NASA
Reference
Publication
1397

1997

Workshop on Strategies for Calibration and Validation of Global Change Measurements: May 10-12, 1995

Bruce Guenther
James Butler
Goddard Space Flight Center
Greenbelt, Maryland

Philip Ardanuy
Hughes STX
Greenbelt, Maryland

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland

This publication is available from the NASA Center for AeroSpace Information,
800 Elkridge Landing Road, Linthicum Heights, MD 21090-2934, (301) 621-0390.

EXECUTIVE SUMMARY

The Committee on Environment and Natural Resources (CENR) Task Force on Observations and Data Management hosted a Global Change Calibration/Validation Workshop on May 10-12, 1995 in Arlington, Virginia. This Workshop was convened by Robert Schiffer of NASA Headquarters in Washington, D.C. for the CENR Secretariat with a view toward assessing and documenting lessons learned in the calibration and validation of large-scale, long-term data sets in land, ocean, and atmospheric research programs. The National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC) hosted the meeting on behalf of the Committee on Earth Observation Satellites (CEOS)/Working Group on Calibration/Validation, the Global Change Observing System (GCOS), and the U.S. CENR.

A meeting-of-experts from the international scientific community was brought together to develop recommendations for calibration and validation of global change data sets taken from instrument series and across generations of instruments and technologies. Forty-nine scientists from nine countries participated. The U.S., Canada, United Kingdom, France, Germany, Japan, Switzerland, Russia, and Kenya were represented.

The Plenary Sessions were convened by Ichtiaque Rasool of the International Geosphere/Canada Biosphere Program (IGBP) Data and Information Systems in Paris, France. The Oceans Sessions were convened by Richard Reynolds of the NOAA/National Meteorological Center in Suitland, MD. The Atmosphere Sessions were convened by John Barnett of the Atmospheric, Oceanic, and Planetary Sciences Clarendon Laboratory in the United Kingdom and Dennis Chesters of the NASA/Goddard Space Flight Center in Greenbelt, MD. The Land Sessions were convened by Philippe Teillet of the Canada Center for Remote Sensing in Ontario, Canada and Forrest Hall of NASA/Goddard Space Flight Center in Greenbelt, MD.

The Workshop began with the presentation of case studies describing experience in developing long-term geophysical data sets such as ozone trends, international satellite cloud climatology, blended sea-surface temperature, and land climate. Splinter sessions (by "traditional disciplines") extended and developed key components from case studies in the following four areas: atmosphere, ocean, land, and sensors. Findings and recommendations (lessons learned) were presented by each discipline in the plenary session to the steering committee for integration. The integrated findings are reported in this Executive Summary. The findings are distinguished into the categories:

- Programmatic Support
- Preflight Calibration
- In-Flight Calibration
- Data Set Continuity and Consistency
- Combining Remotely Sensed and In-Situ Data

For the first listing, each finding is named and briefly described. The Workshop identified 19 key findings. Then the findings are amplified with a paragraph each on finding Description/Rationale, Example/Experience Validating Finding, and Benefit of Adoption.

The findings describe attributes of successful measurement program which are precursor research to the study of global climate change. The Workshop participants find that programs which incorporate most of these attributes will be significantly more successful in their study of global climate change than programs which include few or none of these attributes in their design.

Following the Executive Summary is a comprehensive report which documents the reports of the traditional disciplines to the plenary session, a description of a number of calibration, validation, and data set continuity issues to set the context and perspective of the meeting, and every invited and contributed case study provided by attendees of the Workshop.

RECOMMENDATIONS

Programmatic Support

Finding 1: Programmatic Support for Calibration

Programmatic support for calibration must be an integral part of every operational system:

- Dedicated calibrators and ground truth
- Infrastructure maintenance and upgrades
- Reanalysis of prelaunch calibration data
- Rederivation of radiances from raw data
- Success in developing new calibration techniques and instrumentation requires a firm funding commitment for laboratory calibration development.

Finding 2: Calibration by More Than One Group

Calibration and validation by more than one group will improve accuracy of, and confidence in, the sensor performance and the derived final product.

Finding 3: End-to-End Model Sensitivity Runs

End-to-end sensitivity models are required early in the systems engineering process. The models are needed to establish instrument performance requirements in terms of derived geophysical parameter requirements.

Finding 4: Establishment of Climate Data Set Panels and Supporting Research Workshops

Climate data set panels, similar to the trend panel for ozone, should be established and convened periodically to critically review climate data sets. Critical comparisons and standardizations between similar geophysical variables derived from dissimilar measurement systems and concepts is needed. Periodic workshops should be held to assess, document, and achieve consensus on necessary standards and techniques.

Preflight Calibration

Finding 5: Early Involvement of Users

Users should be involved early in instrument calibration and characterization activities.

Finding 6: Greater Prelaunch Calibration Effort

Greater prelaunch calibration effort is required to:

- Thoroughly explore and suitably quantify instrument behavior
- Calibrate sensors using physical standards

Finding 7: High Contrast Versus No Contrast Scenes

Sensors are calibrated on "no contrast" scenes, while much of the science data come from high contrast scenes—resolution of this discrepancy requires further research.

Finding 8: Round-Robin Measurement Comparisons

Round-robin measurement comparisons are essential components for error budget evaluation—consistency may be more important than "truth" and uncertainty is a more important concept than accuracy.

In-Flight Calibration

Finding 9: Well-Thought-Out On-Board Calibration Systems

Greater investment in well-thought-out on-board calibration systems is worthwhile; these are frequently the optimum method for tracking performance.

Finding 10: Capturing Drifts of On-Board Calibrators

The performance of on-board calibrators must be confirmed through ground-based validation. This information should be incorporated into calibrated radiances in postlaunch processing systems by the data providers.

Finding 11: Potential Benefit of Lunar Observations

Lunar observations may be useful for reflected solar band long-term trending and cross calibration; flight platforms should be designed to accommodate maneuvers.

Finding 12: Calibration of Direct Broadcast Data

Calibration information for direct broadcast data should be made available in a timely manner either through a data header, the World Wide Web, or similarly accessible techniques.

Finding 13: More Systematic and Timely AVHRR Calibration

A more systematic and timely on-orbit calibration effort of the Advanced Very High Resolution Radiometer (AVHRR) series is encouraged. This may extend to include improved airborne sensor systems for congruent underflights and comparisons with better-calibrated satellite sensors.

Data Set Continuity and Consistency

Finding 14: Intersatellite Overlap Periods

Overlap between successively launched satellites is necessary for accurate intercalibration, and should be an essential mission planning factor when seeking accurate long-term records of geophysical parameters. As much as one year of overlap may be required when data are taken from instruments with a dissimilar design.

Finding 15: Multiple Observing Techniques Offer Increased Parameter Confidence

Observations of a specific geophysical parameter by different techniques on different systems give increased confidence in its determination.

Finding 16: Geosynchronous Satellite Utility

Geosynchronous satellites may offer utility for diurnal cycle or intersatellite gap correction, though this technique is not yet fully proven.

Combining Remotely Sensed and *In Situ* Data

Finding 17: Use of *In Situ* Data for Insight, Rather Than Tuning

Satellite and *in-situ* systems and their data should only be combined for comparing end results and giving insight into the satellite measurements, not for tuning to "ground truth."

Finding 18: Ground Test Site Operations

Semi-continuous staffing and instrumentation of ground test sites for validation and radiometric calibration is needed. Systematic use of vicarious data from such ground-based test sites is valuable.

Finding 19: Land Cover Change Detection

Subpixel geometric registration accuracy is required to ensure adequate detection of land-cover properties change.

NARRATIVE DESCRIPTION OF RECOMMENDATIONS

Programmatic Support

Finding 1: Programmatic Support for Calibration

Programmatic support for calibration must be an integral part of every operational system:

- Dedicated calibrators and ground truth
- Infrastructure maintenance and upgrades
- Reanalysis of prelaunch calibration data
- Rederivation of radiances from raw data
- Success in developing new calibration techniques and instrumentation requires a firm funding commitment for laboratory calibration development.

Description/Rationale—Comprehensive knowledge of an instrument's character evolves over time, maturing only years after launch when the observations have been thoroughly studied, all artifacts explained, the explanations implemented in improved calibration algorithms, and the algorithms validated. Such an evolving understanding of an instrument can best be conducted when: (1) comprehensive calibration and characterization data are taken on the ground, both prelaunch during ambient and thermal vacuum chamber testing and postlaunch for anomaly characterization; (2) well calibrated collocated, coangular, and coincident *in situ*, "vicarious," or ground truth observations are obtained; and (3) the commitment is made to reanalyze calibration data and regenerate radiances from the raw data as necessary to meet evolving science requirements. Improved geophysical measurements are derived from studies of new dimensions of a physical scene or improved accuracy in the measurement within an existing dimension. Use of a new measurement technology or improvement of existing technology requires the supporting calibration capability. The calibration capability will be present only when a funded work program is present to develop that capability.

Examples/Experience Validating Finding—Currently, more than 15 years after the launch of Nimbus-7, an optimal global total ozone time series is being produced from TOMS data using Version 7 calibration and retrieval algorithms. The process of working through the set of released and interim algorithm versions has resulted in the continuous, measurable improvement of the quality of this very important climate time series. Similar lessons learned have been experienced in the fields of the Earth's radiation budget, ocean color, and medium/high resolution surface and cloud imaging. TOMS sensors developed after the Nimbus-7 SBUV/TOMS have required improvements in laboratory characterization and calibration of their flight diffuer plates.

Benefit of Adoption—The initial data set errors for these instruments usually are controlled by the uncertainty of sensor performance. Characterization of an instrument in its dynamic orbital environment, and the definitive calibration and validation of its measurement record, yields the most fundamental tool to the global change researcher. The global change signal usually is very small and is easily overwhelmed by noise introduced by the sensors or errors introduced by the algorithms. Consequently, reanalysis of a measurement time series is essential to proper understanding of the climate change signal in the dataset. These translate directly into improved research results, and increased confidence in scientific and policy conclusions drawn from studies using the data. Data sets spanning a decade or greater require several copies of a particular sensor, or multiple sensors of an evolving design. Even in multiple copies of a single design, the sensors have individual differences. Continuation of laboratory calibration support can provide a common calibration of each sensor in the series. This common calibration then assists in providing a common and continuous calibration to the derived climate data set. Implementing this finding requires archival of calibration and test data.

