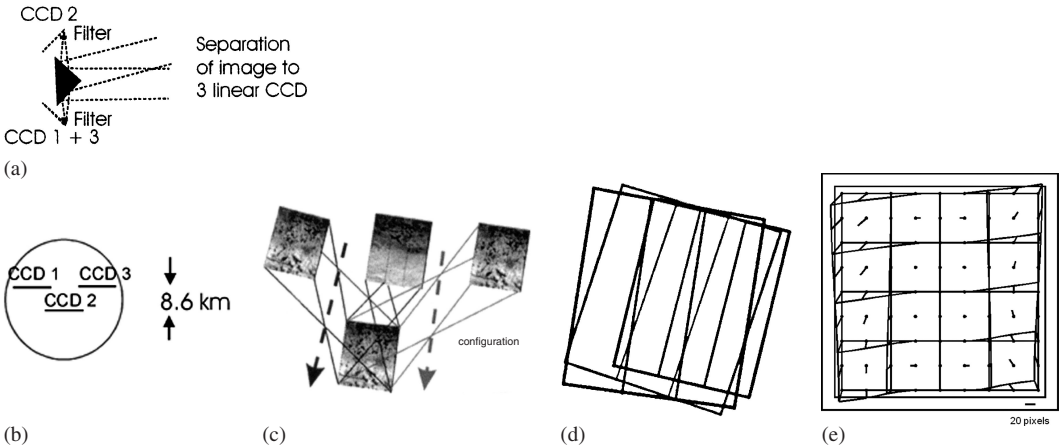


Figure 14. Geometric calibration of IRS-1C pan-camera: (a) configuration of CCD-lines in camera; (b) imaged elements on ground; (c) configuration of test data set; (d) configuration on ground; and (e) image geometry of full scene.

CCD-line sensors would be justified. The computed deformations are just compensating discrepancies of the orbit and attitude information; or in other words, the individual exterior orientation. The values are in the range of a few pixels.

Rectified space images like IKONOS Geo and QuickBird Standard and OrthoReady have shown no systematic effects. For the exact handling of early vintage IKONOS Geo data prior to information about the rectification plane height level and relief displacement, an affinity transformation to control points is necessary. The sub-pixel accuracy obtained showed no systematic effects, confirming the calibration used. Of course the available exterior orientation can be analysed, leading to a system calibration that includes the relation of the imaging sensor to the inertial measurement unit and the star cameras, as well as the focal length. The very narrow field-of-view does not allow a separation between effects of the focal length and the flying height (satellite orbit ellipse). For the boresight calibration, a higher number of controlled scenes in different areas is required.

4 CONCLUSIONS

The geometry of perspective aerial cameras can and should be determined under flight conditions by a bundle block adjustment with self-calibration by additional parameters. The difference of the real geometry against the mathematical model of perspective is shown by the systematic image errors. Only some parts, like radial symmetric distortion, show long-term stability. The focal length and the location of the principle point usually can be determined only by using projection center coordinates determined by relative kinematic GPS positioning at two altitudes. The set of additional parameters used has to be reduced to the required parameters that have a limited correlation. This should be done automatically by the bundle adjustment program using statistical tests. Self-calibration should not include additional parameters that under usual conditions are highly correlated.

CCD-line scanners often use a combination of shorter CCD-lines. Calibration of such systems can be made under flight conditions. For aerial systems, the geometric advantages of crossing flight lines can be used, thus reducing the number of required control points. For space sensors, the number of required control points can be reduced if images are available of the same area taken from different orbits. The calibration of a single image requires a higher number of precise control points.

REFERENCES

- Ebner, H. 1976. Self-calibrating block adjustment. In *ISP congress Proc., Helsinki*, Part 3 1976.
- Gotthardt, E. 1975. Zusatzglieder bei der aerotriangulation mit bündeln. *Bildmessung und Luftbildwesen* 218–221.
- Grün, A. 1979. Self-calibration versus test field calibration. *ISP WG III/3 Seminar, Aalborg*. Helsinki: Inst. of Photogrammetry Technical University Helsinki.
- Heipke, C., Jacobsen, K. & Wegmann, H. 2001. The OEEPE-test on integrated sensor orientation – analysis of results. *OEEPE-workshop on integrated sensor orientation, Hannover, September 2001*. University of Hannover: Institute for Photogrammetry and GeoInformation.
- Jacobsen, K. 1980. Attempt at obtaining the best possible accuracy in bundle block adjustment. *International Archives of Photogrammetry* 22(Part B3): 304–313.
- Jacobsen, K. 1982a. Attempt at obtaining the best possible accuracy in bundle block adjustment. *Photogrammetria* 37 (6): 219–235.
- Jacobsen, K. 1982b. Selection of additional parameters by program. *Int Arch of Photogram and Rem Sen* 24(3): 266–275.
- Jacobsen, K. 1998. Mapping with IRS-1C images. *ASPRS Annual Convention, Proc., Tampa 1998*. Bethesda, MD: ASPRS.
- Koch, R. 1987. Untersuchung von kopierfilmen und -verfahren zur reproduktion hochauflösender luftbilder. diploma thesis 1987. University of Hannover: Institute for Photogrammetry and GeoInformation.
- Meier, H.K. 1978. The effect of environmental conditions on distortion, calibrated focal length and focus of aerial survey cameras. In N.S.S.Gakkai (ed.) *Int. Arch. Photogram. Commission-I, Symp. Data Acquis. & Improv. Image Qual. & Image Geom.*, 29–31 May 1978. Tokyo, Japan.

Intersensor calibration

In-flight intersensor radiometric calibration using vicarious approaches

K.J. Thome

Optical Sciences Center, University of Arizona, Tucson, Arizona, USA

ABSTRACT: Improvements in the protocols and approaches used for in-flight radiometric calibration and validation of imaging sensors has led to using vicarious calibration for intercomparing multiple sensors. The concept presented here is different than the typical cross-calibration techniques used successfully since the 1980s, but instead assumes that results for one sensor can be compared directly to results using the same vicarious approach for another sensor. This is similar to using spherical integrating sources as a calibration standard in multiple laboratories for multiple sensors. Traceability of these sources to a set of specified standards allows sensors calibrated in the same way to agree within the uncertainties of the methods. Extending the idea of a calibration source to that of a ground-reference target readily shows how these concepts can be used in-flight. This paper illustrates this based on results from the reflectance-based approach of vicarious calibration. The reflectance-based method relies on ground-based measurements of surface reflectance and atmospheric transmittance to predict the at-sensor radiance for a selected sensor viewing a test site. The method has been applied to numerous Earth resources sensors of varying spatial resolutions, and improvements in the application of the approach over the past several years have led to precisions better than 2 percent (one standard deviation) in many spectral bands. Thus, results from one sensor can be compared to those of another sensor without the need for coincident acquisitions and disagreements in excess of 2 percent should be indicative of possible biases between the sensors. Examples using the MODerate resolution Imaging Spectroradiometer (MODIS) on Terra and Aqua, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), the Advanced Land Imager (ALI), Hyperion, and the Enhanced Thematic Mapper-plus (ETM+) are presented. Results show that all of these sensors agree for the most part in similar bands except for the ASTER visible and near infrared (VNIR) bands. Also presented is a discussion of the uncertainties in the reflectance-based approach leading to a set of recommended protocols necessary to ensure the accuracy and precision of the reflectance-based method. The ultimate goal of this concept is to extend calibration philosophy to other vicarious methods to allow multiple vicarious methods to be compared directly with each other.

1 INTRODUCTION

Excluding dedicated atmospheric, meteorological, and defense satellites, there are more than 30 sensors in space designed for terrestrial imaging. These sensors are a variety of national and international efforts representing more than 12 countries. They range in spatial resolution from 0.6 to 1000 m with spectral resolution from multispectral to hyperspectral spanning the wavelength regions from visible, near-infrared, thermal infrared, and radar. Swath widths and repeat visits also span a wide range of values all leading to an opportunity to study Earth processes on a global scale. Critical to this approach is the ability to compare data between scales and sensors, and a key part of this is absolute radiometric calibration. Absolute radiometric calibration allows data from individual sensors to be compared directly so long as traceability to known

standards is maintained and estimates of absolute accuracy are themselves accurate.

Absolute radiometric calibration of a sensor begins in the laboratory during the final stages of the sensor's preparation for launch. The typical approach is to have the sensor view an extended radiance source for which the output of the source is known. An extended source is necessary to simulate the type of source that will be measured by the sensor in flight. These radiance sources are either traceable to source- or detector-based standards (Yoon et al. 2002). For a source-based standard, the given output of a source of spectral irradiance that is traceable to a national laboratory is transferred to the radiance source. Detector-based calibrations have the advantage of not relying upon an intermediate step but instead the output of the extended source is calibrated directly using the detector-based standard. Accuracy of sensor calibration in the

laboratory is dependent on the individual calibration facility's skill at implementing a rigorous set of protocols and traceability to the standards chosen.

If a set of well-defined protocols is followed during the preflight calibration process, it should be possible to compare the output of two sensors calibrated in different facilities. The output from these two sensors should agree to within the stated uncertainties when the sensors view a common source. This has been shown to be true for a set of traveling transfer radiometers that have been compared through a series of round-robin exercises at several calibration facilities (Butler et al. 2002, Butler et al. 2003, Johnson et al. 1997). These round-robin exercises were made with carefully crafted, non-imaging radiometers and represent a "best possible situation."

Unfortunately, there are many issues when the sensor is a more complicated imaging system typical of terrestrial remote sensing. These issues can impact the actual accuracy of the radiometric calibration but not the modeled traceable accuracy. That is, even though rigorous protocols are followed, actual errors can be larger due to overlooked aspects of the calibration process. Two examples of this are the size-of-source effect and the spectral difference between the source used in the laboratory and that used in the actual implementation of the sensor. Flittner & Slater (1991) and Slater et al. (2001) show the impact of errors in spectral characterization and the importance of simulating the spectral shape of the in-flight source during the preflight calibration. The issue of size-of-source can be viewed as a stray light issue and is related to the fact that some sensors behave differently when viewing the output of a small extended radiance source in the laboratory, as opposed to the much larger in-flight source. Both of these effects, though well known, are difficult to assess in laboratory calibration. Thus, it should be clear that it is possible to overlook other effects as well, and any overlooked source of errors leads to a bias in the accuracy and/or an underestimate of actual uncertainty.

It is these two effects, size-of-source and spectral difference, that is a partial motivation for the current work. This work proposes that the Earth and Earth's atmosphere can be viewed in a like fashion as a laboratory calibration source (both figuratively and explicitly). The emphasis of this work is on the solar reflective part of the spectrum covering approximately the spectral range from 400 to 2500 nm, but this is done merely for brevity and there are analogous methods for other parts of the spectrum. The Earth and the Earth's atmosphere act as an extended radiance source in the same way that the spherical integrating source acts in the laboratory, hence the use of these laboratory sources. If this Earth-atmosphere source is accurately characterized, then the calibration of the sensor can be determined and the analogy between the laboratory

source and the in-flight calibration source should be straightforward. The primary reason for using scene-based sources for in-flight characterization of a sensor is to understand changes that occurred with the sensor after preflight calibration, but which are not characterized by onboard calibrations. Thus, vicarious calibration can be used to correct for changes that occur in the system's response or other effects such as size-of-source and spectral differences.

There is good justification for including vicarious calibration approaches that are independent of preflight and onboard calibrations. The typical methods used involve either *in-situ* measurements or model-based predictions. Kaufman & Holben (1993) developed one of the first models using a method that predicts the source radiance based on assumptions of molecular scattering and appropriate assumptions of surface reflectance (for water in their case). This method is applied to short wavelength bands for wide angle sensors and was originally used for the visible channels of the Advanced Very High Resolution Radiometer. Vermote et al. (1992) used a similar approach to calibrate the short-wavelength channels of the SPOT-1 HRV sensor where the contributions from aerosols and sea-surface reflection were determined from data at longer wavelengths. Cosnefroy et al. (1996) applied a similar method over land, but in their case the atmosphere did not play a dominant role. Surface reflectance was the key parameter. This approach assumes that test sites in desertic regions are invariant targets and that they provide a well-understood ground target of known radiance. If absolute accuracy is not a requirement, but only inter-band relative calibration is critical, then approaches using clouds and sun glint have been shown to be successful (Hagolle et al. 1999). Lunar calibrations are a more recent approach using no *in-situ* measurements, and several sensors now use the moon in a relative sense for calibration (Barnes et al. 1999, Kieffer & Wildey 1996).

Hovis et al. (1985) made one of the earliest attempts at *in-situ* measurements by measuring the radiance above a ground target from a high altitude aircraft to verify the degradation of the response of the Coastal Zone Color Scanner's (CZCS) shorter wavelength bands. The Remote Sensing Group (RSG) at the University of Arizona followed this work by developing three vicarious techniques of absolute calibration that rely on *in-situ* measurements. These methods are referred to as the reflectance-, irradiance-, and radiance-based techniques (Slater et al. 1987, Biggar et al. 1990). The advantage these *in-situ* approaches have over modeled approaches is better absolute accuracy. However, these methods have suffered from a lack of precision as well as low frequency of data collections. This situation has improved over the years producing results that often rival those from

solar diffuser approaches and with precisions approaching 2 percent in some bands and cases.

It is this improved precision of the reflectance-based approach, and vicarious approaches in general, that is the other motivation for the work described here. Improvements in vicarious approaches, and the confidence in these methods, now make it possible to use them in the same fashion as preflight laboratory calibrations to allow intercomparisons of sensor data. The discussion focuses on reflectance-based methods as an example for implementation, but this is purely for convenience. The concept can be extended to any vicarious approach shown to have sufficient repeatability with limited band-to-band and spatial resolution effects. The method is also suitable for both multispectral and hyperspectral sensors. The paper begins with a brief description of the reflectance-based approach and desired test site characteristics. The recommended minimum set of measurements is included based on past work by the RSG and a discussion as to the importance of following these guidelines. Finally, examples of recent results for several sensors are given demonstrating the current status of *in-situ* vicarious calibration approaches.

2 REFLECTANCE-BASED APPROACH

Past works have described this method in more detail (Slater et al. 1987, Thome 2001). Essentially, the approach relies on measurements of the surface reflectance of a test site at the time of sensor overpass. Atmospheric measurements are made concurrent with surface measurements, and the results from both are used as input to a radiative transfer code to predict at-sensor radiance that is compared to sensor output to provide the calibration.

2.1 Surface reflectance retrieval

The reflectance-based approach relies on surface reflectance measurements of a selected test site. There are two aspects to the size of a test site. The first is the size and orientation of the site measured by *in-situ* radiometers. The size and orientation of a test site, and the ground sampling within it, varies for the sensor type and the spatial resolution of the sensor. Size is a compromise between sampling a large area of the ground for adequate coverage of the sensor's focal plane, yet small enough that the site can be covered in a reasonable amount of time. Measurements extending beyond one hour tend to suffer from changes in atmospheric illumination, solar angle effects, user fatigue, and instrument power limitations. The second aspect of test site size is the overall size of the surrounding area that has similar surface composition to that sampled by ground measurements. As an example,

the RSG collects surface data within a 480 m by 120 m area for the Landsat-series of satellites. This site size is the same at both White Sands National Monument and for Railroad Valley Playa. The size of the entire Railroad Valley Playa, however, is approximately one-fourth the size of White Sands National Monument. Orientation of the test site should be aligned with its long side in the cross-track direction for a pushbroom sensor. Of course, nothing precludes the use of a square site.

Reflectance of the test site is obtained by referencing measurements at the site to those made of a reference of known reflectance. Past work relied on multispectral instruments for these collections and more current work has moved to hyperspectral collections. Three factors to consider in collecting surface reflectance data are the number of samples to collect within the test site, the field of view of the ground-based sensor, and the number of measurements of the reference. These reference measurements require time, but they allow correction for changes in instrument response with time and changing atmospheric conditions. Similarly, greater sampling of the test site offers a more accurate characterization of surface properties, but this is at the expense of time. Finally, a large field of view for the ground instrument will improve spatial sampling but can introduce surface bidirectional reflectance effects. Knowing the bidirectional reflectance properties of the reference negates many of the issues related to time because it allows corrections for sun-angle effects.

2.2 Atmospheric characterization

Atmospheric characterization data are collected at the same time as the surface reflectance measurements. The primary instrument for the RSG in this characterization is a solar radiometer. Data from this instrument are used in a retrieval scheme to determine spectral atmospheric optical depths. Optical depth results are inverted to determine ozone optical depth and a Junge aerosol size distribution parameter. Aerosol size distribution and columnar ozone are used to determine the optical depths at 1 nm intervals from 350 to 2500 nm. Columnar water vapor is derived using a modified Langley approach.

2.3 Radiative transfer code

Atmospheric and surface data are used in a radiative transfer code. There is a variety of codes available that all satisfy the requirements of predicting the at-sensor radiance to the required accuracy. The RSG relies on an historical code based on a Gauss-Seidel iteration approach. Developed at the University of Arizona, this code has been in use with the reflectance-based approach since the first implementation of the

method (Herman & Browning 1965, Slater et al. 1987). Minor modifications have made it more versatile for view angle and to allow for hyperspectral predictions (Thome et al. 1996). More recently, the group has used the DISORT code included in the Moderate Resolution Transmittance model (MODTRAN). Comparisons made between the Gauss-Seidel and MODTRAN show that the result of these two codes agree to better than 0.5 percent when care is taken to ensure that identical inputs are used for each. Past work has shown that similar conclusions can be drawn for other code types such as doubling-adding and the methods used in the 6S code (Thome et al. 1998).

Inherent to the codes are solar irradiance models that allow conversion of relative radiance (can be viewed as reflectance) to absolute radiance. Two solar irradiance standards are currently employed by the RSG. The first is from the World Radiation Council selected for NASA's Earth Observing System (EOS) project. The second is based on the Chance/Kurucz model that is part of MODTRAN4.0. These two solar models have significant differences between them that have to be understood (Thome et al. 2001). In all cases, the RSG takes care to ensure that a solar model consistent with the calibration team of each sensor is used. As long as users take care to convert data to reflectance, comparisons between sensors should be consistent. Once the at-sensor, hyperspectral, absolute radiances are determined, they are band-averaged across the sensor spectral response.

2.4 Test sites

The reflectance-based approach relies on several critical characteristics of an ideal test site (Scott 1996). These include a high-reflectance to reduce the impact of atmospheric errors. In addition, low aerosol loading and predominance of clear skies also reduce these atmospheric errors. The surface reflectance, in addition to having a high value, should have spatial uniformity that allows a large area to be characterized from only a few measurements. Other aspects of the site, such as near-lambertian reflectance, spectral uniformity, and site accessibility are also desired. There is no ideal calibration site that satisfies all of these conditions, but in the Southwestern United States there exist several that have been used for many years. The three most frequently used are the White Sands National Monument in New Mexico, the Railroad Valley Playa in Nevada, and the Ivanpah Playa in California. The most important aspect of these sites is that they are well-understood, and a large number of data sets exists for them.

3 INTERCOMPARISON RESULTS

During the past three years the reflectance-based approach has been used for more than 14 satellite-based

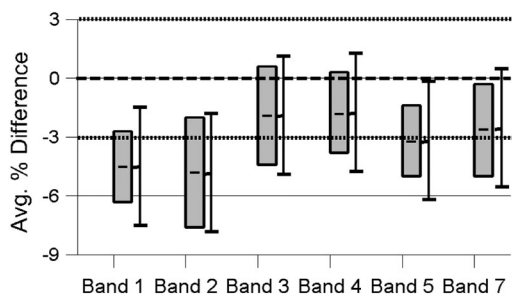


Figure 1. Average percent difference between reflectance-based and pre-flight calibrations for ETM+. Negative value indicates vicarious calibration predicts a higher radiance than the sensor.

sensors. The results from six of these in the visible and near infrared are shown as an example of the proposed intercomparison approach. In addition, five of these sensors had sufficient data sets in the short-wave infrared spectrum to warrant inclusion. Results from the other sensors are not included to allow further evaluation and consultation with the calibration/validation groups from each sensor. The sensors shown here are those for which the author is, or was, a member of the calibration/validation team.

As an example of the precision currently achievable, a summary of the results from Landsat-7 Enhanced Thematic Mapper-plus (ETM+) is shown in Figure 1. The figure shows the average percent difference between the predicted calibration coefficient for ETM+ based on the reflectance-based results of the RSG. These averages include data from all RSG test sites that are 480 m by 120 m in size and are based on 39 data sets between July 1999 and May 2003. The standard deviation of the average is indicated by the size of the box surrounding the average value. In addition, ± 3 percent standard deviation is indicated by the error bars adjacent to each average. The zero percent difference line, as well as the ± 3 percent lines, are also highlighted. Three key conclusions can be drawn from Figure 1. The first is that the vicarious and preflight results agree to within the combined uncertainties of the two methods. The second is that there appears to be a bias between the vicarious and pre-flight calibration. This possible bias is still under investigation, but is not the primary effect considered here. It is the third conclusion that is of primary interest here; that is, the standard deviation for all bands is less than ± 3 percent and less than 2 percent for several bands. This standard deviation can be viewed as a proxy for the accuracy of the reflectance-based retrieval and implies that the accuracy is better than 3 percent.

A second way to evaluate the accuracy is to subset these data. This is shown in Figure 2 for Band 3 of ETM+. The original results for band 3 are shown at

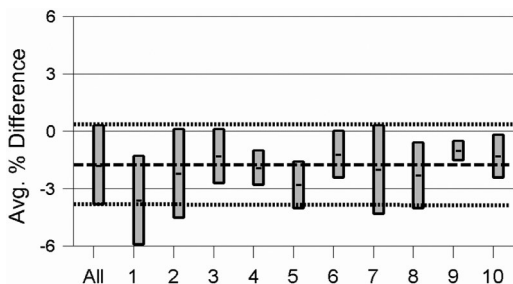


Figure 2. Average percent difference and standard deviation for Band 4 of ETM+. First set is based on all 39 RSG data sets. Other 10 are based on random selection of groups of five data sets.

the left side of the graph. All others shown are averages and standard deviations of randomly-selected sets of five from the original 39. All of these subset results are within the ± 2 percent standard deviation of the original data set. Results were similar for all of the other bands of ETM+. Further analysis regarding the statistical significance is needed, but this result adds confidence to the accuracy of the reflectance-based approach. It should be pointed out that both onboard and vicarious calibrations indicate that the radiometric calibration of ETM+ has remained stable since launch. This fact permits use of the average and standard deviation to examine the accuracy of the reflectance-based method.

Figure 3 shows a plot similar to that shown in Figure 1, except that results for five other sensors are included. The sensors shown are the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua and Terra platforms, the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER), also on the Terra platform, the Advanced Land Imager (ALI) on the Earth Observer-1 (EO-1) platform, and the Hyperion sensor, also on EO-1. The data shown are for the visible and near infrared bands of MODIS and include bands at 250 m, 500 m, and 1000 m spatial resolution. The results of the other sensors are shown with the closest MODIS band. Thus, only results for Terra and Aqua MODIS, and Hyperion, are shown for 412 nm since these are the only sensors with bands at this wavelength. The data include averages and standard deviations for each sensor (39 for ETM+, five for Hyperion, 15 for Terra MODIS, six for Aqua MODIS, nine for ALI, and 20 for ASTER). Results for the same sensors, except Hyperion, are shown for three shortwave infrared bands in Figure 4. Hyperion data are omitted here because only two data sets were available at RSG test sites for the short-wave detectors.

All of the sensors shown in Figure 4 agree to within the accuracy of the approach in the shortwave

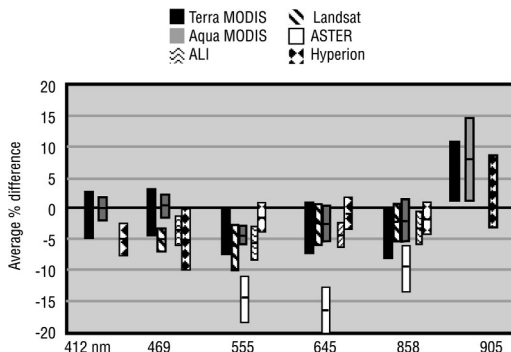


Figure 3. Average and standard deviation of the percent difference between vicarious-based predicted radiances and those reported by six sensors that are part of NASA's Earth Science Enterprise.

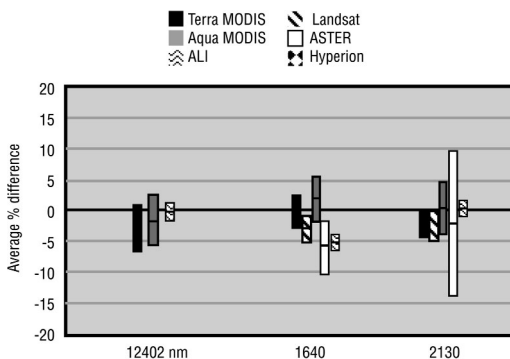


Figure 4. Reflectance-based results for shortwave infrared channels.

infrared except for ALI and ASTER compared with Aqua MODIS at 1640 nm. It also should be noted that the standard deviation of ASTER results is much larger than the others for 2130 nm. This is due to a known optical crosstalk effect where light from the 1640 nm band leaks into other bands including the 2130 nm channel. This effect varies significantly with the test site used, leading to larger standard deviations. For the visible and near infrared, the most noticeable feature is that ASTER appears radically different from the other sensors. Other conclusions are that Aqua MODIS and Hyperion disagree at larger than their combined precision at 412 nm and Aqua MODIS and ETM+ disagree at 469 nm.

The above discussion ignores any possible bias between the reflectance-based results and a given sensor. While the odd behavior at 905 nm relative to the vicarious results is unsettling, the fact that it is consistent in three separate sensors indicates that there is either a bias in the vicarious approach or consistent

errors in the pre-flight calibrations. It seems that all three sensors agree, indicating that there is no bias between these sensors in this particular spectral band. This assumes that the reflectance-based method is repeatable within the RSG and this has been accomplished through attempts to assure consistency in the measurement sets.

4 EXAMPLE PROTOCOLS

Consistency of the measurements over the past five years has been developed through a combination of many collections to gain practical experience, modeling work to examine which measurements are significant, and laboratory characterizations to ensure that the equipment being used is consistent. The primary result of all of this work is that the most important factor in maintaining high accuracy is consistency of a basis set of measurements. The intercomparison method relies on the fact that a single group making vicarious calibration maintains consistency with itself to allow an adequate level of accuracy. The next step would be to allow different groups to compare their results with a different group using the same vicarious calibration approach. In this case, it is important for both groups to use nearly identical collection protocols for a select set of key measurements. The idea is not to preclude different groups from making alternate data collections, but rather, to ensure that two separate groups can compare results with confidence that any differences between them are real.

The analogy to the above in terms of laboratory calibrations is that a group calibrating two sensors in its laboratory should do so using a portion of the same approach for both systems. Thus, if the group uses a spherical integrating source for radiometric calibration, it should use that same source with the same number of lamps, under the same operating conditions for the second sensor. If that second sensor was then taken to another facility, the goal would be to use a similar integrating source with the same output aperture with similar output radiance and color temperature. This would not preclude the second group from using additional measurements, but the ideal would be to have at least some portion of the measurements at both facilities to be "identical."

Thus, it is useful to have a basis set of measurements emphasizing the critical components of a vicarious approach. This set of measurements ultimately would provide a set of protocols that could be followed by all groups to ensure accurate intercomparisons. This can be done for each separate vicarious approach, and a set of proposed measurements is given here for the reflectance-based approach. It cannot be emphasized enough that due to the vagaries of instrumental effects, it is critical that each measurement

group keep a set of measurements that is consistent from campaign to campaign. Thus, the first recommendation is that groups that are currently collecting vicarious calibration data sets, radiance validation data sets, or validation data in general, should continue to collect data in the manner in which they are most familiar. It is also critical that groups maintain this consistency in terms of equipment, processing schemes, and test sites. When any of the three are modified (and they should be from time to time to ensure improvement), there should be a transition period to understand the effects of the changes and, more importantly, provide traceability to the past data sets. Eventually, groups should migrate to an approach similar to that described above for the reflectance-based approach.

While this recommendation may appear self-serving, this conclusion, and the approach used by the RSG have not been reached lightly. Rather, they are based on nearly 20 years of field data measurements. It has included improvements in field reference standards and equipment, changes in test sites, new satellite sensor technology, and most importantly, continued attempts to improve the field measurements and processing schemes. Early attempts to improve the measurements served to improve dramatically the accuracy of the data (from uncertainties >5 percent to values that hopefully can be shown some day to approach 2 percent) but for which a price has been paid in that older data sets are now obsolete and exist only in the literature in a useful form.

5 SUMMARY

In summary, the basic measurement set consists of:

- 1 Surface reflectance measurements that are referenced to a panel of known reflectance. While sampling strategies are important, work by the RSG has shown that if sampling creates uncertainties approaching 1 percent, an alternate site should be selected.
- 2 Solar extinction measurements that include bands to allow for ozone and water vapor characterization and sufficient bands to allow retrieval of aerosol optical depth at 550 nm (or some other reference wavelength), and an Angstrom turbidity coefficient. The frequency of data collection, accuracy of calibration, and specific bands of the radiometer, while important, are of secondary importance if an appropriately bright site can be located. Surface reflectance in excess of 0.3 ensures this to be the case for aerosol loadings giving aerosol optical depths at 550 nm less than 0.15.
- 3 Use of a radiative transfer code including multiple scattering and a method for taking into account

ozone and water vapor absorption. It is strongly recommended that users select a radiative transfer code with which they are familiar and understand the code's limitations. This is far more important than selecting the most complicated and accurate code. In other words, an accurate code that is run improperly is far worse than a code with slightly lower accuracy, but used properly. In addition, if the site has reflectance greater than 0.3, the selection of radiative transfer code becomes less important.

If the above three requirements are met, the data set collected should be of sufficient quality to be compared with results from other groups. This is important in that it increases the number of data sets available, thus increasing the power of vicarious methods. However, if the data sets are of questionable heritage, the results are not useful even if they are extremely accurate.

The final step in the process will be to understand any biases in a given approach relative to a traceable standard. When this is done for multiple types of vicarious approaches, it then becomes possible to compare results from different methods. This is critical to allow vicarious results to be used to evaluate the sensors. For example, assume that the reflectance-based method, invariant desertic scenes, and lunar-based methods are used for a given sensor. Further, assume that careful work has been done to determine the absolute radiometric accuracies of each. Then, for example, if comparisons of results indicate that the desertic scene approach and reflectance-based results agree, but that the lunar-based calibration disagrees by more than the estimated uncertainties, it could be concluded that the sensor may have a significant size-of-source effect.

6 CONCLUSIONS

Recent work by the RSG has indicated that the accuracy of the reflectance-based approach of vicarious calibration by experienced groups can approach 2 percent. Other approaches are also at or better than this level, which can only be achieved if users take care to collect a consistent basis set of data for each collection. It is further felt that this accuracy can be improved through additional field instrument characterization, higher frequency of collections, and separation of results by test site and users. At this level of accuracy, it is possible to use the results from a single group to intercompare results from multiple sensors. This intercomparison does not require joint overpasses of the sensors, but rather consistent processing and collection methodologies for the vicarious approach.

The power of such a method should be clear when considering the lifetime records of Landsat and MODIS. The goal is to continue these data records

through the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project and later through NPOESS itself. However, it is possible that neither Thematic MapperTM nor ETM+ and that both Aqua and Terra MODIS could fail prior to launch of their follow-on missions. The current precision of vicarious methods should allow for a consistent data record across any gaps in data collections.

7 ACKNOWLEDGEMENTS

The author acknowledges the contributions of those who assisted in collecting data used to produce this paper. This work was supported by NASA contract NAS5-31717 and grant NAG5-2448, as well as support from the EO-1 team at Goddard Space Flight Center for the Hyperion and ALI data.

REFERENCES

- Barnes, R.A., Eplee, R.E., Patt, F.S. & McClain, C.R. 1999. Changes in the radiometric sensitivity of SeaWiFS determined from lunar and solar-based measurements, *Appl. Opt.* 38(21): 4649–4664.
- Biggar, S.F., Santer, R.P. & Slater, P.N. 1990. Irradiance-based calibration of imaging sensors. *Int. Geosci. & Rem. Sens. Symp. College Park, Maryland*. 1: 507–510.
- Butler, J.J., Johnson, B.C. & Barnes, R.A. 2002. Radiometric measurement comparisons at NASA's Goddard Space Flight Center: Part 1. The GSFC sphere sources. *The Earth Observer* 14: 3–7.
- Butler, J.J., Brown, S.W., Saunders, B., Johnson, B.C., Biggar, S.F., Zalewski, E.F., Markham, B.L., Gracey, P.N., Young, J.B. & Barnes, R.A. 2003. Radiometric measurement comparison on the integrating sphere source used to calibrate the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Landsat-7 Enhanced Thematic Mapper Plus. *J. Res. Natl. Inst. Stand & Tech.* 108(3): 199–228.
- Cosnefroy, H., Leroy, M. & Briottet, X. 1996. Selection and characterization of Saharan and Arabian desert sites for the calibration of optical satellite sensors. *Rem. Sens. Env.* 58(1): 101–114.
- Flittner, D.E. & Slater, P.N. 1991. Stability of narrow-band filter radiometers in the solar-reflective range. *Photogr. Eng. & Rem. Sens.* 57(2): 165–171.
- Hagolle, O., Galoub, P., Deschamps, P.Y., Cosnefroy, H., Briottet, X., Bailleul, T., Nicolas, J.M., Parol, F., Lafrance, B. & Herman, M. 1999. Results of POLDER in-flight calibration. *IEEE Trans. Geosci. & Rem. Sens.* 37(3): 1550–1566.
- Herman, B.M. & Browning, S.R. 1965. A numerical solution to the equation of radiative transfer. *J. Atmos. Sci.* 22(2): 559–566.
- Hovis, W.A., Knoll, J.S. & Smith, G.R. 1985. Aircraft measurements for calibration of an orbiting spacecraft sensor. *Appl. Optics* 24(3): 407–410.

- Johnson, B.C., Sakuma, F., Butler, J.J., Biggar, S.F., Cooper, J.W., Ishida, J. & Suzuki, K. 1997. Radiometric measurement comparison using the Ocean Color Temperature Scanner (OCTS) visible and near infrared integrating sphere. *J. Res. Natl. Inst. Stand. & Tech.* 102(6): 627–646.
- Kaufman, Y.J. & Holben, B.N. 1993. Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint, and desert reflection. *Int. J. Rem. Sens.* 14(1): 21–52.
- Kieffer, H.H. & Wildey, R.L. 1996. Establishing the moon as a spectral radiance standard. *J. Atmos. & Oceanic Tech.* 13(2): 360–375.
- Slater, P.N., Biggar, S.F., Holm, R.G., Jackson, R.D., Mao, Y., Moran, M.S., Palmer, J.M. & Yuan, B. 1987. Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors. *Rem. Sens. Env.* 22(1): 11–37.
- Slater, P.N., Biggar, S.F., Palmer, J.M. & Thome, K.J. 2001. Unified approach to absolute radiometric calibration in the solar-reflective range. *Rem. Sens. Env.* 77(3): 293–303.
- Scott, K.P., Thome, K.J. & Brownlee, M.R. 1996. Evaluation of the Railroad Valley playa for use in vicarious calibration. In B. Huberty, J.B. Lurie, P. Coppin, J.A. Caylor & P.C. Roberts (eds), *Multispectral imaging for terrestrial applications*. Proc. SPIE 2818: 158–166.
- Thome, K.J. 2001. Absolute radiometric calibration of Landsat-7 ETM+ using the reflectance-based method. *Rem. Sens. Env.* 78(1–2): 27–38.
- Thome, K.J., Gustafson-Bold, C.L., Slater, P.N. & Farrand, W.H. 1996. In-flight radiometric calibration of HYDICE using a reflectance-based approach. In S.S. Shen (ed.), *Hyperspectral Remote Sensing and Applications*. Proc. SPIE 2821: 311–319.
- Thome, K., Schiller, S., Conel, J., Arai, K. & Tsuchida, S. 1998. Results of the 1996 Earth Observing System vicarious calibration campaign at Lunar Lake Playa, Nevada (USA). *Metrologia* 35(4): 631–638.
- Thome, K., Biggar, S. & Slater, P. 2001. Effects of assumed solar spectral irradiance on intercomparisons of earth-observing sensors. In H. Fujisada, J.B. Lurie & K. Weber (eds), *Sensors, systems, and next-generation satellites V*. Proc. SPIE 4540: 260–269.
- Vermote, E., Santer, R.P., Deschamps, P.Y. & Herman, M. 1992. In-flight calibration of large field of view sensors at short wavelengths using Rayleigh scattering. *Int. J. Rem. Sens.* 13(18): 3409–3429.
- Yoon, H.W., Gibson, C.E. & Barnes, P.Y. 2002. Realization of the National Institute of Standards and Technology detector-based spectral irradiance scale. *Appl. Optics* 41(28): 5879–5890.

Validating the MISR radiometric scale for the ocean aerosol science community

C.J. Bruegge, W. Abdou, D.J. Diner, B.J. Gaitley, M. Helmlinger,
R.A. Kahn & J.V. Martonchik

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

ABSTRACT: The Multi-angle Imaging SpectroRadiometer (MISR) is one of five instruments onboard the Earth Observing System (EOS) Terra spacecraft. MISR employs nine cameras that view up to 70 degrees forward and aft of the spacecraft track and enable unique geophysical retrievals. As an example, many on-orbit sensors are able to estimate the amount of aerosol loading present in the underlying atmosphere. MISR, however, is capable of retrieving information on both aerosol amount and microphysical properties. A necessary prerequisite, however, is that the instrument be calibrated to its absolute, band, and camera-relative specifications. Previous work has demonstrated that MISR is calibrated to better than 4 percent absolute uncertainty (1σ confidence level) for bright land targets. This paper validates that MISR has met its calibration specification for dark target (10 percent $\rho_{\text{toa}} = 0.02$). As part of this study, a new look has been taken at the band-relative scale, and a decrease in the radiance reported for the Red and NIR Bands has resulted. The calibration process is now routine, fully developed, and tested. Bimonthly on-orbit calibrations will be continued throughout the life of the mission and will allow MISR to accurately report incident radiances, even in the presence of expected sensor response changes.

1 INTRODUCTION

1.1 Science drivers to accurate radiometry

The Multi-angle Imaging SpectroRadiometer (MISR) (Diner et al. 1998a, Diner et al. 2002) is one of five instruments onboard NASA's Earth Observing System (EOS). Data products include cloud height and albedos, surface bidirectional reflectances, and aerosol parameters, i.e. airborne particles derived primarily from urban and industrial pollution, forest fires, volcanoes, sea spray, and desert dust. These measurements are routinely provided over the globe, and are important in understanding Earth's radiation budget and climate change.

One unique contribution that can be made by MISR is providing aerosol products with improved accuracies and with some degree of distinguishability. Calibration accuracy is particularly important for aerosol retrievals over dark ocean targets. Based upon theoretical studies, Kahn et al. (1998) & Kahn et al. (2001) provide a clear statement of the MISR pre-launch aerosol science objective for such conditions. *"We expect to retrieve column optical depth from measurements over calm ocean for all but the darkest particles, with typical*

size distributions and compositions, to an uncertainty of at most 0.05 or 20 percent, which ever is larger, even if the particle properties are poorly known. The measurements should also allow us to distinguish spherical from nonspherical particles, to separate two to four compositional groups based on indices of refraction, and to identify three to four distinct size groups between 0.1 and 2.0 μm characteristic radius at most latitudes." To achieve this goal in practice requires accurate top-of-atmosphere (TOA) radiances. The validation of MISR radiometry under low light conditions is not only of interest to the ocean aerosol science community, but also to the dense, dark vegetation, and ocean surface color communities.

Top-of-atmosphere equivalent reflectance is defined here as:

$$\rho_{\text{toa}} = (\pi L / E_0) \quad (1)$$

where L is the TOA radiance within a given MISR band, and E_0 is the MISR total-band-weighted exo-atmosphere solar irradiance, derived from the World Climate Research Program (WCRP 1986) published values of solar irradiance. Very low light levels, in the

equivalent reflectance range below 7 percent, are typically found over dark water scenes having aerosol burdens on the order of 0.2 or less at mid-visible wavelengths. Here the desired MISR radiometric calibration accuracy is at the cutting edge of current capabilities. The needed constraint amounts to $\Delta\rho_{\text{toa}} = 0.002$ or better, for equivalent reflectance below 0.02, in all channels. This translates to a 10 percent absolute uncertainty at a scene equivalent reflectance of 0.02.

MISR calibration requirements for bright targets ($\rho_{\text{eq}} = 1$) include 3 percent absolute, and 1 percent band and camera-relative calibrations. MISR radiometric accuracy has previously been documented for homogeneous desert targets (Bruegge et al. 2002). Here vicarious calibration (VC) experiments, in conjunction with sensor cross-comparison studies and onboard-calibrator (OBC) error assessments, have demonstrated that MISR radiances are uncertain to within 4 percent (1σ) for targets that fall mid-range in the sensor's dynamic range ($\rho_{\text{toa}} = 0.3$ to 0.4). Vicarious calibration experiments are intensive field campaigns located at uniform desert sites such as Railroad Valley, Nevada. These are conducted annually for MISR by the Jet Propulsion Laboratory (JPL) staff (Abdou et al. 2002). Unique tools for this JPL operation include AirMISR (Diner et al. 1998b), an ER-2 based aircraft prototype for MISR, and the portable apparatus for rapid acquisition of bi-directional observation of the land and atmosphere (PARABOLA) instrument (Bruegge et al. 2000, Abdou et al. 2000), a surface based radiometer that measures upwelling and downwelling radiance in 5 degree samplings. For these desert VC experiments, the surface reflectance term dominates the TOA radiance. Under clear sky and low aerosol conditions, typical for southwestern sites, radiances are measured within an uncertainty of 3 percent. Vicarious calibrations are used to validate the radiometric scale of some sensors. In the case of MISR, the June 2000 vicarious campaign was used to calibrate the OBC, which in turn produces radiometric gain coefficients for the cameras on a bimonthly basis.

Validation of MISR radiometry over desert targets does not assure accuracy over dark ocean sites. Instrument artifacts, such as additive stray light or electronic biases, if present, would lead to large radiometric errors in the measure of incident radiance. These could be as large, or larger than, the actual radiance to be measured. Although dark water vicarious calibrations can be conducted, they are not routine. For these cases, the atmospheric contribution to TOA radiance dominates the surface term, and the process of computing TOA radiances from *in-situ* measurements is less certain. Cross-comparisons with other sensors provide an alternate validation approach. The Terra/MODIS instrument is one possible cross-comparison source. Co-located on the same platform, MODIS and MISR view a scene

simultaneously and with similar band-passes. Unfortunately, ocean images acquired by nadir-viewing sensors such as MODIS are frequently contaminated with ocean glint. The large radiance gradient of these scenes makes data comparisons less reliable.

The validation of MISR radiometry over dark targets has proceeded with all proposed approaches: (1) an error tree analysis of the potential contributors to low light errors, (2) cross-comparisons with MODIS scenes, (3) use of a lunar calibration experiment, and (4) dark water vicarious calibrations. The first three of these topics are covered in this publication.

A dark water vicarious calibration of MISR has been described by Kahn et al. (2004). Here data from the AERosol RObotic NETwork (AERONET), in conjunction with an ocean reflectance model and radiative transfer calculation, are used to predict TOA radiances, which are then compared to MISR and MODIS. AERONET radiances are found to be systematically lower than MISR, by about 10 percent. MISR and MODIS ocean channels are found to agree within 4 percent, with MISR reporting higher radiances. The agreement of MISR with MODIS ocean channel data over ocean sites is consistent with the agreement with MISR and MODIS land channel data over land sites. These data indicate that MISR ocean data are valid to within the needed uncertainty for these nadir observations.

The validation of MISR radiometry over very bright targets, such as clouds, will be reported elsewhere. Cloud studies have demonstrated that MISR radiance data over very bright targets are consistent with simulations and cross-sensor comparisons, again to within 4 percent uncertainty.

1.2 The MISR instrument

MISR produces global data sets at nine day intervals or less. The effective center wavelengths given in Table 1 have been computed using a moments (centroid) analysis within the region delimited by the 1 percent response points (Bruegge et al. 2002). The effective bandwidths also are given. These parameters are used to define an equivalent square-band response function and to summarize instrument characteristics. Band weighted (both in-band and total-band) exo-atmospheric irradiance values are also shown. As Level 1B radiances have not been adjusted to remove out-of-band response, total

Table 1. MISR spectral parameters.

λ_c in-band (nm)	$\delta\lambda$ in-band (nm)	$E_{0,b}$ in-band ($\text{Wm}^{-2} \mu\text{m}^{-1}$)	$E_{0,b}$ total-band ($\text{Wm}^{-2} \mu\text{m}^{-1}$)
447	41	1871	1867
558	27	1851	1842
672	20	1525	1524
867	38	969.6	977.8

bandpass solar irradiance values are used to compute Level 1B equivalent reflectances. Approximately 3 percent of the camera output comes from signals at wavelengths outside the 1 percent limits, for a spectrally neutral scene.