Finding 2: Calibration by More Than One Group

Calibration and validation by more than one group will improve accuracy of, and confidence in, the sensor performance and the derived final product.

Description/Rationale—Scientific process requires critical assessment and reproducable results. Scientific process hinges on publication of findings, followed by their replication and cumulative refinement by other researchers using other data and/or methods. This is equally true for the calibration discipline and validation of remotely sensed data, and the instruments used to develop these data. The greater the accuracy needed for a measurement, compared to the "state-of-the-art," the more important it is to have multiple teams using independent study techniques to identify and remove biases and ambiguity in the performance of the sensor providing that measurement.

Examples/Experience Validating Finding—The Global Ozone Monitoring Experiment (GOME) flying on ESA's ERS-2 was calibrated by the contractor using flat-plate diffusers. Additional calibrations using NASA's large-aperture integrating sphere validated these comparisons since both techniques provided radiance calibration constants consistent to within 2 percent. In addition, BRDF measurements of diffuse targets by the NASA and ESA facilities gave consistent results. This helps ensure that, given postlaunch changes properly accounted for, ozone data sets produced by GOME are consistent with U.S. UV backscatter measurements.

Benefit of Adoption—The most accurate and valid, and therefore the most accepted, global data will have been examined by more than one research group. Critical assumptions relating to calibration and retrieval algorithms will be independently proven. Biases, nonlinearities, and uncertainties will be identified and corrected.

Finding 3: End-To-End Model Sensitivity Runs

End-to-end sensitivity models are required early in the systems engineering process. The models are needed to establish instrument performance requirements in terms of derived geophysical parameter requirements.

Description/Rationale—Instruments are developed to take measurements, which are in turn fed through a chain of algorithms, resulting in a description of the state of a geophysical variable, and its uncertainty. The inverse of this transformation process relates the sensitivity of the physical variable estimate to any, or a set of, instrument design or implementation details. High fidelity observing system simulation experiments, or focussed sensitivity studies, provide precise relationships between changes in instrument design and observation. This is useful in the conceptual design, in the design refinement process paralleling development, and as a diagnostic tool during fabrication, calibration, and operations..

Examples/Experience Validating Finding—Simulations of orbit, scan pattern and period, instantaneous field of view, clouds and surface, and retrieval algorithms were central to the selection and refinement of the Clouds and the Earth's Radiant Energy System (CERES) instrument by NASA/LaRC for EOS. Temporal, spatial, and angular sampling errors were quantified and minimized in the design process, and accuracy tradeoffs resulting from refinements to the design to enhance instrument lifetime were facilitated. For the Moderate Resolution Imaging Spectroradiometer (MODIS) for EOS, NASA/GSFC was able to define acceptable limits of band-to-band misregistration through sensitivity studies using Landsat TM data, and has used radiative transfer-based simulations to evaluate and approve/disapprove alternative spectral band passes. Dr. Townshend (University of MD) was able to demonstrate the need for subpixel geometric registration accuracies in conducting land-cover global change research.

Benefit of Adoption—In the presence of such end-to-end model sensitivity runs, there exist quantitative links between instrument design and science requirements. Costs would be saved by relaxing requirements in "tall pole" areas, were that performance is not essential to the science product. Furthermore, complete analysis of instrument design changes could result in preservation of data quality due to anticipation of potential consequences.

Finding 4: Establishment of Climate Data Set Panels with Supporting Research Workshops

Climate data set panels, similar to the trend panel for ozone, should be established and convened periodically to critically review climate data sets. Critical comparisons and standardizations between similar geophysical variables derived from dissimilar measurement systems and concepts is needed. Periodic workshops should be held to assess, document, and achieve consensus on necessary standards and techniques.

Description/Rationale—Fusion of data sets from multiple sensors to produce a climate data set merits special and specific attention. It cannot be accomplished by casual mixing of tables and figures from several scientific papers. The successful development of climate data sets requires detailed review and critical assessment of the chracteristics and uncertainties of the individual data sets. Consequently, the construction of important climate data sets should be assigned to a team of scientific researchers selected specifically for this purpose. Frequently this fusion process requires additional measurements, comparison measurements, and in some instances, may require development of a sensor to make measurements during particular physical circumstances. These investigations usually require advice and input from a broader science community in thematic workshops.

Examples/Experience Validating Finding—The Ozone Trends Panel, a committee of peers, has been in place for a decade. The Panel, a mix of theorists and experimentalists, meets every two years and publishes an update to their report. The Panel provides international leadership and recommends changes in research direction (e.g., the need for improved instrument characterization and calibration) when appropriate. The Trend Panel approach has provided a more critical and to-the-point evaluation of measurements and ideas than has been traditionally provided through the normal peer-review process.

Benefit of Adoption—A climate data set derived in this method is more accurate due to bettering the oversight and consensus which is derived from interested scientific researchers working specifically to develop it. The full discussion and resolution of uncertainties leads to broad scientific acceptance of Panel recommendations. International participation in process helps enable appropriate political processes when such action is necessary.

Preflight Calibration

Finding 5: Early Involvement of Users

Users should be involved early in instrument calibration and characterization activities.

Description/Rationale—It is very difficult to fully specify the set of desired instrument characteristics. Even the best attempts are subject to optimistic or divergent interpretations. It is difficult to complement the specifications with more qualitative descriptive information on the intent of the specifications. By involving the data users directly early in the process, the instrument builder receives direct, unaltered, and unambiguous guidance as to the users' requirements. An instrument project scientist with suitable authority and visibility can represent the users.

Examples/Experience Validating Finding—Presence of users in the ATSR calibration program ensured that the objectives were attained and documented. A strong role by the GOES Project Scientist went a long way towards NASA's stunningly successful technical achievements with GOES I and J. The NASA management model of a manager keeping cost and schedule responsibility with an attendant scientist to track conformance to science requirements has been successful for several decades.

Benefit of Adoption—The result of this involvement is the delivery of instruments that most closely meet the users requirements and expectations. It also promotes acceptance of the development process by the users, obtaining their concurrence for all decisions, and their guidance when hard decisions must be made concerning performance versus cost. The learning curve in understanding on-orbit instrument characteristics also is reduced due to the accelerated familiarization of the user with the instrument engineering. Resultant improvements in communications across the engineering-science interface can facilitate the timely and innovative resolution of problems early.

Finding 6: Greater Prelaunch Calibration Effort

Greater prelaunch calibration effort is required to:

- Thoroughly explore and suitably quantify instrument behavior
- Calibrate sensors using physical standards

Description/Rationale—It is insufficient to simply obtain a set of measurements that relate an instrument quantity (e.g., detector output voltage) to a stimulus value (e.g., target temperature). It is important that these measurements be repeated under a wide variety of static and dynamically varying conditions. In this manner, the effects of many individual elements on the system, and their dependence on the external environment can be documented. Some such observations may not yield calibration points, but rather simply a deeper understanding of the instrument itself (e.g., how does the measurement drop off as calibration lamps degrade?). Confidence in calibration of an instrument is enhanced when multiple, independent calibration techniques produce similar results. Greater understanding of the instrument performance is needed when the derived data sets will be subjected to quantitative rather than qualitative interpretation.

Examples/Experience Validating Finding—Two examples where a more extensive calibration effort led to improved knowledge of, and confidence in, a satellite instrument can be found in the prelaunch calibration of Landsat-4. The implementation of a prelaunch calibration test involving the use of a multiwedge collimator revealed gross inaccuracies in the Landsat-4 data processing software. In addition, extensive thermal vacuum testing of Landsat-4 revealed a long time period icing problem.

Benefit of Adoption—Such measurements become invaluable in explaining an instrument's actual on-orbit performance. The use of physical standards permits the fusion of individual data sets, that have been calibrated to the same physical standard, with high confidence.

Finding 7: High Contrast Versus No Contrast Scenes

Sensors are calibrated on "no contrast" scenes, while much of the science data come from high contrast scenes—resolution of this discrepancy requires further research.

Description/Rationale—The process of instrument calibration has traditionally been conducted in a highly controlled environment, where factors that introduce uncertainty are eliminated to the maximum extent possible. Developing a transfer from instrument voltages to radiance units is facilitated through the use of a known, uniform and constant source. Actual observing conditions are highly variable in both space and time. Near-field and out-of-field responses from, say, clouds for an ocean color imager or ocean for a cloud imager have the potential to greatly contaminate an observation's radiometry. This becomes a greater concern for observations of high dynamic range by instruments with complex optics. Careful attention to design and laboratory characterization under a variety of nonuniform illumination sources permits its abatement during fabrication and possible correction by the ground processing system.

Examples/Experience Validating Finding—Laboratory measurements taken with state-of-the-art imagers, such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and MODIS, have demonstrated that 10 and 12-bit optical systems are susceptible to contamination from bright or dark sources near a target of interest. Crosstalk, scatter, ghosting, and other factors introduce data artifacts. Particulate contamination on-orbit contribute to these effects. Design refinement can abate these noise sources, but some artifacts inevitably remain to be corrected, documented, or ignored by the data users.

Benefit of Adoption—By quantifying an instrument's response to high contrast scenes, and by reducing or eliminating such a near-field or out-of-field response, observations may be corrected for these artifacts. Clouds will not appear darker, especially near the edges, and ocean color will not be significantly impacted near large cloud decks or coastlines. Where residual contamination (e.g., near cloud boundaries or coastlines) remains, users, guided by these known characterizations and their unique study requirements, can select with confidence the amount of data to mask out. Software corrections for the artifacts may be possible, with physical models of these characteristics.

Finding 8: Round-Robin Measurement Comparisons

Round-robin measurement comparisons are essential components for error budget evaluation—consistency may be more important than "truth" and uncertainty is a more important concept than accuracy.

Description/Rationale—Earth system science mandates the detection, description, and explanation of change, locally and globally, over extended periods. This extended spatio-temporal domain ensures that multiple sensors are used for full spatial, diurnal, and temporal coverage. To make sense of this heterogeneous data set, we should have high confidence in the consistency of the observations. For change detection, consistency often is more important than the absolute accuracy of the combined data set. Accuracy usually is established in a measurement by the measurement uncertainty. Consequently uncertainty usually is the more fundamental quantity and as such is a more important concept than accuracy.

Examples/Experience Validating Finding—As an example, repeated round-robin campaigns have enabled consistent calibrations across flight and *in situ* calibration sources for SeaWiFS. Inconsistencies of up to 5 percent have been reduced to within 1 percent.

Benefit of Adoption—Benefits of general adoption of round-robin comparisons include improved consistency in the observations (over time and among ground, aircraft, and space-based platforms), more reliable conclusions, and reduced uncertainty over the veracity of the findings. Consistency among various instruments observing Earth science parameters which overlap can be significantly improved if the various instruments are calibrated with the same standard. Consistency of measurements made by several instruments of the same parameter provides confidence in the measurements of other parameters made by these instruments. Benefit of using uncertainty includes clearer use of language. Additionally, when we want a 1 to 5 percent measurement, if we continue to use accuracy, we are frequently in error. As the quality of a measurement improves, the measurement accuracy gets larger, approaching 100 percent and the uncertainty gets smaller, approaching 0 percent.

In-Flight Calibration

Finding 9: Well-Thought-Out On-Board Calibration Systems

Greater investment in well-thought-out on-board calibration systems is worthwhile; these are frequently the optimum method for tracking performance.

Description/Rationale—Well-designed on-board calibrators are demonstrating their utility both for carrying absolute radiometric scales into orbit, and for on-orbit maintenance of the time-series' precision. For many technologies, self calibration is possible. Others require periodic reference to blackbodies, deep space, the Sun or Moon, or on-board lamps. Because such on-board calibrators themselves can often degrade over time, a self-monitoring capability is often a requirement for these systems. Deep space, the Sun, and the Moon are readily available, mainly stable sources than can be used in common among many programs.

Examples/Experience Validating Finding—The solar diffuser plates on the Total Ozone Mapping Spectrometer (TOMS) and Solar Backscatter Ultraviolet (SBUV/2) series of instruments have proven invaluable in maintaining the accuracy and precision of long-term ozone time series. Such diffusers are now routinely designed into many moderate resolution visible imagers, including those designed for ocean color, UV/ozone, and land cover mapping. Similar calibration with respect to the Sun has been employed to good effect with the Earth Radiation Budget Experiment (ERBE) and ScaRaB series of instruments.

Benefit of Adoption—Global change observation is limited by our ability to differentiate changes in instrument character from changes in components of the Earth system. Both are time dependent, and changes in instrument character often mimic geophysical trends. Only through the use of properly designed on-board calibration systems can we capture these changes with high accuracy.

Finding 10: Capturing Drifts of On-Board Calibrators

The performance of on-board calibrators must be confirmed through ground-based validation. This information should be incorporated into calibrated radiances in postlaunch processing systems by the data providers.

Description/Rationale—On-board calibrators (e.g., lamps, solar diffuser panels, blackbodies, silicon photodiodes) enable the on-orbit maintenance of instrument calibration; however, there is evidence documenting the drift of on-board calibrators and their in-flight monitors. Ground-based validation of the on-board calibrators should be designed to quantify the drifts of these systems. On-board calibrator drift can be confused with changes in the satellite sensor.

Examples/Experience Validating Finding—A new solar diffuser stability monitor system was flown on the NOAA-9 SBUV/2 sensor. This on-board calibrator system was very noisy during the mission due to an instability of the on-board light source. Initial ozone data interpretation based on this monitor was wrong due to degradation of the monitor. Understanding of this design error led to an improved design for NOAA-11.

Benefit of Adoption—Early detection on the ground of on-board calibrator drift can lead to improved calibrator design. Realistic uncertainty on calibrator drift improves interpretations of the satellite data set through better understanding of the data set stability and can reduce misdirected effort in the study of sensor stability which is actually on-board calibrator instability. In the instance where the science product is derived from observations of reflected sunlight when a solar diffuser is used, the instrument measurements may be virtually independent of instrument changes on-orbit.

Finding 11: Potential Benefit of Lunar Observations

Lunar observations may be useful for reflected solar band long-term trending and cross calibration; flight platforms should be designed to accommodate maneuvers.

Description/Rationale—The reflectance of the moon remains constant in one part in 10^9 per year, and thus far surpasses the stability of any on-board or terrestrial target. Periodic reference to the Moon at similar viewing/illumination angles, or with suitable BRDF corrections, permit instrument calibration to be tracked in a highly precise manner more than adequate for the most discerning global change research requirements. Similarly, views of deep space provide a dark current check. Cross-calibration by two or more platforms can be effected without the uncertainties introduced by differing paths or observing times through the Earth's atmosphere and off the Earth's surface through simultaneous observation of the Moon. Flight software and hardware must be designed with such maneuvers in mind.

Examples/Experience Validating Finding—To date, no quantitative use has been made of the moon for maintaining the precision of Earth remotely sensed observations. The SeaWiFS sensor has demonstrated its ability to acquire lunar data in a prelaunch ground-based test. Observations have begun at USGS in Flagstaff, AZ towards compiling a spectral lunar reflectance model across one-fourth of a 21-year lunar libration cycle. Precision is obtained through reference to an ensemble average of stellar observations.

Benefit of Adoption—Lunar observations will permit verification of an instrument's stability, serving either as a stand-alone in-flight calibrator or as an independent validation of the precision of a calibrated time series. Lunar observations taken during an instrument lifetime can be used as anchor points tying together time series from multiple instruments years into the future. Views of deep space provide comparable offset verifications.

Finding 12: Calibration of Direct Broadcast Data

Calibration information for direct broadcast data should be made available in a timely manner either through a data header, the World Wide Web, or similarly accessible techniques.

Description/Rationale—Typically, sensor characteristics change with time in orbit. These changes may be in the amount of a few percent yearly; so they can reach 10 percent or more after several years. Scientists getting data in ways other than through the routine data archives for that sensor will not know the proper characterization and calibration to use for that data.

Examples/Experience Validating Finding—The NOAA AVHRR sensors in reflected solar bands are believed to experience 5 to 6 percent degradation per year, and distribution of data sets on tape with current calibration parameters has not occurred. Researchers have expended significant effort in attempting to supply proper calibration values to their data sets.

Benefit of Adoption—Research effort by the science community will be focused on geophysical research rather than on the task of calibration validation. With the accessibility of systems such as the World Wide Web, provision of calibration data sets and coefficients can be provided in a timely manner at very low cost if advanced planning is done.

Finding 13: More Systematic and Timely AVHRR Calibration

A more systematic and timely on-orbit calibration effort of the Advanced Very High Resolution Radiometer (AVHRR) series is encouraged. This may extend to include improved airborne sensor systems for congruent underflights and comparisons with better-calibrated satellite sensors.

Description/Rationale—AVHRR observations have now accumulated from dual time series of observations, each with global coverage and prescribed equator-crossing times extending across multiple instrument copies. Until recently, the data provider has not provided timely calibration information, and data users have been forced to rely on their own resources to remove time-dependent effects. Coefficients obtained through ongoing calibrations using stable ground targets, such as the Libyan desert, can be rapidly disseminated through the World Wide Web. These can be periodically validated through truth measurements obtained through aircraft under flights. A large amount of scientific research is based on AVHRR observations.

Examples/Experience Validating Finding—Pathfinder activities at NASA, NOAA, and academia have demonstrated that it is possible to produce timely and valid radiometric calibrations of the AVHRR instrument series, and that there is great value in the production of radiometrically calibrated AVHRR radiance products for quantitative analysis.

Benefit of Adoption—By rapidly making available to the user community a calibrated AVHRR data record that takes into account time-dependent radiometric effects, the already great value of this instrument will be substantially increased. Users will be able to access the product and perform reliable global change research from its parameters, with much reduced contamination from instrument drift. Similar benefits can be realized for other operations imager data sets, including those in geosynchronous orbit. Scientific effort in the use of the data can focus on scientific understanding rather than estimating sensor performance.

Data Set Continuity and Consistency

Finding 14: Intersatellite Overlap Periods

Overlap between successively launched satellites is necessary for accurate intercalibration, and should be an essential mission planning factor when seeking accurate long-term records of geophysical parameters. As much as one year of overlap may be required when data are taken from instruments with a dissimilar design.

Description/Rationale—Due to finite instrument and satellite lifetimes, and the extended multidecadal nature of global change research, requisite time series of geophysical products must necessarily be composited across data taken by multiple individual instruments. Continuity is assured through comparison among data sets chronologically overlapping each other. For instantaneous observations at identical viewing and illumination angles, surfaces, spectral bands, and times, such comparisons of direct observables are easily made within timespans of several orbits to one month. Where geophysical products are from dissimilar measurement technologies, overlaps of much longer periods may be needed to overcome biases due to seasonal changes and sensor variations.

Examples/Experience Validating Finding—Earth radiation budget measurements taken by the Nimbus-6 and -7 were compared during a year of overlap using a "parallel orbit" technique. These were then tied to data from ERBE, again through extended comparison. ERBE data from NOAA-9, -10, and ERBS were combined only after extensive and prolonged comparisons. SBUV/2 calibration improvements were accomplished with periodic SSBUV shuttle flights of up to 10-day duration, where the SBUV/2 and SSBUV sensors are of nearly identical design, with nearly identical observables. In the case where the ozone trends panel used SAGE and SBUV measurements, the comparison involved many months of data because it was necessary to understand the atmospheric effects which are essential components of the underlying algorithms used to obtain ozone from the SAGE and SBUV observations.