Each of the nine cameras has a unique name, and is associated with a specific view angle. The cameras view a target consecutively in the order Df (70.5° fore), Cf (60.0°), Bf (45.6°), Af (26.1°), An (nadir), Aa (26.1° aft), Ba (45.6°), Ca (60.0°), and Da (70.5°), with 7 minutes from first to last acquisition of a target. Here the first letter of the camera name refers to the lens design and the second designates the fore-, nadir-, or aft-view directions with respect to the spacecraft track. MISR has 14-bit quantization, and therefore has roughly 16,384 gray levels (the finite video offset and square-root encoding reduces this by about 300 counts). A signal of $\rho_{\text{toa}} = 0.02$ results in an output digital number (DN) of from 300 to 800 DN, depending on the detector. For dark targets, errors of 30 DN may begin to affect radiometric accuracy significantly.

MISR cameras acquire data in a pushbroom configuration, using the spacecraft motion to build an image from each of the 36 charge-coupled device (CCD) linear arrays. The spatial resolution of the MISR cameras, established by the size of the detector elements, optical focal length, and spacecraft altitude, is 275 m cross-track (for the off-nadir cameras), or 250 m (for the nadir viewing camera). Down-track instantaneous field-of-view increases due to view angle effects, ranging from 214 m at nadir to 707 m at the most oblique angle. Downtrack sampling is 275 m for all cameras.

In practice, most data are acquired in Global Mode, where pixel averaging is performed to reduce the data rate. Here 24 of the 36 data channels have been 4×4 sample-averaged before transmission from the instrument. For these channels, data are transmitted at 1.1 km resolution. Even in Global Mode (GM), however, high-resolution pixels are maintained for the four nadir channels, and the eight additional Band 3/ Red channels. Complete high-resolution data sets for all 36 channels can be obtained from an instrument configuration called Local Mode (LM). Here specific sites are targeted, such as those where intensive field campaigns are being conducted. The size of a Local Mode region is 300 km down-track by 380 km cross-track. About a dozen Local Mode sites are acquired routinely, including observations over desert calibration sites.

2 THE CALIBRATION PROCESS

2.1 Onboard calibrator

Radiometric data products include geolocated radiance images at nadir and off-nadir Earth view angles.

These are total band-weighted camera incident radiances, in units of $\text{W/m}^2/\text{sr}^1/\mu\text{m}^1$. The MISR radiometric response scale is established by use of an OBC, as well as vicarious calibration experiments (Bruegge et al. 1993a). The strength of the OBC is its ability to provide camera, band, and pixel-relative calibrations, as well as to measure temporal stability. Experiments using the OBC are conducted once every two months. The bimonthly frequency is desirable in that it is prudent to deploy the calibration panels only as needed to capture camera response changes. The MISR cameras have degraded by no more than 2 percent per year, as reported by Bruegge et al. (2002). The OBC consists of two Spectralon diffuse panels and six sets of photodiode detectors. The latter measure solar reflected light from the panels, and provide a measure of the camera incident radiance. These are regressed against the camera output to provide the radiometric response for each of the 1504 CCD detector elements per line array, nine cameras, and four spectral bands per camera. One such photodiode set is on a goniometric arm, and allows panel Bidirectional Reflectance Factor (BRF) degradation to be monitored. Photodiodes are either of a light-trapped design called High Quantum Efficiency (HQE) diodes, or Positive, Intrinsically pure silicon, Negative doped layers (PIN) photodiodes. The latter are constructed with a single diode per package.

Although OBC system degradation can occur in principle, MISR experiment accuracy has benefited from the stability of the calibrator with time. Pre-launch testing (Bruegge et al. 1993b, Stiegman et al. 1993) established Spectralon preparation and handling procedures that would reduce the risk of on-orbit degradation. Hydrocarbon contaminants introduced during manufacture or testing were shown to cause degradation when exposed to on-orbit vacuum ultraviolet light. Knowing this, the MISR Spectralon panels were vacuum baked following laboratory reflectance testing to remove any such contaminants. In addition, the project elected to replace the panels used during instrument integration and spacecraft testing. Prior to launch, the original panels were replaced with panels that had been kept in a nitrogen-purged container following vacuum baking. Degradation analysis on the OBC (Chrien et al. 2002) has demonstrated the success of this plan. The flight Spectralon panels have degraded on-orbit by no more than 0.5 percent.

Degradation of the Spectralon panels would be of concern if the BRF were to change in shape, or if the relative spectral reflectance were to change at MISR wavelengths. Since MISR makes use of in-flight detector standards, a decrease in the panel's hemispheric reflectance would otherwise be inconsequential. The blue-filtered HQE device, a light-trapped three detector radiometer, has remained stable to better than 0.5 percent throughout the mission (Chrien et al. 2002). This diode is therefore used as the primary standard.

Not all of the monitoring photodiodes have remained stable on-orbit. For this reason, all other photodiodes are recalibrated against this standard prior to the bimonthly data analysis.

Reduction of the OBC experiment data begins with an assumption that the instrument response can be modeled as:

$$DN - DN_o = G_1 L_b \quad (2)$$

where:

- L_b is the incident spectral radiance, weighted over the total-band response function,
- DN is the camera output digital number,
- DN_o is the DN offset, unique for each line of data as determined by an average over the first eight “over-clock” pixel elements (output samples that follow clocking of the CCD line array), and
- G_1 are linear gain coefficients that provide the radiometric calibration of a specific pixel.

Originally it was believed that the photodiodes could be used to measure panel-reflected light as the Sun-panel path traversed a varying amount of the Earth’s atmosphere. This would provide radiance and DN points along the sensor’s response curve, including low light levels, and hence, determine camera linearity and offset. Several lines of evidence lead to the conclusion not to utilize atmospherically attenuated data. These are:

1. Photodiode linearity. The linearity of the photodiodes cannot be validated at ρ_{toa} less than 0.03. Levels below this are outside the linearity range of the preamplifiers.
2. Photodiode offset. A finite photodiode output offset is known to exist, but is not well characterized, due to digitization error for these units. Offset knowledge would be required for low light measurements, but can be ignored at higher light levels.
3. Refraction. Much of the signal observed at the low light levels has been deviated by refraction as light traverses the Earth’s atmosphere. The signal is attenuated differently with time depending on spectral band.
4. BRF uncertainties. There is increased uncertainty in the laboratory measured Spectralon BRF database for large illumination and view angles, like those encountered at low light levels. For these cases, Spectralon becomes more non-lambertian and the BRF increases. Extrapolations to geometries outside the limits of the measured BRF database are highly uncertain.

These issues have led the MISR team to reject any calibration data that have been contaminated by the Earth’s atmosphere, and to adopt a linear calibration equation rather than a more complex form. The linear

method is often referred to as a two-point calibration, since only the gain and offset need to be determined. The gain is computed from DN and radiance pairs measured when the Sun-spacecraft path is free of the Earth’s atmosphere. The offset term is set equal to the system electronic bias. Since the dark current is too small to measure, the DN value assumed for a zero incident radiance is equal to the system electronic bias as measured by DN_o . Differences in coefficients derived using the Clear-of-Atmosphere algorithm are typically less than 1 percent. For the An-camera on-axis observations (center of the CCD array), the new algorithm yields a -0.2 , -0.3 , -0.2 , and 0.7 percent change in radiance. For the Blue-Red bands, the new radiances are darker; for the NIR band, the new radiances are brighter.

The results of these and other OBC validation studies (Chrien et al. 2002) have led the team to use the OBC as a stability monitor, with the absolute scale determined by vicarious calibration experiments.

2.2 Process updates

MISR radiometric coefficients are delivered in a file named the Ancillary Radiometric Product (ARP). Over time, the processing algorithm used to derive these coefficients has changed. Table 2 summarizes these revisions. Each change has reduced radiometric uncertainty incrementally. From this table it can be seen that a band adjustment was made to MISR radiances for data processed after November 2003. The analysis that led to this adjustment is given in §2.4, and is based upon studies of data processed with the previous (linear-off-axis) algorithm.

To understand data heritage, MISR data users should document the version number of the products they use. For Level 1B products, the ARP version used to generate the data should be documented. To determine which ARP file was used to produce a Level 1B data product, one would use the Hierarchical Data Format (HDF) browser, hdfscan. This software, written to view MISR data and generic HDF files, is available from the Langley Distributed Active Archive Center (DAAC). Using this browser, one can read the metadata published within the MISR data product. The ARP file name can be found under Annotation Text: Input Data files. This file name can be compared to the latest delivered ARP file name for a specific time period.

2.3 Vicarious calibration

The MISR team has conducted annual vicarious field campaigns, using desert targets in the southwestern United States. As MISR radiances change with processing algorithm updates, the ratio of MISR to VC radiances may also change. To look for consistent

Table 2. ARP algorithm revision history. ARP files are designated by a format Tx_y, where x are the experiment indices (time), and y is the revision number for a given Tx file. T = 1 represents pre-flight computed coefficients.

HQE-Blue.

- The blue-filtered HQE photodiode is used as the primary radiometric standard. This device is selected based upon its stability with time.

Impact: First calibration attempt. Uses pre-flight calibration of the Blue HQE diode – on-orbit validation not yet in place.

Date: T2_4: Aug. 24, 2000

VC scaling.

- The June 11, 2000 vicarious calibration experiment is used to calibrate the HQE-Blue photodiode standard.

Impact: 9 percent increase in MISR radiometric scale.

Date: T2_5: Feb. 24, 2001

Quadratic.

- A quadratic calibration equation is introduced, believed to show an improvement in the radiances reported over dark targets such as oceans.

Impact: Changes in MISR reported radiances are negligible for equivalent reflectances >0.02 , and a few percent otherwise.

Date: T8_1: May 17, 2001

Provisional.

- The South Pole calibration panel is shown to have measured BRF data that agree with the pre-flight determination. The goniometer is used to update the BRF profile for the North calibration panel.

Impact: Aft-camera radiances decreased by a few percent.

Date: T12_1: Dec. 22, 2001

Linear-off-axis (Linear equation & off-axis correction).

- A linear calibration is restored.
- Data ignored that are contaminated by the Earth's atmosphere.
- Fixes error in BRF indexing code which caused an error for radiance reported for off-axis view angles.
- Impact: – Linear equation is less risky.
- Eliminating data contaminated by the Earth's atmosphere increases experiment accuracy.
- Error fix does not impact on-axis radiances. Greatest improvement is for An-camera west edge of swath (10 percent change).

Date: T17_1: Oct. 24, 2002

Band & Clear-of-Atmosphere adjust.

- Decreases Red band radiance by 3 percent and NIR by 1.5 percent.
- Made use of Clear-of-Atmosphere calibration data.
- Impact: – Radiances are more consistent with vicarious.
- Accuracy increased due to elimination of erroneous data

Date: T24_1: Dec. 5, 2003

biases with respect to MISR, a comparison must be made with a common processing algorithm. This was done using the “linear-off-axis” algorithm (see Table 2), which was in place through October 2003. The results are shown in Figure 1.

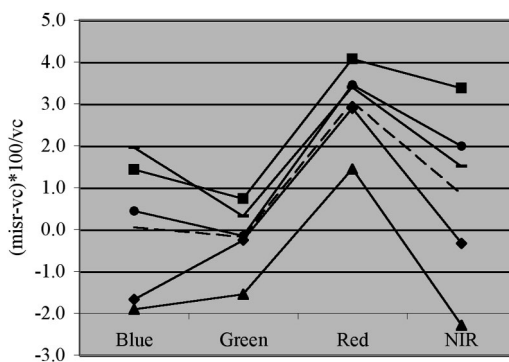


Figure 1. An-camera vicarious calibration results for nadir Terra overpass dates, linear-off-axis algorithm. Lines from top to bottom are: (square) 30-June-2001, Railroad Valley; (circle) 07-July-2003, Black Rock Desert; (dash at inflection points) 22-July-2003, Railroad Valley; (dash line) mean; (diamond) 06-June-2000, Lunar Lake; (triangle) 10-June-2002, Ivanpah playa.

For these experiments, data were acquired at Lunar Lake, Railroad Valley, Ivanpah, and Black Rock Desert (all located in Nevada). Of these, Ivanpah is the smallest playa, roughly 2 km across. Although MISR Level 1B data are corrected for point spread function response as part of the standard processing, the data are exceptionally low, indicating that some residual out-of-field effect may still be present at the 1 percent level. Excluding Ivanpah, the accuracy of the vicarious calibration process appears to be 2–4 percent and smallest for the Green and Red spectral bands. The dashed line in Figure 1 shows the mean of these calibrations. These multi-year data show no systematic change in the reported MISR calibration to within the accuracy of the vicarious calibration methodology. For this reason, no updates have been made to the response coefficient of the OBC primary photodiode standard (Blue-HQE), other than its initial adjustment in February 2001.

2.4 Band-relative adjustment

For these “linear-off-axis” data, the band-relative scale was determined as follows:

1. The June 2000 VC campaign was used to calibrate the Blue-HQE photodiode, the OBC primary standard. This transfer establishes the 2000 VC campaign as the absolute radiometric standard for the MISR experiment. If needed, future VC experiments could be used to update the response of this primary standard, as may be needed if photodiode degradation is measured.
2. For each bimonthly experiment, the measured radiance from the Blue-HQE photodiode is used to

predict the band-weighted radiances for the secondary photodiodes. Differences in field-of-view are taken into account, and the Spectralon reflectance is assumed to be spectrally invariant (Bruegge et al. 2001). The latter assumption is uncertain at the 1 percent level, but impacts the experiment to a lesser degree. For example, Early et al. (2000) published results of a Spectralon BRF measurement round-robin that involved several institutions, including JPL and the National Institute of Standards and Technology (NIST). From this study, it was found that reflectance of Spectralon in the Red band is 0.1 percent higher than that for the Green band. MISR BRF data were acquired at the Helium/Neon (He/Ne) laser wavelength of 632 nm. This result is for a solar illumination angle of 45 degrees and for the nadir viewing pixels, which have the largest viewing angles to the panel. Although the Early report suggests there is some wavelength dependence to Spectralon's reflectance, it is insufficient to explain the wavelength dependent biases shown in Figure 1.

3. The response coefficient of each secondary diode is adjusted such that the measured and Blue-HQE derived radiances agree. This step removes any response degradation from the secondary photodiodes. The photodiodes have degraded differently for the different wavelengths. As an example, the Green and NIR HQE response has changed by 5 percent and 10 percent, respectively.
4. The secondary photodiodes are used to calibrate the cameras with filter-matched diode and camera pairs used in turn. The nadir viewing PIN is used to calibrate the An camera, and Da and Df off-nadir PINs are used to calibrate the eight off-nadir cameras. BRF corrections are required for the C, B, Aa and Af cameras, as there are no photodiodes at these angles.

With this "linear-off-axis" process in place, it can be seen from Figure 1 that there is an inconsistency with wavelength when comparing VC radiances to MISR-measured radiances. In particular, looking at the mean VC comparison, one sees that the MISR radiances are about 3 percent too high in the Red, and 1 percent too high in the NIR.

Figure 2 shows this same mean VC result, but in addition plots lunar comparisons, discussed in §4.2. For this band-relative comparison, a 5 percent bias was removed at all wavelengths from the lunar curve. The consistency of these VC and lunar band-relative results demonstrates the consistency of band-dependent biases. It is noted that these two studies cover a range of target radiances and scene contrasts. The desert targets are roughly 0.3 in equivalent reflectance, the lunar surface approximately 0.04; the desert target is homogeneous and extensive, whereas the lunar target covers only a fraction of the MISR field-of-view.

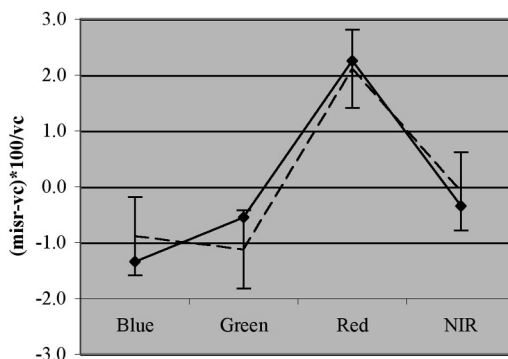


Figure 2. Comparison of mean vicarious calibration (dashed line) results with bias-removed Lunar data (solid line).

As discussed in §2.1, MISR radiances were adjusted slightly when the "Clear-of-Atmosphere" algorithm was implemented (here a two-point calibration method is used, and DN versus radiance data are deleted that are contaminated with the Earth's atmosphere). The Blue-Red radiances are not significantly impacted, whereas the NIR radiances are increased by 0.7 percent, as compared to the linear-off-axis algorithm. This will slightly increase the discrepancy between the MISR NIR band, as compared to the VC and lunar results shown in Figure 2.

Based on these results, supported by Kahn et al. (2004), MISR radiance data products processed after November 2003 have a 3 percent decrease in the red band and 1.5 percent in the NIR. The correction compensates for a systematic error whose origin remains unknown.

3 ERROR SOURCES

Fabrication of a precision camera such as MISR begins with the design of a stable sensor, continues with a component characterization, and is followed with a system-level evaluation. These steps provide: (a) a lower-bound on the experiment uncertainty, (b) a sufficient knowledge of the system such that deviations from the performance specifications, if detected, can be proposed and implemented, and (c) allow design "lessons learned" to be documented for the benefit of future missions. As part of this routine system analysis, potential contributors to radiometric uncertainty, specifically at low light levels, have been investigated. These error sources are discussed in the following sections.

3.1 Electronic offset: baseline stabilization

MISR camera signal chains incorporate a circuitry called BaseLine Stabilization (BLS). The BLS circuitry was incorporated into the MISR design to adjust for

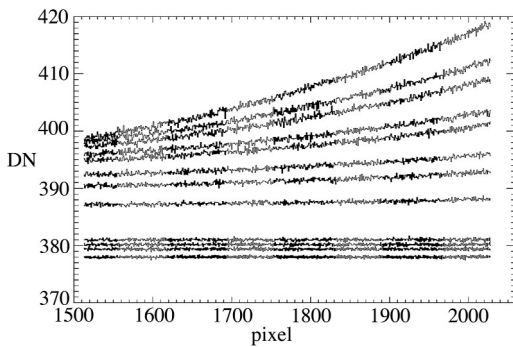


Figure 3. DN from the 512 extended overflow pixels as measured during pre-flight testing. Each line is the response to integrating sphere illumination, spanning the range of equivalent reflectance from 0.05 to 1.0.

sudden changes in the system response, e.g. that which could follow radiation-induced damage to the detector, or undesirable operational amplifier feedback. The circuitry adds a floating electronic pedestal to the signal chain. In the presence of a sudden change in the incident illumination, the BLS stabilizes its output in about 75 lines (3 seconds). To measure the light-induced portion of the signal, this offset must be subtracted. This offset is estimated by use of overflow pixels (Fig. 3). MISR detectors clock out 512 samples of the serial register. These follow reading of the light-sensitive portion of the signal chain. Of these, eight are transmitted from orbit with the active-pixel data. The average of these eight overflow pixels is computed, and used as a measure of the electronic offset, DN. The BLS circuitry also makes use of these overflow pixels to drive the signal chain electronics to -2.9 V, when no optical illumination is present. It is not known whether the BLS circuitry samples the same samples that are transmitted to the ground.

There are several uncertainties associated with the BLS circuitry:

1. Pre-flight testing has shown that the 512 overflow pixels are not constant for a given line of data. Further, it is not known which of the 8 overflow samples is used to establish the BLS output controlling the magnitude of the electronic offset. There is therefore some uncertainty, which can be quantified as the difference in the overflow over the 512 samples.
2. In theory, the BLS should be independent of illumination level. This is found not to be the case, presumably because of light leaking into the serial register even after the CCD signal has been clocked out. This deviation reduces confidence in measuring the electronic offset. A light leak would cause DN to underestimate the bias signal to subtract.

3. The effective offset signal could vary spatially across the active array, and thus no one value of DN would accurately represent the electronic bias. Light leakage into the serial array is one such mechanism by which a spatially variable bias signal could be induced across the CCD array.

With these potential errors, one can bound the radiometric uncertainty due to BLS. If the first 8 overflow samples are used correctly as a representation of the electronic pedestal, then there is no error in the static illumination case. If the electronic pedestal is better represented by the last eight overflow samples, then the error is 20 counts out of 300 DN. This is a 7 percent error for a 2 percent equivalent reflectance signal (worst case error for a homogeneous scene). This BLS uncertainty error would be larger for mixed ocean/cloud or ocean/snow and ice scenes. Here the bright targets would drive up the overflow as well as contribute to leakage into the shield register.

3.2 Electronic offset: dark current

Detector capacity is roughly 1 million electrons and data are digitized to 14 bits. Thus, there are about 60 electrons per DN bin. The CCD readout rate is 40 msec, and the pixel size is 21 mm. The dark current can be expressed as:

$$D = 2.5 \times 10^{15} P N T^{1.5} e - E_g (2kT) \quad (3)$$

where P is the pixel size (cm^2), N is the dark current at 300 K, 26.85 C (nA/cm^2), E_g is the silicon band-gap energy (eV), k the Boltzmann constant, and T the operating temperature (K).

This equation shows that for a -5 degree C operating temperature, the digitized dark current is 1 DN. It is noted that this is an overestimation, as charge is collected only over the integration time (roughly 20 msec), and that the actual pixel area is $21 \times 18 \mu\text{m}$.

In reviewing MISR's measured dark current, a digitized dark current of 0 or 1 DN has been observed for both pre-flight and on-orbit conditions. Thus, this theoretical prediction agrees with the calculations given here. One can conclude that dark current is a minor source of radiometric error, even for dark targets.

3.3 Signal chain: square-root encoding

MISR makes use of square-root data encoding to decrease the data rate required from the spacecraft to the ground station. The algorithm to encode MISR DNs into a compressed number is as follows:

$$DN_{\text{encoded}} = \text{round}(32.0 * \sqrt{\text{float}(DN)}) \quad (4)$$

During data processing, these numbers are restored to their linear representation. To decode these numbers,

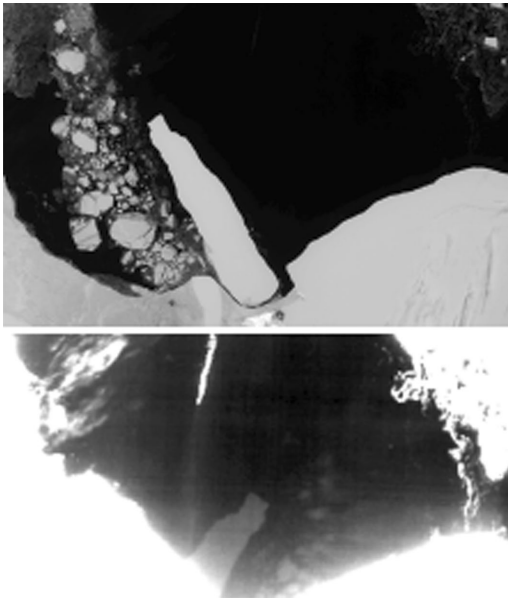


Figure 4. Ghosting in the MISR Bf-NIR band. Data were acquired over the Ross Sea, Orbit 10521, Path 54.

the following operation is performed:

$$DN_L1A = \text{round}(\text{float}(DN_encoded)/32.0)^2 \quad (5)$$

where DN_L1A are DN numbers that are reported to the Level 1A data product. They are only approximately equal to the original camera DN output because the compression scheme cannot be reversed perfectly.

One now can estimate the percentage radiometric error by assuming a typical overclock value:

$$DN_overclock = 350 \quad (6)$$

$$\%_error = (DN_L1A - DN) * 100. / (DN - DN)_o \quad (7)$$

Thus, it is shown that the radiometric error attributable to square-root encoding is 0.5 percent for DN greater than 200 DN above overclock (an equivalent reflectance of 0.005), and decreases with increasing illumination. The error due to square-root encoding is therefore considered to be negligible.

3.4 Optical effects: ghosting

Figure 4 shows an iceberg surrounded by a dark ocean. The image was acquired on December 9, 2001 over the Ross Sea. The lower image shows a highly contrast-stretched image of the normal view, shown above. The

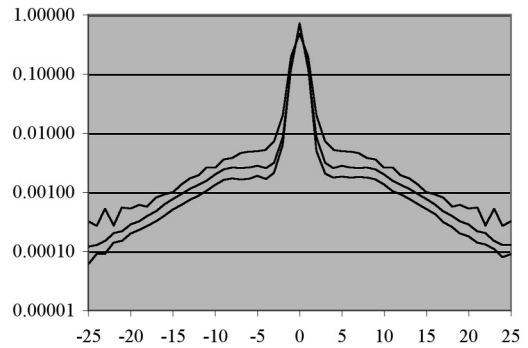


Figure 5. Comparison of pre-flight and in-flight empirical PSF's. Curves from top to bottom are in-flight derived, operational functions, and pre-flight measurements.

iceberg shown in the lower figure is an inverted, blurred ghost image of the original iceberg. The reflectance of the ice is approximately 0.43 in all bands, whereas the ocean reflectance varies from 0.06 to 0.01 from the Blue to NIR wavelengths. In this image, 0.3 percent of the bright target has been reflected into the adjacent dark ocean. For ghosting of this magnitude, a target brighter than $\rho_{toa} = 0.66$ would be required to induce a radiometric error of 10 percent in a dark ocean scene of magnitude 0.02. The dark target would have to be located specifically in the ghost location. It can be concluded that for all but a small number of cases, ghosting will not impact radiometry over dark ocean targets.

No attempt is made to correct for ghosting, as the effect is small, and as the secondary image is not in focus. Any attempt to remove the ghost would result in a shadow where the image had been. We also know that only the A and D cameras were tuned to produce blurred ghost images. This was never done for the B and C cameras. The degree of focus for these ghost images, therefore, is unknown and camera dependent.

3.5 Optical effects: point spread function response

The image of a point object source is always blurred due to diffraction, lens aberrations, and scattering. This output response to a point source is known as the point-spread-function (PSF) for a given optical system. MISR PSF functions have been measured pre-flight.

For an in-flight determination of the PSF, the derivative of the edge response was taken using the iceberg edge of the December 9, 2001 Antarctica scene. The updated response was found to have the same shape, but with a larger halo, as compared to the pre-flight measurement. These pre-flight and in-flight derived PSF kernels are shown in Figure 5. This figure indicates that the pre-flight PSF's underestimate the amount of contrast adjustment needed.

PSF correction is done on all MISR radiance data products. The operational PSF functions were derived by the following procedure:

1. Start with the pre-flight point spread functions.
2. Average right and left about the center to make them symmetrical.
3. Multiply the entire function by a scale factor that adjusts the background halo to that empirically derived from the on-orbit data. A value of 1.5 has been used to produce the operational kernels.
4. Renormalize to unit area by adjusting the energy in the central 3 pixels (which contain ~95 percent of the energy).
5. Take the inverse Fourier transform (FT).
6. Multiply the inverse Fourier transform by the Fourier transform of the central 9 pixels of the PSF (the “core”). This step is performed so that the deconvolved images correspond to the PSF of the core, rather than a delta function, and is done to avoid ringing at contrast boundaries.
7. Inverse transform, take the real part, and average left/right to correct any numerically-induced asymmetries.

It is believed that this procedure minimizes the ringing associated with a sharper PSF core, and improves the contrast. Inspection of MISR Level 1B2 radiance products reveal that sharp radiance discontinuities can be observed in the presence of contrast edges, such as iceberg edges. This is validation of the PSF deconvolution process.

4 CROSS-COMPARISON STUDIES

4.1 MODIS

MODIS is a useful sensor with which to cross-compare MISR radiometric products. The MODIS passband parameters, when derived using the MISR moments analysis algorithm, are listed in Table 3. MODIS Band 4 (land) and Band 9 (ocean) are examples of bands that are well matched for the two sensors. For all bands a radiometric correction is made to predict the radiance that MODIS would have reported, had it been built with MISR bandpasses. This spectral-correction algorithm has been described in Bruegge et al. (2002). MISR to MODIS radiance ratios reported in this Section have all had these spectral corrections performed.

MODIS reports a Reflective Solar Bands (RSB) calibration uncertainty of 2 percent for the reflectance factor and 5 percent for the radiance product. Both land and ocean channel calibrations utilize a solar diffuser (SD). In addition to a direct view of the panel, a second data set is acquired using a 7.8 percent transmission screen deployed in front of the SD. This calibrates those channels that would otherwise saturate.

Table 3. MODIS spectral parameters. The last two columns give spectral correction factors for two scene types.

MISR Band no.	Sensor	λ_c , (nm)	$\delta\lambda$, (nm)	$E_{0,b}$ (Wm^{-2} μm^{-1})	Sensor/MISR radiance scale factor	
					(Desert)	(Ocean)
1	MODIS Band 3	466	21	2015	0.906	1.054
2	MODIS Band 4	554	21	1858	1.002	0.978
3	MODIS Band 1	646	50	1601	0.986	0.903
4	MODIS Band 2	856	45	989.8	0.982	0.997
1	MODIS Band 9	442	11	1865	1.010	0.978
2	MODIS Band 12	547	12	1870	1.012	0.933
3	MODIS Band 14	677	14	1505	1.003	1.027
4	MODIS Band 16	866	19	969.7	1.005	1.029

The MODIS detectors view the calibrated SD to place their data products on a TOA reflectance scale. The measurement precision is about 0.2 to 0.5 percent depending on the bands; or, if the SD screen is used in the calibration (ocean bands use the SD screen). SWIR band uncertainties are higher due to residual cross-talk errors.

The Solar Diffuser Stability Monitor (SDSM) is used to track SD degradation. The SDSM is a small integrating sphere and filtered detectors that look at the Sun and SD respectively. The ratio of the SD to the Sun view provides a measure of the spectral reflectance of the SD, thus tracks the SD degradation. The SDSM makes use of a 2 percent transmission screen when viewing the Sun. This is done to place the incoming signal on the same point on its dynamic range curve.

More recent data use a direct view of the SD for both the land and ocean channels. The MODIS calibration team has found that the difference in using, or not using, the screen is less than 1 percent. Thus, they believe the relative calibration between the ocean and land bands is within 1 percent.

Figure 6 compares the radiances measured by several sensors against the vicarious calibration radiances conducted July 22, 2003. For comparisons over uniform desert playas, MISR is typically 3–7 percent brighter than MODIS, for MISR data of “linear-off-axis” heritage. The agreement is within 4 percent for MISR “band-adjusted” data. The radiometric bias between MISR and MODIS can be traced to utilizing different standards and processes used to establish their respective scales. For example, Thome (University of Arizona) provides VC datasets for the MODIS team,

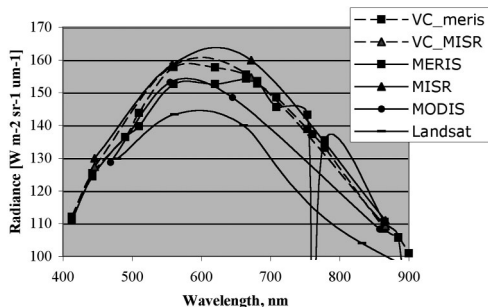


Figure 6. Measured radiances from Vicarious Calibration data, MISR, MERIS, MODIS, and ETM+. Data were acquired July 22, 2003 at Railroad Valley, Nevada.

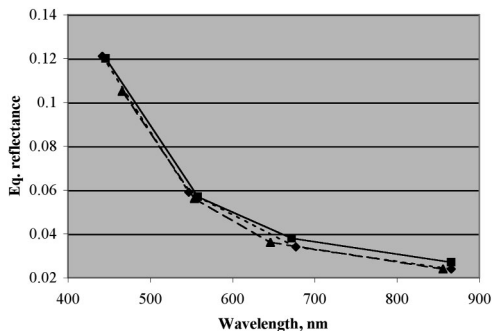


Figure 7. A comparison of radiometry from MISR (squares), MODIS land channels (long dash, triangles), and MODIS ocean channels (dash, diamonds), for a dark ocean target. Data are for a dark ocean target over the Arabian Sea on 06 April 2001 (57.43° E, 17.80° N).

and reports a discrepancy of -1.4 , -0.9 , -3.4 , -2.5 , and -3.4 percent respectively, for the MODIS 412, 469, 555, 645, 858 nm bands. MODIS radiances are lower than his VC observations. As MISR is calibrated against a VC standard, greater consistency between MISR and a VC-adjusted MODIS is found. MODIS VC studies are done using the land channels, since the ocean channels saturate over these bright targets.

MISR and MODIS comparisons were next made over uniform dark oceans. One such example is provided in Figures 7–8. In Figure 7, radiances are compared for a clear, dark ocean target, from MISR and MODIS. In making a large number of such comparisons, one generally sees that MISR data agree with MODIS land-channel data, to within 3 percent, when land targets are observed. For ocean targets, the agreement between MISR and MODIS ocean-channel data is comparable.

Figure 8 shows a discrepancy between MISR and MODIS land channels when an ocean scene is observed. In this figure, the scene is half land and half ocean. The upper panel shows the column averages and standard deviations of radiances, for both MISR

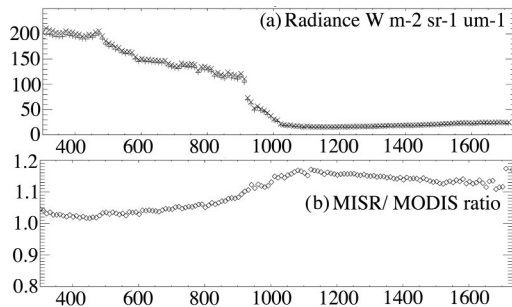


Figure 8. (Top) Plot of mean equivalent reflectance versus sample for the scene compared in Figure 7. The mean is calculated by averaging MISR or MODIS radiances in the vertical (north-south) direction. Standard deviations around the mean are also shown. High standard deviations occur where both land and water contribute to the calculations. (Bottom) Median MISR/MODIS radiance ratio at each sample location, calculated by combining all points in the north-south direction. Standard deviations are also shown. Note the inconsistency between the ratio over water (dark scene) relative to land (bright scene).

and MODIS. The lower panel shows the median ratio. It is noted that the ratio agrees well over land, but increases over ocean. This suggests a non-linearity in one or the other sensor. It is noted that MISR is an all-refractive system, with a Lyot depolarizer as its front element. MISR is insensitive to polarization effects. Since MODIS may have greater polarization sensitivity, this may partially explain the variation in the ratio for the ocean site, as compared to the land. A more detailed validation of MODIS radiometry, using the ocean bands, is on-going by the MODIS ocean community. Further validation studies on the radiometric response of MODIS land channels over dark ocean targets are needed. It is noted that MODIS land channels are used to retrieve aerosol properties over ocean sites. Issues that the MODIS team are investigating include uncertainties in the SD reflected light due to excess radiance (scattering or Earthshine), uncertainties in the effects of the SD attenuation screen, and changing polarization sensitivity.

To summarize, MISR and MODIS radiance data agree to within a 4 percent uncertainty, providing that the MODIS land channels are used over land and the MODIS ocean channels are used over ocean. The agreement of the ocean values, using the MODIS ocean channels, provides a validation of MISR radiometry at these low light levels.

4.2 Lunar observations

Lunar observations are routinely used by the Sea Wide Field-of-View sensor (SeaWiFS) (Barnes et al. 1999), an ocean viewing sensor to track degradation with time.

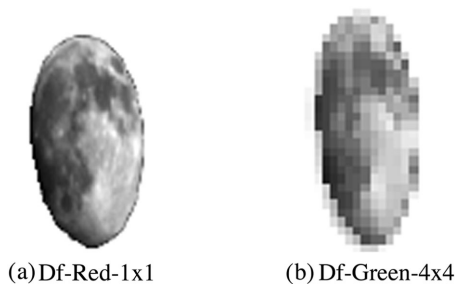


Figure 9. MISR viewed the moon in its baseline Global Mode configuration. (a) Twelve of the 36 MISR channels are configured to high-resolution (no averaging) during Global Mode, including this Df_Red image. (b) Twenty-four channels are in 4×4 pixel averaging mode, such as this Df_Green_4 \times 4 image.

On April 14, 2003, MISR had its first opportunity to view the Moon. A special maneuver of the Terra spacecraft was performed on this date as it traversed the nightside of the orbit. The maneuver entailed a backwards somersault of the spacecraft as it pitched end-over-end, allowing the normally Earth-viewing instruments to look at deep space and the waxing Moon. The purpose of this acrobatic feat was to assist in cross-comparing MISR, MODIS, and ASTER (all Terra instruments), as well as several other sensors on an assortment of platforms. During a 16-minute interval, the lunar disk passed through the fields-of-view of all nine MISR cameras, resulting in a unique set of images. Figure 9 (left) shows one such image for the high-resolution Df-red channel. Familiar lunar features are clearly recognizable. The dark lunar “maria” are vast plains of basaltic lava.

Lunar data were acquired in GM, with 12 high-resolution channels and 24 channels in 4 pixel by 4 line averaging mode, called 4×4 (Fig. 9 right). Varying resolution is also encountered with the four camera designs, due to differences in focal lengths. Examples of the resolution as a function of camera design are shown in Figure 10.

These data were used to derive a measure of lunar irradiance, which was compared to that of the empirically derived Robotic Lunar Observatory (ROLO) model. In computing this irradiance, one must account for a large amount of over-sampling. The equation used for this analysis sums the radiance samples, multiplies by the detector solid angle, and corrects for over-sampling.

$$\begin{aligned} \text{Lunar Irradiance} [mWm^{-2}nm^{-1}] = \\ (\text{fraction of IFOV not in previous samples}) * \\ (\text{detector solid angle} [sr]) * \\ (L [mWm^{-2}sr^{-1}nm^{-1}]) \end{aligned} \quad (8)$$

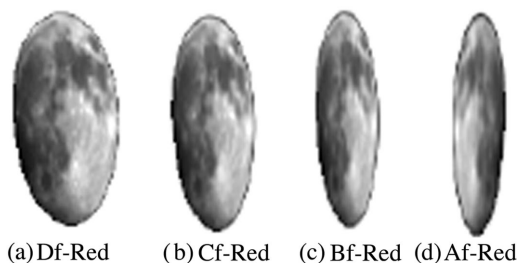


Figure 10. Lunar image resolution degrades progressively from the D to A cameras due to progressively shorter focal lengths from D to A.

Table 4. Lunar calibration parameters for April 14, 2003.

Lunar observation date: April 14, 2003
Orbit: 17672
Time: 2200–2220 UT
Planned Terra pitch rate: D [rad/sec] = 0.002129, or 0.122 deg/sec
Lunar distance: D [1000 km] = $366.5 - 6371 /$ $1000 - 1738 / 1000 - 0.705 = 357.7$
MISR camera sampling rate: tsamp [msec] = 40.8
Camera focal length, Df-Da order: f [mm] = {123.67, 95.34, 73.02, 58.90, 58.90, 59.03, 73.00, 95.32, 123.65}
Cross-track detector dimension: Dxdet [mm] = 21
Along-track detector dimension: Dydet [mm] = 18
Sum of MISR radiance samples: L [$mWm^{-2}sr^{-1}nm^{-1}$]
Fraction of IFOV not in previous sample: tsamp*D*f/Dxdet
Detector solid angle [sr]: Dxdet* Dydet* $10^{-6}/f^2$

The required inputs are given in Table 4.

The results of this comparison are shown in Figure 11. Although there appears to be a 5 percent bias between MISR and ROLO, in fact a similar bias is found between the ROLO model and SeaWiFS, as well as between the model and MODIS land channels (Kieffer 2003). It is concluded that the lunar observations confirm that MISR radiometry is consistent with MODIS land channels, as well as SeaWiFS, for this low light, small extent target.

The accuracy of the MISR measure of lunar irradiance is limited by the over-sampling correction. Its uncertainty is at least 5 percent, based upon the scatter of measurements from the nine MISR cameras. Nevertheless, the experiment proved to be extremely valuable for the MISR community in that it validated the band-adjustment values obtained from the vicarious calibration experiments (Fig. 2). It also confirmed that the absolute radiometry is consistent with comparisons made over land targets. Confidence in MISR radiometry over low light level targets is therefore increased.

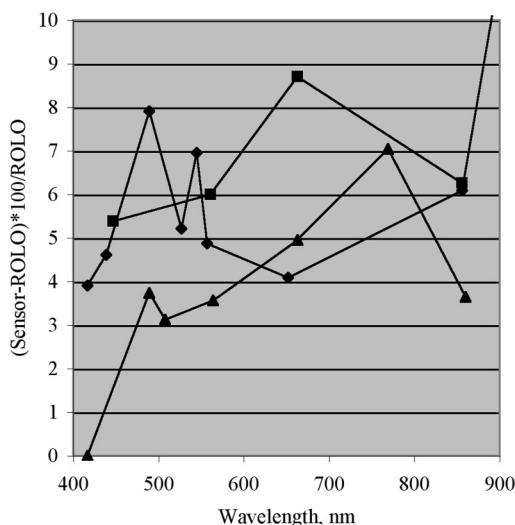


Figure 11. Differences between MISR (squares), MODIS (diamonds), and SeaWiFS (triangles), as compared to the ROLO measure of lunar irradiance.

5 CONCLUSIONS

The MISR calibration and science teams have reviewed carefully the contributors to radiometric accuracy over low light scenes. As with any instrument, radiometry for these conditions is challenging. The impact of electronic bias, dark current, and data compression have been investigated, and there is no evidence that these error sources are degrading the quality of the data. It is felt, however, that system level validation offers the best evidence of radiometric accuracy. Cross-comparison with MODIS over both ocean and lunar observations demonstrate the consistency of data for dark targets, as do dark water vicarious calibrations. For these reasons, it is concluded that MISR is meeting its absolute calibration specification for dark water conditions, including an uncertainty specification of 10 percent at $\rho_{\text{toa}} = 0.02$. Further, for data processed after November 2003, it is concluded that the band-relative requirement for an uncertainty of 1 percent is also being met throughout the dynamic range of the instrument.

This experience leads to a proposal, for future mission designs:

1. a floating electronic bias should NOT be added to a signal chain;
2. Spectralon does not degrade on-orbit, if contamination is avoided;
3. multiple methodologies are required in many calibration programs since not all pathways will succeed;

4. cross-comparisons of sensors need to be made at a range of illumination levels, crosstrack view angles, and scene contrasts; and
5. having a calibration program over desert targets is only the first step in achieving a calibrated instrument. Science communities need to make use of a range of scene brightness and contrast conditions in their validation studies.

ACKNOWLEDGMENTS

The work described in this paper has been carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. MISR data products are processed and made available by the Atmospheric Sciences Data Center, Langley Research Center.

A number of support individuals has assisted in acquiring and analyzing MISR calibration data, including Nadine Chrien, Mike Smyth, David Nelson, Kyle Miller, Tom Thaller, and Tom Nolan. The authors also wish to thank the following individuals who have participated in Peer Reviews of MISR calibration: Roger Davies, Michael Gunson and Cinzia Zuffada, JPL; Brian Markham, Landsat Calibration, GSFC; Stu Biggar, Vicarious calibration, Univ. of Arizona; James Butler, EOS Calibration Scientist, GSFC; Bruce Guenther, MODIS Calibration, GSFC; and Jeff Mendenhall, ALI Design Engineer, Lincoln Labs.

Finally, Jack Xiong, GSFC, and Kurtis Thome are thanked for many useful discussions and data exchanges.

REFERENCES

- Abdou, W.A., Helmlinger, M.C., Conel, J.E., Pilorz, S., Bruegge, C.J., Gaitley, B.J. & Martonchik, J.V. 2000. Ground measurements of surface bidirectional reflectance factor (BRF) and hemispherical directional reflectance factor (HDRF) using the Portable Apparatus for Rapid Acquisition of Bidirectional Observation of the Land and Atmosphere (PARABOLA III). *J. Geophys. Res.* 106: 11, 967–11,976.
- Abdou, W., Bruegge, C., Helmlinger, M., Conel, J., Pilorz, S. & Gaitley, B. 2002. Vicarious calibration experiment in support of the Multi-angle Imaging SpectroRadiometer (MISR). *IEEE Trans. Geosci. & Remote Sens.* 40(7): 1500–1511.
- Barnes, R.A., Eplee, R.E. Jr., Patt, F.S. & McClain, C.R. 1999. Changes in the radiometric sensitivity of SeaWiFS determined from Lunar and solar-based measurements. *Appl. Optics* 38(21): 4,649–4,664.
- Bruegge, C., Duval, V., Chrien, N. & Diner, D. 1993a. Calibration plans for the Multi-angle Imaging Spectro Radiometer (MISR). *Metrologia* 30(4): 213–221.