Benefit of Adoption—Overlapping data sets greatly reduce the possibility of uncertainties in absolute calibration being misinterpreted as variability on the timescale of the individual instrument records. Comparison between measurements taken by an instrument after an extended period in space against one recently calibrated also offers an important data point to verify absolute calibration, and eliminate the potential for confounding on-orbit trends. The net result is the production of global change data sets extending over decades with higher confidence in the time series' calibration and validity.

Finding 15: Multiple Observing Techniques Offer Increased Parameter Confidence

Observations of a specific geophysical parameter by different techniques on different systems give increased confidence in its determination.

Description/Rationale—Remote sensing uses forward or inverse mathematical models to relate measurements to geophysical quantities. A chain of assumptions, simplifications, approximations, and uncertainties is inherent in this process. Both random and systematic errors can result; these are sometimes undetected. When different observing techniques are employed, a different, and often independent set of models are employed. Comparison of the retrieved estimates serves as a strong validation tool.

Examples/Experience Validating Finding—Cloud-top height has benefitted from infrared, lidar, and stereo techniques. Total ozone retrieval benefits from spectral infrared, ultraviolet, and ground-based techniques. Precipitation measurement is improved through comparison of GOES infrared, passive microwave, ground-based radar, and rain gauge and space-based precipitation radar with the launch of TRMM.

Benefit of Adoption—The direct benefit of the use of multiple, partially redundant, technologies for geophysical parameter retrieval lies in the increased opportunity for validation, and the resultant increased confidence in the data products on the part of the consumers. Systematic errors are more easily identified and eliminated through a comparison of such independent techniques. As a result, uncertainties are quantified, and the product can be improved.

Finding 16: Geosynchronous Satellite Utility

Geosynchronous satellites may offer utility for diurnal cycle or intersatellite gap correction, though this technique is not yet fully proven.

Description/Rationale—Geosynchronous satellites cannot achieve global coverage alone because they remain at a near-constant point relative to the Earth's surface, and each can only view some 25 percent of the Earth's surface. They are able to stare at a region continuously and obtain temporal images at frequencies impossible for even multiple satellites in polar orbit. These measurements, though narrow band, could be used to bridge between sun-synchronous satellite observations across the diurnal cycle. In the absence of any or sufficient overlap between polar orbiters, geostationary data may serve as a bridge to transfer calibration from one instrument to another. Their usefulness for global climate change studies must be investigated.

Examples/Experience Validating Finding—Global cloud and precipitation projects are demonstrating the utility of geostationary data for calibration transfer, and conversely, for polar orbiters to transfer calibration among geostationary satellites. Earth radiation budget studies have demonstrated the importance of geostationary data for the generation of robust diurnal correction models, though the measurements have yet to be assimilated explicitly into description of a specific day's diurnal cycle.

Benefit of Adoption—The assimilation of geostationary and polar-orbiter data into a consistent description of the Earth, both in terms of calibration and diurnal cycle, will optimize the information content present in the complementary observing platforms. Geosynchronous platforms, either dedicated or through a "piggy back" approach on commercial communications satellites, can provide a unique measurement platform capable of continuous and diverse environmental monitoring (e.g., lightning, pollution, etc.) at continental scales. Accomplishment of this objective provides better integration of geosynchronous satellite observations into the NASA Mission to Planet Earth climate change program.

Combining Remotely Sensed and *In Situ* Data

Finding 17: Use of *In Situ* Data

Satellite and *in-situ* systems and their data should only be combined for comparing end results and giving insight into the satellite measurements, not for tuning to "ground truth."

Description/Rationale—Ground "truth" is a key element in independent product validation. However, space and ground-based observations often perceive subtly different, though related, attributes of a variable (e.g., surface skin water temperature versus temperature measured by a buoy or at the engine inlet; point measurements of rainfall versus area averages; viewing angle and observer bias differences associated with cloud amount observation). By adjusting (or tuning) one system to fit the measurements of another, independence of the two systems is lost, and a validation tool is wasted.

Examples/Experience Validating Finding—Total ozone measurements are produced using space-based differential UV-reflectance algorithms. Validation against the ground-based Dobson spectrophotometer network provides a measure of uncertainty and bias in the field as a whole and of its trend. Because the data are not combined, the integrity of each data set is preserved.

Benefit of Adoption—In-situ data are best used to indicate potential problems in the remotely sensed data set. Preserving the independence of such related data sets provides investigators with a continuing means for product validation and improvement.

Finding 18: Ground Test Site Operations

Semi-continuous staffing and instrumentation of ground test sites for validation and radiometric calibration is needed. Systematic use of vicarious data from such ground-based test sites is valuable.

Description/Rationale—Use of data from test sites to validate the sensor performance on-orbit is needed to track changes in performance. The data from test sites is most useful when the sites are well instrumented and staffed to certify the characteristics of ground instruments and the test site. Data from test sites which are not carefully instrumented and staffed, but with characteristics that are known, also can be useful either through a network of sites or through periodic revisit and recertification of sites. In the validation climate studies for specific parameters, test sites should be chosen for which the parameter desired can be measured or easily inferred. Test sites which have either stable or well characterized atmospheric conditions above them may be useful to track top-of-the-atmosphere radiances as well. A primary benefit for the use of data from instrumented test sites is that the data from that site can be recalibrated periodically, and thereby used to track the changes in the satellite instrumentation.

Examples/Experience Validating Finding—The Landsat Thematic Mapper and SPOT sensors have been calibrated on-orbit through the use of overflights of instrumented test sites in New Mexico and Nevada. In both instances, there have been on-board calibration devices and a prelaunch calibration, but the best calibration has been provided through a test site which was staffed and instrumented for the occasion. Ozone trends detected by TOMS and SBUV sensors have been tracked and improved through comparison with the Dobson ozone ground network.

Benefit of Adoption—Sensors that operate without requiring continuous maintenance can be placed at sites which are not continuously staffed. By matching ground sensor maintenance to staffing assignments, the best use of money and staff can be used to obtain broadest range of needed science products. A network of sites which are not instrumented, but which have moderately well-understood characteristics also is useful for trend detection of sensor on-orbit performance.

Finding 19: Land Cover Change Detection

Subpixel geometric registration accuracy is required to ensure adequate detection of land-cover properties change.

Description/Rationale—Change detection of derived biogeophysical properties is more valuable than single observations for any region. This requires registration between images taken at different times. Change detection is accomplished on the basis of physical models. Data sets used to initiate and validate those models must be available at a resolution equal to or smaller than the "grid" resolution of the analysis models to be useful input data to those models. Sensors with an intrinsic resolution of 1 X are useful for models with a 1 X resolution only where the data are registered to about 0.25 X (or a factor of four or better subpixel registration). Most scenes are nonlinear; consequently, the quantitative understanding of change within a scene requires data sets with a spatial resolution typical of the decorrelation length of the information contained within that scene. Otherwise, the nonlinear effects within that analysis model resolution grid cell may be interpreted incorrectly.

Examples/Experience Validating Finding—Relatively high spatial resolution systems such as SPOT have demonstrated Landsat images over land are much more complex than appears in the TM products themselves. Similarly, Landsat images demonstrate the heterogeneity of targets in the AVHRR mosaics that have historically been used to monitor change in vegetation over the whole globe. Land cover research has a more strict registration requirement due to the shorter decorrelation lengths for the scenes observed.

Benefit of Adoption—Subpixel geometric registration of land cover products derived from satellite imagery is necessary to drive physical modeling of land-surface change models for models with that geometric registration. Failure to provide that registration limits the models to degraded surface change detection due to the inability to interpret typical nonlinearity within the scene. The global change research program community can obtain results based on physical models at the intrinsic resolution of the satellite data.

TABLE OF CONTENTS

LIST OF ACRONYMS

AATSR	Advanced ATSR
ACRIM	Active Cavity Radiometer Irradiance Monitor
ADEOS	Advanced Earth Observation System
AIRS	Atmospheric Infrared Sounder
AMSR	Advanced Microwave Scanning Radiometer
AMSU	Advanced Microwave Sounding Unit
ANF	Albedo Normalization Factor
ARC	Active Radar Calibrator
ARIES	Airborne Research Interferometer Evaluation System
ARMAR	Airborne Rain Mapping Radar
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ATBD	Algorithm Theoretical Basis Document
ATMOS	Atmospheric Trace Molecules Observed by Spectroscopy
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BNSC	British National Space Center
BRDF	Bidirectional Reflectance Distribution Function
BUV	Backscatter Ultraviolet
CAMPR	Multiparameter Radar
CDR	Critical Design Review
CENR	Committee on Environment and National Resources
CEOS	Committee on Earth Observation Satellites
CLAVR	Clouds from AVHRR
COADS	Comprehensive Ocean-Atmosphere Data Set
COARE	Coupled Ocean/Atmosphere Response Experiment
CP	cross polarization
CRF	cloud radiative forcing
CRL	Communications Research Laboratory
CSIRO	Commonwealth Scientific and Industrial Research Organization
CZCS	Coastal Zone Color Scanner
DEM	Digital Elevation Model
DIEP	diurnal interpolation/extrapolation procedure
DMSP	Defense Meteorological Satellite Program
DOAS	Differential Optical Absorption Spectroscopy
DOD	Department of Defense
EFOV	effective field of view
ENSO	El Niño/Southern Oscillation
ENVISAT	Environmental Satellite
EOF	empirical orthogonal functions
EOS	Earth Observing System
EOS-Chem	Earth Observing System Chemistry Platform
EP	Earth Probe
ERB	Earth Radiation Budget
ERBE	Earth Radiation Budget Experiment
ERBS	Earth Radiation Budget Satellite
EROS	Earth Resources Observation Systems
ERS-2	Earth Remote-sensing Satellite-2
ESA	European Space Agency
EUMETSAT	European Meteorological Satellite

FAPAR	fraction of absorbed photosynthetically active radiation
FGGE	First GARP Global Experiment
FPA	focal plane assembly
GAC	global area coverage
GCOS	Global Change Observing System
GEWEX	Global Energy and Water Exchange Experiment
GHz	GigaHertz
GISS	Goddard Institute for Space Studies
GLI	Global Imager
GOES	Geostationary Operational Environmental Satellite
GOME	Global Ozone Monitoring Experiment
GOOS	Global Ocean Observing System
GSFC	Goddard Space Flight Center
GTOS	Global Terrestrial Observing System
GTS	Global Telecommunications System
GV	ground truth and validation
HALOE	Halogen Occultation Experiment
HIRDLS	High Resolution Dynamics Limb Sounder
HIRS	High Resolution Infrared Sounder
HPBW	half-power full-beam width
HRV	Haute Resolution Visible
I/O	input/output
IASI	Improved Atmospheric Sounding Interferometer
IFOV	instantaneous field of view
IGBP	International Geosphere-Biosphere Program
IOC	Intergovernmental Oceanographic Commission
IR	infrared
ISCCP	International Satellite Cloud Climatology Project
ISLSCP	International Satellite Land Surface Climatology Project
IVOS	Infrared and Visible Optical Sensors
JGOFS	Joint Global Ocean Flux Study
JPL	Jet Propulsion Laboratory
JUWOC	Japanese-U.S. Working Group on Ocean Color
LAI	leaf area index
LDEF	long duration exposure facility
LIMS	LIMB Infrared Monitor of the Stratosphere
LMC	length modulation cell
LST	local standard time
LUP	liquid water path
MAPS	Measurement of Air Pollution from Satellites
MATR	MOPITT Algorithm Test Radiometer
MERIS	Medium Resolution Imaging Spectrometer
MHS	Microwave Humidity Sounder
MIMR	Multifrequency Imaging Microwave Radiometer
MISR	Multiangle Imaging Spectroradiometer
MLE	Maximum Likelihood Estimate
MLS	Microwave Limb Sounder
MODIS	Moderate Resolution Imaging Spectroradiometer
MOPITT	Measurement of Pollution in the Troposphere
MRF	Meteorological Research Flight

MSS	Multispectral Scanner
MSSG	Microwave Sensors Subgroup
MSSL	Mullard Space Science Laboratory
MSU	Microwave Sounding Unit
MTF	Modulation Transfer Function
MTPE	Mission to Planet Earth
NASA	National Aeronautics and Space Administration
NASDA	National Space Development Agency of Japan
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NDSC	Network for Detection of Stratosphreic Change
NDVI	normalized difference vegetation index
NEdt	Noise Equivalent Delta T
NESDIS	National Environmental Satellite, Data, and Information System
NIST	National Institute of Standards and Technology
NMC	National Meteorological Center
NSTC	National Science and Technology Council
NWP	Numerical Weather Prediction
OCTS	Ocean Color and Temperature Sensor
OI	optimum interpolation
OLME	Ozone Layer Monitoring Experiment
OPT	Ozone Processing Team
PMC	pressure modulation cell
PMT	Photomultplier Tube
POAM	Polar Ozone and Aerosol Measurement
POES	Polar Orbiting Environmental Satellite
PR	Precipitation Radar
PRT	platinum resistance thermometer
PTM	platinum temperature monitors
RFI	radio frequency interference
RMS	root mean square
SAGE	Stratospheric Aerosols and Gas Experiment
SAM	Stratospheric Aerosol Measurement
SAR	Synthetic Aperture Radar
SBUV	Solar Backscatter Ultraviolet
SCI	scene/cloud identification
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SME	Solar Mesospheric Explorer
SMM	Solar Maximum Mission
SNR	signal-to-noise ratio
SPOT	Satellite pour L'Observation de la Torre
SSBUV	Shuttle SBUV
SSM/I	Special Sensor Microwave/Imager
SSM/T-2	Special Sensor Microwave/Temperature-2
SST	sea-surface temperature
SSU	Stratospheric Sounding Unit
TAO	Tropical Atmosphere Ocean
TES	Tropospheric Emissions Spectrometer
TM	Thematic Mapper
TMI	TRMM Microwave Imager

TOGA	Tropical Ocean and Global Atmosphere Program
TOMS	Total Ozone Mapping Spectrometer
TOVS	Tiros Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
UARS	Upper Atmosphere Research Satellite
UNESCO	United Nations Educational, Scientific, and Cultural Organization
WFOV	wide field of view
WGCV	Working Group on Calibration and Validation
WMO	World Meteorological Organization
WWW	World Wide Web

1. ATMOSPHERE GROUP SUMMARY REPORT

Participants in the Atmosphere Group included John Barnett, Chair, Clarendon Laboratory, Oxford, UK; Dennis Chesters, Co-chair, NASA/Goddard Space Flight Center, Greenbelt, MD; Phillip Arkin, National Meteorological Center, Washington, DC; William Chu, NASA/Langley Research Center, Hampton, VA; Ernest Hilsenrath, NASA/Goddard Space Flight Center, Greenbelt, MD; Robert Hudson, University of Maryland, College Park, MD; Robert Kandel, Ecole Polytechnique, Palaiseau, Cedex, France; Christopher Mutlow, Rutherford Appleton Laboratory, Oxfordshire, United Kingdom; Ke'ichi, Okamoto, Communications Research Laboratory, Japan; Nagaraja Rao, National Oceanic and Atmospheric Administration, Washington, DC; Roy Spencer, NASA/Marshall Space Flight Center, Huntsville, AL; Jinxue Wang, NCAR/ACD, Boulder, Colorado.

During the previous decade, scientists monitored global change in the atmosphere by developing strategies to determine variations and trends in temperature, moisture, ozone, clouds, and other parameters, using remote sensing systems on satellites originally designed to observe short-term phenomena such as the weather. Each discipline struggled to understand and control the factors that determine our confidence in trend measurements. These factors include:

- Calibration before, during, and after flight
- Observational continuity among overlapping systems
- Validation campaigns and reviews of products
- Reanalysis for corrections and improvements
- A variety of data product-specific issues

With a decade of systematic climate monitoring now behind us, we describe some of the key strategies that have been successful in each of these five areas, and we make some recommendations for improvement in the future.

1.1 Calibration

The usefulness of space-based data for long-term studies of the Earth requires a well-thought-out program for prelaunch and postlaunch instrument calibration. Such calibration must have sufficient accuracy so that long-term changes in the measured parameter can be separated from apparent changes due to differences in instrument calibration or on-orbit degradation.

1.1.1 Prelaunch

A primary requirement is that instruments be calibrated prior to launch in a manner that relates the instrument calibration in a repeatable and quantifiable manner to international standards, and enables all relevant aspects of the instrument behavior to be thoroughly explored and quantified to an accuracy adequate for all purposes for which the instrument data are required. This implies that a considerably greater effort needs to be devoted to prelaunch calibration than has generally been the case for many of the instruments currently flying. Furthermore, it is important to maintain consistency of calibration and instrument or hardware characteristics between different members of a sensor series, and, where practical, between different types or generations of sensor. Whereas intercalibration in a common facility would be ideal, a more realistic approach is to require very high standards of calibration at each facility.

1.1.2 Postlaunch

The instrument design should include calibration subsystems that are inherently accurate and have long-term stability. With some classes of instrument, this can imply substantial additional complexity and cost, but current experience shows that historically a greater investment in on-board calibration systems would have been well worthwhile. In addition, the on-orbit calibration procedure needs careful planning prior to launch to ensure adequate observing opportunities including those over ground calibration sites, lunar viewing, etc., and also an appropriate trade-off between time spent calibrating and time spent taking geophysical data. On-board calibration systems employing sunlight will degrade; therefore, a careful design and deployment strategy must be employed.

1.1.3 Documentation

A vital part of both prelaunch and postlaunch calibration is the formal documentation and archival of the calibration parameters. These documents should be publicly available so that in principle they would enable the postlaunch data processing to be independently verified with no inside knowledge. In addition, the raw prelaunch data, both from the instrument and from the external calibration equipment, need to be archived to enable re-analysis (conceivably decades later) should the effort ever be found necessary. In the case of postlaunch data, the derived calibration quantities (gains, offsets, wavelength, etc.) should be archived in addition to the raw telemetry and made available in the same way as the standard geophysical products.

1.1.4 Program Support for Calibration

Historically, the major emphasis on a satellite measurement system has been on the design, fabrication and operations in space of the system. The prelaunch and on-orbit calibrations received, and still receive less emphasis, especially with operational systems. Once in operation, the association with calibration becomes tenuous and the infrastructure for calibrating (particularly that which depends on people) becomes eroded. A major lesson learned is that this infrastructure must be maintained and even enhanced as instruments become more complex. As with validation against ground "truth," calibration must be part of the operational system for the lifetime of that system. Dedicated calibrators and dedicated validators are as vital as dedicated instrumentalists and scientists!

1.2 Observation Times and Overlap Between Instruments

1.2.1 Orbit Times

The selection of the local time of Sun-synchronous orbits and the launch frequency have had major impacts on the utility of various data sets for climate monitoring. In particular, the extent by which Sun-synchronous orbits have drifted to new times (e.g., the Tiros-N afternoon satellites) has considerably degraded the confidence in and utility of long time series of temperature, ozone, cloud parameters, and rainfall. Unless orbit times can be accurately maintained, derivation of long-term trends for diurnally varying species will need to rely upon knowledge of that diurnal variation in order for corrections to be made. In some instances (e.g., cloud parameters), it may be possible to correct for diurnal variations or intersatellite data gaps with information from geostationary satellites, although this has yet to be demonstrated successfully.

1.2.2 Overlap Between Instruments

In general, it has been found that at least 1 year of overlap between successively launched instruments is necessary for accurate intercalibration of geophysical data from satellites, especially if the satellites are at different times in a diurnal cycle which has a seasonal modulation. This may be difficult or expensive to achieve.