- Bruegge, C., Stiegman, A., Rainen, R. & Springsteen, A. 1993b. Use of Spectralon as a diffuse reflectance standard for in-flight calibration of earth-orbiting sensors. *Opt. Eng.* 32(4): 805–814.
- Bruegge, C.J., Helmlinger, M.C., Conel, J.E., Gaitley, B.J. & Abdou, W.A. 2000. PARABOLA III: a sphere-scanning radiometer for field determination of surface anisotropic reflectance functions. *Rem. Sens. Reviews* 19: 75–94.
- Bruegge, C., Chrien, N. & Haner, D. 2001. A Spectralon BRF data base for MISR calibration applications. *Rem. Sens. Env.* 77(3): 354–366.
- Bruegge, C.J., Chrien, N.L., Ando, R.R., Diner, D.J., Abdou, W.A., Helmlinger, M.C., Pilorz, S.H. & Thome, K.J. 2002. Early Validation of the Multi-angle Imaging SpectroRadiometer (MISR) Radiometric Scale. *IEEE Trans. Geosci. & Rem. Sens.* 40(7): 1477–1492.
- Chrien, N., Bruegge, C. & Ando, R. 2002. Multi-angle Imaging SpectroRadiometer (MISR) onboard calibrator (OBC) in-flight performance studies. *IEEE Trans. Geosci. & Rem. Sens.* 40(7): 1493–1499.
- Diner, D., Beckert, J., Reilly, T., Bruegge, C., Conel, J., Kahn, R., Martonchik, J., Ackerman, T., Davies, R., Gerstl, S., Gordon, H., Muller, J.P., Myneni, R., Sellers, R., Pinty, B. & Verstraete, M. 1998a. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *IEEE Trans. Geosci. & Rem. Sens.* 36(4): 1072–1087.
- Diner, D.J., Barge, L.M., Bruegge, C.J., Chrien, T.G., Conel, J.E., Eastwood, M.L., Garcia, J.D., Hernandez, M.A., Kurzwel, C.G., Ledebøer, W.C., Pignatano, N.D., Sarture, C.M. & Smith, B.G. 1998b. The Airborne Multi-angle Imaging SpectroRadiometer (AirMISR): instrument description and first results. *IEEE Trans. Geosci. & Rem. Sens.* 36(4): 1339–1349.
- Diner, D.J., Beckert, J.C., Bothwell, G.W. & Rodriguez, J.I. 2002. Performance of the MISR instrument during its first 20 months in Earth orbit. *IEEE Trans. Geosci. & Rem. Sens.* 40(7): 1449–1466.
- Early, E.A., Barnes, P.Y., Johnson, B.C., Butler, J.J., Bruegge, C.J., Biggar, S.F., Spyak, P.R. & Pavlov, M.M. 2000. Bidirectional reflectance round-robin in support of the Earth Observing System program. *J. Atmos. & Oceanic Tech.* 17(8): 1077–1091.
- Kahn, R., Banerjee, P., McDonald, D. & Diner, D.J. 1998. Sensitivity of multiangle imaging to aerosol optical depth and to pure-particle size distribution and composition over ocean. *J. Geophys. Res.* 103(D24): 32,195–32,213.
- Kahn, R., Banerjee, P. & McDonald, D. 2001. The sensitivity of multi-angle imaging to natural mixtures of aerosols over ocean. *J. Geophys. Res.* 103(D24): 32,195–32,213.
- Kahn, R., Li, W.H., Bruegge, C., Martonchik, J., Diner, D., Gaitley, B., Dubovik, O., Holben, B., Smirnov, A., Jin, Z. & Clark, D. 2004. MISR low-light-level calibration, and implications for aerosol retrieval over dark water. *J. Geophys. Res.* (in press).
- Kieffer, H. 2003. Characterization and radiometric calibration for remote sensing. *Proc. CALCON; Logan, Utah 15–18 September 2003*. Utah State Univ.: Space Dynamics Laboratory. CD-ROM
- Stiegman, A.E., Bruegge, C.J. & Springsteen, A.W. 1993. Ultraviolet stability and contamination analysis of Spectralon diffuse reflectance material. *Opt. Eng.* 32(4): 799–804.
- World Climate Research Programme (WCRP) 1986. Publication Series No. 7, WMO ITD-No. 149: 119–126, October 1986.

Calibration programs

Forty years of experience with SPOT in-flight calibration

C. Valorge, A. Meygret, L. Lebègue, P. Henry
CNES, Toulouse, France

A. Bouillon & R. Gachet
IGN/CNES, Toulouse, France

E. Breton
IGN Espace, Ramonville Saint Agne, France

D. Léger & F. Viallefont
ONERA/CERT, Toulouse, France

ABSTRACT: Since SPOT 1 launched in February 1986, several teams of engineers from the Centre National d'Etudes Spatiales (CNES), Institut Géographique National (IGN) and Office National d'Etudes et de Recherches Aérospatiales (ONERA) in Toulouse have been involved in on-orbit image quality monitoring. Methods to insure the best quality of imagery delivered to SPOTIMAGE customers have been improved and updated continuously through five commissioning phases and more than 40 years of accumulated life in orbit. These methods had to evolve to be adapted to several on-board improvements in terms of sensor resolution, geometric accuracy, addition of spectral bands, and new instruments. This process has culminated with the on-orbit calibrations phase of the Satellite Pour l'Observation de la Terre-5 (SPOT 5), the most complex, innovative, and demanding satellite in the family. It will be used to illustrate this paper. Traditionally, on-orbit image quality monitoring is divided into two phases depending on the technical domain each covers: the first one (called "calibration") consists of determining and updating the optimal parameters used in the image ground segment to "de-spatialize" the images, i.e. to correct them from all specific onboard defaults affecting raw images. The second one (called "image quality assessment") consists of evaluating the image quality obtained by these final products, quality which is measured through specific figures-of-merit in each technical domain. These two phases are applied in the two technical domains used to divide image quality into: (a) *radiometric quality* – correcting for the different responses of the detectors (normalization) and establishing the rules for converting the digital numbers to radiance (absolute calibration). In terms of image quality, it is quantified by the signal-to-noise ratio (for a given radiance) and accuracy of calibration; and, (b) *geometric quality* – tuning the parameters involved in the geolocation model (geometric calibration). Besides attitude and orbit restitution tuning (not considered here), geometric tuning consists of estimating the biases between the instrument orientation and the attitude and orbit control subsystem (AOCS) reference frame, and in modeling both pointing mirror movement and each detector's line of sight in the focal plane. In terms of image quality, it is quantified by the absolute geolocation accuracy, length alteration, and planimetric and altimetric accuracies. This paper will present the best practices as they have evolved for each theme.

1 INTRODUCTION

Great emphasis has been put on the quality of imagery delivered by the Satellite Pour l'Observation de la Terre (SPOT) satellites to address as many needs as possible. A constant effort has been made during design, development, and on-orbit testing to insure the best quality of imagery in three technical domains: geometry (capacity

to locate the pixels on the Earth), radiometry (capacity to interpret radiometric counts in terms of radiance), and resolution (capacity to identify small objects).

Beyond pre-launch activities (not addressed in this paper), §2 presents the operational organization to monitor image quality of the SPOT constellation, while §3, 4, and 5 describe the methods used for each domain.

2 SPOT SYSTEM OVERVIEW

2.1 SPOT space segment

The SPOT constellation consists of 5 satellites: SPOT 1 (launched in February 1986), SPOT 2 (launched in January 1990), SPOT 3 (launched in September 1993), SPOT 4 (launched in March 1998), and SPOT 5 (launched in May 2002). The first three satellites share the same design and characteristics. They carry two identical instruments called the High Resolution Visible sensor (HRV) with 4 onboard spectral bands: a 10 m panchromatic band with a 6000 pixel linear array; and, three 20 m registered color bands (green, red and near infrared) with a 3000 pixel array.

Thus, each image covers a 60 km swath (for nadir viewing), and both instruments can observe up to 27 degrees on either side of the vertical by virtue of a steerable mirror that gives them access to a 900 km wide corridor on each orbit.

With SPOT 4, significant improvements were made: (a) inclusion of a new 20 m short-wave infrared (SWIR) band onboard the two newly named Haute Résolution Visible et InfraRouge (HRVIR) sensors; (b) onboard registration of these 5 bands; (c) new payload called VéGéTation (VGT), with four 1 km resolution color bands (blue, red, near and short-wave infrared) over a 2200 km swath; and, (d) a new device devoted to real-time onboard orbit determination called Détermination d'Orbite et Radiopositionnement Intègre par Satellite (DORIS).

SPOT 5 also included enhanced equipment: (a) an Haute Résolution Géométrique (HRG) payload to provide 10 m registered color images, a short-wave infrared sensor at 20 m, and 5 m called Haute résolution Monomode (HM) or 2.5 m Très Haute Résolution (THR) panchromatic images, all with the same 60 km swath; (b) a new Haute Résolution Stéréoscopique (HRS) instrument devoted to automatic digital elevation model (DEM) determination consisting of two telescopes allowing a 20 degree forward and aft view over a 120 km swath, respectively; (d) a stellar sensor to improve on-the-ground attitude determination, leading to better geolocation accuracy; and, (e) a yaw steering function used to balance Earth rotation during acquisition. Currently, 3 satellites are used operationally (SPOT 2, SPOT 4 and SPOT 5) in the same 832 km sun-synchronous orbit, with a 26-day cycle giving daily access to any point on the globe. Table 1 lists the names, sensors, and imaging modes discussed in the following sections.

2.2 SPOT ground segment

The SPOT ground segment is operated by different entities: (a) the Satellite Operations and Control Center ensures optimal behavior of the bus and transmits data

Table 1. Names and characteristics for all SPOT sensors and imaging modes.

Satellite	Sensor	Mode	Spectral bands	Resolution (m)
SPOT 1–3	HRV	P	PAN	10
		XS	B1, B2, B3	20
SPOT 4	HRVIR	M	B2	10
		XS	B1, B2, B3	20
		XI	B1, B2, B3, SWIR	20
SPOT 5	HRG	THR	PAN	2.5
		HM	PAN	5
		HX	B1, B2, B3	10
		HI	B1, B2, B3, SWIR	10
	HRS		PAN	5 × 10

acquisition commands; (b) the Programming Center determines the acquisition plans of each Satellite from its clients' image needs; (c) the Processing and Archiving Center archives image telemetry and converts data into products; (d) a network of Direct Receiving Stations that program and process imagery; and, (e) an Image Quality Expertise Center (IQEC) ensures the quality of images.

This IQEC is the operational entity centralizing all methods, means, and people needed to: (a) determine the optimal onboard parameters concerning image quality (focus, radiometric gains, and compression parameters); (b) elaborate, validate, and supply processing centers with optimal ground processing parameters (geometric biases, interior orientation parameters, and normalization coefficients); (c) establish the image quality performance budgets concerning geometry, radiometry and resolution; and, (d) analyze and resolve any identified image quality problems. Thus, this center has the capacity to program the satellites for its own purpose, implement expert methods and algorithms, manage all data produced during its activity, and operate several calibration sites.

3 GEOMETRIC CALIBRATION AND QUALITY ASSESSMENT

3.1 Geolocation model

Calibration of the geometry of Haute Résolution Visible (HRV), HRVIR and HRG images has not changed since 1986. Test sites have been chosen to estimate location model accuracy on a global scale. The availability of accurate maps and distribution of sites are the main selection criteria. Selected test sites are observed during the commissioning phase to calibrate location models. They also are observed during the operational life of the satellite to detect any evolution of sensor performance. The main expected evolutions are related to aging of

attitude restitution equipment and steering mirror axes deformation as a result of repeated off-nadir acquisition.

3.1.1 GCP database management

Since test sites used for SPOT 1 calibration were to be used for subsequent SPOT satellites, a ground control point (GCP) database was created. Map extracts corresponding to each GCP have been digitized, and each corresponding image observation is extracted and stored in the database. Thus, it has been possible to produce a GCP database from the first images acquired on each test site. Once completed, this database participated in reducing workload for SPOT 1–4. Only minor evolutions have occurred during these first ten years of service.

With SPOT 5, resolution was improved, making it difficult to get accurate identification of GCPs in the images. Furthermore, experience accumulated with SPOT 1–4 has shown that for some test sites GCPs lacked accuracy. That was of great concern with the specified improvement in location model accuracy. The objective of decametric location performance requires reliable GCPs, even if only ten GCPs per image are needed. Thus, SPOT 5 was the occasion to refurbish the GCP database, keeping in mind the need for a homogeneous and global distribution of test-sites to ensure daily access to any of them, and various observation conditions to detect unexpected behaviors. Commissioning phase results have shown that such preparation was relevant.

3.1.2 Model calibrations

Since SPOT 1, location model errors have been estimated for each image independently. Bundle block adjustment software for calibration was started with SPOT 5. For each test site, estimates of all available location models can be made. Consistent calibration results are obtained for all instruments with reduced noise level. The software can be used also to improve GCP ground coordinate accuracy, especially for test sites for which there are no GPS measurements. When enough data have been collected, their statistical analysis leads to bias estimation for each HRVIR or HRG look angle.

The steering mirror viewing model is calibrated according to an optical model relying on reflection law on a plane mirror rotating around an axis that is not perfectly in the mirror's plane. Such a model was created after SPOT 1 commissioning and is still in use (Fig. 1).

Once such a model is calibrated, the remaining errors are interpreted in terms of misalignments between the instruments themselves and the AOCS reference frame. Thanks to the global distribution of test sites, it is possible to detect peculiar trends such as a variation of misalignments with latitude (Fig. 2).

Once biases between instruments, AOCS reference frame, and steering mirror viewing model are estimated,

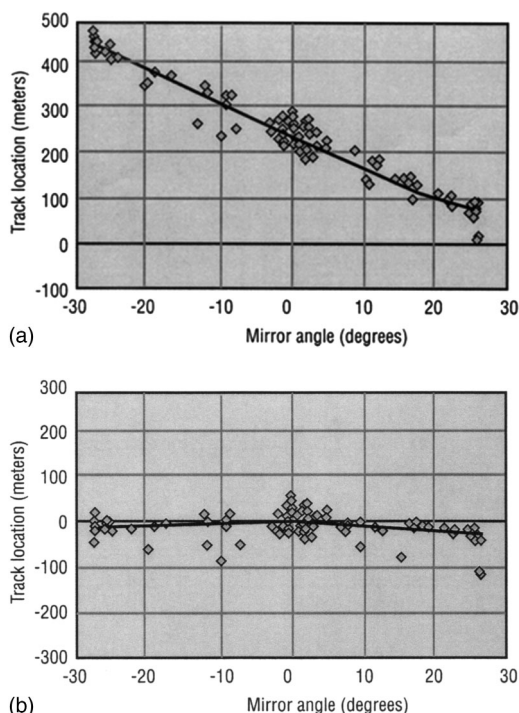


Figure 1. SPOT 5/HRG 2's along-track location performance as a function of mirror angle: (a) before; and (b) after calibration.

they are transmitted to every processing station to process images with the best available calibration coefficients.

3.1.3 Geolocation accuracy assessment

Geolocation accuracy is assessed after geometric calibration simply by looking at the remaining location errors of all images used during the calibration process (Fig. 3). Once again, the global distribution of sites and the various observation conditions is the best way to ensure good confidence in performance estimation.

3.2 Internal orientation

Internal orientation refers to alignment of sensor detectors.

3.2.1 Distortion model calibration

From the beginning of SPOT, it has been a goal to estimate the geometric shape of the charged-couple device (CCD) alignments, first to evaluate their effect on the geometric quality of images (SPOT 1–4), then to calibrate and correct this effect (SPOT 5). For SPOT 1–4, the main geometric defect was due to the optical junction between the four arrays. The possibility

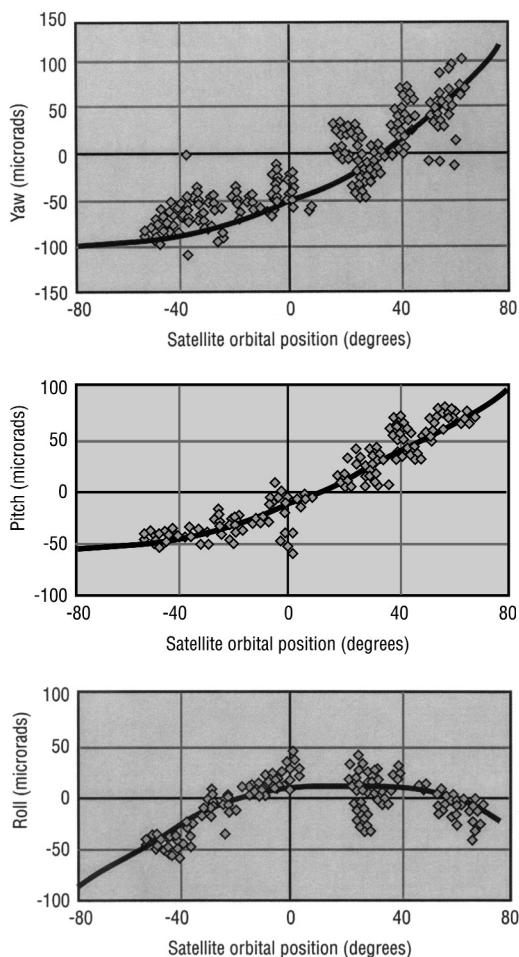


Figure 2. HRS roll, pitch, and yaw performance and orbital position determined by bundle block adjustment over test sites.

of simultaneously acquiring the same scene (slightly shifted across track) with both instruments made it possible to compare images geometrically, with and without the junction. Outside of the junction area, the method brought information about the relative shape between the two instruments and allowed a comparison of the regularity of geometric sampling against its specification. Furthermore, the steering mirror permits one to choose the overlap between images leading, under certain assumptions, to absolute knowledge of the shape of each alignment of detectors. A description of this latter method is given in Gachet (1999). It has been applied experimentally to SPOT 4 instruments. Figure 4 shows the results in pixels for both HRVIR instruments across- and along-track.

For SPOT 5, the improvement of most geometric characteristics of data acquisition led to a decision to

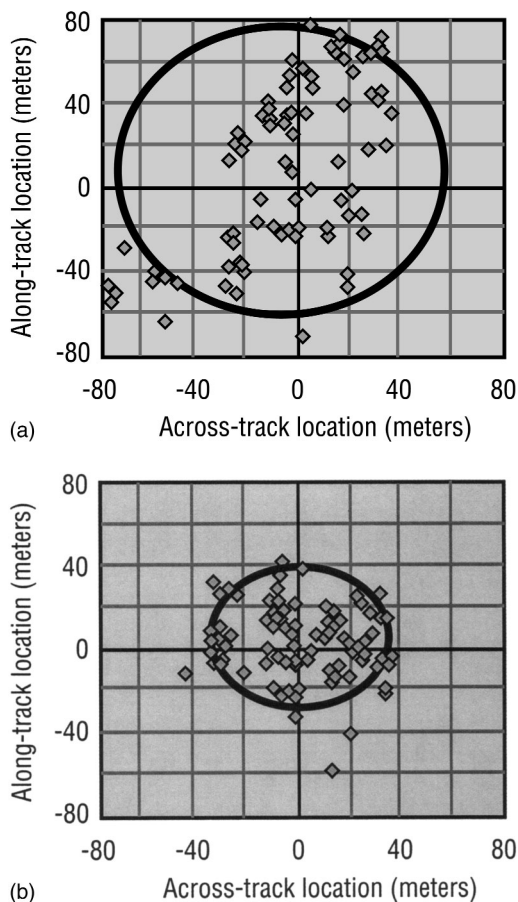


Figure 3. SPOT 5/HRS 2 geolocation performance: (a) before; and (b) after calibration of orbital position.

include improved ancillary data regarding viewing directions, i.e. instead of providing the viewing angles of the first and last detectors, this information is given for each detector and each spectral band. Therefore, the measurement of these viewing directions became a necessity. On the other hand, the monolithic aspect of the SPOT 5 focal plane implied a more regular shape that could be approximated to a low order polynomial.

To achieve these measurements, an “absolute” method was developed. It is based on a “super-site” spreading 60 km east-west and 5 km north-south in Southeastern France, covered with reference airborne digital images having accurate external and internal orientations (0.4 m RMS, 80 percent overlap), as well as a precise DTM. Each airborne image is projected in the SPOT 5 acquisition geometry and automatic 2-D correlation is performed to estimate residual discrepancies between simulated and acquired satellite data

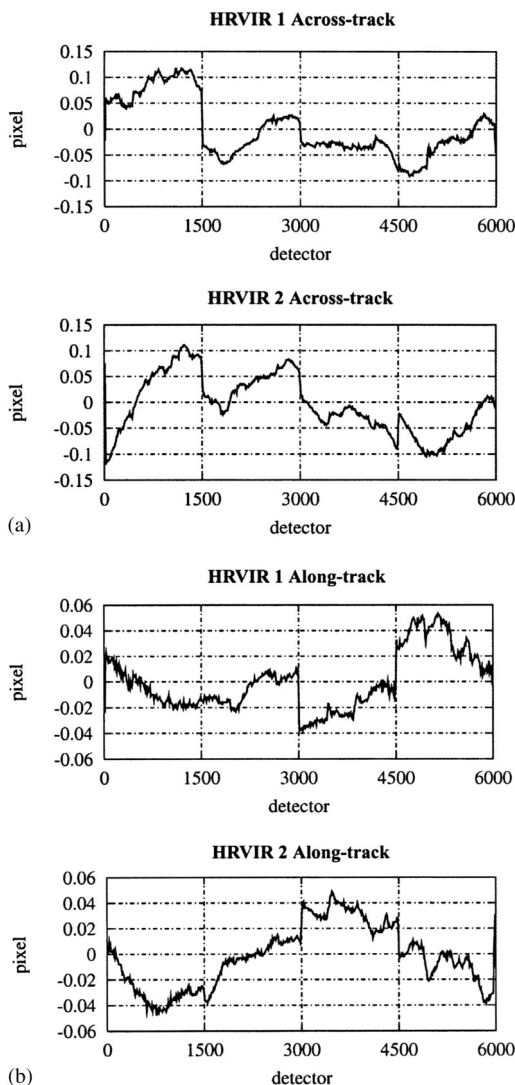


Figure 4. Estimated shape of SPOT 4 HRVIR 1-2, focal planes obtained through “relative” method: (a) across-track; (b) along-track.

both along- and across-track. These discrepancies are further averaged per detector to produce an estimation of the detector’s alignment. Finally, after removal of a bias in both directions, the detector alignments are modeled by up to fifth order polynomials. This method has been applied to each SPOT 5 mode and instrument from HRG’s panchromatic and multi-spectral bands to HRS. Figure 5 shows the results in pixels for both HRG instruments in both modes, across-track.

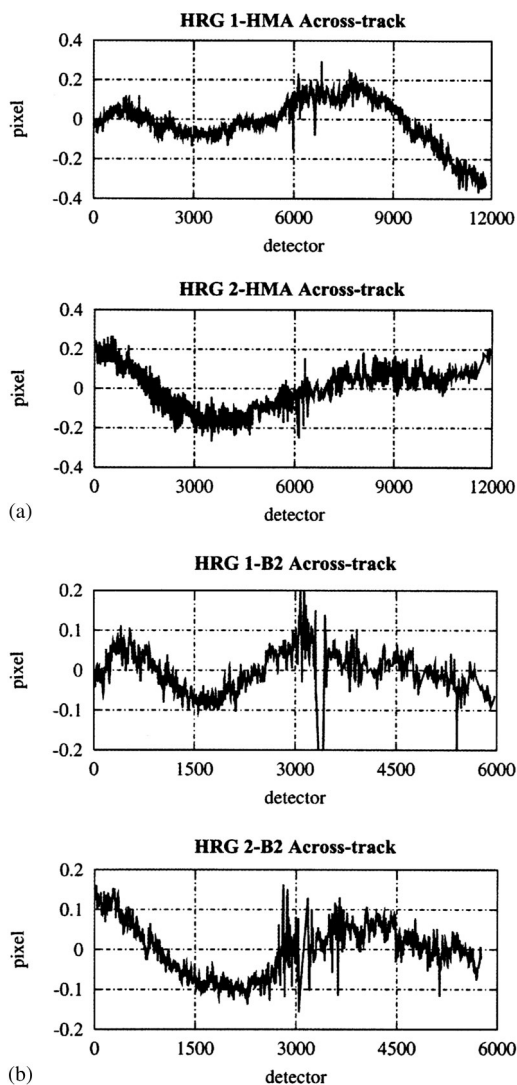


Figure 5. Estimated shape of HRG 1 and 2 focal planes, across-track, obtained through an “absolute” method: (a) HMA; (b) B2.

3.2.2 Registration performances assessment

Since SPOT 1, registration performance between spectral bands has been assessed by acquiring multi-spectral images over desert and other landscapes having as little vegetation coverage as possible. The spectral bands are compared using correlation methods. For SPOT 5, the viewing angles for each detector of each spectral band are obtained relative to the B2 reference band, and registration performance assessment takes into account these viewing angles. Thus, the typical

performance for SPOT 1–4 (0.3 pixel max) is reduced to 0.1 pixel max for SPOT 5.

3.3 Image deformation quality assessment

This § deals with image deformation sources assessment and linked performances. First, methods to assess dynamic perturbation in the images will be presented. This is followed by evaluation of two classical performances directly linked to image deformations: length alteration and planimetric accuracy.

3.3.1 Dynamic perturbation monitoring

An image's ancillary data can be classified into static and dynamic data. Static data are the object of geometric calibration described in §3.1 and §3.2; they depend on each instrument and detection line. Dynamic data are composed of orbit, attitude, and mirror position restitution given by onboard sensors; they depend on the acquisition date.

The objectives are to identify residuals in the images due to dynamic perturbation monitoring phenomena such as (a) satellite micro-vibrations; (b) orbit and attitude restitution quality for different time tags; (c) impact of one instrument's steering mirror displacement while the others are imaging; (d) perturbations due to the magnetic tape recorder (SPOT 1–4); (e) stability and drift of steering mirror; and, (f) steering mirror's stabilization phase.

To assess these factors, specific sets of images are used. Each of them fits with specific dynamic phenomena, including “phased” pairs of images, “coupled” pairs of images, and Autotest images (only available onboard SPOT 5).

3.3.1.1 Assessment of dynamic satellite perturbation

Phased pairs consist of two images acquired with an integer number of SPOT cycles (26 days), over the same area, and with the same observation conditions, i.e. same satellite, same orbit number, same instrument, and same mirror angle. Specific areas are chosen for these phased pairs of images, usually radiometrically stable rocky deserts, offering good correlation results and favorable weather.

Coupled pairs are used to study dynamic satellite perturbations. Indeed, for such a pair, errors due to dynamic ancillary data are not the same, but errors due to static ancillary data are. The superposition performance of the couple thus gives a measure of dynamic errors combined for both images. In case of steering mirror perturbation, the coupled pair is composed of a “reference” image, acquired with no perturbation and a “perturbed” image, acquired during a mirror displacement.

Processing of a phased pair is based on correlation methods. As both images are in close geometry, they

can be compared directly by image matching based on correlation coefficient. The process can take into account a prediction of the shifts based on the physical model of each image. Correlation grids are filtered and synthesised into a mean column to reduce noise due to correlation process and landscape. Each point represents the mean relative filtered shift between the two images for a given image line: this represents the evolution of relative shifts as a function of acquisition time. Spatial and frequential analysis of the mean column allows identification of dynamic phenomena present in the images. Correlation results are also compared to the physical model to assess the quality of dynamic ancillary data. The analysis is divided into two parts: first, bias and linear defaults, i.e. low frequency, which are then removed to identify high frequency residuals (Fig. 6).

The first experiment was conducted on SPOT 1 in 1987 to determine whether image quality of one HRV was affected by the movement of the steering mirror of the other HRV. Conclusions led to improvements in the SPOT 4 imaging system. For the same reasons, an increase in speed of the inertial wheels was successfully tested on SPOT 1 in 1991 (Pausader et al. 1999).

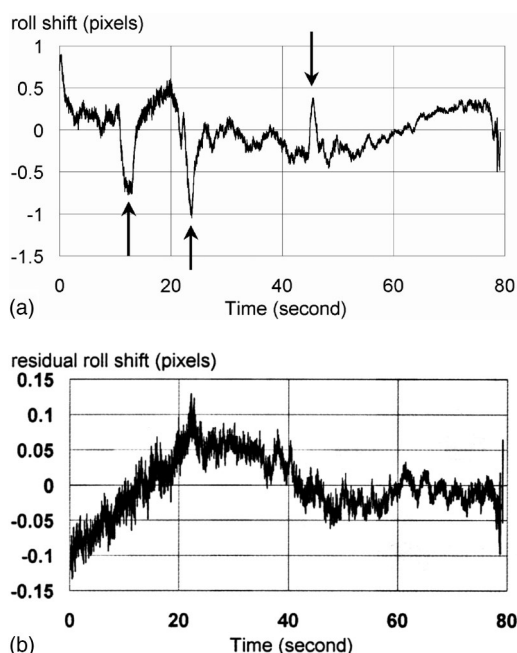


Figure 6. Impact of HRG's mirror rotation on HRS, computed from a phased HRS 1 pair: (a) Across-track roll shift measured by correlating a phased HRS pair (arrows show mirror movements on HRG); (b) Residual roll shift after removing physical models, i.e. no visible impact of mirror movements.

Since SPOT 2, monitoring of dynamic perturbations with these methods has become standard.

3.3.1.2 Assessment of steering mirror perturbations

An image quality (IQ couple) is composed of two simultaneous images acquired on the same satellite with both instruments. IQ couples are used to: (a) study perturbations related to a specific instrument instead of dynamic satellite perturbations (in this case, the two images of the pair share the same satellite dynamic errors but not the same instrument dynamic errors); and, (b) assess mirror stability, drift, and stabilization phase. One image is taken with a “fixed” instrument as a reference, the other with the “mobile” instrument with mirror movement and stabilization phase, as illustrated in Figure 7.

Processing IQ couples is similar to that for phased pairs. An example of such dynamic analysis is shown on Figure 8.

Until SPOT 5, dynamic perturbation monitoring was limited to using ground images subject to problems caused by bad weather and leading to poor correlation results. A new and specific device called Autotest has been installed onboard SPOT 5 that allows acquisition of a test pattern located in each HRG’s focal plane when the steering mirror is in an auto-collimation position. The Autotest image’s characteristic is to have the same pattern reproduced for each image line (see Fig. 9). It is possible to locate finely the pattern in the field by a 1-D correlation. Such analyses have been carried out for SPOT 5 to assess mirror stability, drift, and stabilization phase, and to compare with the previous method.

The advantages of this method with respect to the previous one is its flexibility and processing simplicity. Autotest images can be acquired during night orbits, the 1-D correlation method used is much faster than classic 2-D correlation of two images, and its

success is guaranteed because of the test pattern used. As an example, 1-D correlation of an 800,000 line Autotest image takes 20 minutes (same duration as the acquisition of the image), whereas correlation of a 100,000 line IQ couple every ten points usually takes several hours.

3.3.2 Length distortion assessment

Length distortion performance is the accuracy of distance measures on images using the geometric model. This performance is assessed using the same GCPs on the same images. Differences of distance between

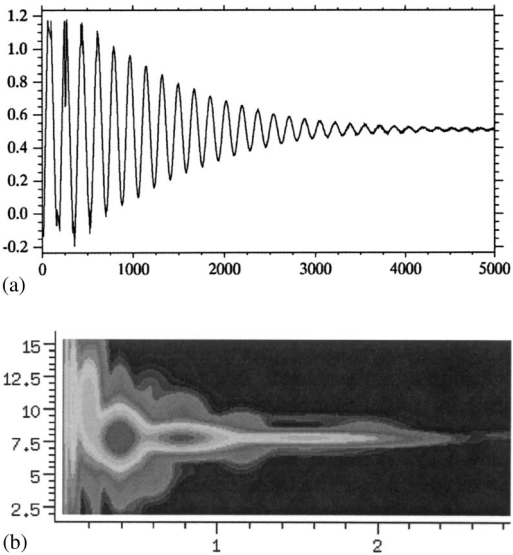


Figure 8. Example of frequency analysis. Mirror’s specific frequency is computed from a coupled THR pair: (a) Result of THR pair correlation during mirror stabilization phase. X-axis is line number; y-axis is the across-track shift (in pixels); (b) Time (x-axis)/frequency (y-axis) analysis.

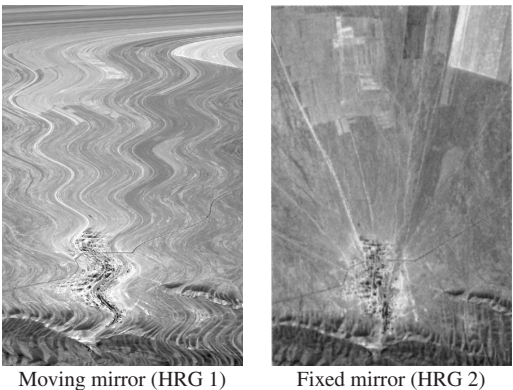


Figure 7. HRG IQ couple during mirror stabilization.

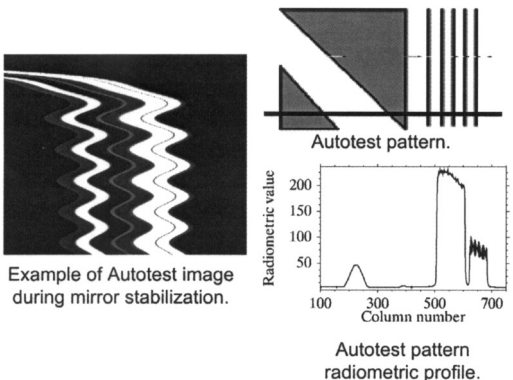


Figure 9. Description of autotest acquisition characteristics.

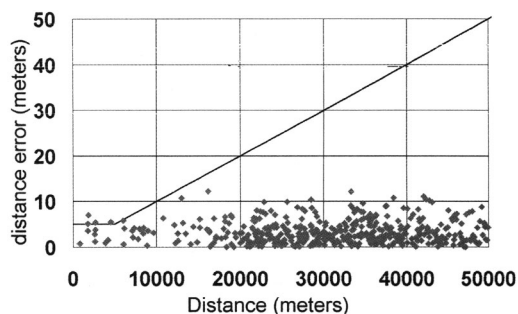


Figure 10. Length distortion evaluation for SPOT 5/HM mode on Johannesburg site during SPOT 5 commissioning phase. Black line represents RMS specification.

two GCPs and the two corresponding points in the image and processed for each pair of GCPs in a given image. The ratio between the distance difference and the measured distance gives the length alteration. A statistical analysis of the measures is realized as a function of orientation and length of each couple of points (Fig. 10).

3.3.3 Planimetric accuracy assessment

This performance quantifies the remaining location errors after correcting system physical geolocation model using perfect GCPs. This performance is the result of high frequency residual distortion in the images. The operational method to estimate it consists of compiling the residuals over GCPs after bundle block adjustment. This method is applied many times over the same sites, which are also used for calibration purposes, leading to several observations of the same GCPs and refining their geographic coordinates, if needed.

Another method is to produce geocoded images using their adjusted geolocation models, and to compare their superposition either to a reference ortho-image, or between themselves (multitemporal superposition). Such a method can only be applied on a site with a precise reference DTM available to be independent from distortion due to relief knowledge. Thus, such an evaluation is performed along with altimetric accuracy assessment.

3.4 Altimetric accuracy assessment

Altimetric accuracy represents stereoscopic restitution capacity of the system. It is based on the exploitation of stereoscopic pairs allowing altitude estimation (usually base-to-height (B/H) ratio > 0.6) with adjustment of the pair's geolocation using GCPs. Pairs are then exploited into a DTM production process. Altimetric accuracy results from combined location errors of each image comprising the stereoscopic

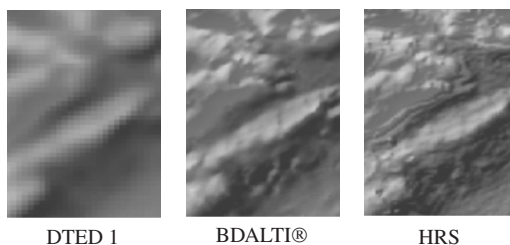


Figure 11. HRS DTM compared with digital terrain elevation data (DTED) 1 and IGN BDALTI® over an area in Corsica.

pair. Thus, performance focuses every potential non-linear error source for each image of the pair, and is linked with planimetric accuracy.

As CNES is not a producer of DTMs, it contracts with value-added specialists to assess not only altimetric accuracy but also operational DTM production capacity (Nonin 2003, Rudowski 2003). Concerning HRS and DTM production, a joint international program was organized with the International Society for Photogrammetry and Remote Sensing (ISPRS), providing data to international experts to generate DTMs over given reference sites (Baudouin et al. 2003). Some encouraging preliminary results are shown in Figure 11.

4 CALIBRATION AND QUALITY ASSESSMENT

SPOT instruments observe the at-sensor radiance from the Earth + Atmosphere system. For each spectral band, the digital count of an image pixel depends on the observed radiance. Because of the image sampling principle, the detector number is used in the linear radiometric model (equation 1):

$$X(k, n, b, m) = R \left[\frac{A(k) \cdot G(m, k) \cdot g(k, n, b) \cdot \gamma(k, b)}{L(k, n, b) + C(k, n, b, m)} \right] \quad (1)$$

where, the raw digital count $X(k, n, b, m)$ of pixels from detector n of read-out register b in spectral band k to input radiance $L(k, n, b)$ when imaging is performed with amplification gain number m . R is the "round" function modeling the analog-to-digital converter (ADC) (Fig. 12).

The parameters of the radiometric model are divided into two families: *normalization* coefficients, i.e. dark currents $C(k, n, b, m)$, and inter-detector $g(k, n, b)$ and read-out register $\gamma(k, b)$ coefficients; and *absolute calibration* coefficients, i.e. programmable amplification gain $G(m, k)$ and absolute calibration coefficient $A(k)$. After estimation of these coefficients, the main radiometric performances are assessed through signal-to-noise ratio measurements.

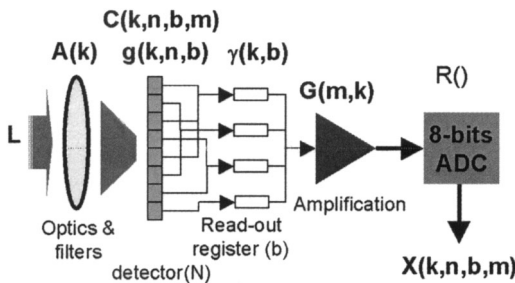


Figure 12. Radiometric model coefficients vs onboard image chain.

4.1 Normalization

4.1.1 Normalization principles

The aim of normalization is to correct images for relative inter-detector sensitivities so that a uniform landscape gives a uniform image. The normalization equation for each pixel is derived from (1). It gives the ratio between the normalized radiometric count $Y(k,n,b,m)$ and the raw radiometric count $X(k,n,b,m)$:

$$Y(k,n,b,m) = \frac{X(k,n,b,m) - C(k,n,b,m)}{g(k,n,b) \gamma(k,b)} \quad (2)$$

Equation (2) can be written according to equation (3):

$$Y(k,n,b,m) = A(k).G(m,k).L(k,n,b) \quad (3)$$

Level 0 SPOT images are 8-bit raw radiometric counts that show inter-detector dispersion. Full dynamics (0–255 last significant bit [LSB]) is tuned in red, green and blue channels to fit image data with B3 (101–131 LSB), B2 (170–201 LSB) and B1 (164–192 LSB) respectively, so that image dynamics are about 30 LSB wide for each spectral band. Images show high frequency inter-detector dispersion (thin stripes), low frequency inter-detector dispersion and inter-readout-register dispersion. Level 1A (or higher level) SPOT images contain 8-bit digitization of normalized radiometric counts, so that residual normalization faults are due to instrument inter-detector sensitivity instabilities. Full dynamics (0–255 LSB) are tuned in red, green, and blue channels to fit image data with B3 (103–122 LSB), B2 (175–192 LSB) and B1 (164–184 LSB) respectively, so that image dynamics are about 20 LSB wide for each spectral band.

4.1.2 Darkness calibration: $C(k,n,b,m)$

4.1.2.1 Method

For the HR, HRVIR, and HRG instruments, dark images are acquired by setting the steering mirror in

the calibration position, which closes the instrument aperture. For HRS instruments, dark images are acquired at night over oceans to avoid interference light. Dark currents are measured by averaging the signal for each detector.

The quality of normalization depends on the stability of the dark currents. Short-term stability (several minutes) is estimated by measuring dark levels over long segments (around 10 minutes). Medium term stability (several days) is estimated by taking one dark acquisition per week for the visible channels. This frequency is reduced to one per month for routine monitoring. Since SPOT 4, an onboard offset control is employed to reduce the dark current instability. Dark current evolution is lower than 0.05 LSB per year for SPOT 4 and SPOT 5. It rises to 0.30 LSB for earlier SPOT satellites.

4.1.2.2 SPOT 5 haute résolution monomode (HM) model

On SPOT 5 HRG, electronic deregistration was applied between two HM channels (HMA and HMB) to induce particular patterns on dark current profiles. The position and amplitude of these patterns depends on the deregistration (Pa and Pb) for both channels, and on onboard image chain amplification gain $G(m,k)$. The combined number of configurations from Pa–Pb (varying from 0 to 15) and “m” (varying from 1 to 10) gives 310 different dark currents for each HMA or HMB retina of each HRG instrument. To avoid in-flight calibration of every configuration, a model was established during ground tests and validated on-orbit to measure only the dark currents for deregistration configurations for a reference amplification gain $m = 0$, and for all amplification gains when both channels are registered (Pa = Pb = 0). Hence, dark currents are measured for only 40 configurations for each HRG instrument and are derived from these with the model for the others. The model is therefore very useful for making calibration less stringent.

Dark currents also are corrected from the digitization effects when the noise on the D/AC input signal is too low (low amplification gain configurations).

4.1.2.3 SWIR dark currents

The SWIR detection chain is different from the visible channels. Detectors are not CCDs but Indium Gallium Arsenide (InGaAs) photodiodes assembled into 10 modules of 300 detectors. The dark levels of the detectors are deducted onboard to encode only the effective part of the signal. The 3000 dark level values are measured and updated once a month during calibration sequences, called dark level acquisition (DLA). SWIR dark currents experience high increases, extending to saturation in certain cases. These events occur randomly, both in time and along the retina, and are consecutive to solar proton collisions (Barde et al. 2000). The mean occurrence is about 3 and 2 detectors

per month for SPOT 4 and SPOT 5 respectively. For these detectors, the corresponding radiometry in the delivered products is computed using a combined spectral and spatial interpolation from all available spectral bands, giving much better results than a simple linear interpolation between regular columns.

4.1.3 Inter-detector coefficients: $g(k,n,b)$ and $\gamma(k,b)$

Determination of inter-detector coefficients requires observation of a uniform input radiance. On-the-ground uniformity is obtained using an integrating sphere. In flight, one can use the calibration lamp unit onboard SPOT 1–4 to detect evolution. However, as the lamp signal may vary in the field-of-view according to the mirror position, and as this device was not kept onboard SPOT 5, the best method used in flight is the acquisition of quasi-uniform landscapes such as snowy expanses of Greenland in summer and Antarctica in winter. The cumulative experience on these landscapes has led to defining accurate targets that provide the best uniformity.

Normalization campaigns are conducted once a year for each satellite. The method consists of making many acquisitions to select cloud-free scenes having low-frequency uniformity. The ratio between acquired and selected scenes is about 10 percent. For each spectral band, five good scenes are needed. Therefore, about 50 acquisitions are required. Thanks to the simultaneous acquisition capacity between the various modes (B1,B2,B3 and SWIR channels in multispectral mode, HMA and HMB channels in THR mode, the HRV/HRVIR/HRG “Image Quality” mode, and the HRS channels in stereo mode), fewer than 200 acquisitions are needed in a campaign for one satellite.

For each scene, the process consists of correcting the influence of solar incidence angle for every pixel to avoid low-frequency variation at the edges of the swath, computing the average line over the most uniform part of the image based on at least 1000 lines, and deducting the dark currents from the average line. Then, for each channel, the best average lines are sorted and filtered according to a likeness criterion that minimizes the low-frequency difference. Finally they are averaged to provide the global average line $I(k,n,b)$ from which the inter-detector and readout-register coefficients are derived.

$$g(k,n,b) = \frac{I(k,n,b)}{\langle I(k,n,b) \rangle^{n \in b}} \quad (4)$$

$$\gamma(k,b) = \frac{\langle I(k,n,b) \rangle^{n \in b}}{\langle I(k,n,b) \rangle} \quad (5)$$

For SPOT 5, readout-register coefficients are cross-calibrated between HMA and HMB channels for each

HRG instrument to enhance normalization of THR products.

4.1.4 Normalization quality assessment

Before delivering the new coefficients to the ground segment, changes since the last campaign are assessed to measure the normalization residuals on the corrected images. Differences between the new and last coefficients are quantified through five criteria (Fig. 13): High Frequency (HF), Low Frequency (LF), Inter-Array (IA), HMA and HMB profile (HMA/B), and even and odd detectors (E/O).

An amplitude greater than 0.30 percent on any of these criteria may cause visible normalization defaults (vertical stripes) on level 1A images until the new coefficients are delivered.

Normalization quality is assessed computing residuals derived from the same criteria. Average lines of raw images (level 0) are normalized by the new coefficients according to equation (2) and then the criteria are computed. The residuals do not include the 8-bit digitization noise (included in level 1A images) and differ on this point from line-wise noise described later.