1.3 Validation

Full confidence in the determination of a parameter will generally be achieved only when another determination of the same parameter is made by a different system, preferably using different techniques. Comparisons of the results by diverse systems and techniques should reveal any flaws in particular systems which can then be addressed by their advocates and so lead to a better estimate of the parameter. Ideally, all systems should produce the same quantitative values (within their error bars), arrived at independently, without any one system being tuned to another. In particular, comparisons between satellite systems and ground "truth" can give information about the long-term operation of satellite systems if, and only if, the ground truth system has already produced a credible and well proven data set. This was well demonstrated by the way that the Dobson/Brewer ozone network was used to help validate the TOMS/SBUV satellite measurements.

1.3.1 Comparison Campaigns

In addition to long-term measurement sequences at individual sites, field campaigns should be held periodically to compare satellite systems with a variety of others, including at least some which are well-proven. Campaigns need to become an integral part of any satellite program to track the time histories of observational differences, whether due to atmospheric, instrumental, algorithm changes, etc. However, flexibility should also be retained to enable quick and detailed study immediately after large volcanic eruptions or other unexpected events. The Upper Atmosphere Research Satellite (UARS) validation effort is a good example of what should be implemented for future observing systems.

1.3.2 Trend Panels

Trend panels should be established and convened periodically to validate the time-dependent geophysical products. A prime example of this was the Ozone Trends Panel wherein expert instrumentalists, modellers and others jointly analyzed all aspects of the basic measurements and retrievals to establish the credibility of trends. This paradigm should be followed for any major parameter of climatic significance and importance.

1.4 Processing Strategies

From the beginning, it must be understood that processing algorithms are mathematical models based on current knowledge of the instrument, the atmosphere, and also basic physics (such as spectral information) and contain assumptions and approximations. However, an objective of the EOS and other environmental observing programs is to improve the knowledge of the atmosphere. Thus, in the course of data interpretation, some assumptions may be found to be incorrect, and better approximations will be formulated. In the case of the instrument model, new information or analyses may become available which necessitate rederivation of radiances from the raw counts or even reanalysis of the prelaunch calibration data.

This implies two things: the processing algorithms will inevitably need to change with time or the primary data product to be archived must be the raw counts from the instrument. We are particularly concerned that space agencies, in their emphasis on getting geophysical data to the users in a timely fashion, might accept on-board processing such that the original data cannot be recovered. Any move in this direction could preclude reprocessing for long-term trend analysis.

It is also important that the processing teams document a complete record of the processing algorithms and any changes that take place. It was suggested that it might be equally important to archive the personnel who had performed the processing and any reprocessing; otherwise, the corporate memory that is associated with the data products will be lost. Recognizing that processing teams cannot perform quality assurance on every piece of data, and that in some areas the basic assumptions will be questioned, this information about the algorithms should be made available to the general scientific community in a timely and open fashion, along with the instrument and calibration data and the data products.

1.5 Product Issues

1.5.1 Temperature

Climate monitoring of temperature variations has been successfully demonstrated with the Microwave Sounding Unit (MSU) and Stratospheric Sounding Unit (SSU) infrared radiometers, and progress is being made with the High Resolution Infrared Sounder (HIRS-2) infrared sounders, all of which are carried on the Tiros-N operational satellites. Future instruments (Advanced Microwave Sounding Unit (AMSU) and Atmospheric Infrared Sounder (AIRS)) should provide for a long-term record of global deep-layer temperatures and the limb sounders on the EOS-Chem satellite will provide higher vertical resolution.

In the case of the MSU, the solar time of observation has emerged as the most troublesome problem, especially the drift by up to a few hours of the orbit times through the diurnal cycle. Radiosonde data have provided useful independent validation information, although changes in the sonde system have caused problems and need to be either avoided altogether (probably impractical), or accurately quantified.

There is a definite lack of upper stratospheric validation information, and this will need to be alleviated to validate the upper level data from the AMSU, from when it is first launched in 1996. Systems which previously provided such data (rocketsondes) are no longer operational, and these or comparable systems (e.g., lidars) need to be reinstated and integrated into an overall satellite global observing system.

1.5.2 Water Vapor

Climate monitoring of water vapor has been demonstrated for total precipitable water over the oceans with the 7-plus-year record of 22 GigaHertz (GHz) Special Sensor Microwave/Imager (SSM/I) data. Infrared-based climate data sets are being developed (e.g., HIRS-2) for a few deep tropospheric layers, but lack of sensor intercalibration and high data volume problems have hindered progress. Over 3 years of Special Sensor Microwave/Temperature-2 (SSM/T-2) 183 GHz data appear very promising for monitoring tropospheric layers analogous to the HIRS measurements, but with the advantage of negligible cloud contamination. In addition, the Stratospheric Aerosols and Gas Experiment (SAGE) occultation instruments have provided information on climatology and variability above cloud level, and the UARS results are providing stratospheric measurements spanning several years and EOS-Chem will continue this sequence after 2002.

The planned future instrumentation (HIRS-3, AIRS, Microwave Humidity Sounder (MHS)) should provide the necessary information for monitoring water vapor in the troposphere. Validation of climatically critical upper tropospheric water vapor, however, remains a serious problem. Various *in-situ* measurement techniques exist for measurements at very low temperatures and humidities, but routine measurements in selected locations are still required.

Because the infrared and microwave water vapor channels are also sensitive to atmospheric temperature change, the long-term record of retrieved water vapor depends upon an accurate measurement of temperature trends so that the two effects can be separated.

1.5.3 Precipitation

Precipitation is notoriously difficult to measure. Passive microwave sensors now flying (SSM/I), and planned for the future (TRMM Microwave Imager (TMI), Advanced Microwave Scanning Radiometer (AMSR) Multifreqency Imaging Microwave Radiometer (MIMR)), will provide a necessary source of routine precipitation information over the oceans. The less-direct passive microwave estimates over land will require all available rain gauge data to provide most of the land coverage. The precipitation radar flown on the 3-year Tropical Rainfall Measuring Mission (TRMM) mission will provide essential validation data for the passive microwave methods.

Diurnal sampling issues continue to plague the polar orbiting microwave measurements, with drifting orbits and differing orbit times adding to the uncertainty in the monitoring of rainfall trends. The geostationary visible and infrared imagers will help to fill in the diurnal gaps left by the polar orbiters, although quantification of cloud features as surrogates for rainfall continues to be problematic (although less so in the tropics than at high latitudes). Continuing field campaigns, with rain gauge and radar observations, will be necessary to resolve the many remaining rain retrieval issues. All data sources intelligently combined together, however, should provide a reasonably complete record of climatic rainfall variability for most uses. At this point absolute accuracy of any of these methods, even for globally averaged rainfall, is probably no better than estimates based upon global heat budget calculations.

1.5.4 Ozone

The measurements of total ozone trends since 1978 have been reasonably successful using back scatter ultraviolet (BUV) techniques. Problems with instrument degradation were overcome with a variety of techniques in a very detailed study by the Ozone Trends Panel in the late 1980's. This has led to better management of on-board calibration systems for the current ozone measurement program to minimize solar degradation. It should be noted that long-term instrument characterization is not operational, and analysis iterations are required using data intervals of 1 year or more to derive accurate time dependent corrections to instrument performance. The ozone data series has the advantage of coming from one class of instruments that will continue with sounders on the Tiros-N satellites, Earth Remote-sensing Satellite-2 (ERS-2), Environmental Satellite (ENVISAT-1), and EOS-Chem. Any systematic biases that might be introduced by the processing algorithm should not, therefore, affect the ability to detect long-term changes. An effort is now under way to ensure that these new instruments have calibrations which can be related to those of previous instruments. The need to be able to reprocess after many years referred to previously is exemplified by these measurements, since there are now moves to reprocess the Nimbus-4 Backscatter Ultraviolet (BUV) data dating back to 1970 with a calibration that is consistent with that of the later instruments.

The above considerations also apply to measurements of the ozone profile in the stratosphere which is also obtained by the BUV technique. However, the confidence in the trend measurements of the profile is substantially inferior to that of the total column. In addition, a second long-term data set is available from the series of SAGE instruments. There are other shorter data sets from infrared (IR) and microwave limb sounders, and such instruments are scheduled to fly on the EOS-Chem and ENVISAT-1 satellites. Because of the variety of instrument techniques, cross-validation will be an important consideration in the derivation of long-term trends.

1.5.5 Clouds and Aerosols

The International Satellite Cloud Climatology Project (ISCCP) cloud analysis is based on two channels (visible and thermal IR window) common to the geostationary and polar orbiting satellites. Significant effort went into normalization of geostationary radiometers to the afternoon orbiter to tie all instruments to a common standard. Succeeding polar orbiters were normalized to this standard. This relative calibration standard was tied to an absolute calibration standard using congruent aircraft flights. The ISCCP is the best calibrated data set of its kind. However, despite the significant efforts expended, the climate change signal we are searching for is still smaller than the instrument variability and calibration uncertainty of the currently operating instruments.

In the case of aerosols, a long-term data set exists for stratospheric aerosols (and also the cloud-free upper troposphere) from the measurements of Stratospheric Aerosol Measurement (SAM II), Stratospheric Aerosols and Gas Experiment (SAGE), and SAGE II solar occultation instruments from 1978 to the present. Future stratospheric aerosol measurements of the same type will be provided by the EOS SAGE III instruments into the next century. Occultation measurements are difficult to validate because of poor coverage. Other data for the stratosphere have been obtained from the Polar Ozone and Aerosol Measurement (POAM) occultation instrument launched in 1993 on the Satellite pour L'Observation de la Terre (SPOT-3) satellite (with further launches planned in 1997) and from UARS launched in 1991. There will be other instruments which map aerosol on the Advanced Earth Observation System (ADEOS-1), ENVISAT-1 and EOS-Chem satellites. Cross validation of the different aerosol products derived from measurements at different wavelengths and from different techniques is crucial. Balloons and ground lidar also play critical roles.

Tropospheric aerosol measurements will be provided by the EOS AM-1 platform by the MISER and MODIS instruments and by instruments on ERS-2, ADEOS-1 and ENVISAT-1. These tropospheric aerosol measurements will be primarily research products and there is at present no clear long-term monitoring strategy. In addition, since 1990, NOAA has generated an operational product of columnar aerosol over the global oceans derived from the Advanced Very High Resolution Radiometer (AVHRR) visible radiance measurements, and the derivation of valid long-term trends is possible given the ability of operation meteorological satellites to deliver long time series.

1.5.6 Radiation

Accurate determination of the earth radiation budget (ERB) and the cloud radiative forcing (CRF) is essential to constrain climate models. Monitoring the ERB and CRF components is needed to check how well models simulate climate anomalies (e.g., El Niño/Southern Oscillation (ENSO) events) and to help us judge whether model estimates of climate sensitivity are believable. If the requirements for calibration and overlap stated in Sections 2.1 and 2.2 can be assured for operational instruments (e.g., AVHRR and HIRS), there will be a chance for continuity in the record, even if based on various narrow bands. With broad band research instruments, continuity (and overlap) are unlikely, but traceability to absolute standards makes this less critical.

Whether narrow bands (given a sufficient number) or broad bands (with imperfectly flat spectral response) are used, the determination of "unfiltered" longwave and shortwave radiances requires a spectral correction which depends both on the instrument and on the scene. This spectral correction procedure must be fully documented, archived, and made available, and the sensitivity of the results to details of the procedure studied. Similarly, the radiance-to-flux conversion depends both on scene characteristics and on the bidirectional sampling by the instrument/satellite combination.

Sensitivity of the results to the algorithm, to assumed scene anisotropy, and to the observing system parameters must be assessed.

Changes in the observation local times (due to new satellites or orbit drift) introduce jumps or drifts in the record if diurnal variations are not correctly accounted for, particularly in the shortwave domain where diurnal variation is also sensitive to the bidirectional reflectance distribution function (BRDF). Moreover, the mean diurnal cycle is itself a climate parameter of interest. More use should be made of data from the geostationary satellites, and the sensitivity of products to the diurnal correction algorithm must be assessed.

The above discussion applies to the ERB (and CRF) components at the top of the atmosphere. Within the atmosphere and at the surface, the problem is no longer one of sampling (of spectrum, angles, times) but rather a remote sensing question, dependent on determinations of temperature, humidity and cloud properties as a function of space and time.

Important to our understanding of the radiation balance is knowledge of the changes in the solar radiation at the top of the atmosphere, both in total and spectrally resolved. The Active Cavity Radiometer Irradiance Monitor (ACRIM) and ERB series of satellite instruments have obtained an unbroken series of measurements since 1978, but satisfactory trend determination depends upon overlap since the differences between instruments, even after careful precalibration, is typically comparable with or larger than the trend to be measured. Long-term measurements are, therefore, vulnerable to a gap in the sequence.

1.5.7 Trace Species

An examination of the instruments which measure trace species, scheduled for launch in the next decade, indicates that they are mainly for research rather than monitoring. Thus, the scientific community will be faced with obtaining long-term trends from disparate instruments, with limited overlap periods and data gaps for many species. These instruments will have different spectral ranges, wavelength resolution, altitude and spatial resolution. For example, the Measurement of Pollution in the Troposphere (MOPITT) on EOS AM-1 will provide global measurements of tropospheric CO profiles and total column amounts of CO and CH_4 from 1998 to 2004. From 2002, the Tropospheric Emissions Spectrometer (TES) on EOS-Chem will make measurements of the same species using a different technique, and there will inevitably be differences which may be inseparable from any trend, even after cross-validation. NO_2 provides another example: there is a sequence of profile measurements since the 1970's including from the LIMB Infrared Monitor of the Stratosphere (LIMS,) SAGE I, SAGE II, Solar Mesospheric Explorer (SME), Atmospheric Trace Molecules Observed by Spectroscopy (ATMOS), and UARS instruments, but at a variety of local observation times. NO_2 exhibits a very strong diurnal variation, which makes it very difficult to merge these data into a single homogeneous time series.

This committee recommends that a working group be formed to determine to what extent precise long-term data records can be obtained for trace species under these circumstances. It is particularly important for this study to be made as soon as possible. Part of this activity should be to address for which species global space-based monitoring is needed and for which a more limited observation set of ground-based profiles or total columns (e.g., through the Network for Detection of Stratospheric Change (NDSC)) is sufficient.

2. OCEANS GROUP SUMMARY REPORT

Participants in the Oceans Group included Richard Reynolds, Chair, NOAA/National Centers for Environmental Prediction, Camp Springs, MD; Curtiss Davis, Naval Research Laboratory, Washington, DC; David Halpern, Jet Propulsion Laboratory, Pasadena, CA.

2.1 General Ocean Recommendations

Ocean parameters measured from satellites include sea level, winds, ocean color, and temperature. The measurement of sea-surface temperature (SST) is presently the most important variable because it has been observed from satellites for the longest period and is of direct significance in determining long-term global temperature trends. Because of this importance of SST, it has been given greater attention. Thus, this section is divided into general recommendations for all sensors and specific recommendations for SST. The final section includes specific recommendations for ocean color.

For any climate study, biases in satellite data are unacceptable and must be eliminated. Because of these and other errors, it is important that the data be maintained in an archive which includes the original observations and any meta data. Our experience has clearly demonstrated that errors occur to the satellite data and that it is necessary to be able to correct them. This archive would allow and require reprocessing.

Another possible error in longer records of ocean quantities may occur as instruments age and are replaced by new instruments on different satellites. To maintain the consistency of the satellite observations, it is important that the old and new instruments be cross-calibrated and validated for an extended period. A 1-year period is suggested to allow comparisons over an annual cycle.

To improve the accuracy of satellite observations, it is also necessary to verify and, if necessary, to recalibrate these data against *in situ* data. However, it is important to recognize that *in situ* observations have their own errors. These errors can be reduced by careful comparison of the *in situ* data with other *in situ* measurements. This is of particular importance in ocean measurements where *in situ* data are often deployed and never recalibrated. It is recommended that *in situ* instruments be designated as critical in several locations and carefully maintained over multiple satellite missions.

It is also important to continue to develop instrument designs which optimize the desired measurement. Often calibration and validation by more than one group can improve the accuracy of the final product. In addition, it is important that instrument climate drift be minimized. Early user involvement can help optimize instrument design and define meta data requirements.

The best product for climate purposes may be a product combined from different satellite instruments and different *in situ* data. Careful combination of these fields can lead to an improved product which can utilize the strengths of different sensors and minimize the effects caused by satellite instrument changes. For this purpose, the archives of different instruments should be designed to augment the combination of data.

2.2 SST Recommendations

An example of the size of the biases for SSTs is shown in Figure 1. The two curves in the figure show the average weekly *in situ* and nighttime AVHRR SST anomaly. The drop in the satellite SST anomalies in July 1991 is due to the impact of the stratospheric aerosols from the June 1991 eruptions of Mt. Pinatubo on the AVHRR operational SST retrievals. This resulted in negative satellite SST biases. The figure shows that the magnitude of the average satellite bias exceeded 1°C from July to September 1991. In the beginning of October, a new satellite algorithm was implemented which allowed a partial correction of the negative satellite biases. The remaining satellite biases gradually disappeared by April 1992 as the aerosols dissipated. (Additional details can be found in Reynolds, 1993.)

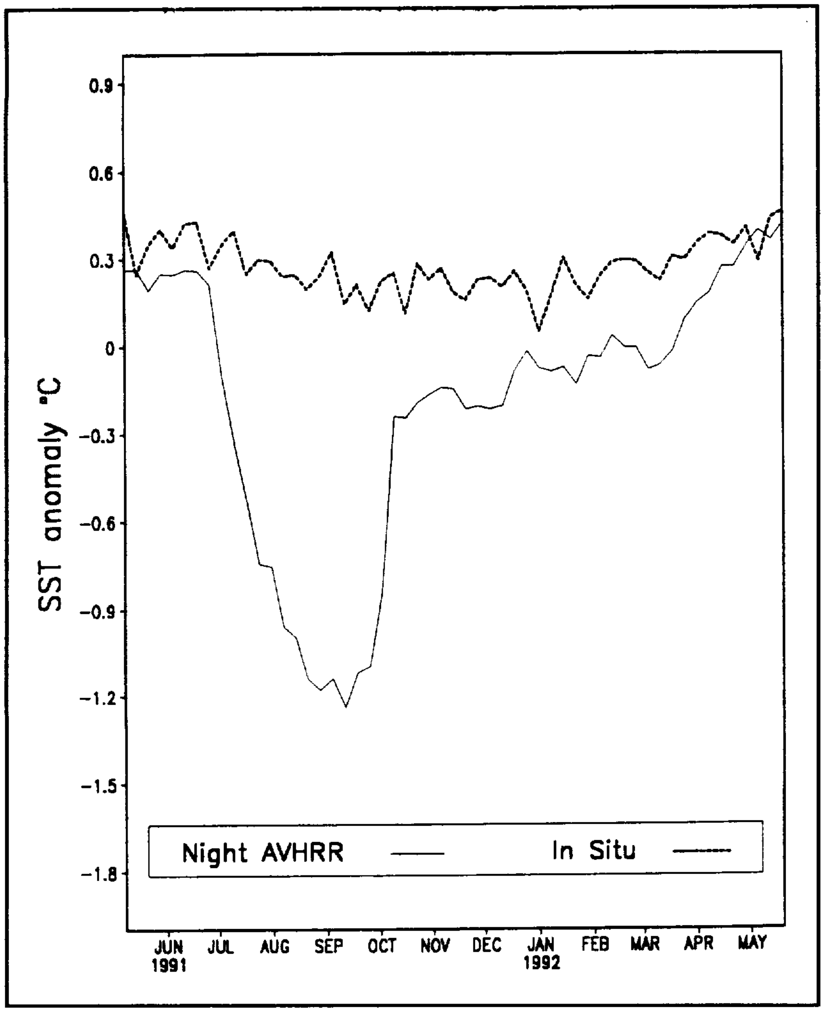

Figure 1. Time series of in-situ (ship and buoy) and nighttime AVHRR SST anomaly. The anomalies are averaged from weekly data between 20°S and 20°N.