4.2 Absolute calibration

4.2.1 Overview

Accurate conversion of image digital counts to radiance values has become a challenge for space agencies wanting to compare responses of different sensors, to monitor changes through time, and to provide users with physical data that do not depend on the sensor. Since the launch of SPOT 1, CNES has placed particular emphasis on developing calibration techniques (Begni et al. 1986, Dinguirard & Henry 1993) to improve the in-flight radiometric characterization of sensors. SPOT 1 was one of the first high-resolution satellites to have an onboard Sun sensor for absolute calibration. The experience gained with onboard absolute calibration systems, the difficulty of ensuring their stability on-orbit, and their high cost have led CNES to study calibration techniques over

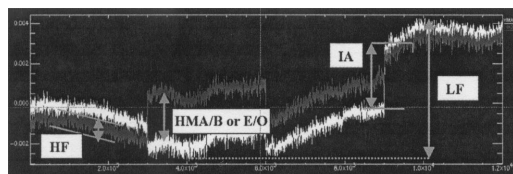


Figure 13. Example of difference profile for SPOT 5 HRG 2 high frequency (HF), low frequency (LF), inter-array (IA), HMA and HMB profile (HMA/B), and even and odd detectors (E/O).

natural targets. Natural targets enable CNES to calibrate a sensor under operational conditions. Following this logic, SPOT 5 was launched with no onboard calibration system. However, such systems give access to calibration data that cannot be provided by ground methods. Both techniques are complementary. Thus SPOT 4, which flies a calibration unit benefiting from SPOT 1–3 experience, and which takes advantage of CNES experience in calibration over ground targets, is the best compromise.

4.2.2 Pre-flight calibration

Before launch, the cameras are calibrated by viewing a large-aperture integrating sphere. This sphere is regularly calibrated and provides a first estimate of the camera sensitivity. This measurement is completed by estimating the spectral sensitivity variation under vacuum conditions. The accuracy of pre-launch calibration is typically 5 percent for the visible channels and 8 percent for the SWIR channel.

4.2.3 Onboard calibration systems

HRV and HRVIR instruments fly a lamp and a Sun sensor.

4.2.3.1 The onboard lamp

An onboard tungsten lamp computes the detectors' relative sensitivity. But the lamp also has shown to be stable on-orbit and is used to monitor instrument sensitivity. Lamp stability is validated by comparing its temporal changes to measurements provided by other methods.

An experiment performed on SPOT 1 at the end of 1991 allowed the acquisition of images from a redundant lamp, which to that time had never been used. This experiment showed that the lamps agreed to within 3 percent. Figure 14 shows typical variations of sensitivity measured with the lamp for SPOT 4 HRVIR. The sensitivity variations are stronger for short wavelength channels because of spectral filter out-gassing and optics yellowing (due to UV irradiation). For the SWIR channels, spectral filter out-gassing increases the spectral sensitivity.

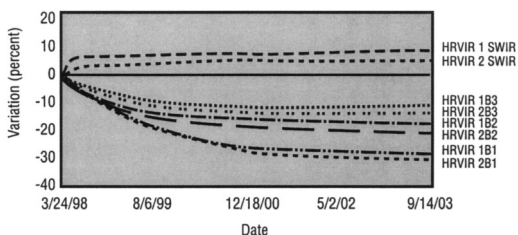


Figure 14. SPOT 4 sensitivity variation measured with the lamp.

4.2.3.2 The Sun sensor

The Sun sensor consists of an optical fiber unit (24 fibers for HRVIR and 48 for HRV) that projects solar irradiance onto some of the detectors of each spectral band. The fibers are grouped and oriented in 3 directions so as to always receive solar illumination when getting out of the eclipse. They have a 12 degrees aperture (± 6 degrees), and 8 degrees separate the 3 directions. In the focal plane, the fibers are set in a staggered line so as to always cover the multi-spectral CCD line arrays regardless of rotation or translation between the calibration unit focal plane and the camera focal plane during the launch. The system is calibrated before launch for bands PA, B1, B2 and SWIR. In orbit, each Sun observation provides a known irradiance for comparison to the measured radiometric counts to compute the absolute calibration coefficient (equation 6):

$$L_k = \frac{u(t) \cdot T_{FIB}(j) \int_0^\infty E_o(\lambda) \cdot T_{BE}(\lambda) \cdot s_k(\lambda) \cdot d\lambda}{\omega_o(j) \int_0^\infty s_k(\lambda) \cdot d\lambda} \quad (6)$$

where, $E_o(\lambda)$ = spectral solar illumination; $u(t)$ = Earth-Sun distance variation; $\omega_o(j)$ = solid angle of fiber j ; T_{FIB} = transmission of fiber j ; $T_{BE}(\lambda)$ = spectral transmission of the calibration unit; and $s_k(\lambda)$ = spectral sensitivity of channel k .

Because of its complexity and the required accuracy, the difficulty in this system lies in characterizing it perfectly before launch (T_{FIB} , $T_{BE}(\lambda)$, $\omega_o(j)$). The fibers also have shown sensitivity to solar irradiation. These two problems made the system unusable on SPOT 1–2. Efforts to better characterize it on SPOT 3 made it operational during the first six months after launch, after which the fibers began to loose sensitivity. Better success came with SPOT 4 whose Sun sensors fly new optical fibers with improved pre-launch characterization. After more than four years, the fibers have shown to be quite stable on orbit (Fig. 15).

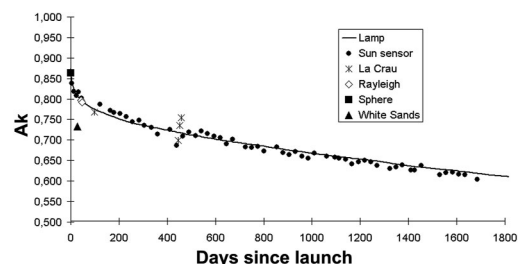


Figure 15. SPOT 4 HRVIR1 B1 calibration measurements.

4.2.4 Rayleigh scattering

The Rayleigh scattering method is valid for short wavelengths (B1, B2). It is one of the most attractive calibration methods since it provides a standard radiance to calibrate sensors in the short wavelength domain without human operation. It is based on the idea that the top-of-atmosphere apparent radiance observed over a clear ocean comes mainly from atmospheric molecular scattering. This scattering is well understood and modeled as a function of pressure and viewing angle. However, the measured signal also comes from atmospheric scattering by aerosols and from reflections over the sea surface (foam, sun glint, in-water scattering). The difficulty of the method lies in evaluating these terms, which contribute highly to the error budget. Viewing conditions must be defined to reduce them. Geographic areas having an *a priori* well-known weak and stable chlorophyll content, with no clouds, a low wind speed, and few aerosols are used. Aerosol optical thickness is estimated using the near infrared channel (B3), which is supposed to be calibrated. A large west-viewing angle increases the atmospheric path and avoids sun glint. Gaseous absorption is corrected using meteorological data (water vapor content, wind speed, and surface pressure) and TOMS data (ozone content). The accuracy of this method is typically 5 percent for SPOT and is improved for wide field-of-view sensors like VGT.

4.2.5 Calibration over deserts

Calibration over deserts consists of cross-calibrating two sensors using site stability and atmosphere characterization to avoid simultaneous viewing. The comparison is made between similar or reciprocal viewing conditions to limit directional effects. The measurements made with the sensor used as a reference are corrected for atmospheric effects and used to fit a spectral model of surface reflectance, which is then integrated in the bands of the sensor being calibrated. These reflectances are finally transferred at the top of the atmosphere to be used for comparison. Although the nominal method uses POLDER as a reference because of its multispectral and multiangular capabilities, the VGT sensor for the SWIR channel on SPOT also is used. Twenty deserts located in North Africa and the Middle East have been selected and are viewed regularly by SPOT satellites. For each SPOT acquisition, similar viewing conditions of the reference sensor are sought to take into account the atmosphere and to correct for spectral sensitivity differences. SPOT measured reflectance can then be compared with an estimated reflectance. Assuming that the reference sensor is well calibrated, one can deduce an absolute calibration. A temporal analysis also provides the sensitivity variation of the instrument as a function of time. This method is nominally used to monitor SPOT 5 sensitivity variation since there is no

onboard lamp. The accuracy of this method applied to SPOT is better than 4 percent.

4.2.6 Vicarious calibration over test sites

This technique consists of characterizing the reflectance and the atmosphere of ground targets simultaneously with the satellite's overpass to estimate the top-of-atmosphere observed radiance. These ground campaigns were performed by the Remote Sensing group at the University of Arizona over the White Sands National Monument in New Mexico from 1986 to 1998, and, by both the Laboratoire d'Optique Atmosphérique (LOA) from Lille and the Laboratoire Interdisciplinaire en Sciences de l'Environnement (LISE) from Wimereux over La Crau (France) from 1989 to 1999. These ground campaigns were replaced in 2000 by an automatic radiometer ground based station over La Crau test site. A CIMEL radiometer, mounted on top of a 10 m high mast, provides continuous ground and atmosphere characterization to calibrate any sensor every time it observes the site. The data are transmitted to CNES by a phone link and then processed. This autonomous station is now fully operational and the processing software is being commercialized.

4.2.7 HRV 1, 2; HRVIR 1, 2; and HRG 1, 2 cross-calibration

Only one instrument on the same satellite is calibrated for each spectral band. The other one is cross-calibrated with regard to the first one using simultaneous viewing. This cross-calibration is very accurate (better than 2 percent), and varying the spectral responses of the viewed landscapes eliminates the slight spectral sensitivity difference between the two instruments.

4.2.8 HRVIR and HRG cross-calibration with VGT 1, 2

Nearly simultaneous images taken by HRVIR or HRG and VGT (1 or 2) cameras on the same platform provide an opportunity to cross-calibrate them and to ensure the radiometric homogeneity of SPOT and VGT data. Accurate registration of the two images is made easier by their common geometric reference. A first "coarse" registration based on the two sensors' geolocation model is refined using a correlation technique.

4.2.9 Synthesis of calibration coefficients

The different techniques used to calibrate SPOT cameras provide a set of measurements. The absolute calibration coefficient at a given time is estimated by matching a curve of sensitivity variation, either given by the lamp for SPOT 1–4 or over deserts for SPOT 5. By using different methods, systematic errors that each of them can induce are avoided. The discrepancies

between them give an estimate of calibration performance: typically 6 percent for visible channels and 7.5 percent for the SWIR channel.

4.2.10 Calibration and use of electronic gains

A good calibration also is essential for the automatic procedure to choose an onboard electronic gain ($G(m,k)$) that fits the viewed landscape. Thanks to work carried out to ensure good calibration of the cameras during their orbital lifetimes, a large amount of credible information is now available. In 1987 (one year after SPOT 1 launch), CNES decided to store the histograms of all cloud free images acquired and to organize them into a database. Built by SPOT 1–4 (and now 5), it gives a significant statistical estimate of observed radiances on a worldwide basis along with their monthly variations, at least in the five SPOT spectral bands. Estimation of the maximum observed radiance over a grid element makes it possible to determine the viewing gain that optimizes image dynamics and avoids image saturation. Thus, the IQEC regularly provides the Programming Center with an updated viewing gain file extracted from the SPOT histogram database. This file gives a monthly and projected annual estimate of the viewing gain for each spectral band of each satellite and for each grid element of a world map. The elementary period is one month so as to take seasonal variations of reflectance into account. The grid-sampling interval is 120 km. This grid size is a good compromise between the spatial resolution of the information extracted from SPOT images (for which the nadir field-of-view is 60 km), and the reliability of the statistics obtained over the grid. For areas with no information, spectral, temporal, or spatial reflectance interpolations are performed. The improvement of SPOT 5 resolution with regard to SPOT 1–4 is taken into account through margins on the reflectances.

4.3 Signal-to-noise ratio (SNR) assessment

The radiometric quality of corrected images is quantified by several SNR measurements. Digital counts observed on images over uniform landscapes are not constant. Apart from the onboard 8-bit digitization that induces maximum fluctuation of ± 0.5 LSB, noise is caused by two separate phenomena: column-wise noise (σ_c) and line-wise noise (σ_l). For each image, these two types of noise are combined (quadratic sum) in an “image noise” (σ_i) that quantifies the variations of digital counts on a uniform landscape:

$$\sigma_i^2 = \sigma_c^2 + \sigma_l^2 \quad (7)$$

4.3.1 Column-wise noise

For SPOT 1–4, column-wise noise is calculated directly on the calibration sequences obtained with

the lamp unit. Although not uniform, the spatial profile of the lamp is stable throughout imaging so that each detector sees a uniform landscape. The images are divided into 50×50 pixel chips. For each column of every chip, the standard deviation of the digital counts is computed. The column-wise noise is the root mean square of the different standard deviation.

For SPOT 5, which has no lamp unit, images acquired over the quasi-uniform landscapes of Antarctica and Greenland are used. However, uniformity of these landscapes is not sufficient to measure column-wise noise accurately. Another experimental method has been used during SPOT 3 in-flight commissioning. It consists of using two images I_1 and I_2 of the same landscape, acquired simultaneously to eliminate the landscape contribution. The two images can be obtained with any of three combinations: the two HRG instruments in IQ mode for multispectral images, the HMA and HMB images for one HRG instrument, or the two HRS instruments.

Differences between images I_1 and I_2 give a new noise image (D) without landscape contribution. The instrument noises noted B_1 and B_2 for images I_1 and I_2 , respectively, are assumed to be separate and landscape independent. For each triplet I_1, I_2, D , the standard deviation of noise noted $\sigma_{I1}, \sigma_{I2}, \sigma_D$ can be calculated. Then, the resolution of the following three equation system (equation 8) gives the instrument standard deviation noises noted σ_{B1} and σ_{B2} , and the standard deviation of the landscape contribution noted σ_L :

$$\begin{aligned} \sigma_{I1}^2 &= \sigma_L^2 + \sigma_{B1}^2 \\ \sigma_{I2}^2 &= \sigma_L^2 + \sigma_{B2}^2 \\ \sigma_D^2 &= \sigma_{B1}^2 + \sigma_{B2}^2 \end{aligned} \quad (8)$$

These equations are valid for column-wise and image noises. The two images must be registered with a precision better than 0.1 pixel. Image 2 is re-sampled for each pixel in image 1 geometry. The line and column shift of the re-sampling grid are computed by the images' correlation or by application of the images' geometric models. In the end, the noises calculated in the re-sampled image are corrected from the re-sampled effects. Indeed, re-sampling reduces the noise by a factor that depends on the re-sampling filter (cubic splines) and line or column shift.

4.3.2 Line-wise noise

Normalization defects (radiometric model deviations) may cause visible “columns” on a uniform landscape. This “line-wise” noise is estimated to be less than 0.1 percent of signal for SPOT 5 and 0.3 percent for the other SPOTs. For all the instruments,

line-wise noise can be determined as a statistical criterion computed on average lines of normalized images (level 1A). The lines are divided in 50 pixel intervals on which standard deviation is computed. The root mean square (RMS) of the standard deviations gives the line-wise noise. For SPOT 5, line-wise noise can be deduced also from column-wise and image noises thanks to equation 7.

4.3.3 Column-wise noise model

Column-wise noise represents onboard image chain performances and is independent of normalization operations. It can be considered as instrument noise.

4.3.3.1 Instrument noise

Column-wise noise may be divided into three main contributions: signal noise caused by the Poisson fluctuation of the signal delivered by the detector, darkness noise caused by various constant chain noises like electronic noise and detector thermal noise, and 8-bit digitization noise. A model can be given to express the standard deviation of each contributor in radiance units ($W/m^2/sr/\mu m$). Note that the factor conversion between digital count unit and radiance unit is $A(k) \cdot G(m,k)$ (from equation 3). The standard deviation of signal noise (σ_s) is proportional to the signal square root (Poisson distribution):

$$\sigma_s^2 = KL \quad (9)$$

where, L represents the signal in units of radiance.

For all instruments, darkness noise standard deviation is linear versus the programmable onboard image chain gain $G(m,k)$ when expressed in digital count unit. In radiance unit, it gives:

$$\sigma_{DARK} = \frac{\alpha}{A(k)} + \frac{\beta}{A(k) \cdot G(m,k)} \quad (10)$$

The coefficients α and β are estimated on the darkness calibration sequences.

Eight-bit digitization induces a maximum fluctuation of ± 0.5 LSB so that the standard deviation of digitization noise (σ_q) expressed in radiance unit is equal to:

$$\sigma_q = \frac{1}{A(k) \cdot G(m,k) \cdot \sqrt{12}} \quad (11)$$

Finally, the model of column-wise noise is the quadratic sum of the contributors:

$$\sigma_C = \sqrt{\sigma_{DARK}^2 + \sigma_q^2 + \sigma_s^2} \quad (12)$$

For each detector and each spectral band, the column-wise noise model (standard deviation σ_C) depends

on the programmable onboard gain $G(m,k)$ and the observed radiance L . For a given detector, amplification gain, and spectral band, the column-wise noise model can be written:

$$\sigma_C = \sqrt{a + KL} \quad (13)$$

and the only unknown parameter remains K .

4.3.3.2 Compression noise

The new image compression algorithm, discrete cosine transform (DCT) implemented on SPOT 5 is used to increase the compression ratio significantly (from 1.33 for SPOT 1–4 to 2.81 for SPOT 5) while complying with SPOT user image quality requirements (Lier et al. 1998). With fixed rate telemetry (100 Mb/s/sec), a variable rate linked to landscape complexity is optimised to eliminate residual block effects in uniform areas. Nevertheless, a compression noise model has to be added to the instrument noises described above to assess system image noise. This model depends on the digitization factor (F), which in turn depends on landscape complexity (Lier 1998):

$$\sigma_{comp} = \frac{F}{4\sqrt{12}} \quad (14)$$

4.3.4 Common reference configuration for SNR

Once the models have been established, one can deduce the SNR for reference values of radiance $L_{ref,k}$ and amplification gain $G(m_{ref},k)$. The reference radiance is representative of the typical landscape acquired by the SPOT family and corresponds to the same target, i.e. where reflectance = 0.5, solar zenith angle = 60 degrees, standard atmosphere visibility = 23 km, $O_3 = 0.35$ cm.atm, and $H_2O = 2$ g.cm $^{-2}$ as seen by the different SPOT instruments.

The reference amplification gain is the lowest operational gain for SPOT 1–3 and the neutral gain ($G(m,k) = 1$) for SPOT 4–5, which is consistent with the gain that would be used in operational conditions to observe such a radiance. The example in Figure 16 presents results for SPOT 4 HRVIR2 in M mode (B2 spectral band) obtained in May 2003. The lamp measures are the dots, the models fitted on the measures are the curves. The column-wise noise ($\sigma_C = 0.64$ W/ $m^2/sr/\mu m$) is deduced for reference radiance ($L_{ref,B2} = 102$ W/ $m^2/sr/\mu m$) and amplification gain $G(2,B2) = 1$. The SNR is 159 for this configuration.

This approach enables a comparison of every SPOT instrument's radiometric quality with a common reference. This is what is published in the Spotimage-edited "SPOT Image Quality Performances" document, which provides SPOTIMAGE customers with the latest estimations of image quality in terms of geometry, radiometry and resolution, for all operational satellites.

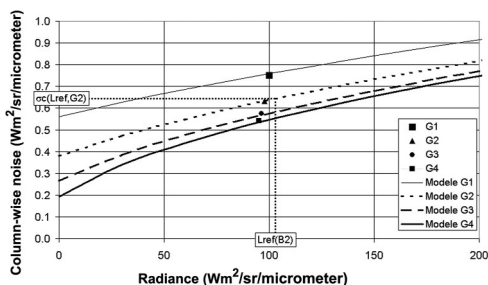


Figure 16. Column-wise measures and model for SPOT 4 HRVIR 2/M, May 2003.

5 CONCLUSIONS

Monitoring image quality of five satellites implies constant effort and involvement of many people. Furthermore, a dedicated and independent sub-system, the IQEC, has been designed to manage these tasks operationally during the lifetime of the five satellites. These two factors are the major contributors in the effectiveness of SPOT on-orbit image quality activities. Needed improvements in calibration methods are linked in general to a gain in accuracy. The trend is to simplify instrument design (no calibration device and no onboard registration), which implies increasing complexity of both image processing methods and calibration methods. For instance, for Pléiades (SPOT high-resolution follow-on), the methods currently being studied are:

1. For geometric image quality: automatic matching to identify GCPs, elaboration of new test sites with high-resolution GCPs, and improvement of attitude restitution using image correlation,
2. For radiometric image quality: normalization of non-linear detectors, use of specific steering modes, and use of artificial neural networks for noise measurements.

Concerning test sites used for geometry, radiometry or resolution, SPOT is interested in sharing and exchanging reference data with other entities to facilitate intercomparisons of methods and results within the ISPRS framework.

ACKNOWLEDGMENTS

The authors wish to thank their predecessors whose involvement in successive in-flight calibration activities has allowed the authors to achieve this success-story. To name a few:

- SPOT 1: G. Begni, B. Boissin, M. Leroy, D. Pradines, M. Dinguirard, D. Léger, V. Rodriguez, J.P. Darteyre, P. Gigord.

- SPOT 2: P. Henry, M. Dinguirard, X. Briottet, D. Léger, A.C. De Gaujac, P. Gigord.
- SPOT 3: O. Hagolle, A. Meygret, M. Dinguirard, D. Léger, F. Viallefont, A.C. De Gaujac, R. Gachet.
- SPOT 4: Ph. Kubik, A. Meygret, E. Breton, F. Masson, M. Pausader, D. Léger, L. Poutier.
- SPOT 5: A. Meygret, C. Latry, L. Lebègue, V. Pascal, F. Cabot, D. Léger, F. Viallefont, E. Breton, A. Bouillon, R. Gachet, F. De Lussy.

REFERENCES

- Barde, S., Ecoffet, R., Costeraste, J., Meygret, A. & Hugon, X. 2000. Displacement damage effects in InGaAs detectors: experimental results and semi-empirical model prediction. *IEEE Trans. Nuc. Sci. & Rad. Eff. on Components*. 47(6): 2466–2472.
- Baudoin, A., Schroeder, M., Valorge, C., Bernard, M. & Rudowski, V. 2003. The HRS-SAP initiative; a scientific assessment of the High Resolution Stereoscopic instrument of SPOT 5 by ISPRS investigators. *High resolution mapping from space 2003, Hannover, 6-8 October 2003*. Hannover: Univ. of Hannover, Inst. Photogr. & GeoInfo.
- Begni, G., Léger, D. & Dinguirard, M. 1984. An in-flight refocusing method for the SPOT HRV cameras. *Photogr. Eng. & Rem. Sens.* 50(12): 1697–1705.
- Begni, G., Dinguirard, M., Jackson, R.D. & Slater, P.N. 1986. Absolute calibration of the SPOT 1 HRV cameras. In P.N. Slater (ed.), *Earth remote sensing using the Landsat TM and SPOT sensor system*. 660: 93–97.
- Dinguirard, M. & Henry, P. 1993. Calibration of the SPOT HRV cameras. *Workshop on remote sensing of soil and vegetation; Tempe, AZ 6-8 January 1993*.
- Gachet, R. 1999. Caractérisation des directions de visée. *Bulletin SFPT* 159: 73–79.
- Léger, D., Viallefont, F. & Déliot, P. 2004. On-orbit MTF assessment of satellite cameras. In S.A. Morain & A.M. Budge (eds), *Post-launch Calibration of Satellite Sensors*. Rotterdam: Balkema.
- Lier, Ph., Moury, G., Latry, C. & Cabot, F. 1998. Selection of the SPOT 5 image compression algorithm. In W.L. Barnes (ed.), *Earth observing systems III*. Proc SPIE 3439: 541–552.
- Meygret, A. & Léger, D. 1996. In-flight refocusing of the SPOT 1 HRV cameras. In A.E. Iverson (ed.), *Algorithms for multispectral and hyperspectral imagery II*. Proc. SPIE 2758: 298–307.
- Nonin, P. & Piccard, S. 2003. Performance analysis of DEM automatic extraction from SPOT 5 sensors. *IGARSS 2003; Toulouse, 21–25 July 2003*.
- Pausader, M., De Gaujac, A.C., Gigord, P. & Charneau, M.C. 1999. Mesure des micro-vibrations sur des images SPOT. *Bulletin SFPT* 159: 80–89.
- Rudowski, V. & Bouillon, A. 2003. Altimetric restitution assessment of SPOT 5 stereo pairs. *IGARSS 2003; Toulouse, 21–25 July 2003*.

The Joint Agency Commercial Imagery Evaluation (JACIE) team and product characterization approach

V. Zanoni

National Aeronautics and Space Administration, John C. Stennis Space Center, Mississippi, USA

M. Pagnutti & R. Ryan

Lockheed Martin Space Operations – Stennis Programs, John C. Stennis Space Center, Mississippi, USA

D. Helder

South Dakota State University, Brookings, South Dakota, USA

W. Lehman & S. Roylance

Booz Allen Hamilton, Reston, Virginia, USA

G. Snyder

U.S. Geological Survey, Reston, Virginia, USA

ABSTRACT: The Joint Agency Commercial Imagery Evaluation (JACIE) team is an interagency group focused on the characterization of commercial remote sensing data products. The team members include the National Aeronautics and Space Administration (NASA), the National Geospatial-Intelligence Agency (NGA), and the U.S. Geological Survey (USGS). JACIE characterizes delivered image products, focusing on three primary areas: geopositional accuracy, image quality, and radiometric accuracy. For geopositional accuracy, the team utilizes sites containing several “image-identifiable” targets and compares their known locations with those defined by the image product. For image quality, spatial response and modulation transfer function (MTF) are characterized using edge targets and pulse targets. Imagery is characterized also through the National imagery interpretability rating scale (NIIRS). Radiometric accuracy is characterized using reflectance-based vicarious calibration methods at several uniform sites. The JACIE team performed characterization of IKONOS and QuickBird products, and will evaluate OrbView-3 products.

1 INTRODUCTION

In recent years, U.S. Government agencies have begun purchasing commercial remote sensing data to support Government research and operational objectives. The National Geospatial-Intelligence Agency (NGA) purchased remote sensing products from, among others, Space Imaging, Inc.; SPOT Image Corporation; Digital Globe; and Intermap Technologies Corporation to support Department of Defense (DoD) customers. The National Aeronautics and Space Administration (NASA) purchased commercial remote sensing products from Space Imaging, DigitalGlobe, and others to support Earth science research. The U.S. Geological Survey (USGS) has experience in validating and applying new sources of airborne and spaceborne data,

including commercial data, to support science and mapping applications.

Historically, the U.S. Government has had extensive insight and oversight in developing, launching, operating, and characterizing space-based remote sensing systems. In many cases, the scientific users of the resulting data are intimately involved with detailed system design and calibration. For commercial remote sensing systems and products, however, this paradigm changes. Government science and applications objectives often require highly characterized sensors and products. However, commercial requirements for calibration and characterization may not be as stringent as government users desire. Since government-built satellite systems cannot compete with commercial industry (National Research Council 2002),

commercial providers must fill government requirements for high-resolution data. Such procurement from the commercial sector requires independent data characterization.

Recognizing the importance of characterizing commercial remote sensing data for government use, NASA, NGA, and the USGS formed the Joint Agency Commercial Imagery Evaluation (JACIE) team in February 2000 to capitalize on mutual interests and to leverage resources for characterizing commercial remote sensing data. The three agencies formalized this partnership through a JACIE Interagency Space Act Agreement signed in June 2002.

2 APPROACH

Since commercial systems are built and operated with no government insight or oversight, the JACIE team provides independent product characterization of delivered image and image-derived products. End product characterization differs from the systems calibration approach that is typically used with government systems, where detailed system design information is available. Although the focus is on on-orbit characterization, some important information required to characterize system performance fully can be acquired only in a pre-launch laboratory setting. The JACIE team interacts with individual vendors through technical interchange meetings to understand vendors' pre-launch and post-launch calibration methods and results. Results of vendor pre-launch and early orbit measurements for parameters such as spectral filter response, linearity, dynamic range, polarization, spatial resolution, geometric calibration, and bad pixel maps are reviewed and discussed with the vendor. When the commercial system is on-orbit and operational, JACIE team members characterize imagery and derived products in terms of spatial resolution, geospatial accuracy, and radiometric accuracy.

Each JACIE agency provides unique expertise and capabilities for characterizing commercial data products. The USGS, with heritage in mapping and cartography, conducts geospatial accuracy assessments that include evaluations of georectified, orthorectified, digital elevation model (DEM), and stereo image products. NGA and its customers perform DoD and intelligence-based product evaluations, including image interpretability, feature extractions, photogrammetry assessments, and geospatial assessments. NASA performs overall product performance characterization by evaluating commercial remote sensing products' spatial resolution, geospatial accuracy, and radiometric accuracy. NASA has partnered also with experts from the Earth science community to support system characterization. These experts from South

Dakota State University, the University of Arizona, and the University of Maryland have years of experience in characterizing and validating NASA-developed sensors.

The JACIE characterization approach relies on multiple experts using different methods and test sites to achieve a thorough assessment of commercial high-resolution image products.

2.1 Geospatial accuracy

Geospatial accuracy is evaluated using sites containing many features that are identifiable in the image products under evaluation and whose locations are known within an accuracy better than that of the image product. Identifiable features typically consist of road intersections, permanent structures, or geodetic targets (Helder et al. 2004, Dial & Grodecki 2004 this volume). Exact latitude and longitude locations of such features are measured to sub-meter accuracy with global positioning system (GPS) receivers for the products evaluated by the team. The features or targets used must be distributed evenly throughout the image and must be of a sufficient number to perform valid statistical analysis. The Federal Geographic Data Committee (FGDC) standards for the number and distribution of points used for geospatial accuracy evaluation states that a minimum of 20 points distributed across the image footprint should be used, that the points should be separated by 10 percent of the diagonal distance across the image, and that at least 20 percent of the points should fall in each quadrant (FGDC 1998). For non-orthorectified products, changes in terrain elevations within the image footprint can affect the image-defined position of the points. Therefore, terrain displacements must be taken into account when assessing geospatial accuracy of non-orthorectified products.

The true locations of identifiable features are compared with the locations of the features as defined by the image product. The differences in the X-directions (Easting) and Y-directions (Northing) are determined for each point, and the root mean square error (RMSE) in both directions is calculated for all points in the image using the following equations (Greenwalt & Shultz 1968):

$$RMSE_x = \sqrt{\frac{\sum (x_{img} - x_{cont})^2}{n}} \quad (1)$$

$$RMSE_y = \sqrt{\frac{\sum (y_{img} - y_{cont})^2}{n}} \quad (2)$$

where x_{img} and y_{img} = X and Y positions as defined in the image product, x_{cont} and y_{cont} = known X and Y positions, and n = number of points used in the analysis.

The net, or radial, RMSE (*RMSEr*) is then determined for the image product under evaluation using

$$RMSEr = \sqrt{(RMSE_x)^2 + (RMSE_y)^2} \quad (3)$$

Commercial remote sensing imagery geospatial accuracy is often specified in terms of circular error at a 90 percent confidence level (CE90). To calculate CE90, team members again follow FGDC methods (FGDC 1998) using the following equation:

$$CE90 = 1.5175 \bullet RMSEr \quad (4)$$

In addition to the CE90 accuracy value, team members also report geospatial accuracy according to FGDC standards, which require a circular error at 95 percent confidence level (CE95) (FGDC 1998). This value is calculated using the following equation:

$$CE95 = 1.7308 \bullet RMSEr \quad (5)$$

Equations (4) and (5) assume that the errors in the X and Y directions represent a bivariate normal distribution and that the X and Y errors are equal, independent, and zero-mean, with no systematic bias. If a bias exists, the above approach can be modified to account for known biases (Greenwalt & Shultz 1968). Alternatively, the bias can be removed using the methods defined by Mikhail and Gracie (1997) so that only random error is considered.

An alternate and possibly more intuitive method for determining CE90 involves plotting the RMSEr values on a radial plot. The radius of the circle that includes 90 percent of the points in the plot determines the CE90 accuracy value.

JACIE team members employ these approaches to evaluate georectified and orthorectified products having various mapping accuracies. Differences in geospatial accuracy resulting from different geometric resampling algorithms have also been explored.

Additionally, derived products such as stereo imagery and DEMs have been evaluated also by team members (Constance 2002). Vertical accuracy of DEMs is evaluated by comparing the actual vertical location of an image-identifiable feature to the location of that feature as derived from the DEM, similar to the method for horizontal accuracy. Horizontal and vertical accuracy of stereo images are also evaluated by comparison against actual feature locations. JACIE members have evaluated the accuracy of multiple stereo images through block adjustment techniques (NASA et al. 2002).

2.2 Spatial response and image quality

Spatial resolution is often discussed in terms of a system's ground sample distance (GSD). Although GSD

is an important factor, image quality and the effective resolution of a system or data product are driven by several parameters, including GSD, point spread function (PSF), and signal-to-noise ratio (SNR). The PSF represents the amount of blur caused by the imaging system when viewing an ideal point source. The Fourier transform (FT) of the PSF is known as the optical transfer function (OTF). This function measures the response of the system to sinusoidal inputs at all spatial frequencies. A more commonly used metric is the modulation transfer function (MTF) which is simply the modulus of the OTF normalized such that the first value is unity. (Hecht 1998). In principle, the MTF is a measure of a system's ability to reproduce accurately the content of a ground scene within the image. However, the relative edge response (RER) is a more intuitive parameter that is easier to measure. The RER of an image is a measure of the system's ability to clearly distinguish a straight, high-contrast edge. The general image quality equation (GIQE) (Leachtenauer et al. 1997) provides a mathematical relationship between GSD, RER, image enhancement, and SNR to estimate image quality objectively. The GIQE also includes ringing and edge sharpness effects associated with an edge response. One commonly accepted form of the GIQE expresses this relationship in terms of the National imagery interpretability rating scale (NIIRS) (Ryan et al. 2003). The NIIRS is a 10-level rating scale that defines one's ability to identify certain features or targets within an image (IRARS 1996).

To evaluate RER, an ideal edge target is imaged by the sensor to obtain the edge response of the system. The "slope" of the edge is a measure of the amount of blur in the resulting image and is used by team members to measure RER. These targets may be "permanent" as in the case of large painted concrete areas, or may be deployable as in the case of tarps commonly used by the team. Target size is critical to ensure that enough pixel samples are available for analysis. The JACIE typically uses 20 m edge targets for evaluation of 1 m GSD or finer resolution imagery. Tilted edge techniques are employed that require the edge targets to be placed slightly off-axis relative to the image pixel grid. These techniques minimize aliasing and ensure that the edges are properly sampled by the image pixels (Reichenbach et al. 1991).

In the edge target method, an analysis area is selected from the image of the edge target. Each line of pixels across the edge is scaled by the cosine of the tilt angle to provide a single edge response. The individual edge responses are fitted to a smooth analytical function. A superposition of three sigmoidal functions is used to minimize effects of noise and limited sampling (Tzannes & Mooney 1995). The slope and amount of overshoot of the RER can then be measured. Additionally, numerical differentiation of the edge response produces the line spread function (LSF). The full-width

half-maximum (FWHM) value of the LSF provides an estimate of edge sharpness that is often compared to the system's GSD. Example edge responses and LSFs for panchromatic imagery of tarp targets are shown in Figure 1. A FT and normalization of the LSF will yield the MTF. The value of the MTF at the Nyquist frequency is often the parameter used to specify system spatial performance. RER and MTF analyses are performed for both the across-track and the along-track directions of image products.

For larger GSD products, it is difficult to find edge targets large enough to estimate edge response. Thus, for the commercial multispectral products evaluated by JACIE, the pulse target method provides a better estimate of image quality (Ryan et al. 2003). In the pulse method, long "strip" or pulse targets are employed.

An ideal pulse input (the pulse tarp target) is imaged by the sensor and the sharpness of the resulting image is evaluated. From the image of the pulse target, exact

edge locations are determined on a line-by-line basis using model-based parametric methods that detect sub-pixel edge locations. The sub-pixel locations are then fitted to a mathematical function to define the edges of the pulse output response (Helder & Choi 2003). Both the input pulse and the output are Fourier transformed. The MTF is then calculated by dividing each output value by its corresponding input and by normalizing the result as described previously (Ryan et al. 2003). Figure 2 illustrates input and output pulse responses, their FTs and the resulting MTF estimate.

The pulse method balances the need for smaller targets against SNR considerations for determining the MTF. A target that is too narrow will create insufficient contrast for estimating the transfer function. If the target is too wide, the zeros of the input sinc function will lead to infinities in the resulting MTF.

Team members utilize a variable width, 60 m long tarp target that is comprised of multiple 3 m by 30 m tarp strips. The particular target used is suitable for estimating image quality in the blue and near infrared (NIR) bands of 3–4 m GSD multispectral imagery. Again, tilted edge techniques are employed when deploying the pulse target to ensure proper sampling (Fig. 3).

Analyses have also considered effects of MTF compensation (MTFC), a post-processing image restoration technique that effectively enhances image sharpness but increases image noise. Images processed with and without MTFC applied were analyzed. MTFC was found to be beneficial when image interpretation or feature recognition are of primary importance. When performing radiometric analysis, MTFC will affect image digital number (DN) values for certain features; when working with large, uniform areas, MTFC does not appear to have adverse effects (Zanoni et al. 2003, Pagnutti et al. 2003, Ryan 2001, 2002).

Team members also evaluated image interpretability to support military and civilian applications. The NIIRS and Essential Elements of Information (EEIs) are used by NGA to assess imagery. EEIs are certain features and targets, e.g. railcars, aircraft, that correspond to

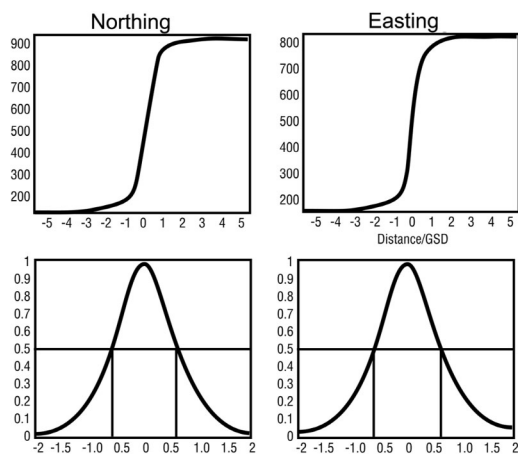


Figure 1. Edge responses and fitted sigmoidal functions for IKONOS panchromatic images (MTFC-off) (top row). LSF from fitted edge responses for IKONOS panchromatic images (MTFC-off) (bottom row).

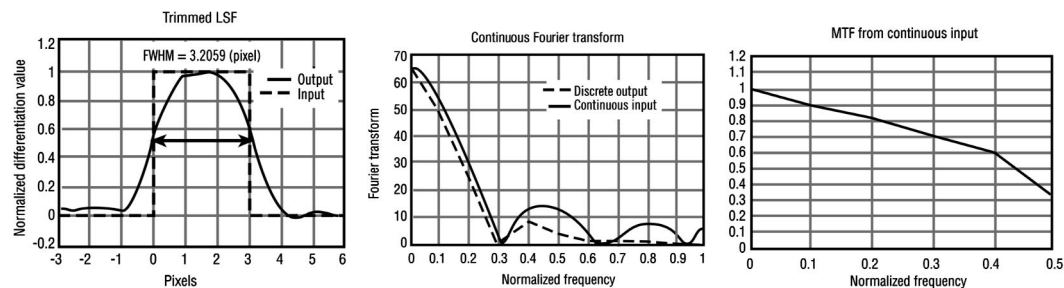


Figure 2. Plots illustrating the pulse method for spatial response. Examples of pulse input/output response (left), FTs of input and output responses (middle) and the resulting modulation transfer function (right).

the various NIIRS levels. Several image chips are extracted from images acquired over a given time period and over multiple locations. The image chips are evaluated by a group of NGA-certified image analysts. The analysts each evaluate the same set of images under the same conditions, i.e. using the same computer and amount of image magnification with no additional image processing or enhancement. The analysts assign NIIRS ratings and confidence ratings associated with identification of EEs. Statistical analyses are performed on analysts' results to understand the consistency and reliability of the different analysts and to identify any outlier image chips used in the assessment. Good correlation among the analysts provides confidence in the average NIIRS and EEs generated (Ryan et al. 2003).

2.3 Radiometric accuracy

Radiometric accuracy characterization is performed on delivered image products. Radiometric assessments do not include methods that are dependent on special satellite operations or maneuvers, such as use of onboard calibration sources, or star or moon calibrations. The team employs a reflectance-based vicarious calibration approach that utilizes several sites across the United States (Dingirard & Slater 1999, Kaufman & Holben 1993, Meygret et al. 2000, Slater et al. 1987, 1994, 1996, Thome et al. 1993).

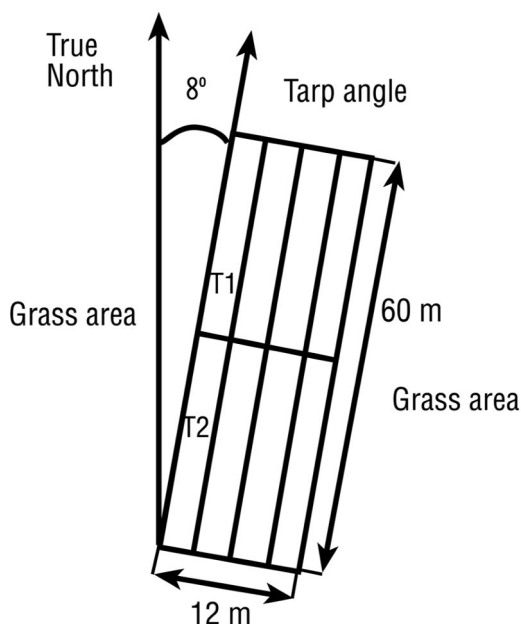


Figure 3. Illustration of tarp orientation with respect to true north and corresponding example IKONOS image of pulse target tarp.

The reflectance-based approach requires uniform target areas that are at least four times larger than the GSD of the system under evaluation. Exact target size depends on the MTF of the system being characterized. Additionally, knowledge of the target's surrounding area surface albedo is necessary to account for adjacency effects. Typical sites used by JACIE include: (a) large, dry lake beds such as Lunar Lake Playa, NV; (b) Railroad Valley Playa, NV; (c) Ivanpah Playa, CA; (d) White Sands National Monument, NM; (e) large grass fields such as those in Brookings, SD; and (f) artificial deployable reflectance tarps such as those used at NASA Stennis Space Center (SSC), MS (Fig. 2). Dry lake beds in arid climates are ideal targets because of their nearly Lambertian, high-surface albedo; their clear atmospheric conditions; and their simplified adjacency analysis. Lower reflectance targets in non-arid environments are more challenging but are often more realistic in terms of how the data may be examined by an end user. Targets of multiple reflectance values also allow radiometric characterization along a system's full dynamic range, which is important in understanding system linearity and in identifying any systematic biases.

At or near the time of sensor overpass, *in situ* measurements of target radiance or reflectance are acquired using analytical spectral devices (ASDs) hand-held spectroradiometers that have been previously characterized against sources traceable to the National Institute of Standards and Technology (NIST). The spectroradiometer collects radiance spectra within the visible–shortwave infrared wavelengths and at high spectral resolutions, e.g. 10 nm. Several hundred spectra can be acquired with each measurement, and measurements are typically acquired by walking the instrument over the large uniform area to increase spatial averaging. Averaging the thousands of resulting measurements reduces overall measurement uncertainties. Throughout this data collection period, ASD measurements of a 99 percent reflectance Lambertian Spectralon panel are also acquired to generate reflectance values (Pagnutti et al. 2003).

Atmospheric measurements are also acquired before and during the sensor overpass time. Optical depth measurements are acquired using a calibrated solar radiometer. Diffuse and direct solar irradiation are measured with a multi-filter rotating shadow-band radiometer (MFRSR). Radiosonde balloons are launched to measure temperature, pressure, and humidity profiles through the atmosphere (Pagnutti et al. 2003). Because the atmospheric modeling procedures used by various team members are slightly different, different team members may not utilize, or require, the same amount and type of atmospheric instrumentation, i.e. MFRSRs and radiosondes are not used at all vicarious calibration sites.

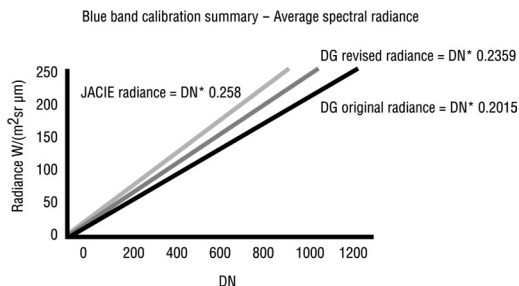


Figure 4. Example of JACIE-derived blue band radiometric calibration curve and coefficient compared to original and revised DigitalGlobe (DG) calibration curves/coefficients. The JACIE-derived curve was the result of several vicarious calibration activities conducted by team members during the 2002 acquisition season.

The *in situ* surface reflectance and atmospheric measurements, site location, time, sensor viewing geometry, and sensor and spectral response received from the commercial operator are used within a radiative transfer and atmospheric modeling software environment to estimate at-sensor radiance (Pagnutti et al. 2003).

When the target albedo differs significantly from the surrounding area albedo, adjacency effects can be significant. For the non-playa sites used by JACIE, adjacency effects must therefore be taken into account. The Air Force Research Lab's moderate resolution transmittance (MODTRAN) radiative transport code, which is used to predict at-sensor radiance, requires the specification of an area averaged surrounding surface spectral albedo in addition to the target surface spectral albedo to properly account for adjacency effects. Calibrated spectroradiometer radiance measurements from fully characterized Spectralon panels are used to verify the atmosphere and scene properties in MODTRAN. Agreement between MODTRAN predicted radiance and radiance measured using a calibrated spectroradiometer gives confidence in the model.

At-sensor radiance estimates are produced for each multispectral band (blue, green, red, and near-infrared) using several reflectance values at several sites through an acquisition season. JACIE radiance estimates for a given season are plotted against image DN values, and a least-square curve fit is performed to generate a calibration curve and a JACIE-derived radiometric calibration (gain) coefficient (Fig. 4). JACIE coefficients are compared to calibration coefficients provided by the commercial vendor and percent differences are calculated. This method evaluates the absolute radiometric accuracy of a commercial image product.

Relative (pixel-to-pixel) radiometric accuracy has also been evaluated using extremely large, uniform sites, such as Sarahan desert areas and Antarctic ice sheets. Images of large, highly uniform areas that fill

the system's entire image footprint can indicate relative radiometric differences within the focal plane and can help identify bad pixels. Using this approach, team members have examined the relative radiometric accuracy of selected image products (Pagnutti et al. 2003).

These same large, uniform areas have also been used to approximate SNR. The ratio of the mean radiance across the entire uniform image to the standard deviation of the radiance is an approximation of the system's SNR. JACIE has used this method to examine the SNR of a small number of images (NASA et al. 2001).

3 CONCLUSIONS

Since the start of the JACIE collaboration, the team's product characterization approach has been used to evaluate Space Imaging's IKONOS and DigitalGlobe's QuickBird. Team evaluations have resulted in several improvements to IKONOS image products (Zanoni et al. 2003), including an update to initially released radiometric calibration coefficients (Peterson 2001, Pagnutti et al. 2003). QuickBird evaluations are ongoing but already have resulted in updates of DigitalGlobe's radiometric calibration coefficients (NASA et al. 2003, Krause 2003). In general, JACIE assessments have found the geopositional accuracy and image quality of both IKONOS and QuickBird products to be quite good. To date, assessments also have found the radiometric stability of these systems to be very good. Future activities for the team include continued assessment of QuickBird data products, evaluations of products from Orb-Image Inc.'s OrbView-3 system, and temporal stability assessments of IKONOS.

Initial results are presented and discussed with vendors prior to release outside of the JACIE team. NASA, NGA, USGS, and the commercial data vendors communicate openly to resolve any differences between JACIE findings and those of the company. The interaction between member agencies and commercial industry is a critical component to the success of the effort. Government agencies and industry have developed trust and a collaborative spirit that benefits the U.S. government, the commercial remote sensing industry, and users of these high-resolution image products.

The JACIE effort and product characterization approach provides a model for characterizing all commercial remote sensing data purchased by the government, as well as for other cases where access to detailed system design and calibration information are not available.

ACKNOWLEDGMENTS

This work was prepared in part as an official duty of an employee of the U.S. Government and was supported

by the NASA Earth Science Applications Directorate under contract number NAS 13-650 at the John C. Stennis Space Center, Mississippi. Special thanks to Marcia Wise for assistance in editing and formatting this manuscript.

REFERENCES

- Constance, E. 2002. IKONOS DEM evaluation. In *Proceedings of the 2002 High Spatial Resolution Commercial Imagery Workshop, Reston, VA, USA, 25–27 March 2002*. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team, CD-ROM.
- Dinguirard, M. & Slater, P.N. 1999. Calibration of space-multispectral imaging sensors: A review. *Rem. Sens. Env.* 68(3): 194–205.
- Federal Geographic Data Committee (FGDC) 1998. *Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy, FGDC-STD-007.3-1998*. Washington: Federal Geographic Data Committee.
- Greenwalt, C.R. & Schultz, M.E. 1968. *Principles of error theory and cartographic applications*. ACIC Technical Report No. 96, St. Louis: Aeronautical Chart and Information Center, U.S. Air Force.
- Hecht, E. 1998. *Optics*. (3rd ed.) Reading, MA: Addison-Wesley.
- Helder, D. & Choi, J. 2003. On-orbit Modulation Transfer Function (MTF) measurement of QuickBird. In *Proceedings of the 2003 High Spatial Resolution Commercial Imagery Workshop, Reston, VA, USA, 19–21 May 2003*. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team, CD-ROM.
- IRARS 1996. *Civil NIIRS reference guide*. Imagery Resolution Assessment and Reporting Standards Committee, http://www.fas.org/irp/imint/niirs_c/guide.htm, accessed 9 March 2004.
- Kaufman, Y.J. & Holben, B.N. 1993. Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert reflection. *Int. J. Rem. Sens.* 14(1): 21–52.
- Krause, K. 2003. *Radiance conversion of QuickBird data*. DigitalGlobe Technical Note RS_TN_radiometric_radiance_4002, 2003-07-07, <http://www.digitalglobe.com/downloads/Radiance> Conversion of QuickBird Data.pdf.
- Leachtenauer, J.C., Malila, W., Irvine, J.M., Colburn, L.P. & Salvaggio, N.L. 1997. General Image-Quality Equation: GIQE. *Appl. Optics* 36(32): 8322–8328.
- Mikhail, E.M. & Gracie, G. 1997. *Analysis and adjustment of survey measurements*. New York: Van Nostrand and Reinhold Company.
- Meyret, A., Briottet, X., Henry, P. & Hagolle, O. 2000. Calibration of SPOT 4 HRVIR and vegetation cameras over Rayleigh scattering. In W.L. Barnes (ed.), *Earth observing systems V*. Proc. SPIE. 4135: 302–313.
- NASA, NIMA, & USGS 2001. *Proceedings of the 2001 High Spatial Resolution Commercial Imagery Workshop, Greenbelt, MD, USA, 19–21 March 2001*, NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team. CD-ROM.
- NASA, NIMA & USGS 2002. *Proceedings of the 2002 High Spatial Resolution Commercial Imagery Workshop, Reston, VA, USA, 25–27 March 2002*, NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team. CD-ROM.
- NASA, NIMA & USGS 2003. *Proceedings of the 2003 High Spatial Resolution Commercial Imagery Workshop, Reston, VA, USA, 19–21 May 2003*, NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team. CD-ROM.
- National Research Council 2002. *Toward new partnerships in remote sensing: Government, the private sector and Earth science research*. Washington: National Academy Press.
- Pagnutti, M., Ryan, R., Kelly, M., Holekamp, K., Zanoni, V., Thome, K. & Schiller, S. 2003. Radiometric characterization of IKONOS multispectral imagery. In S.N. Goward & V. Zanoni (eds), *Rem. Sens. Env.* 88(1–2): 53–68.
- Peterson, B. 2001. *IKONOS relative spectral response and radiometric calibration coefficients*. Space Imaging, Inc., Doc. #SE-REF-016, Revision A, 26 April 2001.
- Reichenbach, S.E., Park, S.K. & Narayanswamy, R. 1991. Characterizing digital image acquisition devices. *Opt. Eng.* 30(2): 170–177.
- Ryan, R. 2001. Initial IKONOS modulation transfer function compensation (MTFC) evaluation. In *Proceedings of the 2001 High Spatial Resolution Commercial Imagery Workshop, Greenbelt, MD, USA, 19–21 March 2001*. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team. CD-ROM.
- Ryan, R. 2002. IKONOS signal-to-noise estimation (MTFC-on versus MTFC-off). In *Proceedings of the 2002 High Spatial Resolution Commercial Imagery Workshop, Reston, VA, USA, 25–27 March 2002*. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team. CD-ROM.
- Ryan, R., Baldrige, B., Schowengerdt, R., Choi, T., Helder, D. & Blonski, S. 2003. IKONOS spatial resolution and image interpretability characterization. In S.N. Goward & V. Zanoni (eds), *Rem. Sens. Env.* 88(1–2): 37–52.
- Slater, P.N., Biggar, S.F., Holm, R.G., Jackson, R.D., Mao, Y., Moran, M.S., Plamer, J.M. & Yuan, B. 1987. Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors. *Rem. Sens. Env.* 22(1): 11–37.
- Slater, P.N., Biggar, S.F., Thome, K.J., Gellman, D.I. & Spyak, P.R. 1994. In-flight radiometric calibration of ASTER by reference to well-characterized scenes. In W.L. Barnes & B.J. Horais (eds), *Platforms and systems*. Proc. SPIE 2317: 49–60.
- Slater, P.N., Biggar, S.F., Thome, K.J., Gellman, D.I. & Spyak, P.R. 1996. Vicarious radiometric calibration of EOS sensors. *J. Atmos. & Oceanic Tech.* 13(2): 349–359.
- Thome, K.J., Gellman, D.I., Parada, R.J., Biggar, S.F., Slater, P.N. & Moran, M.S. 1993. In-flight radiometric calibration of Landsat-5 Thematic Mapper from 1984 to present. In P.S. Chavez & R.A. Schowengerdt (eds), *Recent advances in sensors, radiometric calibration, and processing of remotely sensed data*. Proc. SPIE 1938: 126–131.
- Tzannes, A.P. & Mooney, J.M. 1995. Measurement of the modulation transfer function of infrared cameras. *Opt. Eng.* 34(6): 1808–1817.
- Zanoni, V., Stanley, T., Ryan, R., Pagnutti, M., Baldrige, B., Roylance, S., Snyder, G.L. & Lee, G. 2003. The Joint Agency Commercial Imagery Evaluation team: Overview and IKONOS joint characterization approach. In S.N. Goward & V. Zanoni (eds), *Rem. Sens. Env.* 88(1–2): 17–22.

Camera calibration program in the United States: past, present, and future

G.Y.G. Lee

U.S. Geological Survey, Menlo Park, California, USA

ABSTRACT: The U.S. Geological Survey (USGS), which has responsibility for camera calibration in the United States, has been performing this function for the past 3 decades. The agency provides calibration services for traditional film cameras to other government agencies, as well as to the private sector in North America. Camera calibration has been performed with a multicollimator instrument designed and built in the early 1970s. Although the relative merits of the various laboratory and *in-situ* calibration methods have been debated for years, all of these methods have provided the camera calibration parameters required for photogrammetric work. As cameras evolved beyond film to digital technology and were integrated with existing airborne global positioning system (GPS) and inertial measurement unit (IMU) systems, the USGS could no longer provide necessary camera calibrations for the mapping community in its Optical Science Laboratory. Furthermore, its laboratory calibration method could not provide camera performance measures in conjunction with GPS and IMU components, which are commonly used in modern aerial photographic systems. The USGS is currently addressing this issue on three fronts: (1) procuring an interim laboratory calibration system for small format digital cameras, (2) developing an *in-situ* calibration range and software to compute camera calibration parameters similar to the analog calibration model, and (3) establishing an independent product characterization activity. Clearly, new calibration standards need to be developed, and new approaches for certifying aerial photographic systems and procedures must be established. Future strategy may include some form of International Organization for Standardization (ISO) certification that establishes a quality assurance program and review process.

1 INTRODUCTION

The responsibility for camera calibration in the United States was transferred from the U.S. National Bureau of Standards (NBS now renamed NIST), to the U.S. Geological Survey (USGS) in 1973. Since then, the USGS has been providing camera calibration services by using a multicollimator instrument custom built in 1951. These services enabled both government and private sector users to be assured that acquired aerial photographic services were compliant to specifications.

Over the years, additional collimators were added to the multicollimator instrument to provide an angular field-of-view of 120 degrees with uniform spacing along each bank of collimators. Furthermore, new resolution targets in the collimators were installed to extend the resolving power to accommodate the improved aerial cameras being marketed. The multicollimator instrument performed well for its designed purpose, and was extremely reliable. However, although several modifications were made to accommodate better performance cameras, it remained an instrument for calibrating only analog cameras.

The USGS established a fee schedule for its camera calibration services based on reimbursement for the actual costs of labor and materials. In earlier years, because the USGS contracted for its photographic data and performed the subsequent map production internally, it also subsidized part of the cost of camera calibration, which was considered a quality assurance step. However, no funds were collected to create a capital fund for equipment modernization or replacement. Furthermore, no resources were allocated to research or develop new calibration techniques.

2 BACKGROUND

Digital technology changed the photographic world, and photogrammetry. The early 1990s witnessed development of film scanners as the first step toward softcopy photogrammetry. Soon major camera manufacturers began designing digital aerial cameras that would eliminate analog-to-digital conversion in film scanning.

A major problem facing the USGS was its inability to calibrate digital cameras using the multicollimator instrument designed for analog cameras with a

230 × 230 mm focal plane. The multicollimator instrument provided insufficient targets for calibrating small format digital cameras having narrow fields of view. Subsequently, more complex digital camera designs presented another problem. For example, the Z/I Digital Mapping Camera (DMC), which nearly approximates an analog aerial camera, uses eight cameras and software to capture and compose full frame panchromatic and multispectral images. Leica Geosystems' Advanced Digital Sensor-40 (ADS40) camera, which uses a line scanning design, presents an impossible case for traditional camera calibration.

The manufacturer's combined hardware/software solution, as well as the variation in camera types, presents numerous new problems for traditional analog calibration methods. Furthermore, new digital cameras are being integrated with existing technologies, such as airborne global positioning system (GPS) and inertial measurement unit (IMU) equipment, to form a complete data collection system. Beyond digital cameras, a multitude of nonoptical sensors also is being used for acquiring geospatial data. Such technologies as synthetic aperture and interferometric radar, and light detection and ranging (LiDAR) have been in use for years and will become increasingly important to the mapping community. As already mentioned, both optical and nonoptical sensors are commonly integrated with other support systems, such as GPS/inertial navigation system (INS) airborne surveying. Yet another calibration issue is the data obtained by satellite sensors. Note that in these types of systems, the orientation and motion of the platforms carrying the sensor must be precisely determined.

Finally, the USGS is changing how it conducts its mission of providing base map data. The agency is transitioning from entirely inhouse map production to increasing dependence on private sector contractors. The evolution of how the USGS will accomplish its mapping work in future is further demonstrated with *The National Map* concept (Groat 2003, Ogrosky 2003), which outlines a greater dependence on partners to provide data. The underlying challenge is for the USGS to develop appropriate standards and quality assurance procedures to ensure that users of *The National Map* are satisfied with the data obtained. Furthermore, the USGS needs to establish a quality assurance procedure for maps prepared by commercial vendors. Thus, this changing environment has implications for the entire mapping community beyond camera calibration alone.

3 PAST

In 1973, the USGS provided calibration of both lenses and cameras used for aerial mapping. Now, its Optical Science Laboratory (OSL), performs only analog

camera calibrations and serves as the official calibration facility of aerial cameras for the United States and its North American Free Trade Agreement (NAFTA) partners.

The USGS has performed aerial camera calibration by using a multicollimator instrument modeled after an instrument introduced and used at the NBS. The original instrument contained 37 collimators arranged at 7.5 degree intervals along the diagonals of the photographic frame for calibrating cameras with a maximum of 90 degree field-of-view. Each collimator is focused at infinity and photographically exposes a position reticule and resolution target onto glass or film. Later, the instrument was modified to add additional collimators and higher-resolution targets that allowed for calibration of both wide-angle and super-wide-angle cameras and improved lenses. The calibration method is based on measured versus theoretical positions of the collimators' reticules to determine the calibrated focal length of the lens and the location of the principle point and fiducial marks. Values of resolving power are also read from the high-contrast-resolution targets to determine radial and tangential distortion characteristics. In addition, the area weighted average resolution (AWAR) is computed and reported. Camera performance, such as stereo model flatness, platen flatness, shutter performance, and filter traces are also determined (Tayman 1974).

The OSL's "Report of Calibration" is recognized as the official aerial camera certification in the United States (Tayman 1984). The USGS requires a camera to have been calibrated within a 3-year period for a company to compete for USGS contracts. This requirement has become the de facto requirement of other Federal agencies, as well as State and county governments.

4 PRESENT

The USGS acknowledges its role and responsibility for establishing standards, and accordingly, to provide camera calibration services to the mapping community. The USGS also believes that these functions are "inherently governmental" and will continue its sensor calibration efforts (ASPRS 2000).

4.1 Analog camera calibration

The USGS' OSL is the only active camera calibration site in North America. The OSL calibrates approximately 100 cameras per year from an active pool of approximately 400 aerial cameras. Because analog cameras are usually well maintained and kept operating for years, and new aerial cameras are still being sold, the USGS plans to continue calibrating them. For now, this service will be provided in the same

manner as in the past. To ensure continued performance of this important function, and because of the age of system components, steps are being taken to eliminate major weaknesses in both the instrument and its supporting software that might cause a service interruption. Thus, the multicollimator instrument has been modified to upgrade its lighting system and supporting electronics, and a new comparator for plate coordinate measurements was added. The simultaneous multiframe analytical calibration (SMAC) software (Light 1992) was ported to a newer computer system, and new database management software and data handlers also were implemented. No further capital expenses are planned currently for the analog calibration equipment and its existing support infrastructure.

Although the multicollimator instrument has operated satisfactorily and has been stable for many years, deficiencies exist in its inability to calibrate both digital and analog cameras. The age and technology of the instrument date to the 1950s and should be improved. Recently, preliminary inquiries were made to estimate the cost of designing and building a new instrument that could calibrate both digital and analog cameras. The preliminary estimate was judged to be cost prohibitive. Clearly, other alternative approaches will be pursued.

4.2 *Small format digital camera calibration*

In 2002, the USGS awarded a contract to Pictometry International to design, develop, and implement a digital calibration system for small format digital cameras. One of the major requirements was that the system would provide the same calibration parameters as those from the SMAC software being used for analog cameras. The rationale for this requirement was that it would allow researchers to compare the performance of analog and digital cameras and provide practitioners a common processing path for both types of cameras.

The digital calibration system was delivered for testing in June 2003. The system is based on a multi-camera convergent method using nine images of a target "cage" taken from different camera-exposure locations, orientations, and rotations; the "cage" is 3.6 m wide by 3.0 m high by 2.4 m deep, with 207 reflective targets distributed on three planes. The computation does not require known object-space coordinates of the targets and uses the Australis photogrammetric software, which permits automated digital close range photogrammetric image measurement, orientation/triangulation, and sensor calibration (Fraser 2001).

4.3 *In-situ camera calibration*

In 2002, the USGS awarded a research grant to the Ohio State University (OSU) to develop and test an

in-situ camera calibration system. The motivation for performing this work was to find a method for calibrating complex sensor systems beyond just digital cameras. The method would include the influences of in-flight operating conditions; and furthermore, might provide a future alternative procedure to laboratory calibration of analog cameras on an aging multicollimator instrument. Again, the camera calibration parameters computed were required to be similar to those from the SMAC software. The resulting software developed from this research is to become the property of the USGS for use by the public.

4.4 *Product characterizations*

The main purpose of USGS involvement in camera/sensor calibration is to ensure the quality of its map products. Over the years, the systems used to acquire imagery for geospatial products have become more varied, as have the ways in which the USGS has accomplished its mission of providing base map data to mapping communities. The complexities of data acquisition systems, the numerous map production methods, and the number of data providers force the USGS to consider alternative methods of quality assurance and control beyond camera calibration alone. Such product characterization or testing is a change to the traditional camera certification paradigm.

In a parallel effort, the USGS has been actively participating in the Joint Agency Commercial Imagery Evaluation (JACIE) efforts for commercial satellite data (Zanoni 2003). This evaluation is based on product characterization. In support of this effort, the National Aeronautics and Space Administration (NASA)'s Stennis Space Center (SSC) has developed a targeted testsite for performing geometric and spatial resolution evaluations. In early 2003, the USGS and NASA signed an agreement to use SSC for characterizing products developed from imagery acquired by digital camera technology. NASA modified their testsite to include a more densely targeted area that would accommodate airborne sensors.

The USGS, with the support of NASA, is currently scheduling test flights for all its contractors, subcontractors, and cooperators providing data to the USGS. Final products are submitted to NASA for evaluation. The USGS will review the results and approve a specific contractor/camera combination. The USGS plans to provide this testing capability to others in the future.

5 FUTURE

The current USGS effort to develop alternative calibration and validation methods for sensors and map products needs to be expanded to meet the long-term needs of both the agency and the mapping community.

New calibration standards need to be developed, reviewed, and accepted.

5.1 Analog camera calibration

The USGS will continue providing its analog camera calibration services until new standards are developed that encompass both analog and digital cameras. At that time, *in-situ* calibration of analog cameras may be an alternative method to present laboratory procedures.

5.2 Small format digital camera calibration

More work will be performed on small format cameras using the laboratory digital camera facilities. As new customers emerge, and as new cameras appear, the USGS will perform camera calibration and make comparisons with cameras already calibrated in the OSL. Repeatability testing will be done to determine the reliability and stability of individual cameras and to compare results from other laboratory sensor calibration methods. In addition, the results will be compared with those obtained by *in-situ* methods.

5.3 In-situ camera calibration

As for *in-situ* camera calibration, the next step is to perform operational testing. Laboratory versus *in-situ* calibrations will be performed, and comparisons will be made for both digital and analog cameras. The software will be generalized, if needed, to perform camera calibrations over any site configuration and location. The plan is to enhance the software to eliminate the need for object-space coordinate control, similar to the techniques developed by Leica Geosystems for their ADS40 camera (Tempelmann 2003).

If the *in-situ* calibration approach is going to be part of the solution for sensor calibration, more strategically located test sites will be required for operational use. These additional test sites should be easily accessible and provide alternatives for weather and seasonal conditions. They also should be representative of various terrain and vegetation cover types, and permit testing of different sensor, imaging system, and geospatial product types. In addition to the SSC and the OSU test range, the USGS also has performed extensive field control efforts over areas in Colorado, Missouri, and South Dakota. The USGS plans to extend its calibration development to include standards and specifications for *in-situ* calibration ranges so that other government agencies and the private sector can establish alternative test sites that better suit their needs.

5.4 Product characterization

The USGS plans to extend its product characterization capability beyond that currently provided by NASA.

Three potential test sites that are located near existing USGS facilities will be part of its regionalized quality assurance program. Certainly, part of the quality assurance function can be performed by contractor personnel, as in some of the current satellite monitoring operations at the EROS Data Center in South Dakota. Furthermore, product characterization by independent, certified, private consultants should be acceptable in lieu of an evaluation by the Government.

Current product characterization work at the USGS focuses only on geometric and spatial characteristics of digital cameras used to generate digital-orthophoto products. Calibration methods need to be extended to radiometric and other calibration parameters. The USGS and NASA have been discussing an extension of their product characterization capability to include the radiometric performance of digital sensors, both laboratory and *in-situ*. The USGS wants to pursue development of radiometric product characterization techniques because their application extends to data collected by multispectral and hyperspectral sensors.

NASA plans to extend its product characterization capability at SSC to include LiDAR data. The USGS believes that characterization of LiDAR data is its next priority, and will be discussing possible cooperative efforts.

5.5 Quality assurance for The National Map

The USGS plans to implement a more comprehensive quality assurance program for data provided by *The National Map* partners. In general, these data will have been produced by various contracting mechanisms within State and local agencies. Quality control and review methods, as well as quality of the metadata, will vary among potential partners. Product characterization may be the only viable alternative for validating these data.

6 CONCLUSIONS

The ultimate solution to camera calibration may not be a single approach but a combination of different approaches and methods that evolve into “best practices” for performing end-to-end quality assurance. As demand for verification and validation mechanisms increases, USGS will need to consider alternative approaches for how it provides independent validation of a new sensor, camera, production procedure, or map product.

6.1 Acceptance of manufacturer's calibration plan

Major camera manufacturers are likely to have specialized equipment to test the quality of their products.

Clearly, they have intimate knowledge of how each camera operates and how it was built. Test equipment is probably custom built and applicable only to their particular product. It would be cost prohibitive for the Government to acquire such specialized equipment for every camera in the market. Certifying the manufacturer's testing and calibration procedures would be an effective way to satisfy users of any particular type of camera. Periodic inspection of facilities and records will still be needed after initial approval of the test procedure. To ensure objectivity, an oversight committee composed of qualified experts from government agencies, the private sector, professional societies, and academia can oversee the evaluation.

6.2 Calibration by independent testing organizations

As mapping technology evolves, new methods and procedures should emerge from research and development activities. Comparisons of these new methods and procedures will be needed in a research environment. As they mature, they also will need to be evaluated on an operational basis to verify and validate any new performance claims. An independent testing organization, either in the public or private sector, would serve the needs of practitioners, manufacturers, and users.

6.3 ISO 9001 certification

The best approach to overall quality assurance of products and services offered to the mapping community may be greater emphasis on some form of International Organization for Standardization (ISO) certification for companies that would require them to establish, document, monitor, and enhance on a

continuing basis a quality assurance program for their products and services. Generally, ISO certification also would require proper qualifications for workers, as well as for the equipment and procedures being used. Specifically, ISO certification would allow acceptance of calibration methods established by data providers, and ultimately, would ensure quality products and services to every day users of geospatial data.

REFERENCES

- ASPRS 2000. Camera calibration panel report: *Photogr. Eng. & Rem. Sens.* 66(3): 239–245.
- Fraser, C. 2001. *Australis: software for close-range digital photogrammetry users manual, version 5.05*, Victoria: Univ. of Melbourne.
- Groat, C.G. 2003. The National Map – a continuing critical need for the Nation. *Photogr. Eng. & Rem. Sens.* 69(10): 1087–1090.
- Light, D. 1992. The new camera calibration system at the U.S. Geological Survey, *Photogr. Eng. & Rem. Sens.* 58(2): 185–188.
- Ogrosky, C.E. 2003. National mapping examined: an introduction. *Photogr. Eng. & Rem. Sens.* 69(10): 1081–1085.
- Tayman, W.P. 1974. Calibration of lenses and cameras at the USGS, *Photogr. Eng.* XL(11): 1331–1334.
- Tayman, W.P. 1984. User guide for the USGS aerial camera report of calibration, *Photogr. Eng. & Rem. Sens.* 50(5): 577–584.
- Templemann, U. et al. 2003. ADS40 calibration & verification process. *Optical 3-D measurement techniques, 6th conference, 22–25 September 2003*. Zurich: Inst. Geod. & Photogra.
- Zanoni, V., Stanley, T., Ryan, R., Pagnutti, M., Baldrige, B. & Roylance, S. 2003. The Joint Agency Commercial Imagery Evaluation (JACIE) Team: Overview and IKONOS joint characterization approach, *Rem. Sens. Env.* 88(1-2): 17–23.

Building and characterizing calibration test ranges

In-flight characterization of spatial quality using point spread functions

D. Helder, T. Choi & M. Rangaswamy

Electrical Engineering Dept., South Dakota State University, Brookings, SD, USA

ABSTRACT: With the advent of high spatial resolution commercial satellite imaging systems, it is important to assess the spatial quality of these systems. Prior to flight, laboratory measurements of spatial quality typically depend on the system point spread function (PSF) or, equivalently, modulation transfer function (MTF) using various targets. However, the majority of the life of these sensors is spent in flight, and it is during this time that the sensor is actually acquiring data for which it was designed. Hence, it is important to measure the spatial performance during the most important part of their life cycles to determine if any system degradations occur over time. This paper reports on knife-edge, pulse, and point techniques that have been developed to accomplish this task. Each of these has been optimized for a particular range of spatial resolutions, and for both one-dimensional and two-dimensional PSF estimations.

1 INTRODUCTION

1.1 *Rationale for measuring spatial quality*

An important function of any imaging system is to resolve spatial objects. A system's ability to do this is determined by many design attributes including optics, focal plane, detectors, and electronics. However, one of the most difficult things to define is an imaging system's ability to resolve spatial objects or its "spatial resolution." Often this term is improperly defined and poorly understood. Also, it is sometimes confused with the concept of instantaneous field-of-view (IFOV). The problem is compounded because there is no standard method for measuring spatial resolution, and there is no single metric that adequately measures it.

When spatial quality is measured in the laboratory, several targets and techniques are used. These measurements quantify spatial quality of the imaging system to a high degree, but in many cases only for a short period of time in the life of the sensor. While airborne sensors can be reevaluated periodically, spaceborne sensors cannot. Thus, for most of the useful life of sensors, spatial quality can be difficult, if not impossible to measure. Conversely, for airborne sensors, laboratory measurements may not be an accurate reflection of a sensor's performance when it is airborne. Therefore, it is important to assess the spatial resolution performance of a sensor while in flight.

1.2 *Typical measures of spatial quality*

If an imaging system can be considered as a linear, shift-invariant system, then its ability to resolve spatial objects can be described mathematically by a point spread function (PSF). The PSF is a direct measure of the system's ability to reproduce an infinitely narrow source of radiance. However, this two-dimensional function is difficult to obtain because of difficulties in generating an infinitely narrow source. The problem is compounded by the digital nature of currently used systems. Edge targets are often used because they are easy to generate; however, the drawback is that only a one-dimensional slice of the PSF is obtained. The normalized magnitude of the Fourier transform of the PSF, or modulation transfer function (MTF), is used widely as a measure of the spatial quality of an imaging system. Specifications are written to obtain minimum values of the MTF at critical spatial frequencies. While these metrics are mathematically precise, they do not reflect adequately the human contribution to image quality assessment. To address this issue, other systems have been developed. For example, the National image interpretability rating scales (NIIRS) combines several attributes to assess the overall quality of an image (Leachtenauer 1996, Irvine 1996). However limiting they may be, this work focuses on the PSF and MTF representations of image spatial quality because they are mathematically precise and widely accepted.

1.3 Laboratory measurements

Laboratory measurement of image spatial quality has been extensive (Feltz 1990, Samei 1998, Droegge 1979, Kaftandjian 1996, Tzannes & Mooney 1995, Cao 2000, Hong 1996). One of the early attempts by Coltman (1954) used a sinusoidal input illumination pattern for evaluation of the MTF:

$$\text{Input} = 0.5[1 + \cos(s)] \quad (1)$$

where s is a spatial parameter. The sinusoidal input was changed in amplitude and phase to obtain the transfer function from the resulting sinusoidal output. MTF could be calculated by the ratio of output to input:

$$\text{MTF} = \frac{B_{\max} - B_{\min}}{B_{\max} + B_{\min}} \quad (2)$$

where B_{\max} = neighboring maximum; B_{\min} = neighboring minimum of the output.

An MTF for a linear X-ray detector was calculated by using ball, wire, edge, and bar/space patterns as stimuli (Kaftandjian 1996). The ball target was a very small sphere of high attenuation, which led directly to the PSF of the system. A thin wire (smaller than the pixel size) was used to derive the line spread function (LSF), which was the 1-D PSF orthogonal to the wire orientation. Lastly, instead of using a sine wave input, a set of rectangular pattern inputs of diminishing spatial width was used to measure the contrast transfer function (CTF). The MTF was not directly calculated from the bar/space pattern target; however, it could be derived by the mathematical relationship:

$$\text{MTF}(u) = \frac{\pi}{4} \left[\text{CTF}(u) + \frac{1}{3} \text{CTF}(3u) - \frac{1}{5} \text{CTF}(5u) + \frac{1}{7} \text{CTF}(7u) \dots \right] \quad (3)$$

where u = the frequency domain variable. To avoid noise influence on the MTF evaluation, the edge spread function (ESF) was modeled by an error function $e(x)$ defined:

$$e(x) = A \cdot \text{erf}(B \cdot x) + C \quad (4)$$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt; \quad (5)$$

where A , B , C = constants depending on the experimental profile. This parametric method resulted in smooth ESF and LSF estimates from the differentiation process; however, the profile could not follow obvious ringing effects presented on both sides of the LSF.

Another approach used a sharp edge positioned on the detector array with a small angle to obtain a finely sampled ESF (Tzannes & Mooney 1995). A PtSi infrared camera's spatial quality was characterized by applying a modified knife-edge technique to fit the ESF to a summation of Fermi functions, $F(x)$, of the form:

$$F(x) = D + \sum_{i=0}^2 \frac{a_i}{\exp\left[\frac{x - b_i}{c_i}\right] + 1} \quad (6)$$

where D , a_i , b_i , c_i = constants for $0 \leq i \leq 2$. Therefore, the LSF was completely specified by analytical differentiation of the ESF and the values determined by a_i , b_i , and c_i . This model was used to avoid noise from the data and noise amplification caused by the numerical differentiation process. The target angle with respect to the sampling grid was not critical because the number of scan lines ranged from 20 to 40, which was large enough to obtain good sub-pixel accuracy.

From a carefully located slight angle to the edge line, a complete uniformly sampled ESF was achieved by using the trigonometric relationship between sampling grid and angle (Cao 2000). Even though the target angle and length were carefully designed, this level of accuracy is not possible to achieve in satellite imagery. Because satellite output images contain many sources of uncertainties from original scene acquisition to post-processing, a direct technique is needed to extract edge angle and finely sampled sub-pixel locations with a reasonable error margin. After obtaining the 1-D ESF, a boxcar mask and low-pass filter were applied to suppress the noise in differentiation and in the ESF itself.

As an alternative approach, random targets were used to measure discrete MTF beyond the pass band of the detector array (Hong 1996). Printed random transparency targets aligned in the row direction were designed by use of a microscan technique. The targets may be easier to use and to construct than previously mentioned targets (such as sinusoidal patterns), but this approach is obviously limited to a laboratory environment.

1.4 In-flight measurements

On-orbit spatial quality estimation for the Landsat 4 Thematic Mapper (TM) utilized scenes containing the San Mateo Bridge over the south end of San Francisco Bay (Schowengerdt 1985). The 18.3 m wide bridge was used as a pulse input to the sensor at an angle of 31.1 degrees to the TM scan direction. Since the width of the bridge was smaller than the TM resolution (30 m), the Nyquist frequency of the bridge input spectrum occurred before its first zero crossing point. But the small target size caused noise problems, which were

solved by interpolation of the power spectrum values. This work extracted image profiles perpendicular to the bridge. Therefore, MTF plots did not quite follow the conventional across-track direction.

Another paper (Rauchmiller 1988) presented a method for measuring the TM imaging system PSF using a 2-D array of black squares on a white sand surface. The 16 square targets were shifted 0.25 pixel throughout sub-pixel locations within a 30 m ground sample distance (GSD) in along- and cross-track directions. This superimposed TM pixel grid enabled sub-pixel phasing to avoid aliasing in the MTF plot, but four sub-samples were not enough to reconstruct the original PSF in 2-D space.

Measurement results also were compared to a spatial response model given by (Markham 1985). Markham's model described the TM and multispectral scanner (MSS) PSF by pre-launch component and system measurements. The PSF was separately described by the optics, detector, and electronics that all affected the net spatial response. The LSF in the scan direction was described by the convolution of the three components:

$$LSF_x(x) = OSF(x) * DSF(x) * ESF(x) \quad (7)$$

where $OSF(x)$ = the optics line spread function; $DSF(x)$ = the detector line spread function; and $ESF(x)$ = the electronics line spread function. These results were generated by using pre-launch design documents and test measurements.

By using modeling concepts, Storey (2001) implemented a PSF model fit to predict the Enhanced Thematic Mapper Plus (ETM+) optical system's MTF performance change. The least error model fit method was developed to monitor the along-scan MTF by using on-orbit data of Lake Ponchartrain Causeway in Louisiana. The parallel, two-lane bridge structure made a small angle to the along-track orbital path, which allowed construction of an over-sampled target profile for the along-scan direction. Model parameters were adjusted to yield the best fit between the output scenes and simulated profiles. Spatial performance degradation over time was observed in the panchromatic band by comparing between on-orbit estimated parameters and pre-launch values.

This paper focuses on in-flight estimation of PSF and MTF for both spaceborne and airborne sensors. Emphasis is given to target development, constraints, and advantages/disadvantages of three types of targets found to be particularly useful: edge, pulse, and point targets. The second part of estimating spatial quality involves proper processing of data recorded by the sensor. Unless care is taken to extract the estimates from noisy data properly, inevitable biases occur that degrade the result. Thus, proper data handling methods are presented that are heavily dependent on target

type. Lastly, examples will be presented for each of these methods using commercially available, high spatial resolution satellite sensors.

2 TARGET TYPES AND FIELD CAMPAIGN

2.1 General target attributes

A basic assumption for PSF/MTF characterization is that the imaging system is linear and shift invariant. If this is accepted, then it is possible, in theory, to characterize any system simply when the input function is known and the output is measured. Thus, in principle, almost any target may be used. However, practice indicates that the target must also have several other features. One of the foremost is that the target must be characterized, i.e. it must be accurately known in terms of reflectance, size, and orientation. This must be true at least at a scale equal to, or greater than, the IFOV of the sensor.

Often orientation becomes critical due to the imaging system being discrete, or sampled. As a result, with only samples of the system response available on a regular grid of locations, it is often required to orient the target so that sub-sample reconstruction of the system response to the target is possible. For example, an edge target may need to be oriented such that the edge can be reconstructed at a much finer resolution than the sampling distance of the sensor.

Targets also need to be well controlled in terms of homogeneity and contrast. In many cases the targets consist of a sharp transition separating two or more homogeneous regions. The homogeneous regions should be uniform and isotropic. Any variation in the regions will affect the accuracy of the PSF estimate. In addition to this, the contrast across the edge must be as large as possible. This is limited only by the dynamic range of the sensor. A reasonable figure of merit for the target in terms of a signal-to-noise ratio (SNR) may be built upon the amount of contrast between the bright and dark portions of the target divided by the variance in the homogeneous regions. Simulations have shown that an $SNR > 50$ is a reasonable lower threshold.

Another important property for all PSF targets is that they must be time-invariant. This is important for measuring the degradations of a sensor's spatial quality as a function of time. Changes with respect to time should be due to change in the sensor rather than change in the target. This means that the target must be maintainable, which means it should be accessible on a regular basis and be small enough to be painted, paved, mowed.

Another important aspect of PSF targets is whether they are designed for estimating 1-D or 2-D PSF's. The PSF is always a 2-D spatial function. However, a 2-D target is difficult to develop. This often arises

because one spatial dimension must be used to provide a phased set of image samples so that the target can be reconstructed with enough resolution to approximate the continuous PSF. Many times this is adequate if a circular PSF is assumed to be an accurate representation for the imaging system. However, a PSF function may not only be non-circular, but also vary with location in the focal plane. Hence, it is desirable to design targets such that a 2-D PSF can be estimated.

Figure 1 illustrates the three types of targets discussed in this section: edge, pulse, and point. These targets were deployed at the Brookings, SD test site.

The edge target is a set of tarps developed by Stennis Space Center (SSC). The blue tarps and mirrors were developed at South Dakota State University.

2.2 Edge target

Edge targets are the most common type of PSF target. The concept is to present a sharp, high-contrast edge to the imaging system so that an accurate edge response function can be obtained. The edge response can then be differentiated numerically to obtain a 1-D slice of the PSF. Important aspects of this target are size, orientation, uniformity, and homogeneity. In many cases, an existing feature can be used for an edge target, e.g. a bright concrete parking lot next to a darker uniform grassy area. While these types of targets can often produce acceptable results, a better approach is a target where both sides of the edge can be maintained. Examples of this are dedicated targets painted on concrete or other surfaces, or tarps that can be deployed as needed. An example of a tarp-based edge target is shown in Figure 2.

A good edge target consists of a 'dark,' or low-reflectance side, and a 'bright' or high-reflectance side. The high-reflectance side ideally has a reflectance level that provides a sensor response near the high end of its dynamic range; a rule of thumb is >75 percent of the dynamic range of the sensor. This estimate must include atmospheric effects as well. The dark side of the edge should be a reflectance level as low as possible, yet coordinated with the surrounding background. If the coverage region of the target is small enough that the imaging system is affected by the surrounding area, the low-reflectance side of the target should have a reflectance similar to the surrounding region. If the target is large enough that the PSF measurement is not affected by the surrounding region, then the target material should have as low a reflectance as possible, say less than 5 percent. There is no one measure to determine whether the PSF measurement will be affected by the surrounding area. However, a reasonable rule of thumb is that the target should extend 7–10 IFOV's beyond the edge. Most high quality image systems will not have a PSF that extends beyond a

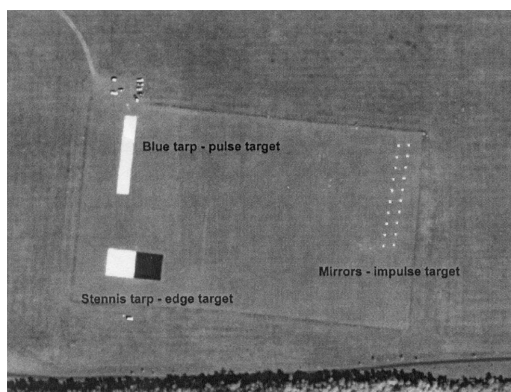


Figure 1. Panchromatic band QuickBird target area imaged on August 25, 2002.

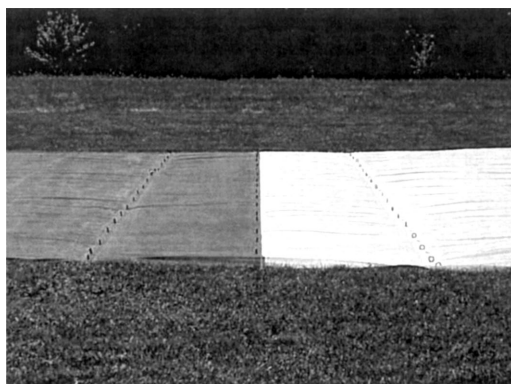


Figure 2. SSC tarp on a flat grassy area.

distance of a few IFOV's. The SSC tarps shown in Figure 2 have a reflectance of 3.5 percent on the dark side, and 52 percent on the bright side.

Spectral characteristics of the edge target should be well known, and the target should be spectrally flat. A target that is spectrally flat over a broad range of wavelengths can be used for PSF estimation for a broad variety of sensors. This enhances the usefulness of the target with comparisons across sensors. However, many targets may be used with less than ideal spectral characteristics if they are well characterized with a spectroradiometer. This should be done on a regular basis in the field while the target is deployed. Characterization is important to insure that the target has the properties mentioned above for the spectral band of interest, e.g. a target may have excellent reflectance in the blue region, be unusable in the red, and be quite bright again in the near infrared. Figure 3 shows sample spectra for an edge target.

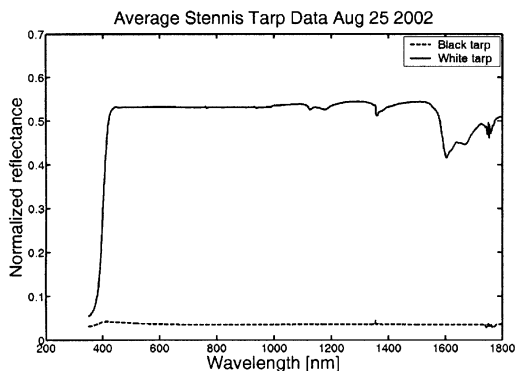


Figure 3. Spectral response of both sides of an SSC tarp edge target.

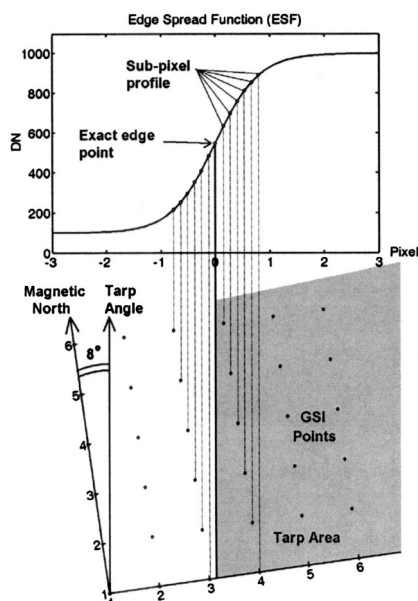


Figure 4. GSI projection to ESF with angle.

Orientation of the target edge is important for discrete imaging systems. If the edge is aligned with the imaging system's sampling grid, then samples of the edge response will be aligned with each other no matter how long the edge is. However, if the edge is oriented at a slight angle with respect to the sampling grid, then samples of the edge response can be obtained at a resolution much finer than the resolution of the imaging system (Fig. 4). The angle to use is dependent on the imaging system's ground sample interval (GSI) and the length of the edge. If the angle is chosen incorrectly, the samples across the edge will align

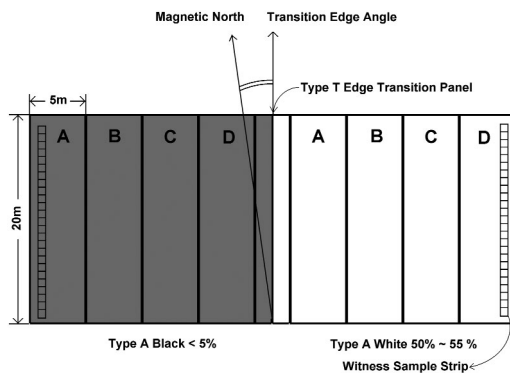


Figure 5. SSC tarp layout configuration.

with one another to form 'clumps' of data rather than uniform sample spacing.

Figure 5 shows a typical edge target layout. This target consists of four tarps forming the dark side of the edge, four tarps forming the bright side of the edge, and a smaller transition tarp that forms the actual edge. For this target system, an angle of 8 degrees with the image system grid is nearly ideal.

The length of the edge should be as long as possible. However, the upper end of the length is driven by the IFOV of the sensor, the size of the area available for target deployment, the cost of the target material, the amount of time and effort needed to deploy or maintain the target, and the uniformity of the surrounding area. The minimum length of the target should be selected such that the desired level of sub-pixel resolution is obtained when reconstructing the edge response. A minimum value that works well is to ensure that the edge is long enough to obtain at least 20 samples of the edge response. This should allow reconstruction of the edge at a resolution at least 10 times greater than the resolution of the imaging system, e.g. tarps with a length of 20 m are excellent for use with 1 m or less GSI sensors.

2.3 Pulse target

A second type of useful target is a pulse target. This consists of a uniformly bright region with two homogeneous dark regions such as grass on either side. The boundaries between the light and dark regions are straight edges. An example of such a target is shown in Figure 6.

This target is designed to obtain a 1-D cross-section of the system PSF. Its length is designed to obtain several 'slices' of the PSF in the orthogonal direction, similar to the length of the edge target. The orientation of the pulse target must be maintained so that the over-sampled pulse response can be obtained from the imagery.

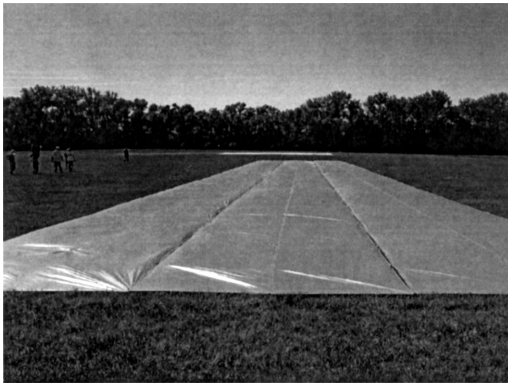


Figure 6. Blue tarp as a pulse input.

Usefulness of the pulse target is based on the GSI of the sensor whose PSF is being estimated. For sensors with a GSI in the 1–2 m range, an edge target is probably a better target to use since both sides of the edge are controlled more easily than a pulse type target. For a larger GSI in the 3–10 m range, it is difficult to develop a target that is large enough and yet deployable and maintainable, e.g. a 4 m GSI sensor requires an edge target 40 m in width (approximately 10 GSIs) to ensure there is no contribution due to edge effects. This width is achieved easily in natural objects such as asphalt parking lots or uniform grassy areas. Deployment of bright tarps in the center of such regions is often an attractive pulse target.

There are some disadvantages to pulse targets. One is that the imaging system produces a pulse response to the target and it is difficult to obtain the PSF directly from the pulse response. Also, when calculating system MTF, a Fourier transform of the pulse response function is required. Since the well-known Fourier transform of a square pulse function is a sinc function, it is necessary to deal with zero-crossings. The pulse target width must be designed carefully to place the zero-crossings at locations where the value of the MTF is not critical to evaluating system performance. In particular, since many systems are specified in terms of the MTF value at Nyquist, care must be taken to insure that a zero-crossing does not occur near this frequency. Figure 7 illustrates the amplitude of the input function at the Nyquist frequency as a function of tarp width measured in pixels (or GSI).

From the figure, it is apparent that if zero-crossings are to be avoided, the width of the pulse should be as narrow as possible, say one GSI or less. However, the difficulty is that the strength of the signal received by the sensor from the narrow pulse width decreases linearly as the width of the pulse. As a result, the SNR is compromised and a good estimate cannot be obtained. Practice has shown that a pulse width of

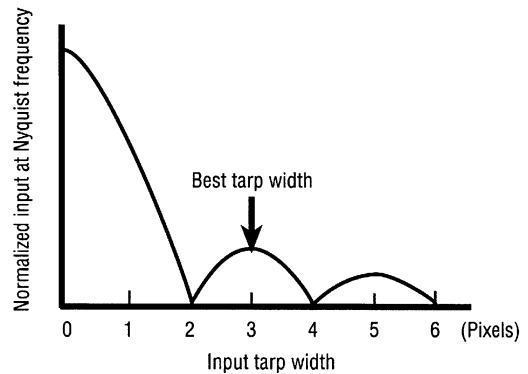


Figure 7. Nyquist frequency position on the input sinc function based on tarp width.

three GSI is optimal for this type of target. With this width, a good compromise is reached between obtaining a strong signal and maintaining ample distance from placing a zero-crossing at the Nyquist frequency.

Deployment of pulse targets is more difficult than edge targets. With an edge target, only one straight edge is needed. With a pulse target two straight edges are needed. Moreover, the two edges must be parallel and separated by a precise distance. With both edge and pulse targets, several pieces of target material are often used to create the desired target size. In these cases, it is important to join the sections so there is no change in reflectance at the seams. Lastly, pulse targets should deliver a high reflectance across the wavelengths of interest.

2.4 Impulse target

To measure the PSF of an imaging system, it is desirable to use a point source as the target. This has proven to be difficult, but attempts have been made using a passive source (Evans 1974). An advantage is that a true 2-D PSF can be produced. A complicating factor is that many high-spatial-resolution sensors are pointable, so it is difficult to determine where the sensor will be when it images the target. This becomes a very difficult problem if the target is directing a radiance source towards the sensor.

One approach has been to deploy a convex mirror system as an array of point sources. These mirrors can be designed to reflect the Sun's energy over a given solid angle so that it can be imaged by a pointable sensor. They can also be designed to reflect the right amount of energy so that a strong signal is received by the imaging system without saturation. An array is required for sampled data imaging systems much like a long edge or pulse target is needed for 1-D estimation of the PSF. In this case, each mirror produces a

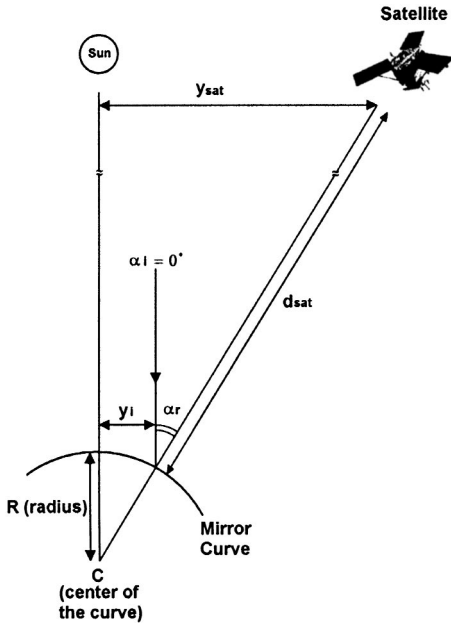


Figure 8. An example that relates incident energy at the convex mirror and at sensor.

sampled response of the PSF. By properly placing the mirrors in an array, an oversampled PSF can be reconstructed.

The relation between incident energy at the convex mirror and incident energy at the satellite sensor can be developed using geometric optics (Hecht 1998). The viewing geometry is shown in Figure 8. Assume, without loss of generality, the incident angle of the sun with respect to the mirror is zero, $\alpha_i = 0$.

From geometric optics, the following relationship can be derived:

$$A_s = \pi y_i^2 \left[\frac{d_{sat}}{f} + 1 \right]^2 \quad (8)$$

where, y_i = radius of the circle made by circular cross-section of the convex mirror, d_{sat} = altitude of the satellite from the Earth's surface, f = focal length of the convex mirror, and A_s = area of the circle through which the convex mirror projects light at a distance d_{sat} from the mirror.

From the law of conservation of energy, power received at the satellite sensor (P_s) is equal to the power incident (P_i) on the mirror.

$$P_s = P_i \quad (9)$$

$$E_s A_s = E_i A_i \quad (10)$$

$$E_s = \left(\frac{f}{d_{sat}} \right)^2 E_i \quad (11)$$

where, E_s is at-sensor irradiance [W/m^2], E_i is solar irradiance [W/m^2].

As the reflected energy passes through the atmosphere, a part of it is reflected, some is scattered, and some is absorbed before reaching the sensor. Furthermore, only a fraction of this energy is captured by the sensor. The sensor has an internal gain that amplifies this energy to a larger quantity. All of these components can be represented in a composite way by a scaling factor, ' k ', and the energy relationship can be written as

$$E_s = k \left[\frac{f}{d_{sat}} \right]^2 E_i \quad (12)$$

Thus, the above relationship relates the focal length of the convex mirror to the orbital height of the satellite and the amount of energy the satellite receives. Unfortunately, it is difficult to accurately determine the scaling factor, ' k '. One approach is to use an atmospheric model to predict the amount of attenuation expected at the site of interest, and to use the published gain of the sensor to estimate a nominal focal length for the mirror. Then, an inexpensive mirror of this size can be deployed and the instrument response can be measured. Based on this result, an accurate estimate of the scaling factor can be determined and the mirror focal length can be calculated precisely. The focal length should be designed so that the peak sensor response is on the order of 75 percent of its dynamic range. This exercises the sensor over a broad portion of its dynamic range yet insures adequate protection from saturation of the sensor. As an example, this process led to a focal length of 0.67 m for a convex mirror array designed for the QuickBird sensor.

An array of mirrors must be deployed to produce an over-sampled estimate of the system PSF. This implies that both the mirrors, and the background area surrounding the mirrors, must be uniform. The number of mirrors used should be as large as is practical for the sensor and site. Experience has shown that 15–20 mirrors can produce good results.

Deployment of the mirror array depends on the GSI of the sensor. The following guidelines must be followed to place the mirrors so they are properly phased for reconstruction of the PSF. First, the individual mirrors must be far enough apart so that the PSF region of support from one mirror does not overlap any adjacent mirrors. For many systems, the PSF does not extend over 3-GSI. Thus, a minimum mirror separation of 5-GSI should be adequate.

A second criterion to consider is how many mirrors to deploy. This will drive the sampling grid layout. Consider using a 20-mirror array to over-sample a PSF function. One possible uniform sampling layout is shown in Figure 9a. Here a 0.25-GSI spacing is used in the cross-track direction and a 0.2-GSI spacing in the along-track direction. To insure that the mirrors are far enough apart, the physical spacing between each mirror should be at least 5-GSI, or in this case:

$$\Delta x = 0.25M > 5 - \text{GSI}, \Delta y = 0.2N > 5 - \text{GSI} \quad (13)$$

where M, N are integers. Choosing $\Delta x = 1.25\text{-GSI}$ and $\Delta y = 10.2\text{-GSI}$ insures the above constraint is met and defines the angle of orientation of the mirror array with respect to the sampling grid of the sensor,

$$\tan \theta = \frac{\Delta x}{\Delta y} \quad (14)$$

which yields an angle of 7 degrees. The resulting mirror layout is shown in Figure 9b. The distance between rows of mirrors is arbitrary, but must satisfy the above constraints. The number of mirrors per row and number of rows are primarily determined by physical size and orientation of the target site.

Since the size of the mirrors can be physically large compared to the IFOV of the sensor, care must be taken that the sensor is really imaging a point source (Slater 1980). In practice, the convex mirror design results in an image of the Sun that is physically very small (on the order of a few millimeters). This can be shown in the following development. From basic optical theory:

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q} \quad (15)$$

where p = Object distance (approx. 1.5×10^{11} m); q = image distance; and f = focal length. For convex mirrors:

$$f = -R/2 \quad (16)$$

where R = the radius of curvature of the mirror. Substituting this relation in Equation 15 yields

$$\frac{-2}{R} = \frac{1}{p} + \frac{1}{q} \quad (17)$$

Now, since $q \ll p$, this relationship can be approximated as

$$q \approx -\frac{R}{2} \quad (18)$$

and it can be seen that q is on the order of 1 m in size for mirrors for this application. From geometric optics, magnification ' M ' can be defined as,

$$M = \frac{-q}{p} \quad (19)$$

Magnification can also be represented as:

$$M = \frac{D_i}{D_o} \quad (20)$$

where D_i = the diameter of the image, i.e. diameter of the image of Sun inside the mirror; D_o = the diameter of the object, i.e. diameter of Sun = 1.4×10^9 m. Rearranging this yields

$$D_i = -\frac{q}{p} D_o \quad (21)$$

Using the values for these parameters yields an image diameter for the Sun on the order of millimeters to

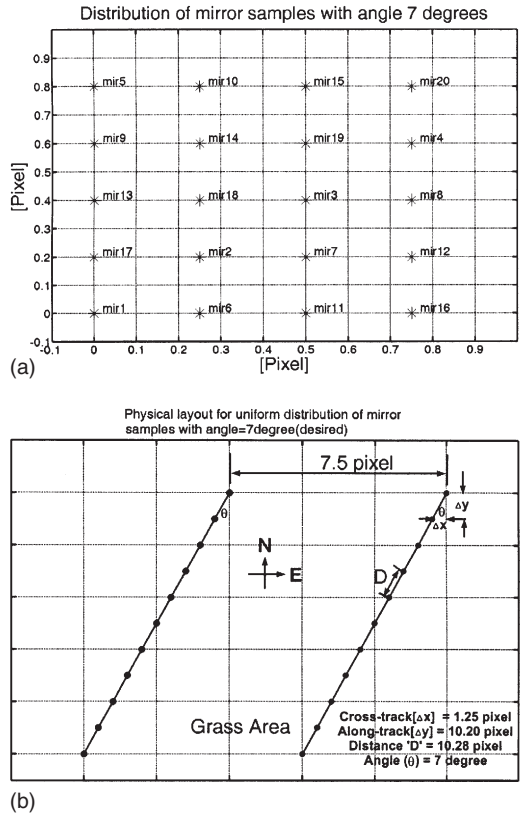


Figure 9. Physical layout of mirrors for a desired uniform sampling distribution: (a) Distribution in one GSI; (b) Physical mirror array layout.

centimeters in size. Thus, the Sun's image is significantly smaller than the IFOV of typical high spatial resolution sensors.

However, the mirror itself may be 1 m or so in diameter. Thus, the target seen by the imaging system consists of three components: the Sun; the surrounding mirror area; and the surrounding background region beyond the mirror. The surrounding mirror area is a reflection of the diffuse radiance obtained from the sky. The surrounding background region beyond the mirror is the combined direct and diffuse radiance components from the Sun and sky, respectively, reflected from that surface. Therefore, it is important for the surrounding background region beyond the mirror to have an upwelling radiance component similar to the amount of radiance produced by the surrounding mirror area in order for this target system to be viewed by the sensor as a bright point surrounded by a uniform darker background.

3 PROCESSING TECHNIQUES

Once the targets have been deployed and imaged, it is necessary to process the imagery to obtain estimates of the PSF/MTF. The type of processing required is largely dependent on the type of target used, e.g. an edge target yields an edge response function that can be differentiated to obtain a PSF. However, a pulse target gives a response function that is analyzed easily in the Fourier domain to obtain an MTF.

3.1 Non-parametric approaches

A decision to make when processing image data to obtain a PSF estimate is whether or not to follow a parametric or non-parametric approach, i.e. will the approach assume an underlying analytical model that accurately describes the imaging system? Both approaches have advantages and disadvantages.

3.1.1 Underlying model assumptions

One general approach is to assume no knowledge of the imaging system PSF. In this sense limitations on the shape of the PSF are minimized. As an example, many imaging systems can be modeled with an approximately Gaussian PSF. It is easier to estimate a few parameters that describe the Gaussian function than it is to estimate the value of the PSF at every point of its surface. However, if the underlying PSF is decidedly non-Gaussian, the parametric model will be inaccurate. Disadvantages of the non-parametric method include the need for additional processing and a susceptibility to noise. Thus, in practice it is generally best to follow a parametric method when an underlying model exists, and a non-parametric approach when such information is unavailable.

3.1.2 SNR definition

One important figure of merit for estimating the accuracy of PSF estimation is the SNR. Since there is no standard definition for SNR in PSF estimation, one is suggested in Figure 10 that is helpful. For an edge target, the difference between the dark and bright area pixel values (often measured in terms of digital numbers, or DN) is divided by the one-sigma standard deviation noise level to get SNR. For a pulse target, the difference between maximum average value and the background area brightness level were divided by the one-sigma standard deviation noise estimate of the background areas. Research on the relationship between SNR and MTF values at Nyquist was done by Choi (2003). Results of this work suggested a minimum SNR value existed ($\text{SNR} > 50$) for estimating stable MTF values at Nyquist with edge target methods.

3.1.3 Edge detection and alignment

Since both the edge and pulse targets produce 1-D PSF estimates based on aligning several 'slices' of image samples to obtain over-sampled reconstructions

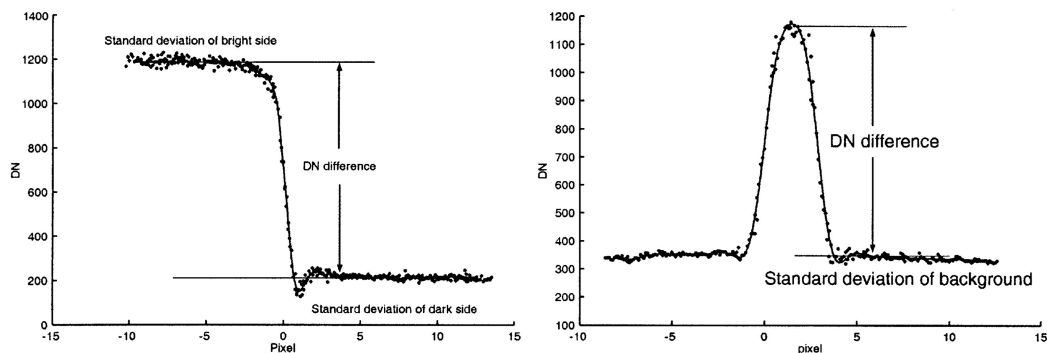


Figure 10. SNR for edge and pulse targets.

of the target, it is necessary to locate accurately the position of the edge in each of the slices to align them. When noise is present in the imagery, this can be difficult.

Instead of using a numerical edge detection method based on differentiation, a more accurate model-based parametric method was developed to detect sub-pixel edge locations. The Fermi function was chosen (Tzannes & Mooney 1995) to fit to the ESF:

$$f(x) = \frac{a}{\exp\left[\frac{(x-b)}{c}\right] - 1} + d \quad (22)$$

Another underlying assumption is that the edge of the target is a straight line. Any deviation from the line represents errors in the geometry of the image and a potential contribution to the overall MTF of the system. With this thought in mind, all edge cross-sections were forced to lie along a straight line. This final edge alignment process was achieved by fitting a line through the sub-pixel edge locations obtained from the previous step and declaring the actual edge locations to be on the line instead of where the previous algorithm step indicated the location to be. Results of the process are illustrated in §4.

3.1.4 Interpolation scheme

Three different data interpolation algorithms were developed: spline, sliding-window interpolation, and modified Savitzky-Golay (mSG) filtering. Spline interpolation was developed as an initial PSF interpolator for IKONOS MTF measurements (Helder 2001, 2002). Cubic spline interpolation was used to resample the profile of each line with a uniform spacing of 0.05 pixel. A second sliding-window interpolation method was designed to obtain uniformly spaced data points for an averaged profile. Both of these techniques smoothed the edge transition area, essentially averaging the profiles so that sharp transitions were not captured. In other words, the profile did not fully follow the data trends. This phenomenon significantly affected MTF estimates at higher frequencies. By using a generic sensor modeling approach (Choi 2003), static MTF estimation errors were calculated from these two methods. System frequency response from spline and sliding-window interpolation with edge targets under-estimated the system's original spatial resolution capability by as much as 55 percent and 13 percent, respectively. A novel interpolation method was needed that did not dampen high frequency components in the edge or pulse profile.

The aim of Savitzky-Golay (SG) filtering is to find filter coefficients that preserve higher moments to approximate the underlying function within a moving window, not by a constant whose estimate is an average, but by a polynomial of higher order, typically

quadratic. At a given filter location, a least-square fit polynomial is found for the data points in the moving window. The value of the polynomial at the center point in the window is then chosen as the output of the filter at that location. When the window is shifted to the next point, a whole new least-square process is performed for the new window. This is repeated at each location in the data set. Unfortunately, this method was only defined for equally spaced data.

A major problem is that data points after edge location alignment are unequally spaced, as shown in Figure 11. Thus, the SG filter was modified to accept unequally spaced data points as input, and produce equally spaced data points as output. This modified technique, termed the modified Savitzky-Golay filter (mSG), still fits a polynomial to the data within the window, but uses an alternative optimization method to find the polynomial coefficients. Once the polynomial coefficient values are known, the polynomial is evaluated at the center of the window to obtain an output value regardless of whether an input data point exists at that precise location. Figure 11 illustrates the mSG filter methodology. This approach was the most accurate of all approaches mentioned, but still underestimates the PSF slightly (on the order of 5 percent).

Once the edge or pulse response function has been aligned and interpolated, it is possible to determine the PSF or MTF. Only a differentiation step is needed for the edge-based method. The differentiated ESF is the LSF. This step was performed by a simple numeric differentiation described by

$$LSF(n) = ESF(n) - ESF(n+1) \quad (23)$$

As the differentiation process always reduces SNR, smoothing is sometimes employed in this step (Greer 2000). However, convolving with a spatial mask always smooths the ESF profile around the transition area,

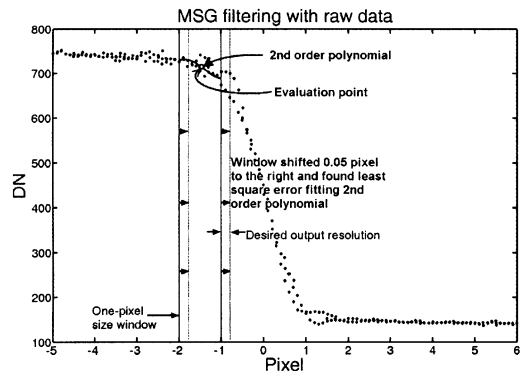


Figure 11. Modified SGolay filtering with 1-pixel window and polynomial fitting.

which will cause degradation of the final PSF/MTF estimate.

To avoid this undesirable smoothing and increase SNR, another SG filter was applied to the LSF. Normally a 1-pixel wide window and second order polynomial filtering are used. Lastly, the LSF profile was trimmed to reduce the noise present in the uniform background areas on either side of the edges. The overall length of the LSF should be significantly larger than the region of support for the PSF. Fourier transformation is applied to the trimmed LSF to obtain the MTF by dividing the absolute transformed function values by the first absolute value, i.e. the DC value.

Processing for pulse type targets is similar to edge targets with a couple of exceptions. First, one side of the pulse target must be chosen to perform the edge alignment step. Since the pulse response function is obtained after interpolation, the LSF cannot be found directly. Instead the function may be transformed via Fast Fourier transform and divided by the input sinc function to obtain the MTF after proper normalization.

3.2 Parametric approaches

Parametric approaches will be illustrated by point source targets. The algorithm first determines peak locations of each synthetic point source data set to sub-pixel accuracy using a parametric two-dimensional Gaussian model. The individual point source data sets are aligned along their model-estimated center positions to a common reference. Then, the 2-D model is applied again to estimate the overall PSF. A Fourier transform is applied to the PSF and normalized to obtain the corresponding MTF, as shown in Figure 12.

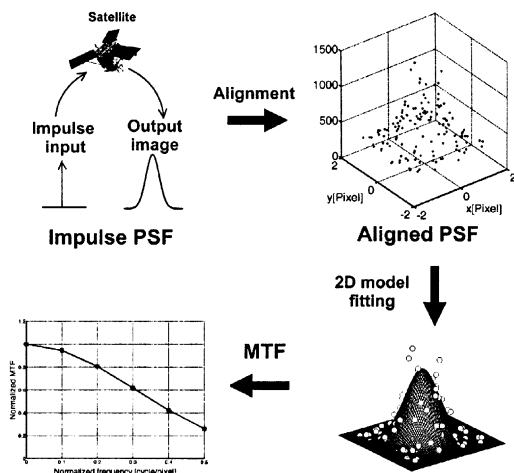


Figure 12. Impulse technique using parametric two-dimensional Gaussian model.

The various steps of the algorithm development are described in additional detail in the following sections.

3.2.1 Underlying model assumptions

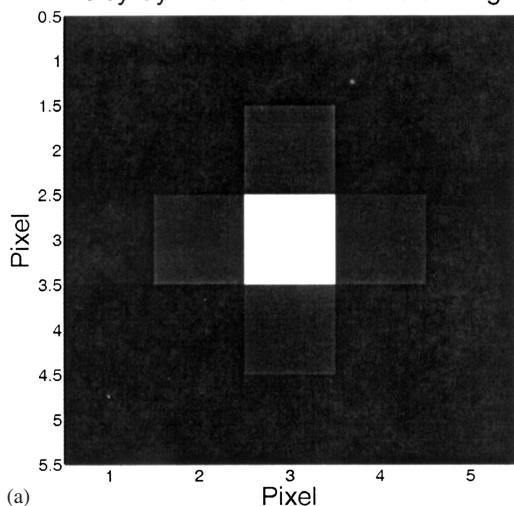
Often, the form of the PSF of an imaging system can be assumed to be known. This may be due to accurate measurement of the PSF prior to flight, or through well-understood modeling of the system. When this is the case, and in-flight validation of the PSF is required, an underlying model can be utilized to obtain more accurate estimates of the PSF during flight. This greatly simplifies the problem since only a limited number of model parameters need to be estimated, as opposed to estimating the value of the PSF at every location in its region of support. The effects of system noise are reduced using this approach. It is very useful also when the amount of over-sampling that can be obtained is limited. Thus, this method was found to be especially appropriate for use with point source targets. In this case, the over-sampling that can be obtained is limited by the physical size of the target array, and, at the same time, the dimensionality of the function being estimated is doubled. This process is illustrated in the following discussion. The underlying PSF model is assumed to be Gaussian, which is a commonly used PSF for many imaging systems.

3.2.2 Alignment of target samples

As in the 1-D cases discussed previously, the first step is to align the individual mirror responses. This can be difficult since the mirror array is designed precisely not to be centered or aligned with the imaging sensor sampling grid. An example is shown in Figure 13a, b. Figure 13a shows the pixels for a synthetic image of one mirror. Figure 13b shows a plot of the values of these pixels with respect to their image locations. With this sparse data set, it is difficult to determine that the underlying PSF is Gaussian. Since there is no guarantee that the brightest pixel represents an image sample precisely centered on the mirror location, it is necessary to estimate where the mirror was located with respect to the sampling grid. This is done by fitting a Gaussian surface to the data. The particular algorithm used to accomplish this is not critical. A commonly used method is the Nelder-Mead downhill simplex algorithm (Lagarisa 1998). An example of the resulting surface is shown in Figure 13c. The location of the peak of this surface is an estimate of the location of the mirror with respect to the imaging system's sampling grid.

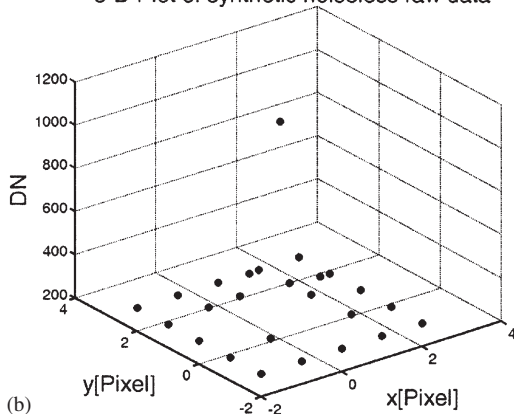
Successful use of any optimization algorithm requires initial estimates of model parameters. These can be obtained from inspection of the pixel samples. Results of a Gaussian fit to the data in Figure 13. Table 1 lists actual and model predicted peak locations of the example data.

Noisy Synthetic Point Source-5 Image



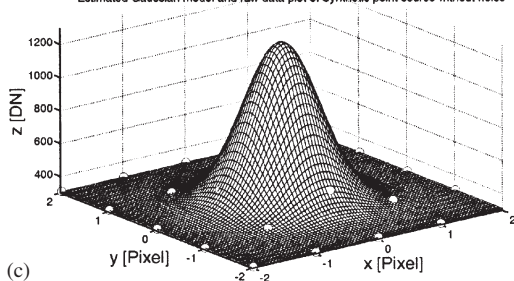
(a)

3-D Plot of synthetic noiseless raw data



(b)

Estimated Gaussian model and raw data plot of Synthetic point source without noise



(c)

Figure 13. Example estimating peak positions of mirror data: (a) mirror image; (b) raw data; (c) 2-D Gaussian model estimation.

Table 1. Comparing actual and model parameters.

Noiseless synthetic point sources	X-mean position (cross-track) [pixel]	Y-mean position (along-track) [pixel]	Peak DN
Actual	2.96	2.95	1196
Model	2.97	2.95	1200
Initial values	3.0	3.0	1000

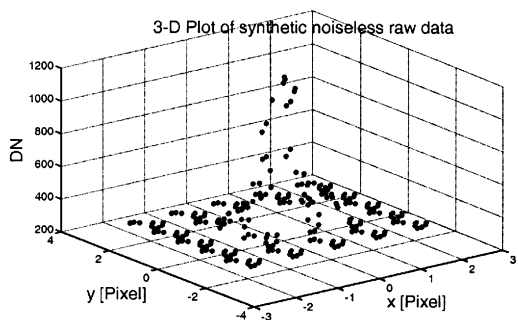


Figure 14. Aligned point source data.

Once model estimates of peak location were obtained from all of the available point source data sets, the point source data sets can be aligned such that the estimated model peak locations are at a common reference location as shown in Figure 14. Twenty simulated noiseless data sets were used and the results were limited to a 3×3 window. While the data are still rather sparse, an overall surface shape is apparent.

3.2.3 PSF/MTF estimation

Once the data from all mirror responses are aligned, a final 2-D Gaussian surface can be fitted using the same approach described for individual mirror responses. The resulting model should be centered at the origin. The standard deviations in the 'X' (across-track) and 'Y' (along-track) directions are indications of the width of the marginal 1-D PSF functions. One-dimensional slices of this function are shown in Figure 15a and 15b. Full-width half-maximums (FWHM) values for the marginal PSF's can be calculated directly from these standard deviations. Fourier transformation of the surface leads to a 2-D MTF function.

4 HIGH-SPATIAL-RESOLUTION SENSOR EXAMPLES

Examples and processing methods of the use of various targets are presented for both the IKONOS and QuickBird sensors. A target site has been developed in Brookings, SD, consisting of gently rolling grassy

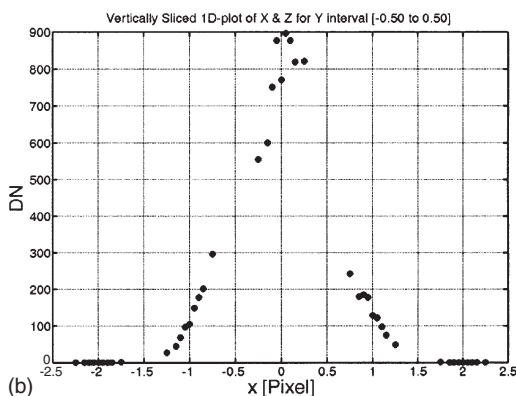
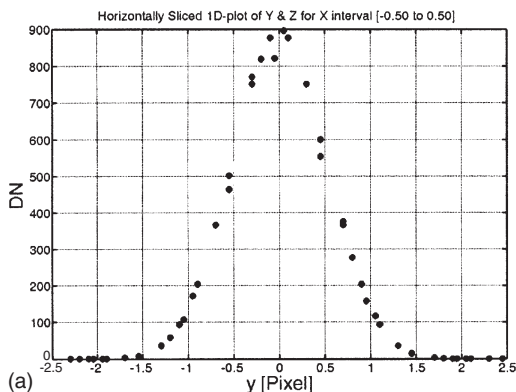


Figure 15. Estimated model PSF of synthetic point source without noise in the X and Y directions: (a) 1-D slice in X direction; (b) 1-D slice in Y direction.

region approximately 2.8 hectares in size. For PSF/MTF characterization of high-spatial resolution-sensors, a variety of targets are deployed as shown in Figure 16. These consist of an edge target (lower left), pulse target (upper left), and a point source array (right edge).

4.1 Edge method procedure

Four plots illustrating the processing procedure are shown in Figure 17 (a–d). Figure 17a shows edge detection and the least square error line. Edges for each image ‘slice’ are designated with a circle. The center of the circle is aligned where the estimated center of that slice is located. The line intersecting the circles is the least squares estimate of the complete edge. It also gives the estimated edge angle.

Each pixel from Figure 17a contributes a data point in Figure 17b to construct the over-sampled aligned edge profile. Second order polynomial fitting and a one-pixel window mSG filter was applied to

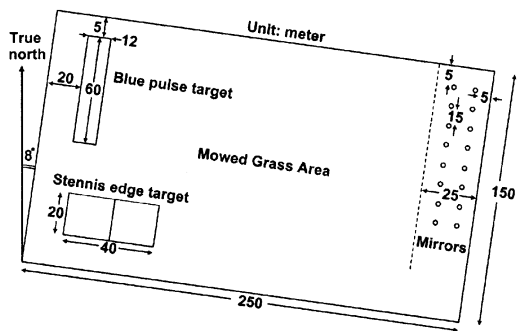


Figure 16. Target layout plan.

this non-uniformly distributed data to produce a uniformly sampled profile, which is shown as a solid line in Figure 17a. Differentiation and normal fourth order Savitzky-Golay filtering was applied to obtain the smooth LSF shown in Figure 17c. Finally, Fourier transformation was applied to obtain the MTF plot in Figure 17d. This particular example had an excellent SNR of 141.

Overall results using this method on all three dates from 2002 show that FWHM values varied from 1.43 to 1.57 pixels. Larger FWHM values contributed the smaller MTF values at Nyquist, and similarly, smaller FWHM values produced larger MTF values. Conversely, sharper LSF profiles in the spatial domain contain more high frequency components in the Fourier domain. The process is repeatable using this technique. FWHM values ranged from 1.44 to 1.57 pixels and MTF at Nyquist ranged from 0.13 to 0.18.

4.2 Pulse method procedure

A processing method sequence similar to the edge method was performed also for a pulse target. The target consisted of eight tarps 3 m by 30 m in size and arranged in a 4×2 pattern to produce a width optimized for the IKONOS sensor’s multispectral bands. This sensor has a multispectral GSI of approximately 4 m. While these tarps are not spectrally flat, they are very bright in the blue ($0.5 \mu\text{m}$ reflectance) and the near infrared ($>0.6 \mu\text{m}$ reflectance) bands. The data from an IKONOS overpass of the Brookings test site on July 3, 2002. Figure 18a shows the sub-pixel location of the left edge of the tarps in a row-by-row fashion as identified by the circles. The complete pulse profile after alignment of all rows ($=102.03$) (Fig. 18b) showed excellent SNR using the mSG interpolation filter. Superimposing the interpolated pulse response function with the input pulse function and FWHM resulted in an FWHM of 2.8689 pixel (Fig. 18c). The MTF at Nyquist was 0.4511. Figure 18a–e illustrates the process sequence. Figure 18d, e shows

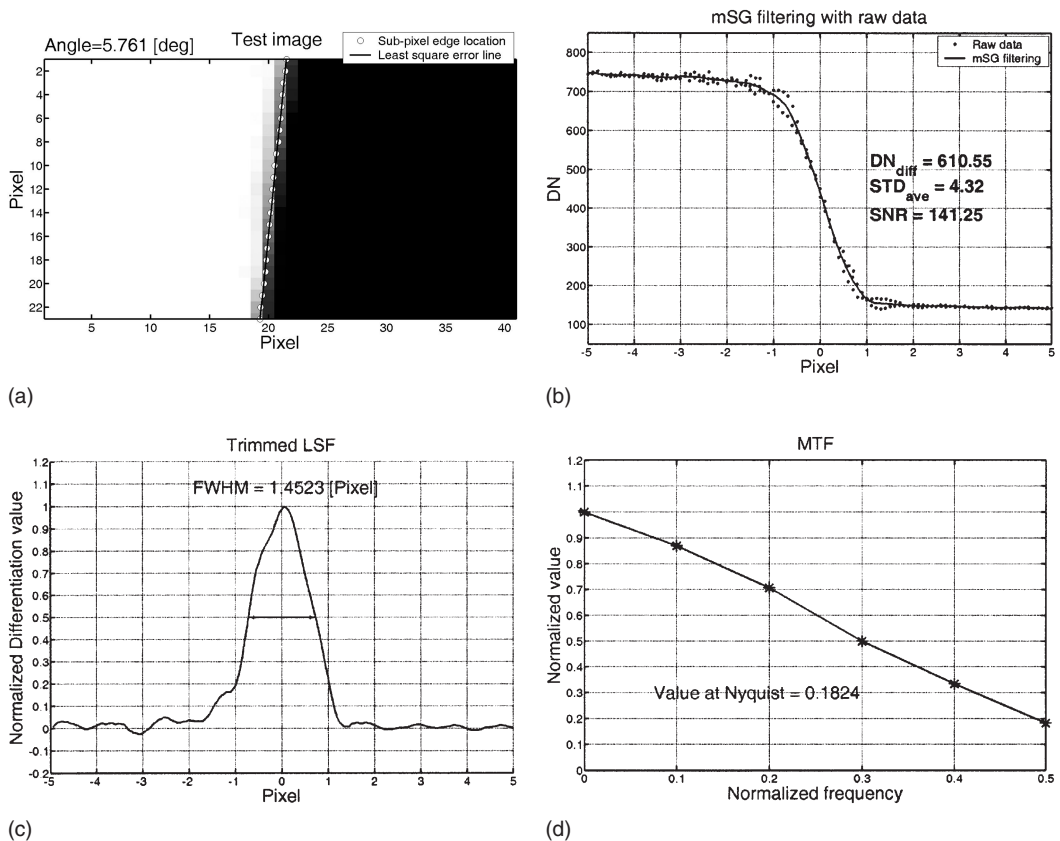


Figure 17. Example of panchromatic band analysis procedure: (a) edge detection; (b) mSG interpolation of edge (c) 4th order smoothed SG-derived LSF; (d) Fourier transform of MTF at Nyquist.

the Fourier domain input and output, and the resulting MTF function, respectively.

Results from analysis of the IKONOS Blue band is shown in Table 2 for three dates in 2002 for the Brookings site. Note that the SNR varied over a broader range of values for these three dates, as did the FWHM and MTF estimates. This is typical of a pulse type target since there are two edges, and the background region on either side is not controlled precisely.

4.3 Point target using QuickBird data

As an example of point source targets and parametric modeling, two QuickBird images are presented from August 25 and September 7, 2002 overpasses of the Brookings, SD site. On both dates, an array of convex mirrors was deployed. A 2-D Gaussian model was applied to the individual mirror responses to estimate the peak position and their corresponding peak DN values.

Figure 19a–d show the data for mirror #7 from the September 7 scene. It is apparent, visually and from examination of the DN values within the cropped window, that the blurring is asymmetric in both the cross- and along-track directions. Application of the proposed model to this data set yields peak position estimates that are shifted about 0.36 pixels right and about 0.72 pixels above the center of the brightest pixel.

Similarly, a 2-D Gaussian model was applied to all the mirror responses of August 25 and September 7, 2002 QuickBird data to estimate their peak positions. Once estimates of peak location were obtained for all mirrors on both days, the mirror data were aligned to obtain the oversampled PSF of the mirror responses shown in Figure 20a,b.

Figure 20c,d display the least square error Gaussian surface and the aligned mirror data of August 25 and September 7 respectively. The Gaussian surface model does not pass through the data from mirrors #4, #15 and #17 from the August

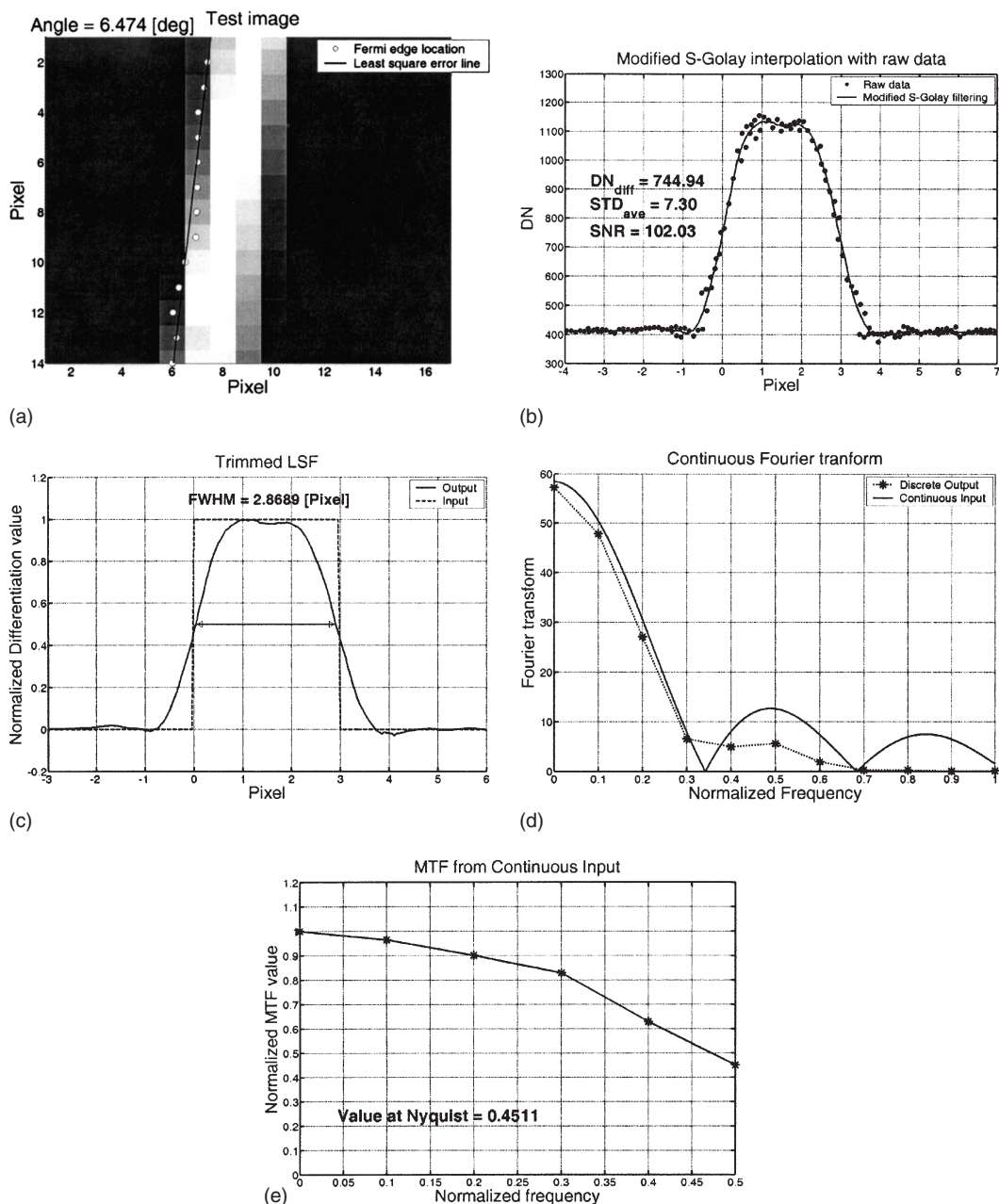


Figure 18. IKONOS blue band pulse method procedure: (a) subpixel left edge of the blue tarp; (b) pulse profile (dots = raw data, solid line is mSG interpolation); (c) superimposed input/output data and calculated FWHM; (d) input/output Fourier domain; (e) MTF at Nyquist.

25 scene. Those mirror data were most likely degraded due to errors in mirror placement, as well as effects due to a spectrally non-uniform and spatially uneven background. Similarly poor fits to the model

surface were observed with the data from mirrors #2, #3 and #7 in the September 7 scene. The model parameters and RMSE values of both dates are shown in the figures and are listed in Table 3. The peak

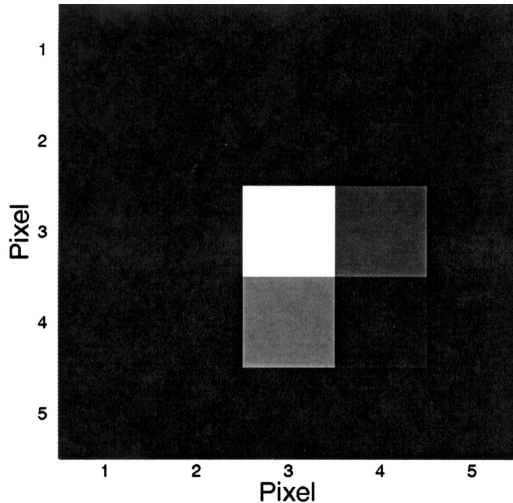
position on both days is at the center (0,0). The ‘Ystd’ (0.59) is greater than ‘Xstd’ (0.58) on August 25, indicating more blur in the along-track compared to cross-track. This is attributed most likely to the motion of the satellite in the along-track direction. A similar increase in blur is observed on September 7 (Xstd = 0.57, Ystd = 0.58). On both dates, the error between the surface model and the actual data, as given by the

Table 2. IKONOS blue band analysis on 2002.

Date	FWHM (pixel)	MTF at Nyquist	SNR
6/27/2002	2.9149	0.4722	55.7
7/03/2002	2.8689	0.4511	102.0
7/22/2002	2.8336	0.3347	82.1

RMSE, is fairly small. The mean DN value of the grass background from August 25 (335) is greater than September 7 (306). The grass background on August 25 was more reflective than on September 7 indicating that the grass was brighter and probably drier on August 25. The standard deviation (Noise Stdev) of the grass was less on September 7 than on August 25, indicating slightly more uniformity on September 7.

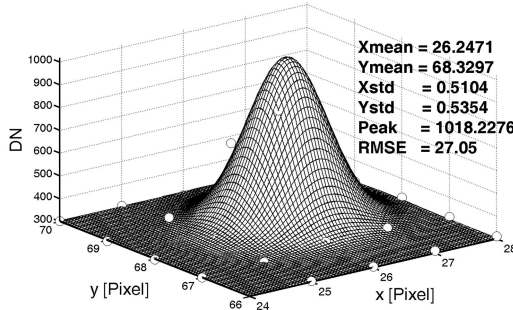
Marginal PSF’s in the along-track and across-track directions were obtained from the two-dimensional PSF to compare the FWHM measurements with the results obtained from using a 1-D edge target. First, Figure 21a shows the PSF plots of August 25 (solid line) and September 07 (dashed line) in cross-track and Figure 21b in the along-track direction. The FWHM value of August 25 data is greater than September 07 data, as shown in Table 4. The relative



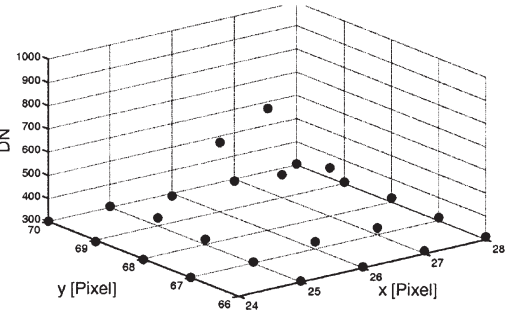
(a)

305	296	255	294	324
312	371	781	674	290
317	410	1315	972	211
306	325	385	358	304
303	310	315	310	296

(b)



(c)



(d)

Figure 19. Peak estimation of Sept 7, 2002 mirror #7 data: (a) mirror #7 image; (b) pixel values; (c) raw data; and, (d) 2-D Gaussian model.

error between August 25 and September 7 FWHM values for cross-track and along-track is shown also in Table 4.

The increased blurring on August 25 may result from atmospheric scattering and manual errors during mirror placement. The humidity on August 25 was 60 percent, compared to 48 percent on September 07, which may be responsible for the increased blurring. From Table 4 it is evident that the FWHM in cross-track and along-track direction on either day are about the same except for a slight increase in the along-track direction compared to their respective cross-track FWHM. The difference (0.001 pixel on August, 0.002 pixel on September 07), even though small, may be attributable to blur caused by satellite motion in the along-track direction. Note that there is excellent agreement between the edge target approach FWHM = 1.45 from Figure 17c) and the point source target approach (average 2-date FWHM = 1.41) in estimating the cross-track FWHM for the QuickBird sensor.

The measured MTF values also indicate similar trends in along-track and cross-track directions as the FWHM values, except that, as the FWHM value increases, the MTF value decreases. Figure 22a shows the MTF plots of August 25 (asterisks) and

Table 3. Comparison of model parameters, background data and SNR of August 25 and September 7, 2002 mirror data.

Parameters	August 25, 2002	September 7, 2002
Xmean [pixel] [cross-track]	0	0
Ymean [pixel] [along-track]	0	0
Xstd [pixel] [cross-track]	0.58	0.57
Ystd [pixel] [along-track]	0.59	0.58
Peak [DN]	1390	1191
RMSE	61.04	66.79
Noise Mean [DN]	335.18	305.64
Noise standard deviation	10.16	9.85
SNR	103.67	90.24

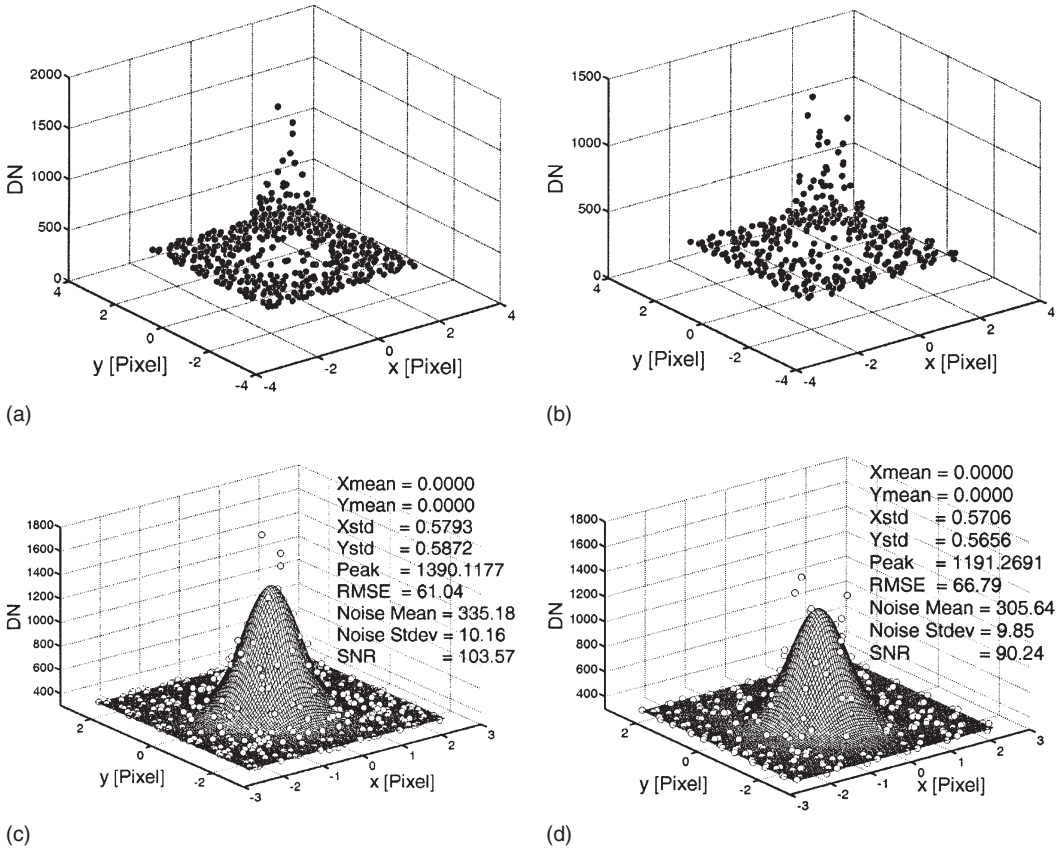


Figure 20. Gaussian Surface fit for aligned mirror data of August 25 and September 7, 2002, QuickBird images: (a) August 25 mirror data; (b) September 7 mirror data; (c) August 2-D PSF; and, (d) September 2-d PSF.

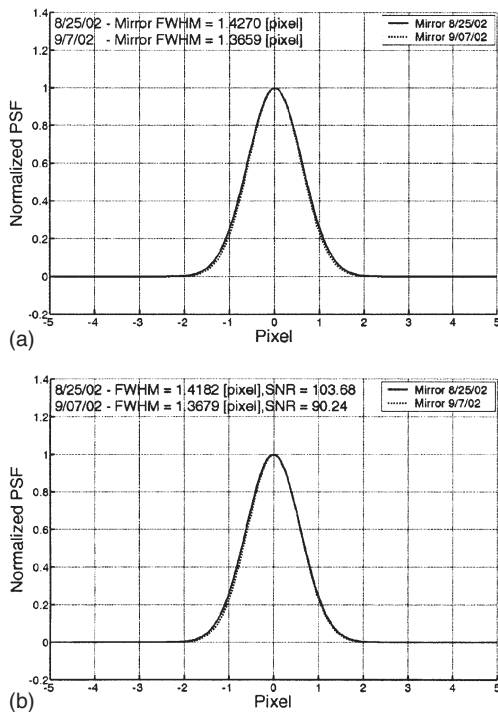


Figure 21. August 25 and September 27, 2002 PSF plots for QuickBird mirror data: (a) cross-track direction; and (b) along-track direction.

Table 4. Comparison of FWHM from August 25 and September 07, 2002 QuickBird mirror data.

Mirror data	Full-Width at Half-Maximum measurement [FWHM]	
Overpass date	Cross-track [pixel]	Along-track [pixel]
August 25, 2002	1.427	1.428
September 7, 2002	1.396	1.398
Relative Error (%)	2.17	2.10

September 07 (diamonds) in the cross-track direction. In Figure 22b for the along-track direction, the August 25 data are shown as diamonds; the September 7 data, as asterisks. Table 5 compares MTF values at the Nyquist frequency, which is a commonly used value for specifying system performance.

5 CONCLUSIONS

In-flight estimation of PSF and MTF is possible with suitably designed targets well adapted for the type of

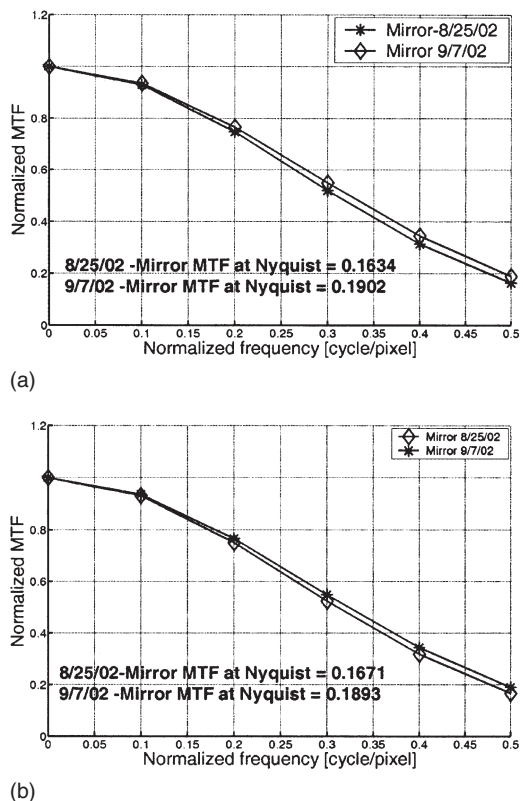


Figure 22. 1-D MTF estimates from point sources: (a) August 25 and September 07 MTF plots in cross-track direction; and (b) August 25 and September 07 MTF plots in along-track direction.

Table 5. Comparison of MTF from August 25 and September 7, 2002 QuickBird mirror data.

Mirror data	MTF values @ Nyquist	
Days/direction	Cross-track	Along-track
August 25, 2002	0.163	0.162
September 7, 2002	0.190	0.189
Relative Error (%)	16.60	16.70

sensor under evaluation. The spatial and spectral properties of the sensor are major considerations when targets are developed. In general, sensors with larger GSI's are more difficult to evaluate. With the current progress in high spatial resolution systems, it may also become more difficult to build appropriate targets. Point source approximations will necessarily become smaller, and target material uniformity may be more difficult to obtain. However, with current

resolutions in the sub-meter to several meter range, appropriate targets can be fielded to obtain PSF and MTF estimates.

Edge targets are most commonly used. An advantage of this is that the imaging system's response to an edge is intuitively attractive and can be well understood by laymen. In addition, it is easy to derive a 1-D PSF. Targets are easy to maintain and deploy, very high SNR can be obtained leading to accurate PSF estimates, and a variety of materials and approaches can be used to develop such a target. Drawbacks are that only a 1-D estimate is possible and larger sensor GSI make this target difficult to deploy. In many ways, this is probably the best 1-D in-flight target for most applications.

Pulse targets are useful for 1-D PSF estimates in situations where the GSI of the sensor makes an edge target difficult to deploy. While good results can be obtained, limiting factors include controlling background regions on either side of the pulse, avoiding zero-crossings when designing the target, and directly deriving the PSF. However, this target may be optimal for a range of spatial resolutions that are larger than what is feasible for an edge target.

Point source targets have not been as widely used historically, but show great promise for sensors in the sub-meter to several meter GSI range. These targets can be developed from convex mirrors and placed in an array that can be shaped to best fit the available test site. These targets provide a true 2-D estimate of the PSF. This promising approach agrees well with results from 1-D edge target approaches. More work is warranted for optimizing this method.

Processing techniques are important to obtain accurate PSF and MTF estimates. Improper processing can result in estimates that have a large systematic bias due to the PSF of the processing system itself. Thus, it is important to use processing techniques that contribute minimally to the PSF estimate.

Processing techniques can be divided largely into two forms: parametric and non-parametric methods. Non-parametric methods are most advantageous when not enough information is known about the imaging system to assume an underlying model. In these cases, non-parametric processing can provide a truer picture of the PSF. In fact, a non-parametric method can be used potentially in all PSF estimation situations. It is clearly sub-optimal when an underlying model exists and can be used confidently. In these situations, the advantages of a parametric approach clearly outweigh the versatility of the non-parametric method. Parametric approaches are significantly less sensitive to noise, and estimation of a few parameters is far easier than estimation of an entire function. In some situations, a non-parametric method may be an effective initial estimate of the PSF. Results from this first analysis may indicate that an underlying model is

appropriate. From there a parametric technique can be employed to estimate the model parameters.

While in-flight PSF estimation techniques have been available for several decades, the advent of high spatial resolution commercial sensors require their further refinement so that accurate estimates of image quality can be made during the operational mode of these systems. With properly designed targets and processing techniques, these estimates can be obtained accurately to ensure adequate quality of the imagery.

ACKNOWLEDGMENTS

The authors would like to thank the JACIE team at Stennis Space Center for their support. This project was funded under NASA grant NAG1303023.

REFERENCES

- Cao, X., Huang, H.K. & Shyhliang, A.L. 2000. A novel algorithm for measuring the MTF of a digital radiographic system with a CCD array detector. In J.T. Dobbins III & J.M. Boone (eds), *Medical imaging 2000: physics of medical imaging*. Proc. SPIE 3977: 580–589.
- Choi, T. 2003. Generic sensor modeling. In *Proceedings of the 2003 High spatial resolution commercial imagery workshop*, Reston, VA, USA, 19–21 May 2003. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team, CD-ROM.
- Coltman, J. W. 1954. The specification of imaging properties by response to a sine wave input. *J. Opt. Soc. Am.* 44(6): 468–471.
- Droege, R.T. 1979. A megavoltage MTF measurement technique for metal screen film detectors. *Med. Phys.* 6(4): 272–279.
- Evans, W. 1974. Marking ERTS images with a small mirror reflector. *Photogr. Eng.* 40(6): 665–671.
- Feltz, J.C. 1990. Development of the modulation transfer function and contrast transfer function for discrete systems, particularly charge-coupled devices. *Opt. Eng.* 29(8): 893–904.
- Greer, P.B. 2000. Evaluation of an algorithm for the assessment of the MTF using an edge method. *Med. Phys.* 27(9): 2048–2059.
- Hecht, E. 1998. *Optics*. (3rd ed.) Reading, MA: Addison-Wesley.
- Helder, D. 2001. Techniques for on-orbit spatial characteristics of IKONOS. In *Proceedings of the 2003 High spatial resolution commercial imagery workshop*, Reston, VA, USA, 19–21 May 2003. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team, CD-ROM.
- Helder, D. 2002. On-orbit spatial characteristics of IKONOS. In *Proceedings of the 2003 High spatial resolution commercial imagery workshop*, Reston, VA, USA, 19–21 May 2003. NASA/NIMA/USGS Joint Agency Commercial Imagery Evaluation Team, CD-ROM.

- Hong, Q.H., Lettington, A.H. & Macdonald, J. 1996. Measuring the MTF for focal plane arrays using random noise targets. *Meas. Sci. Technol.* 7(7): 1087–1091.
- Irvine, J.M. 1996. A methodology for developing image interpretability rating scales. *ASPRS/ACSM Annual Convention & Exposition Technical Papers* 1: 273–281. Bethesda: ASPRS/ACSM.
- Kaftandjian, V. 1996. A comparison of the ball, wire, edge, and bar/space pattern techniques for modulation transfer function measurements of linear x-ray detectors. *J. X-ray Sci. & Tech.* 6(6): 205–221.
- Lagarias, J.C. 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optimiz.* 9(1): 112–147.
- Leachtenauer, J.C. 1996. National imagery interpretability rating scales overview and product description. *ASPRS/ACSM Annual Convention & Exposition Technical Papers* 1: 262–272. Bethesda: ASPRS/ACSM.
- Markham, B. 1985. The Landsat sensor spatial response. *IEEE Trans. Geosci. & Rem. Sens.* 23(6): 864–875.
- Rauchmiller, R. 1988. Measurement of the Landsat Thematic Mapper modulation transfer function using an array of point sources. *Opt. Eng.* 27(4): 334–343.
- Samei, E. 1998. A method for measuring the presampled MTF of digital radiographic systems using an edge test device. *Amer. Assoc. Phys. Med.* 25(1): 102–113.
- Schowengerdt, R. 1985. Operational MTF for Landsat thematic mapper. In L.R. Baker & E.M. Granger (eds), *Image quality: an overview*. SPIE 549: 110–116.
- Slater, P.N. 1980. *Remote sensing optics and optical systems*. Reading, MA: Addison-Wesley.
- Storey, J. 2001. Landsat 7 on-orbit modulation transfer function estimation. In H. Fujisada, J.B. Lurie & K. Weber (eds), *Sensors, systems, and next-generation satellites V*. Proc. SPIE 4540: 50–61.
- Tzannes, A.P. & Mooney, J.M. 1995. Measurement of the modulation transfer function of infrared cameras. *Opt. Eng.* 34(6): 1808–1817.

Test ranges for metric calibration and validation of high-resolution satellite imaging systems

G. Dial & J. Grodecki

Space Imaging, Thornton, Colorado, USA

ABSTRACT: Metric test ranges for satellite imaging systems have photo-identifiable features with surveyed coordinates. Images of these test ranges are used to determine calibration parameters and to test system accuracy. The size of the test range, layout of the ground control points (GCP), and choice of photo-identifiable features determines the utility of the test range and the accuracy of calibrations derived from it. In this paper, the historical development of test ranges is traced for IKONOS™, as well as the application of those ranges verifying design requirements, improving calibrations, and testing system accuracy. Lessons learned regarding ground control points also are discussed. Test ranges used for IKONOS will be exhibited and the utility of those ranges for various calibrations discussed.

1 INTRODUCTION

The IKONOS satellite launched September 24, 1999 provides high-resolution panchromatic and multispectral satellite imagery. At nadir, the swath width is 11.3 km and the ground sample distance is 0.82 m for the panchromatic sensor and 3.28 m for the multispectral sensor. The multispectral blue, green, red, and NIR bands approximate Landsat bands 1–4. From its 681 km high orbit, IKONOS provides 3-day revisit to mid-latitude locations.

Of particular interest is the metric accuracy of the IKONOS system. IKONOS design requirements for metric accuracy without ground control were 25 m circular error 90 percent (CE90) horizontal and 22 m linear error 90 percent (LE90) vertical accuracy for a stereo pair and 50 m CE90 for a monoscopic image exclusive of terrain displacement. After recalibration and test, IKONOS achieved 6 m CE90 horizontal and 10 m LE90 vertical accuracy for stereo and 8 m CE90 exclusive of terrain displacement for monoscopic images without ground control. Relative accuracy is similarly impressive: IKONOS imagery with ground control achieves 2 m CE90 horizontal accuracy.

Specialized test ranges with high quality ground control points were required to calibrate and validate the accuracy of the system. The paper begins with the test ranges established before launch, continues with specialized tests used for in-flight calibration and test, and concludes with a discussion of ground control points (GCPs).

2 PRE-LAUNCH TEST RANGES

Two test ranges were established pre-launch, one in San Diego, California, and the other in rural Western Australia. The purpose of the San Diego test range was primarily to verify that the satellite met its design requirements. These test results, reported in Cook et al. (2001), were part of the sell-off of the satellite from Lockheed to Space Imaging. The Southern Test Range in Australia was designed specifically to test satellite calibration. This pair of test ranges tested performance in the northern and southern, as well as the eastern and western, hemispheres.

2.1 *Planning a test range*

Considerations for planning a test range include (1) geodetic control, (2) weather, (3) availability of suitable ground control features, and (3) access.

If the test range is designed to test the absolute accuracy of satellite imagery, it is necessary to know exactly where the test range is. Geodetic grade GPS equipment can measure relative displacement quickly and accurately from a known point. Advanced countries have high quality geodetic control networks, such as Australia's continuously-operating global positioning system (GPS) network shown in Figure 1. Ideally, the test range should be located near to a high accuracy geodetic control point.

Weather is also a primary consideration. The most exquisite test range is of little value if it cannot be seen

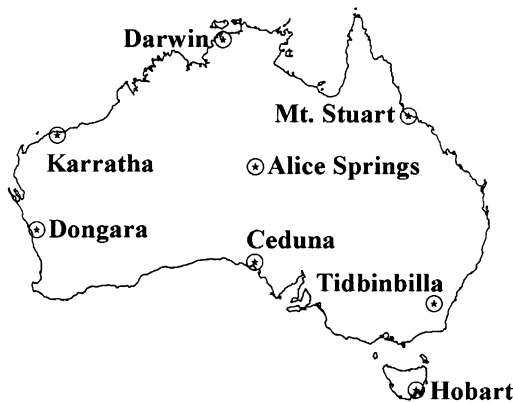


Figure 1. Continuous GPS stations in Australia.

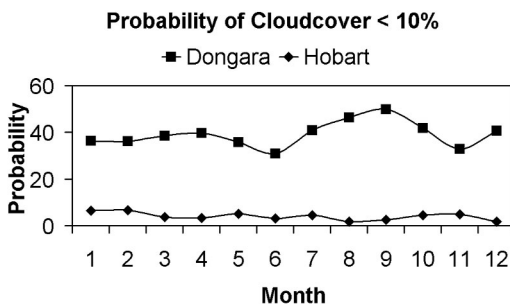


Figure 2. Cloud cover prediction for Dongara and Hobart Australia.

through persistent cloud cover. Compare the predicted weather for Dongara and Hobart shown in Figure 2. The probability that clouds will cover less than 10 percent of Dongara is about 40 percent, whereas the probability for less than 10 percent cloud cover in Hobart is nearly zero. The predictions from the weather database are not completely reliable; indeed, an image was successfully acquired from a research test range in Hobart on the first attempt (Fraser et al. 2003, Fraser & Yamakawa 2003). Nonetheless, sites should be selected for high probability of clear acquisition to minimize the time required to calibrate and test the satellite.

A further consideration is availability of suitable control points. It is expensive to place permanent markers on the test range and to re-deploy those markers for every imaging access, so availability of permanent features for ground control becomes important. It is difficult to use natural features for ground control. Bushes, trees, streams, and rocks do not possess the well-defined geometric features permitting sub-pixel positioning. Because paneling is not generally feasible, and natural features typically lack well-defined, photo-identifiable features, cultural features become

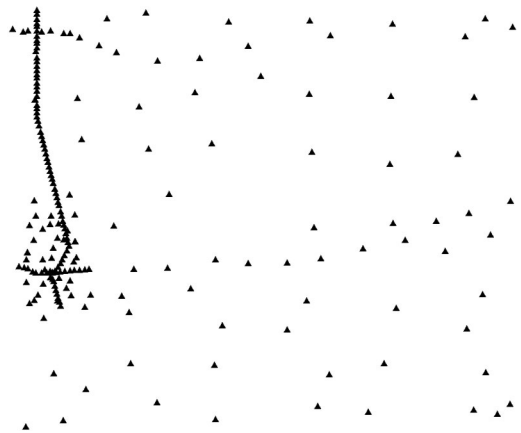


Figure 3. GCP layout for Southern Test Range.

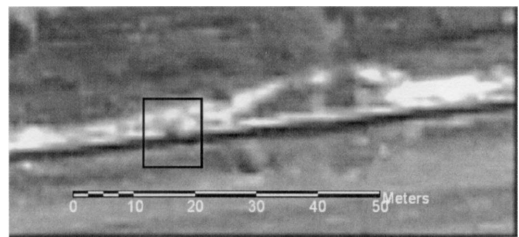


Figure 4. 1 m by 2 m black rectangle on 1 m white pipe.

the ground control features of choice. Availability of suitable ground control features leads one to use large city and suburban areas for test ranges. Cities in the desert southwest of the United States are particularly suitable, considering their weather and gridded road networks.

2.2 Southern test range

The Southern Test Range in rural Western Australia is 100 km by 100 km with 200 ground control points (GCP)s laid out as shown in Figure 3. The intent of this test range was to test large area mapping with a minimum of control points. The west-most strip includes two strong east-west lines of ground control straddling a north-south line. The east-west lines were intended to test IKONOS interior orientation while the north-south line was to test exterior orientation. Unfortunately, the features chosen were mostly too small or had insufficient contrast to be reliably identified in test range imagery. An example of such a feature consisting of a 1 m by 2 m black paint spot on a white pipeline of 1 m diameter is shown in Figure 4.

The rectangle in Figure 4 identifies the black paint mark finally located by a photo interpreter with eight

years of experience working at the National Geospatial and Intelligence Agency (NGA), formerly the National Imaging and Mapping Agency (NIMA). The interpreter reasoned that the black line was the shadow of the pipeline, and that the white line above it was the pipeline itself, and found the black marks spaced 1 km apart along the pipeline. This experience convinced the authors of the importance of human sensory perception in spotting GCPs. The southern test range pipeline marks made an important contribution to the interior orientation recalibration described below.

2.3 Northern test range

Space Imaging established its northern test range in San Diego, California. Figure 5 is an overview of the 20 km NS by 30 km EW area. It is populated with 140 cultural features. The width of the test range was selected to be three IKONOS strips wide.

Features were selected and measured from a 1:600 precision aerial orthophoto with 0.5 m CE90 horizontal accuracy and an accompanying digital elevation model (DEM) with 1 m LE90 vertical accuracy.

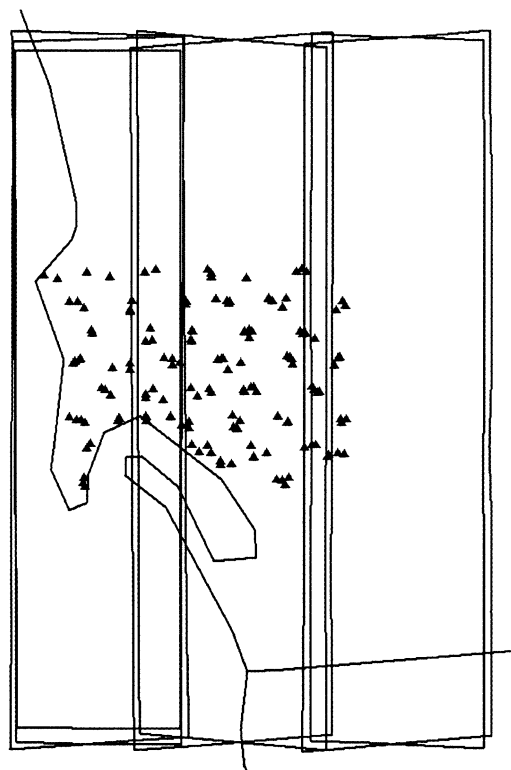


Figure 5. San Diego Test Range.

Selecting features from large-scale photography was convenient because markings on playgrounds, streets, and helicopter pads could be used without exposing the surveyor and equipment to the hazards presented by children, automobiles, or helicopters.

Initial imagery of the northern test range was acquired between January 8 and February 7, 2000. A stereo triplet, consisting of a stereo pair and an interleaved monoscopic image acquired on the same orbital pass, was acquired for the western strip. Same-pass stereo pairs were acquired for the central and eastern strips. Analysis of these data, reported in (Grodecki & Dial 2001), showed that IKONOS met all design requirements for metric accuracy. These same images, known as the “San Diego 7” have on occasion been made available to university researchers and software developers to verify that user software measurements are consistent with IKONOS ground station results.

3 TEST RANGES FOR CALIBRATION OF INTERIOR AND EXTERIOR ORIENTATION

The concepts of interior and exterior orientation of a simple imaging system are illustrated in Figure 6.

Interior orientation comprises everything interior to the camera from the perspective center of the optical system to the image detectors. Physically, interior orientation results from mechanical alignments and optical parameters such as focal length. The Field Angle Map (FAM) describes the calibrated relationship between detectors and field angles external to the camera.

Exterior orientation comprises the position and orientation (pointing direction) of the camera. For

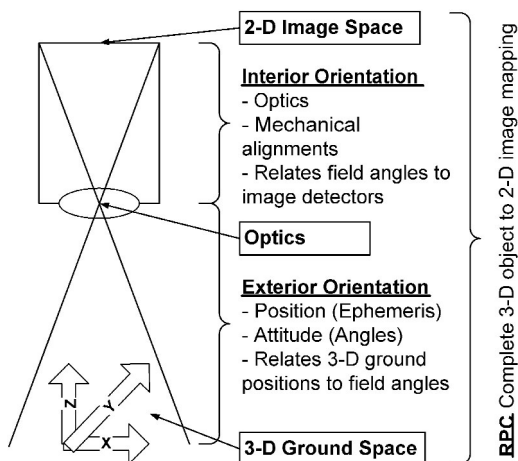


Figure 6. Interior and exterior orientation.

satellite imaging systems, position is derived from the satellite ephemeris and orientation is derived from the attitude profile. Both position and attitude are functions of time that must be correlated with the imagery.

3.1 Test range for interior orientation

It first became known that the original IKONOS calibration for interior orientation could be improved when studying a block of six side-by-side stereo strips. When the imagery was adjusted with just one GCP to correct absolute error, the residual error pattern was strongly suggestive that there was a small error in pixel scale of about 200 parts-per-million. This scale error accumulated from the GCPs outward to the edges of the block. Figure 7 shows the image block and Figure 8 shows the residual error pattern.

A special test range was established near the main office of Space Imaging. An area of gridded roads was selected. Space Imaging personnel selected and surveyed the GCPs. The points were arrayed along an east-west line to be parallel to the sensor array. All points would be acquired at nearly the same instant of time, minimizing the impact of attitude changes during image acquisition. The result was the "128th street" test range shown in Figure 9.

Analysis of the data from the 128th Street test range resulted in a new FAM calibration (Grodecki & Dial 2002) put online on November 19, 2001. After recalibration, the test range residuals were less than one pixel (worst case), a remarkable achievement considering the difficulty of correctly locating GCPs and the near-perfect calibration required. The calibration was verified with new imagery of the pipeline from the southern test range. Reprocessing the six stereo pairs from Figure 7 that originally showed the problem did the final verification. Residual errors after recalibration clearly are evident by comparing Figure 8 with Figure 10.

3.2 Test range for exterior orientation

Exterior orientation is defined by the position and orientation of the camera. For a satellite imaging system, position is determined by the satellite ephemeris, while orientation is a combination of satellite attitude and camera interlock angles. Attitude is determined by combining star tracker and gyro data. Interlock angles orient the camera to the attitude determination coordinate system. While ephemeris is determined by onboard GPS measurements and attitude by stellar reference, the interlock angles must be calibrated by camera observations.

After initialization, IKONOS camera interlock angles were refined with ground observations. Because exterior orientation errors are somewhat random from image to image, many observations are required to

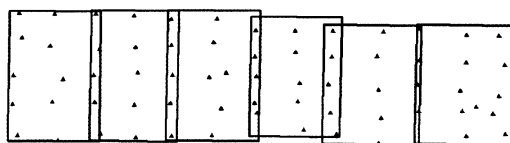


Figure 7. Six IKONOS stereo pairs (rectangles) shown with 60 GCPs (dots).

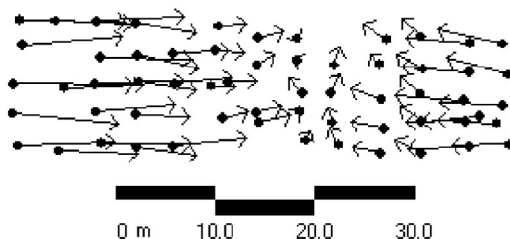


Figure 8. Residual errors, shown as magnified arrows. Worst case residual error is about 12 m.

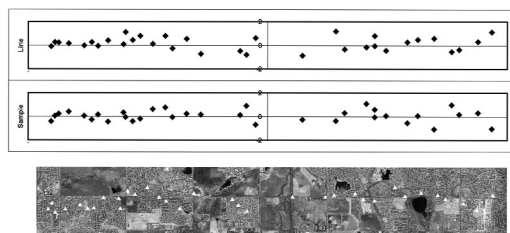


Figure 9. The 128th Street test range (bottom) shown with line and sample residual error plots (top) after FAM recalibration.

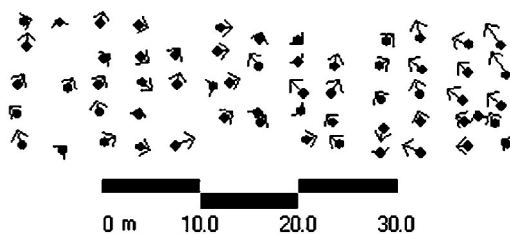


Figure 10. Residual errors from the 6-stereo strip mapping block after the FAM recalibration of November 19, 2001. Worst case residual error is 2 m horizontal and 3 m vertical.

separate any bias errors from the random errors. To eliminate survey accuracy and GCP observability as issues, it is useful to use an easily visible target, accurately surveyed, with many repeat observations.

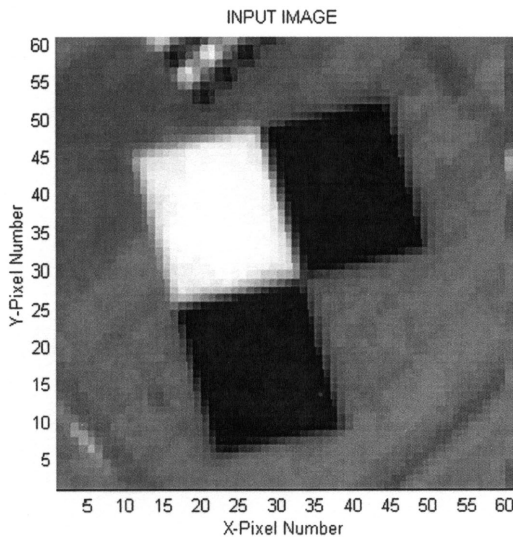


Figure 11. MTF target painted on asphalt used to measure image edge response and exterior orientation. Excessive zoom and nearest-neighbor resampling show the edge response. The vertex of the three squares was used as the GCP.

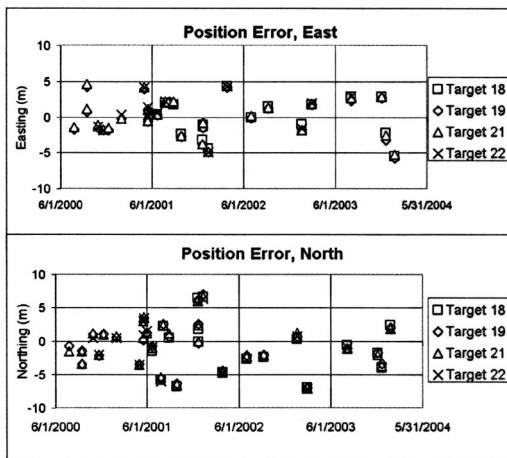


Figure 12. Historical trend of image position errors (exclusive of terrain displacement) from May 2000 to December 2003, processed without ground control after the interlock angle update.

Space Imaging painted the large target shown in Figure 11 to test the edge response of satellite images. This target was accurately surveyed, is easily measured, and so was used to test exterior orientation.

Historical imagery of the modulation transfer function (MTF) target was processed using the target

coordinate as a GCP. The interlock angle adjustments show considerable scatter during the early, initialization period. Statistical control with less scatter was established by May of 2000. The median average of the interlock adjustments from May of 2000 to November of 2001 was $+4.5$ micro-radian roll and -2.5 micro-radian pitch (Grodecki & Dial 2002). Calibrations were adjusted by this amount on January 31, 2002 to correct for biases in all data processed after that date. The median, rather than mean, average was used for a more robust estimation of the bias error.

Figure 12 shows the historical trend of position errors of the MTF and other nearby targets after the interlock angles were adjusted. The data do not appear to show long-term or seasonal trends, indicating that the interlock angles have been stable. Within a single image, all points show a consistent bias error of <10 m, showing good relative accuracy.

Key attributes of a test range to calibrate and monitor exterior orientation are frequent observation of a high quality GCP.

4 TESTING SYSTEM ACCURACY

While calibration test ranges require only a small number of high quality GCPs, tests of system accuracy should be globally distributed and cover the entire time span of system operation, thus they will require large numbers of GCPs. Using GCPs from customer orders has proven to be an economical approach to system accuracy testing. Imagery collected for customer orders can be reprocessed without using GCPs, so they can be used to assess system accuracy. Even low cost GCPs with 1 m CE90 horizontal and 2 m LE90 vertical accuracy are substantially better than uncontrolled system accuracy and so can be used for system accuracy tests.

Two studies illustrate the use of imagery and GCPs collected for customer orders to test system accuracy. The first study (Dial & Grodecki 2002) used imagery and GCPs collected to prepare a precision orthomosaic of Puerto Rico to determine the absolute and relative accuracy of long IKONOS image strips. The second study (Dial & Grodecki 2003) tested the stereo accuracy of IKONOS without ground control using stereo imagery and GCPs from around the world.

4.1 Testing absolute and relative accuracy with long image strips

The Puerto Rico orthomosaic project provided an opportunity to test IKONOS accuracy in relatively long strips. Eighteen strips more than 50 km long with five GCPs or more were selected for study. The strips were reprocessed without ground control. The RPC camera model (Grodecki 2001, and Tao & Hu

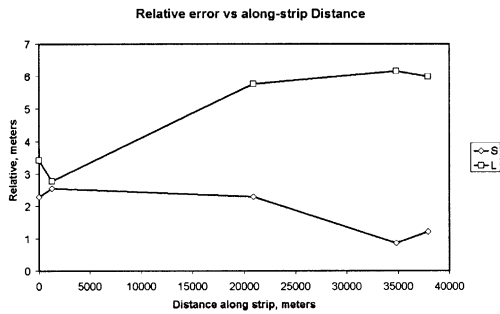


Figure 13. Relative errors within an image strip. Diamonds show line errors; squares show sample errors.

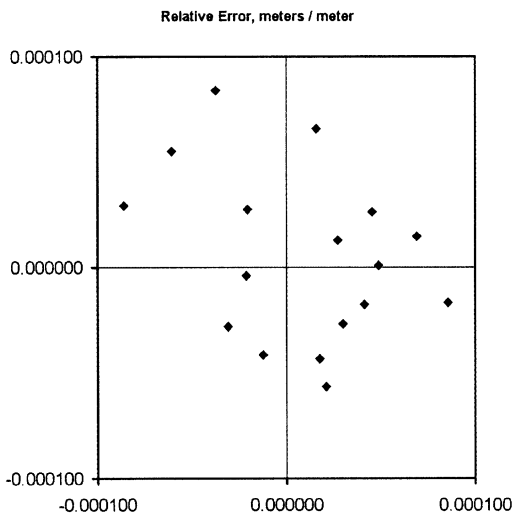


Figure 14. Absolute image error measurements from 18 image strips in Puerto Rico.

2001) was used to predict the image line and sample coordinate from the GCP latitude, longitude, and height. Differences between the predicted and measured line and sample were analyzed to determine the absolute and relative accuracy of long IKONOS strips. Figure 13 shows relative image coordinate errors for one such strip. Figure 14 shows absolute image errors of the GCPs on the 18 strips.

The average error over the strip is the absolute error of that strip. The slopes of the line and sample errors, expressed in parts per million, is a measure of the relative error of the system. Results showed that IKONOS absolute accuracy is 4 m root mean square (RMS) for line and sample, and 50 parts-per-million RMS per axis relative error.

Stereo Position Errors

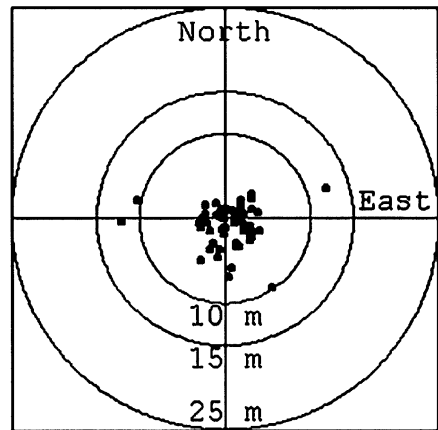


Figure 15. Horizontal errors of stereo pairs processed without GCPs showing 6 m CE90 horizontal accuracy. Vertical errors were 10 m LE90 from the same images.

4.2 Testing stereo accuracy

Testing stereo accuracy by reprocessing commercial orders was similarly straightforward (Dial & Grodecki 2003). Stereo images were collected and processed without using ground control from 24 locations around the world. Available GCPs were used as check points. Operators using commercial stereo measurement software made 3-D coordinate measurements of the checkpoints. Figure 15 shows the horizontal errors from the 24 images. Results showed IKONOS stereo to have 6 m CE90 horizontal and 10 m LE90 vertical accuracy without ground control. Similar accuracy results were found by Ager (2003).

5 QUALITY GROUND CONTROL POINTS

GCPs should be surveyed to sub-pixel accuracy for use in calibration test ranges. Lower accuracy GCPs can be used to assess uncontrolled accuracy. Surveying GCPs to sub-pixel accuracy is wasted effort unless the feature being surveyed can be identified within the image to sub-pixel accuracy. Ground control accuracy and choice of features is discussed in (Main 2000).

Painting targets on roadways or deploying panels for ground control is commonly done for aerial photography with 0.25 to 0.1 m pixels. Painting targets of 5 to 10 pixel size is hardly practical for 1 m pixels. We must seek natural or cultural features that can be accurately surveyed on the ground and precisely measured in the imagery. The reluctance of surveyors to stand in busy roadways limits the availability of road markings, while difficulty measuring other

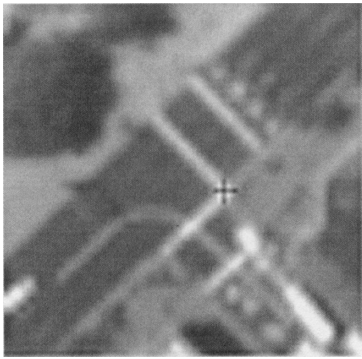


Figure 16a. Road marks for GCP.



Figure 16b. Intersecting sidewalks.

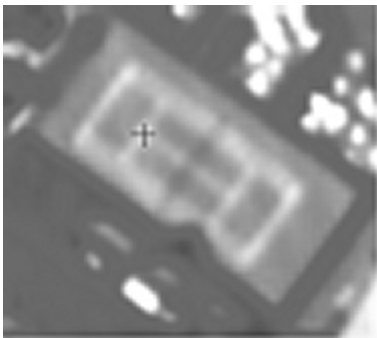


Figure 16c. Tennis court lines.

features limits their precision. Ideally, GCPs are well-defined, photo-identifiable features surveyed with great accuracy, but sometimes the practicalities of collecting ground control overwhelm concepts of absolute perfection as described by Andrews (pers. comm. 2000) for sites in Kenya.

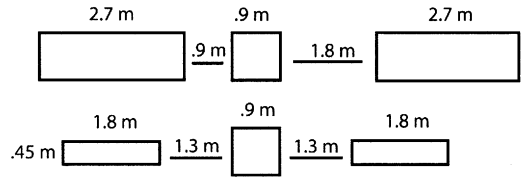


Figure 17. Panel design used by EDC: (top) original design and (bottom) revised design.

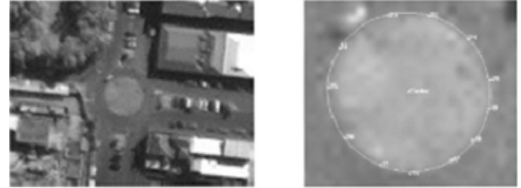


Figure 18. Roundabout for GCP. IKONOS image on left and best-fit ellipse on right.

5.1 Examples of good ground control points

The best photo-identifiable features are perpendicular intersections of thin, high-contrast lines such as sidewalks, road markings, or sports fields. Narrow, sub-pixel wide stripes found on roads and sports fields can be seen because of their high contrast. For wider features, such as sidewalks, the center of the intersection, rather than a corner, should be used. Examples of good GCPs are shown in Figures 16a–16c.

Deploying panels for ground control is generally not practical because of the large size required even for high-resolution satellite imagery. Panels were successfully used by the EROS Data Center (EDC) to test IKONOS accuracy (Helder et al. 2003). Figure 17 shows the panel design used by EDC. EDC sometimes encountered difficulty resolving the central 90 cm by 90 cm panel from the adjoining leg separated by 90 cm. They recommend increasing the separation between tarps from 90 cm to 130 cm in future work. The original and revised panel designs based on DEC's experience are shown in Figure 17 (Coan pers. comm. 2003).

While narrow, intersecting linear features can be picked reliably to the nearest pixel, combining multiple point measurements can further improve GCP pick accuracy. Figure 18 from (Fraser et al. 2001) shows a "roundabout" road pattern. To determine the ground coordinate of the center point, multiple points were surveyed around the circumference of the roundabout and the center was calculated. To determine the image coordinate of the center point, multiple points were measured around the image of the roundabout and the center of a best-fit ellipse calculated. Accuracy to 0.2 pixels is ascribed to measurements by this technique.



Figure 19. Dirt road intersection.

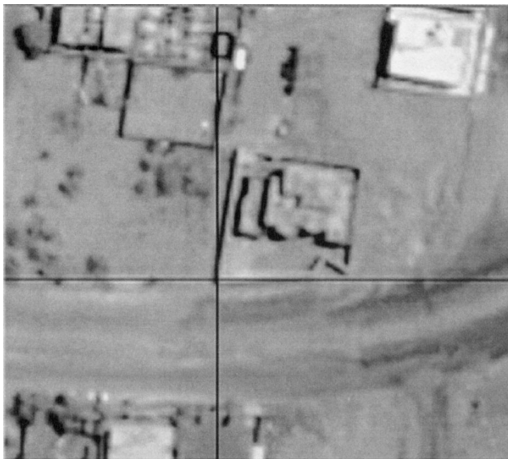


Figure 20. Walled compound.

5.2 Examples of poor ground control points

Ill-defined features, features that can be easily obscured, and corners make poor GCPs. As an example of an ill-defined feature, consider the dirt road intersection shown in Figure 19.

Although the authors do not favor fences for ground control, the intersecting fences in Figure 19 would have been better than the dirt roads. The problem with fences is shown in Figure 20. Depending on the sun angle and camera perspective, the visual appearance of fences and walls can change substantially, leading to changes in image measurements made by the operator.

The difficulty with corner features can be seen in the MTF target image in Figure 11. To the surveyor on the ground, the exterior corners appear to be

distinct points, but the operator will have difficulty picking the corner to the nearest pixel, much less to sub-pixel accuracy, because the optical system blurs the corner over several pixels.

6 CONCLUSIONS

The test ranges and control points used to calibrate and validate IKONOS accuracy have been described. Test ranges were used to update IKONOS interior and exterior orientation. Commercial GCPs were used to validate the global accuracy of IKONOS mono and stereo imagery. Test ranges were selected for good weather, geographic distribution, and ample control points. Cultural features with intersecting, high-contrast lines are preferred for control points. Others have used panels and circular features with excellent results.

REFERENCES

- Ager, T.P. & Bresnahan, P. 2003. Evaluation of the geometric accuracy of IKONOS imagery. In S.S. Shen & P.E. Lewis (eds), *Algorithms and technologies for multispectral, hyperspectral, and ultraspectral imagery IX*. Proc. SPIE 5093: 613–620.
- Cook, M., Peterson, B., Dial, G., Gerlach, F., Hutchins, K., Kudola, R. & Bowen, H. 2001. IKONOS technical performance assessment. In S.S. Shen & M.R. Descour (eds), *Algorithms for multispectral, hyperspectral, and ultraspectral imagery VII*. Proc. SPIE 4381: 94–108.
- Dial, G. & Grodecki, J. 2002. IKONOS accuracy without ground control. In *Pecora 15/land satellite information IV/ISPRS Commission I/FIEOS 2000; Proc. intern. symp., Denver, 11–14 November 2002*. Bethesda: Amer. Soc. Photogr. & Rem. Sens. (CD ROM).
- Dial, G. & Grodecki, J. 2003. IKONOS stereo accuracy without ground control. In *ASPRS 2003 annual conference; Proc. ann. conf., Anchorage, 4–5 May 2003*. Bethesda: Amer. Soc. Photogr. & Rem. Sens. (CD ROM).
- Fraser, C.S., Hanley, H.B. & Yamakawa, T. 2001. Sub-metre geopositioning with IKONOS geo imagery. In *Joint international workshop on high resolution mapping from space; Proc. intern. wksop, 19–21 September 2001*. Hannover: Univ. of Hannover, Inst. Photogr. & GeoInfo. (CD ROM).
- Fraser, C., Dare, P., Hanley, H., Yamakawa, T. & Cronk, S. 2003. Advances in geometric processing of high-resolution satellite imagery for three dimensional scene analysis. In *Spatial sciences 2003 conference; Proc. Canberra, September 2003*.
- Fraser, C.S. & Yamakawa, T. 2003. Applicability of the affine model for IKONOS image orientation over mountainous terrain. In *Joint international workshop on high resolution mapping from space; Proc., Hannover, 6–8 October 2003*. Hannover: Univ. of Hannover, Inst. Photogr. & GeoInfo. (CD ROM).
- Grodecki, J. (2001). IKONOS stereo feature extraction—RPC approach. *ASPRS 2001 annual conference; Proc.,*

- St. Louis, 23–27 April, 2001*. Bethesda: Amer. Soc. Photogr. & Rem. Sens. (CD ROM).
- Grodecki, J. & Dial, G. 2001. IKONOS geometric accuracy. In *Joint international workshop on high resolution mapping from Space; Proc., Hannover, 19–21 September 2001*. Hannover: Univ. of Hannover, Inst. Photogr. & GeoInfo. (CD ROM).
- Grodecki, J. & Dial, G. 2002. IKONOS geometric accuracy validation. In *Pecora 15/land satellite information IV/ISPRS Commission I/FIEOS 2000; Proc. intern. symp., Denver, 11–14 November 2002*. Bethesda: Amer. Soc. Photogr. & Rem. Sens. (CD ROM).
- Helder, D., Coan, M., Patrick, K. & Gaska, P. 2003. IKONOS geometric characterization. *Rem. Sens. Env.* (88)1-2: 69–79.
- Main, J.D. 2000. Precise ground control is essential for spatial accuracy. *Imaging Notes (July-August)* 15(4): 6–7.
- Tao, C.V. & Hu, Y. 2001. A comprehensive study of the rational function model for photogrammetric processing. *Photogr. Eng. & Rem. Sens.* 67(12):1347–1357.

Locating potential vicarious calibration sites for high-spectral resolution sensors in the Negev Desert by GIS analysis

U. Gilead & A. Karnieli

*The Remote Sensing Laboratory, Jacob Blaustein Institute for Desert Research,
Ben Gurion University of the Negev, Israel*

ABSTRACT: Deserts, and especially dry playas, are preferable regions for implementing reflectance-based vicarious calibration of multi- and hyper-spectral sensors. This paper uses a Geographic Information System (GIS) to locate potential vicarious calibration sites for small footprint sensors, i.e. analyzing spatially mapped, satellite, and tabular data, with respect to certain criteria such as spectral reflection, temporal stability, size, and elevation. This methodology was applied to the Negev Desert in Israel and resulted with 8 sites selected for further *in situ* characterization and measurements for validation. Results of one validation campaign are presented. The GIS analysis has proven to be a useful tool for saving time and effort by searching through vast and remote regions.

1 INTRODUCTION

Optimal operation of any air or spaceborne multi-spectral imaging system for remote sensing applications depends on accurate calibration of the system. Different methods and approaches are used commonly for calibration, namely preflight, onboard, and vicarious techniques. Regardless of the quality of the preflight calibration, once launched, any spaceborne sensor is susceptible to significant changes in its performance (Teillet et al. 2001). Consequently, in order to ensure the best performance of the system, all the above-mentioned complementary methods should be used independently to determine if systematic errors exist in one or more techniques and, if possible, to identify, examine, remove, or account for them in the calibration results.

Several approaches are usually used for vicarious calibration. These are the radiance-, irradiance-, and reflectance-based approaches. The latter involves extensive ground truth measurements of the surface spectral reflectance and the atmosphere during satellite overflight. After applying a radiative transfer code, it is possible to compare the sensor's top-of-atmosphere (TOA) measurements with those conducted independently *in situ*. This absolute calibration method produces a new set of radiometric coefficients that can be used instead of those derived during the preflight process (Thome 2001).

Deserts, and especially dry playas, are preferable regions for implementing vicarious calibration.

Well-used and documented sites are the White Sands National Monument, Railroad Valley Playa, Lunar Lake, and Roach Lake Playa, all in the southwestern part of the United States (Thome 2001), and Lake Frome in South Australia (Barry et al. 2001). In order to minimize calibration uncertainties, Scott et al. (1996) and Thome (2001) listed crucial conditions for selecting vicarious calibration sites. Such sites should be characterized by:

1. High reflectivity values that reduce the impact of errors in determining the radiance due to atmospheric scattering. A site reflectance greater than 0.3 ensures that the radiance of the direct solar irradiance from the surface is the dominant contributor to the at-sensor radiance.
2. High spatial uniformity, relative to the pixel size, over a large area that minimizes the effects of scaling the reflectance data to the size of the entire test site.
3. Nearly lambertian surface to decrease uncertainties due to changing solar and view geometry. These are flat sites that reduce bidirectional reflectance factor (BRF) effects and eliminate shadow problems.
4. Spectral uniformity over as wide a spectral region as possible. This simplifies band integrations and decreases the effects of spectral mismatch between ground-based measurements and the sensor.
5. High probability of cloud-free days.
6. Location in remote areas, far from anthropogenic (urban and industrial) aerosols.

7. Extremely low precipitation providing minimum change in soil moisture.
8. Lack of vegetation so seasonal dynamics are minimal.
9. Location far from the sea to minimize the influence of atmospheric water vapor.
10. High accessibility for those who perform the *in situ* measurements.

Scott et al. (1996) describe a search algorithm used to identify suitable sites for vicarious calibration based on Landsat-MSS imagery covering a vast portion of the southwest USA drylands (including southern California, Arizona, southern Nevada, and southern Utah). This search program mainly focuses on locating areas that have large spatial extent, spatial uniformity, and high surface reflectance. The search algorithm identified many promising sites, however only two of them (Railroad Valley Playa and Lunar Lake) were selected for further ground investigation.

The State of Israel, which has already manufactured, launched, and operated several Earth observation satellites (the EROS and Ofek series) (Steinberg 2001), is seeking independent calibration sites for such small footprint sensors. The current research explores the Israeli Negev desert by Geographic Information System (GIS) technology to locate and characterize potential sites that may be used for vicarious calibration.

2 MATERIALS AND METHODS

2.1 GIS data

The search for potential calibration sites was conducted in accordance with the criteria described above. That is, the site should be characterized by high spectral reflectance, high elevation, high spatial uniformity, minimal slope, spectral uniformity, minimal seasonal changes, and be readily accessible to the organization that will perform the calibration. The site should also be large enough relative to the pixel size of the sensor (Clark et al. 2002).

Locating potential sites was performed, for the most part, by GIS techniques, i.e. analyzing spatial mapped data and satellite imagery along with tabular data. This facilitated the location procedure by eliminating areas that were not suitable, leaving only a small number of appropriate ones. The idea behind this methodology is to save time and effort by searching through a vast and remote region while checking a number of critical variables such as spectral reflection, temporal stability, size, and elevation. Nevertheless, one should keep in mind that as promising as the results may be, there is still a need to perform *in situ* examinations to verify the results and appraise the site's suitability.

The general concept underlying the search for sites was to join information from relevant GIS layers,

creating one polygonal layer with an inclusive attribute table. The variables attributed to each polygon are key parameters considered necessary to characterize calibration sites. Building such a layer enables finding ideal sites by selecting polygons that have specific parameters.

2.2 Study site – the Negev Desert

The study area lies in the southern part of Israel in the Negev Desert. The eastern (Jordan) and western (Egypt) borders of this area were delineated according to political considerations. The northern border of the Negev is considered to be the 200 mm isohyetal line, thought to be the beginning of the dryland (Stern et al. 1986). The overall size of the study area is 10,000 km².

Summer in the Negev is characterized by hot (mean annual temperatures of 25 degrees C), dry (relative humidity ranges from 20 to 50 percent), and clear days with little day-to-day variation. Rain falls almost exclusively in the cool (9 degrees C) winter. Mean annual rainfall decreases gradually from 200 mm at the northern boundary to about 25 mm at the southern edge. The number of rainfall days per annum varies between 40 in the north to 5 in the south. Since the Negev has a relatively small size, spatial changes in solar radiation are minor. The horizontal global total radiation is generally around 2000 kWh/m²/year. The highest radiation is in June (220–250 kWh/m²/year) and the lowest in December (90–115 kWh/m²/year) (Faiman et al. 2003). The main soils found in the Negev are coarse desert alluvium, calcareous serozems, sands, and bare rocks. Perennial vegetation is sparse (less than 20 percent) in the northern Negev, and concentrated only along the ephemeral streams in the southern part (Danin 1983). Annual phenology is limited to the spring months (Karnieli 2003).

For this report, the Negev can be divided into three zones that vary mainly in topography, mineral composition, precipitation, and radiation levels:

1. Northwestern Negev. Average elevation ~150 m, precipitation 100–200 mm per year, 20–25 percent cloud cover per year, annual horizontal global total radiation ~1900 kWh/m²/year. The area is comprised of sandy regosol, sand dunes, and arid brown soils.
2. Eastern Negev along the Rift Valley. Elevation ranges between –400 m to sea level (increase southwards), precipitation 25–100 mm per year (increase northwards), 15–20 percent cloud cover per year, annual horizontal global radiation changes with elevation, between 1900–2100 kWh/m²/year. The area is comprised of regosols and coarse desert alluvium.
3. The central Negev highlands – average elevation 500 m, precipitation ranges from 0 to 100 mm per year, 15–20 percent cloud cover per year, annual

horizontal global radiation $\sim 2100 \text{ kWh/m}^2/\text{year}$. The area is comprised of calcareous serozems (higher areas), bare rock, desert lithosols, regosols, and coarse desert alluvium.

2.3 Building the GIS

The first step in constructing the GIS was to gather input layers that include relevant data. All components of the GIS are described in Table 1 and are illustrated in Figure 1.

The digital elevation model (DEM) of the Negev is based on the Geographical Society of Israel (GSI) digital terrain model (Hall 1993). Forty sheets, each consisting of a matrix of 801×801 pixels, were mosaicked. The spatial resolution of the DEM is 25 m while the height resolution is 10 cm. The DEM was used for slope retrieval, after being converted to slope map (Fig. 2a). Later on, since the site should be as flat as possible, all areas with slope higher than 0 degrees were excluded. The final product was a binary raster map where 1 = 0 degrees slope and 0 = no data. Afterwards, the raster map was converted to vector polygons. This enabled building topology for the slope 0 degree areas, i.e. size, perimeter and *xy* coordinates for each polygon. An attempt to include more areas, by including all pixels with value of slope = 1 degree, presented similar results with no significant increase in polygon size.

A major advantage of converting raster to vector is the better handling of excess polygons. This occurs

since each pixel with 0 degree slope can become a polygon. In the current research, more than 87,000 polygons were created for a size smaller than $10,000 \text{ km}^2$. When not clustered, these polygons are useless since they do not qualify as potential sites due to their small size. One of the advantages of working in vector format is the possibility of assigning each aggregation of pixels (now polygon) an area dimension, whereas in raster format this operation is difficult, since pixels maintain the value of each single

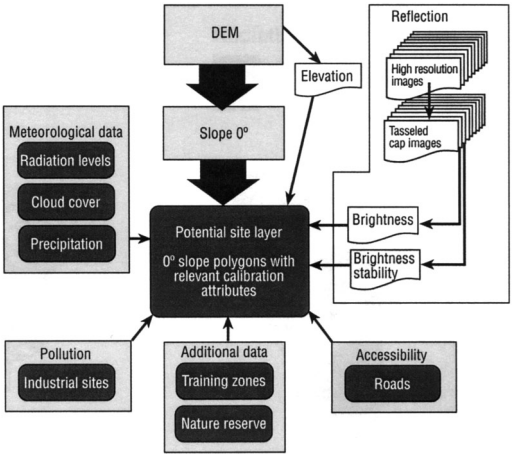


Figure 1. Input layers of the GIS (graphic scheme).

Table 1. Input layers of the GIS.

Input	Details	Data use
DEM	25 m spatial resolution 10 cm height resolution Raster	Obtain elevation and slope
Landsat TM and ETM+ images	Frame 174-39 (South Israel) Dates of acquisition: 18 Jan 1987; 14 Aug 1987; 29 Mar 1995; 21 Sep 1995; 6 Apr 1998; 7 Aug 1999; 10 Oct 1999; 30 Jan 2000; 21 May 2000; 27 July 2001 30 m resolution	Tasseled cap index was performed to obtain brightness values (represent high reflection)
Precipitation	Annual mean precipitation in Israel Polygonal vector layer 100 mm isohyets	Predicting precipitation amounts at the selected site
Cloud cover	Mean yearly cloud cover of the sky (percent)	Assess probability of clear days throughout the year
Radiation	Mean monthly radiation levels ($\text{kWh/m}^2/\text{year}$) Polygonal vector layer	Assess probability of clear days throughout the year
Air polluting industry	Air pollution sources Polygon point layer	Knowledge of air polluting industries nearby sites
Roads	Main roads of Israel Line vector layer	Distance information and accessibility
Military training zone	Training zones of the Israeli Army Polygonal vector layer	Identify potential sites which are inside military training zones
Nature reserves	Nature reserves in Israel Polygonal vector layer	Additional information

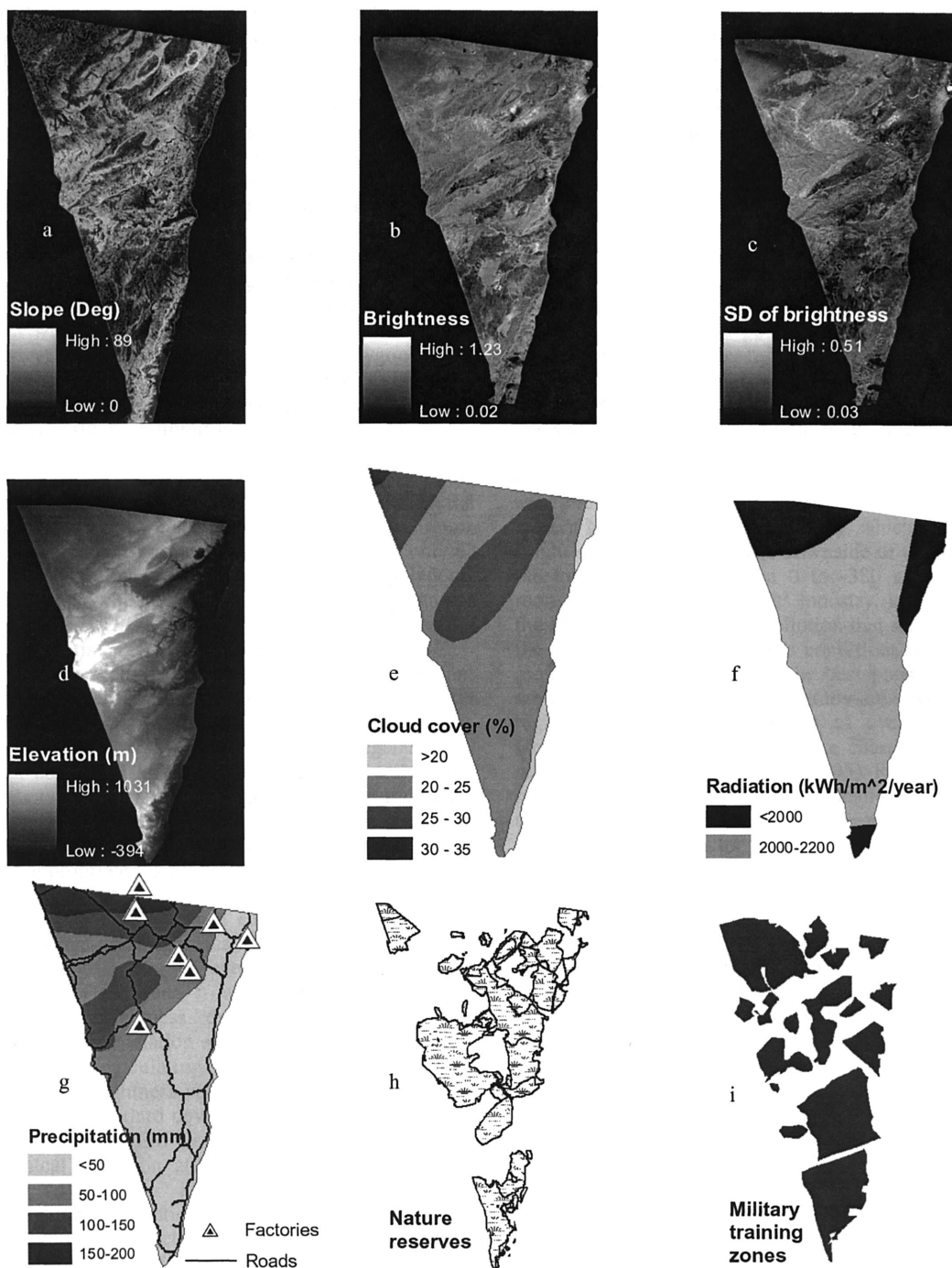


Figure 2. Input layers of the GIS. (a) slope; (b) brightness; (c) standard deviation (SD) of brightness; (d) DEM – elevation; (e) cloud cover; (f) radiation; (g) precipitation, roads, and factories; (h) nature reserves; (i) military training zones.

pixel and do not aggregate to one object with area dimension. The conversion to vector format therefore enables deletion of polygons under a specific size, resulting in a smaller file with relevant polygons that may be used as potential sites for vicarious calibration. Another advantage of working in vector is that it permits integration of polygons that are close to each other and could actually be considered as one large site. After deleting all polygons smaller than 1 km² (minimum size determined for potential calibration site) only 87 polygons remained, creating a reduced polygonal layer on which the rest of the analysis was performed. These polygons are referred to hereafter as 'potential sites layer' or PSL.

The next step was associating the PSL polygons with tasseled cap brightness values to locate sites with high soil reflectance. The tasseled cap transformation (Crist & Ciccone 1984) is a vegetation index used to disaggregate the amount of soil brightness, vegetation, and moisture content in individual pixels. The transformation was performed on five Landsat TM and four Landsat ETM+ images using the appropriate tasseled cap coefficients of each sensor (Huang et al. 2002). Later, all images were overlaid and a mean value of each pixel was calculated to produce one brightness image (Fig. 2b). Since the nine Landsat images were acquired at different months of the year, it was possible to assess the temporal brightness dynamics due to seasonal changes of vegetation and/or wetness. This was performed by calculating the standard deviation from the mean brightness value. As seen in Figure 2c, the lower the standard deviation, the higher the stability of the site. A DEM was used to create a mean topographical elevation attribute for each PSL polygon (Fig. 2d).

Spatial distribution of meteorological data, such as cloud cover (Stern et al. 1986), radiation (Faïman et al. 2003), and precipitation (Survey of Israel 1995), were extracted from different maps. Industrial air polluting facilities were mapped as points, and roads as vector lines (Figs 2e, f, g). Polygonal vector layers of military training zones and nature reserves were obtained and intersected with the PSL, to gain additional information involving potential restrictions for each potential site (Figs 2h, i). Finally, the accessibility of each site was calculated using a line-vector layer of roads. The 'distance' attribute was calculated from the remote sensing laboratory at Sede Boker to the centroid of each potential site. The final product is an inclusive attribute table that enabled selecting, by means of a query with desired parameters as input, optimal sites out of the 87 sites that initially composed the PSL.

3 RESULTS

The final step, after completing the construction of the GIS, was to locate potential sites for vicarious

calibration in the Negev by performing a query with desired parameters on the PSL layer. Doing so revealed, as expected, that no site has ideal characteristics that meet all the attribute requirements. Consequently, the chosen sites are a function of the flexibility of the stipulations set in the query. The area with the highest brightness values (>1.0) is the eastern Negev along the Rift Valley, which is rich in bright surface minerals. The downside of this area is the low elevation, between 0 to -380 m. Furthermore, the 'Dead Sea Works' industry, located near the Dead Sea, produce air pollution that might affect the visibility and atmospheric conditions in this region. In terms of elevation, the best potential sites are at the western parts of the study area, where the highest points reach elevations of over 1000 m. However, these potential sites are characterized by relatively low brightness (0.85-0.95). Radiation and precipitation do not change dramatically over the entire study area due to its relatively small size. Similarly, the distance factor in the Negev is negligible, since every point can be reached in less than a 3-hour drive from any other point. In some regions, such as along the Rift Valley, one principle wind direction (north to south) is dominant throughout the entire year. In other regions, industrial air pollution should be dealt with individually in each case since the wind regime changes temporarily. A site within a distance of 2 km from a polluting factory can have aerosol-free air in a certain month at certain hours, while at different times the situation may be reversed.

To reduce the number of potential sites, a brightness threshold was set at 0.95. This resulted in 8 potential sites rather than 87. These 8 selected sites should be carefully examined by remote sensing methods and by *in situ* measurements as well. Although the PSL and its attributes may be comprehensive, there are other factors that may prevent the potential site from being used as a vicarious calibration site, e.g. sparse vegetation, surface roughness. Examination by spectroscopic methods can spot unknown parameters of the site, e.g. soil mineralogy with pronounced absorption features. Therefore, ground verification is an essential condition to complete the GIS study.

Of the final eight sites, the most appropriate one appears to be site number 5, with a size of almost 6 km², elevation of 370 m, and high brightness values of 0.96. It also has low cloud cover and high radiation levels. The disadvantage of this site is that it is located in a military training zone, which may restrict access and mobility. Site number 2 also is inside a military training zone. Site numbers 7 and 8 have high brightness values and are located outside military training zones. Their setbacks are low elevation (-220 m and -261 m, respectively) and relative proximity to the Dead Sea Works industry (3.84 km and 13.94 km, respectively).

Some regions were disqualified due to unsuitable conditions. The northwestern Negev, with relatively

low brightness values (0.5–0.75) and medium elevation (~200 m), is an example of such an area. Furthermore, precipitation levels are higher in the northwestern Negev than in other regions and radiation is lower. All of these conditions led to the removal of this region from the potential site list.

4 AMIAZ PLAIN – CHARACTERISTICS

Within the framework of the GIS validation efforts, each of the selected sites was visited to verify and characterize its suitability for vicarious calibration. This section presents one example. Site number 8, Amiaz Plain, is located at 31°4'41"N, 35°22'10"E. The total size of the plain, which is a part of the 'Judea Desert Nature Reserve', is about 6 km². About 1 km² has a 0 degree slope; the rest of the area has slopes ranging up to 10 degree. The elevation of the site ranges between –260 to –270 m below mean sea level (MSL), with mountain ridges on the western and eastern edges (–50, –180 m below MSL, respectively).

Climatically, the site is found in an extremely arid area. The mean annual rainfall is 47 mm, with large variations between years. Ninety-nine percent of the rainfall occurs between October and April. The probability for more than 50 mm/year is 45 percent. Cloud cover is about 27 percent from October to May and 4 percent from June to September. Horizontal global total energy ranges between 81 kWh/m² in December to 231 kWh/m² in June. The prevailing wind directions are south (December–February), north (March–May, October–November) and northwest (June–September) (Faiman et al. 2003). The site is located approximately 4 km northwest of the Dead Sea Works industry; therefore aerosol pollution should be taken into account, particularly during December–February. The site is within a one-hour drive from the remote sensing lab at



Figure 3. Bright surface consists of marls at Amiaz Plain.

Sede Boker. The plain's substrate consists of bright marls of the Lisan formation that contribute to its high surface brightness (Fig. 3). Figure 4 presents the spectral reflectance of the site as measured by a portable analytical spectral device spectrometer (ASD).

The graph depicts the relative advantage of this site – very high reflectance throughout the reflective range (350–2500 nm) with minimal fluctuations.

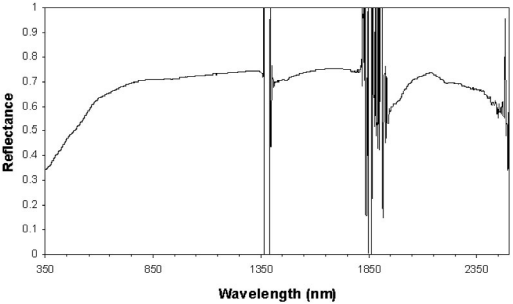


Figure 4. Spectral reflectance of Amiaz Plain.

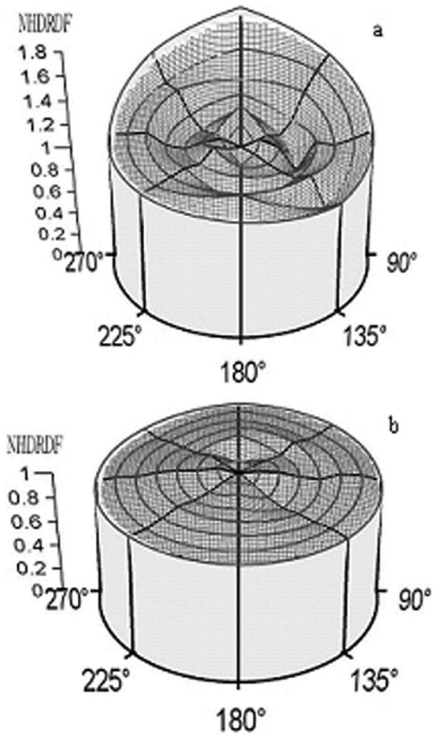


Figure 5. NHDRDF for a smooth surface at Amiaz Plain for the 850 nm waveband and for solar zenith angle of 75 degrees (a) and 15 degrees (b). The basis of each diagram represents the view angle with respect to the solar principle plane.

The normalized hemisphere-directional reflectance distribution function (NHDRDF) of the surface is presented in Figures 5a, b. Measurements conducted by a CIMEL radiometer (Cierniewski, in press) reveal near-Lambertian characteristics of the surface.

5 CONCLUSIONS

The advantage in using GIS analysis for locating potential sites for vicarious calibration is the ability to search through large areas and identify suitable sites by intersecting various crucial parameters such as reflectance levels and elevation. Furthermore, the temporal dynamics of these variables can be checked over time using time series techniques. However, this GIS analysis is by no means a replacement for *in situ* measurements; validation of the suitability of sites is also essential. As extensive as the GIS characterization may be, one must keep in mind that it provides general results, which are dependent on the resolution used in the GIS, and may not provide micro-characterization of the site. A 25 m resolution DEM, as used in this research, may not detect gullies, rough soils, and other types of micro-morphologies that can affect the distribution of incoming radiation. Such is also the case for other parameters, e.g. precipitation and radiation values, that should be extracted at site resolution and not by regional interpolation. The brightness values for each site are the average of all pixels within the site, and do not necessarily reflect a homogeneously bright site. Consequently, further research should be conducted to assess site reflectance homogeneity as suggested by Bannari et al. (2002).

Another reason for *in situ* validation is the potential lack of information concerning land use. While some land use maps are easily obtained and inserted in the GIS, it is hard to obtain all possibilities. Generally, the resolution of the GIS input is dependent on the spatial resolution of the sensor intended for calibration. For high spatial resolution sensors, the higher the resolution used in the raster input, the better the results. However, it should be taken into consideration that the higher the ratio between the study area and the spatial resolution of input, the more computer power is needed for processing, especially when working in vectors. Altogether, the GIS analysis has proven to be a useful tool for saving time and effort by searching through vast and remote regions.

ACKNOWLEDGMENTS

The authors thank Prof. Jerzy Cierniewski, Adam Mickiewicz University in Poznan, Poland, for providing indispensable BRDF data and assistance in performing the field measurements.

REFERENCES

- Bannari, A., Omari, K., Teillet, P. & Fedosejev 2002. Potential of Getis Statistic to characterize the spatial and temporal uniformity of test sites used for radiometric calibration of earth observation sensors. In J.A. Sobrino (ed.), *International symposium on recent advances in quantitative remote sensing, Valencia, Spain*. Valencia: Univ. of Valencia.
- Barry, P.S., Jarecke, P. & Pearlman J. 2001. *Radiometric calibration of the Hyperion instrument using ground truth at a site in Lake Frome, Australia*. TRW Redondo Beach, CA.
- Cierniewski, J., Gdala, T. & Karnieli, A. 2004. A hemispherical-directional reflectance model as a tool for understanding image distinctions between cultivated and uncultivated bare surfaces. *Rem Sens Env.* 90(3): TBD.
- Clark, R.N., Swayze, G.A., Livo, K.E., Kokaly, R.F., King, T.V.V., Dalton, J.B., Vance, J.S., Rockwell, B.W., Hoefman, T. & McDougal, R.R. 2002. Surface reflectance calibration of terrestrial imaging spectroscopy data: a tutorial using AVIRIS. *U.S. Geological Survey Spectral Laboratory*. <http://speclab.cr.usgs.gov/PAPERS/calibration/tutorial>
- Crist, E.P. & Cicone, R.C. 1984. A physically-based transformation of Thematic Mapper data – the TM Tasseled Cap. *IEEE Trans. Geosci. & Rem. Sens.* 22(3): 256–263.
- Danin, A. 1983. *Desert vegetation of Israel and Sinai*. Jerusalem: Cana.
- Faiman, D., Feuermann, D., Ibbston, P., Zemel, A., Ianetz, A., Liubansky, V. & Setter, I. 2003. *Data processing for the Negev radiation survey: ninth year*. Sede Boker: The Jacob Blaustein Inst. Desert Res.
- Hall, J.K. 1993. The GSI digital terrain model (DTM) project completed. *GSI Current Res.* 8: 47–50.
- Huang, C., Wylie, B., Yang, L., Homer, C. & Zylstra, G. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. *Int. J. Rem. Sens.* 23(8): 1741–1748.
- Karnieli, A. 2003. Natural vegetation phenology assessment by ground spectral measurements in two semi-arid environments. *Int. J. Biomet.* 47(4): 179–187.
- Scott, K.P., Thome, K.J. & Brownlee, M.R. 1996. Evaluation of the Railroad Valley playa for use in vicarious calibration. In B. Huberty, J.B. Lurie, P. Coppin, J.A. Caylor & P.C. Robert (eds), *Multispectral imaging for terrestrial applications*. Proc. SPIE 2818: 158–166.
- Steinberg, G.M. 2001. Commercial observation satellites in the Middle East and the Persian Gulf commercial observation satellites. In Baker, J.C., O'Connell, K.M., Willimason, R.A. (eds), *Commercial observation satellites*; Santa Monica: RAND.
- Stern, E., Gradus, Y., Meir, A., Krakover, S. & Tsoar, H. (eds) 1986. *The Atlas of the Negev*. Be'er-Sheva: Keterpress Enterprises.
- Survey of Israel 1995. *The new Atlas of Israel*. Tel Aviv: Survey of Israel Pubs.
- Teillet, P.M., Fedosejevs, G., Gauthier, R.T., O'Neill, N.T., Thome, K.J., Biggar, S.F., Ripley, H. & Meygret, A. 2001. A generalized approach to the vicarious calibration of multiple earth observation sensors using hyperspectral data. *Rem. Sens. Env.* 77(3): 304–327.
- Thome, K.J. 2001. Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance based method. *Rem. Sens. Env.* 78(1–2): 27–38.

Acronyms

2MASS	2 Micron All-Sky Survey	CIPM	Comité International de Poids et Mesures
ABIR	Ambient Background InfraRed (facility at NIST, USA)	CNES	Centre National d'Etudes Spatiales
ADC	Analog-to-Digital Converter	CS	Coordinate System
ADS	Airborne Digital Sensor (sensor by Leica Geosystems, Inc.)	CTF	Contrast Transfer Function
AERONET	AErosol RObotic NETwork	CZCS	Coastal Zone Color Scanner
AFGL	Air Force Geophysics Laboratory (USA)	DAAC	Distributed Active Archive Center (NASA, USA)
AGL	Above Ground Level	DCT	Discrete Cosine Transform
ALI	Advanced Land Imager	DGIWG	Digital Geographic Information Working Group
ANN	Artificial Neural Network	DIN	Deutsches Institut für Normung
AOCS	Attitude and Orbit Control Subsystem	DIS	Draft International Standard
ARP	Ancillary Radiometric Product (refers to MISR)	DLA	Dark Level Acquisition
ASD	Analytical Spectral Device	DMC	Digital Mapping Camera
ASTER	Advanced Spaceborne Thermal and reflection Radiometer	DN	Digital Number
AWAR	Area-Weighted Average Resolution	DoD	Department of Defense (USA)
BASC	Bundle Adjustment with Self-Calibrator	DoE	Department of Energy (USA)
B/H	Base-to-Height (Ratio)	DORIS	Determination d'Orbite et Radiopositionnement Integre par Satellite
BIIF	Basic Image Interchange Format	DTED	Digital Terrain Elevation Data
BIPM	Bureau International des Poids et Mesures	DTM	Digital Terrain Model
BLS	BaseLine Stabilization	DTEM	Digital Terrain Elevation Model
BLUH	BundLe adjustment University Hannover	ECR	Earth-Centered Rotating (coordinate system)
BRDF	Bidirectional Radiance Distribution Function	ECT	Earth Calibration Source
BRF	Bidirectional Reflectance Factor	EEl	Essential Elements of Information
CCD	Charge-Coupled Device	EOS/Terra	Earth Observing System/Terra platform
CCPR	Consultative Committee of Photogrammetry and Radiometry	ERBE	Earth Radiation Budget Experiment
CCRI	Climate Change Research Initiative (USA)	ESA	European Space Agency
CD	Committee Draft	ESF	Edge Spread Function
CDR	Climate Data Record	ESR	Electrical Substitution Radiometer (aka Cryogenic Radiometer)
CE	Circular Error	ETM+	Enhanced Thematic Mapper-plus
CEOS	Committee on Earth Observing Satellites	EU	European Union
CERES	Clouds and Earth's Radiant Energy System	FAM	Field Angle Map
CGPM	Conference Generale de Poids et Mesures	FAO	Food and Agriculture Organization (United Nations)
CHRIS	Compact High Resolution Imaging Spectrometer	FASCAL	Facility for Automated Spectroradiometric Calibrations (at NIST, USA)
		FDIS	Final Draft International Standard
		FFT	Fast Fourier Transform
		FGDC	Federal Geographic Data Committee (United States)

FT	Fourier Transform	IPCC	Intergovernmental Panel on Climate Change
FWHM	Full-Width, Half Maximum	IQ	Image Quality
GCP	Ground Control Points	IQEC	Image Quality Expertise Center (CNES, France)
GERB	Geostationary Earth Radiation Budget	IRAS	InfraRed Astronomical Satellite
GIANTS	Global Instrumented and Networked Test Sites	IRS	Indian Remote Sensing satellite (India)
GIQE	General Image Quality Equation	IS	International Standard
GIS	Geographic Information System	ISCCP	International Satellite Cloud Climatology Project
GM	Global Mode (refers to MISR)	ISO	International Organization for Standardization
GML	Geography Mark-up Language	ISPRS	International Society for Photogrammetry and Remote Sensing
GSD	Ground Sample Distance	ITS	International Temperature Scale
GSE	Ground Support Equipment	ITU	International Telecommunication Union
GSI	Ground Sample Interval	JCT	Joint Technical Committee (as in ISO/IEC-JCT 1)
GUM	Guide to the expression of Uncertainty in Measurement	JPL	Jet Propulsion Laboratory
HDF	Hierarchical Data Format	JScript	JAVA Script
HDRF	Hemispherical Directional Reflectance Factor	LANL	Los Alamos National Laboratories (USA)
HM	Haute résolution Monomode (SPOT 5 m mode; consisting of HMA and HMB images to produce a 2.5 m THR)	LBIR	Low Background InfraRed (facility at NIST, USA)
HQE	High Quantum Efficiency	LISE	Laboratoire Interdisciplinaire en Sciences de l'Environnement
HRG	Haute Résolution Géométrique (SPOT 5 high-resolution cameras)	LM	Local Mode (MISR high-resolution mode)
HRS	Haute Résolution Stéréoscopique (SPOT 5 stereoscopic cameras)	LOA	Laboratoire d'Optique Atmosphérique
HRSC	High Resolution Stereo Camera (DLR, predecessor to ADS40)	LSB	Last Significant Bit
HRV	Haute Résolution Visible (SPOT 1–3 cameras)	LSF	Line Spread Function
HRVIR	Haute Résolution Visible et InfraRouge (SPOT 4 cameras)	MBIR	Middle Background InfraRed (facility at NIST, USA)
HTML	HyperText Markup Language	MFRSR	Multi-filter Rotating Shadow-band Radiometer
IAG	International Association of Geodesy	MISR	Multi-angle Imaging Spectroradiometer
ICR	Inertial Coordinate Reference (system)	MOBY	Marine Optical Buoy
IEC	International Electrotechnical Commission	MODIS	MODerate-resolution Imaging Spectroradiometer
IFOV	Instantaneous Field of View	MODTRAN	Moderate Resolution Transmittance (refers to atmospheric correction software)
IGN	Institut Géographique National (France)	MRA	Mutual Recognition Arrangement
IHO	International Hydrographic Organization	mSG (filter)	modified Savitzky-Golay (filter)
IM	Image Measure (software program)	MSG	Meteosat Second Generation (Eumetsat)
IMU	Inertial Measurement Unite	MSL	Mean Sea Level
INCITS/L1	InterNational Committee for Information Technology Standards/Technical Committee L1	MSX	Midcourse Space Experiment
InGaAs	Indium-Gallium-Arsenide	MTF	Modulation Transfer Function
INS	Inertial Navigation System	MTFC	Modulation Transfer Function Compensation
InSb	Indium Antimonide	MTFC-off	MTFC (without compensation)
IORD	Integrated Operational Requirements Document		

MTFC-on	MTFC (with compensation)	PRT	Platinum Resistance Thermometer
NAFTA	North American Free Trade Agreement	PSF	Point Spread Function
NAS	National Academy of Science (USA)	PSL	Potential Site Layer (as in a GIS coverage)
NASA	National Aeronautics and Space Administration (USA)	PTB	Physicalische Technische Bundesanstalt
NBS	National Bureau of Standards (United States now NIST, see below)	PTFF	Polytetrafluoroethylene
NGA	National Geospatial-Intelligence Agency (USA, formerly NIMA, see below)	QA	Quality Assurance
NHDRDF	Normalized Hemisphere-Directional Reflectance Distribution Function	RER	Relative Edge Response
NIIRS	National Imagery Interpretability Rating Scale (USA)	RMSE	Root Mean Square Error
NIMA	National Imagery and Mapping Agency (now NGA)	ROLO	Robotic Lunar Observatory
NIR	Near InfraRed	RPC	Rational Polynomial Coefficient
NIST	National Institute of Standards and Technology (NMI for USA)	RSB	Reflectance Solar Bands
NLRF	National Laser Radiometer Facility (UK, NPL)	RSG	Remote Sensing Group (University of Arizona)
NMI	National Metrology Institute (each nation has its own specific named institution)	SABER	Sounding of the Atmosphere using Broadband Emission Radiometry (facility at SDL)
NOAA	National Oceanic and Atmospheric Administration (USA)	SBRS	Santa Barbara Remote Sensing
NPL	National Physical Laboratory (NMI for UK, India)	ScaRaB	Scanner for Radiation Budget
NPOESS-IPO	Nation Polar-orbiting Environmental Satellite System Integrated Program Office (future satellite platform)	SD	Solar Diffuser
NRC	National Research Council (research arm of NAS)	SDL	Space Dynamics Laboratory (Utah State University)
NWIP	New Work Item Proposal	SDSM	Solar Diffuser Stability Monitor
OBC	Onboard Calibrator	SeaWiFS	Sea-viewing Wide Field-of-view Sensor
ODOT	Ohio Department of Transportation	SensorML	Sensor Model Language
OGC	Open GIS Consortium	SG (filter)	Savitzky-Golay (filter)
ONERA	Office National d'Etudes et de Recherches Aérospatiales	SI (Units)	Système International (d'Unités) (as in metric)
OSL	Optical Sciences Laboratory (United States Geological Survey)	SIRCUS	Spectral Irradiance and Radiance Responsivity Calibrations with Uniform Sources (facility at NIST, USA)
OSU	(The) Ohio State University	SMAC	Simultaneous Multiframe Analytical Calibration
OTD	Optical Technology Division (of NIST, USA)	SMIL	Synchronized Multimedia Integration Language
OTF	Optical Transfer Function	SNR	Signal-to-Noise Ratio
PARABOLA	Portable Apparatus for Rapid Acquisition of Bidirectional Observation of the Land and Atmosphere	SPOT	Satellite Pour l'Observation de la Terre
PIN	p+ (Positive), i (Intrinsically pure silicon), n ⁻ (Negative) doped layers	SRIPS	Spectral Radiance and Irradiance Primary Scales
POLDER	Polarization and Directionality of the Earth's Reflectances	SSMI	Special Sensor Microwave Imager
		SST	Sea Surface Temperature
		SSTL	Surrey Satellite Technologies (UK)
		SURF III	Synchrotron Ultraviolet Radiation Facility
		SWIXR	ShortWave Infrared Transfer Radiometer
		TC	Technical Committee (as in ISO/TC 211)
		TCL/TK	Tool Command Language/Tool Kit
		THR	Très Haute Résolution, refers to SPOT 5 2.5-m imaging mode
		TIFF	Tagged Image File Format

TIR	Thermal Infrared	UV	Ultraviolet
TM	Thematic Mapper (sensor on Landsat)	VBS	Visual Basic Script
TOA	Top of Atmosphere	VC	Vicarious Calibration
TRUTHS	Traceable Radiometry	VG	VéGéTation (SPOT 4–5 low resolution cameras)
	Underpinning Terrestrial- and Helio-Studies	VIIRS	Visible/Infrared Imager/Radiometer Suite
TSARS	Transfer Standards Absolute Radiance Sources	VIM	International Vocabulary of Basic and General Terms in Metrology
TXR	Thermal-infrared Transfer Radiometer	VRML	Virtual reality Markup Language
UHTBB	Ultra High Temperature BlackBody	VXR	Visible Transfer Radiometer
UM	University of Miami (Florida)	WCRP	World Climate Research Programme
UML	Unified Modeling Language	WD	Working Draft
USAF	United States Air Force	XML	Extensible Mark-up Language
USGS	United States Geological Survey	ZnSe	Zinc-Selenium
UTH	Upper Tropospheric Humidity		