Because of these biases, the National Environmental Satellite, Data, and Information Service (NESDIS) is undertaking the development of procedures to correct SST retrievals for aerosol radiative effects. These procedures will be implemented to reprocess the AVHRR satellite retrievals during periods of high volcanic stratospheric aerosols. These periods would include the El Chichón (1982-83) and Mt. Pinatubo (1991-92) periods. The procedures for these corrections include determination of aerosol characteristics such as size, composition, etc. The aerosol correction techniques would be maintained in state of readiness so that they could be operationally implemented when needed. This effort should be encouraged.

A combined SST product using *in situ* and AVHRR satellite data has already been developed at the National Meteorological Center (NMC) (see Reynolds and Smith, 1994). A further improvement of this analysis can be done by using Along Track Scanning Radiometer (ATSR) data. This instrument uses a dual scanning technique to eliminate sensitivity to the aerosols shown in Figure 1. However, both of these instruments use infrared detectors which cannot provide SSTs in cloud areas. Thus, the use of a microwave SST sensor, which can provide SSTs in cloud-covered regions, could provide more continuous coverage in both space and time. However, the microwave cannot provide the resolution of the other products and the ATSR cannot provide the coverage of the AVHRR. Thus, a new combined product could provide the best SST product.

Before combining different satellite SST products, it is crucial to carry out intercomparisons to determine how best to balance the strengths and weakness of each type of data. For these intercomparisons, it is important that the SSTs be the same type. At this time, the AVHRR retrievals produced by NESDIS provide a "bulk" temperature while the ATSR retrievals provide a "skin" temperature. All satellites, including microwave, actually measure a skin temperature. The NESDIS AVHRR retrieval is converted to a bulk temperature by regression against *in situ* SST from selected buoys. Accurate conversion between skin and bulk can be done with a model which requires knowledge of the wind stress and the net heat flux at the air-sea interface. Comparisons between different satellites should be done using the directly measured skin temperatures. Intercomparisons of satellite SSTs by first converting the skin to a bulk temperature is less direct and may introduce errors because of the conversion. However, satellite SSTs must also be converted to bulk temperatures to facilitate direct comparison with *in situ* bulk SSTs. This is necessary because the number of *in situ* skin measurements is very much sparser than the number of *in situ* bulk measurements. Furthermore, the conversion allows the satellite measurements to be included in the long-term climate record of *in situ* bulk SSTs.

An estimate for the present accuracy of SSTs was determined from the weekly 1-degree NMC SST product using independent buoy SSTs in the Equatorial Pacific. These results showed errors of 0.3°C in the western Pacific and 0.6°C in the eastern Pacific. Reynolds and Smith (1994) produced an error estimate for the first-guess of the analysis. Although this is an error parameter for the analysis, it can be used to provide an estimate of the relative increase of SST error that can be expected between the tropics and the extratropics. In particular, errors in the Gulf Stream and Kuroshio can be expected to exceed the western tropical by a factor of between three and four. This increase is due to the greater SST in the extratropics which is more difficult to sample in space and time.

To obtain accuracy SSTs from satellites it may be necessary to balance requirements for bias and scatter. Thus, for example, the SST bias error may be improved during periods of atmospheric aerosols. However, the correction of the bias is often accompanied by greater scatter in the individual retrievals. This occurs because of errors estimating the aerosol composition and particle size. For some studies, for example analysis of ocean features, this increase in the scatter may not be acceptable. However, for climate studies, biases should be kept as below 0.3°C in the tropics to avoid obscuring climate signals.

2.3 Ocean Color Recommendations

The goal of ocean color is to make a decade long data set of ocean optical properties and phytoplankton biomass and productivity to assess the impact of global change on the ocean food web and to assess the effect of changes in energy absorption, CO_2 uptake, or dimethyl sulfide production by the marine food web which may dampen or accelerate global change. The observable changes will be in the range of a few percent per decade, and to observe those features requires a data set that is reasonably accurate and highly consistent. However, the global data set must be pieced together from data from a series of international satellites each with a unique design and with its own calibration and validation procedures. To achieve the required combined data set—production of one or more products for the decades long time period which can be used to assess global change—requires excellent calibration and validation, reference to a common standard, and standardized processing.

Comparison of data from different satellite instruments should be at the radiance level. A long-term radiance data set could be used for the determination of phytoplankton variability by applying the same model to the radiance data. Many models can be tested using a large, long time series data base of *in situ* measurements as a reference for assessing the success of the models. The models may not be perfect, but if it is applied uniformly they should show trends over time.

A key point for developing a long-term global ocean color data base is that all of the sensors be calibrated against the same radiance standard or standards that are intercompared and shown to agree to within a percent or two. *In situ* instruments that will be used to collect validation data sets should be calibrated against the same standards. The Sea-viewing Wide Field Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer (MODIS) projects have initiated such an intercalibration program, and are working to include international partners who are building ocean color sensors or collecting validation data. This is a good approach that should be continued and expanded. Ideally the standard should be the Sun. This may require better diffusers, or at a minimum a better characterization of the present diffusers and standardization of the methods for using them.

Large long time series reference data sets are essential for the validation of radiance data and calculated products such as chlorophyll and primary production. Again, there must be a concerted effort to compare techniques, evaluate blind standards and otherwise show that all of the data in the international data base is comparable. The Joint Global Ocean Flux Study (JGOFS) program is developing a standardized data base for chlorophyll and primary production measurements that will be useful for this purpose.

2.4 Other General Comments

This discussion on ocean recommendations has been strongly motivated by climate requirements. However, in the present environment of restricted funding, it is important to combine climate and weather requirements as much as possible to increase chances for financial support.

3. LAND SURFACE GROUP SUMMARY REPORT

Participants in the Land Surface Group included Philip Teillet, Chair, Canada Centre for Remote Sensing, Ottawa, Ontario, Canada; Forrest Hall, Co-Chair NASA/Goddard Space Flight Center, Greenbelt, MD; John Townshend, University of Maryland, College Park, MD; Joseph Cihlar, Canada Centre for Remote Sensing, Ottawa, Ontario, Canada; Christopher Brest, NASA/Goddard Institute for Space Studies, New York, NY; and Gideon Kinyodah, Kenya Meteorological Department, Nairobi, Kenya

3.1 Introduction

This paper summarizes the deliberations of the land panel that met during the workshop. It outlines some of the key issues and makes recommendations with respect to the calibration and validation of global, long-term, land surface data sets based on satellite observations of the Earth. Global monitoring by satellite sensors combined with a network of surface observations may be considered to be the only feasible approach for the measurement and long-term monitoring of terrestrial parameters needed by scientific investigators and decision makers around the world. The challenge is to ensure that such measurements yield self-consistent and accurate geophysical parameters over time, even though the measurements are made with a variety of different instruments under different observational conditions. Hence, sensor calibration and data product validation are critical aspects of Earth observations if they are to show terrestrial processes as they really are and if they are not to be compromised by sensor and data processing effects.

While satellite data requirements of the land surface are often driven by the need for biophysical and climatological information, the scope of the land panel discussions focused primarily on relative and absolute sensor calibration to top-of-the-atmosphere quantities in physical units and subsequent correction to surface reflectance or temperature. The discussions emphasized radiometric calibration and it was beyond the scope of this particular forum to address the important topic of spectral calibration and characterization. Other papers in these proceedings address derived product validation issues. After an outline of calibration requirements, the paper goes on to summarize discussion highlights under the headings of existing sensors and data sets, future data sets from existing sensors and follow-on sensors from the same series, and anticipated sensors and data sets. Finally, some general recommendations are made in addition to any specific ones made in earlier sections.

3.2 Requirements

Land satellite data requirements are largely driven by the need for long-term monitoring of land surface parameters to determine surface condition and to detect change, and inputs to regional and global carbon, energy and water process models. An initial documentation of these requirements was made by remote sensing and modeling groups at the 1992 workshop of the International Satellite Land Surface Climatology Project (ISLSCP) (Sellers et al., 1995; Hall et al., 1995). A compilation of parameter requirements encompassed parameters such as land-cover type, biophysical characteristics (biomass, fraction of absorbed photosynthetically active radiation (FAPAR), leaf area index (LAI), surface roughness), albedo, surface radiation, and surface energy balance. Spatial and temporal resolution requirements vary, ranging from 250 m to 100 km and from every 6 hours to annual, respectively. Specific accuracy requirements for these parameters were also defined at the ISLSCP workshop based on the sensitivity of models to the various parametric inputs. Additional end-to-end model sensitivity analysis is needed to translate the parameter accuracy requirements into requirement specifications for surface reflectance.

The land panel feels that it should be within the capability of current calibration and atmospheric correction techniques to resolve relative changes at the pixel level on the order of 20 percent in vegetation visible reflectance (1 percent absolute change) and about 5 percent in near-infrared reflectance (2 percent absolute change). This would translate to a change in Normalized Difference Vegetation Index (NDVI) of approximately 0.1 or 10 percent and, hence, to accuracies of about 10 percent for FAPAR and LAI. With known emissivities and accurate atmospheric temperature and humidity profiles, it should be possible to recover land surface temperatures to within 0.5 Kelvin at 30 degrees Celsius. While aggregations of pixels can further improve the aforementioned accuracies, there will still remain the issue of how accuracies scale from specific values of individual pixels to the regional and global levels used by modelers.

An additional requirement is that of subpixel geometric registration accuracy to ensure accurate detection of change (Townshend et al., 1992). The fine scale of land cover properties means that, without such fine scale, many of the observed differences will be spurious. It is also important to note that, even with perfect geometric registration, there will still remain the issue of variable pixel size as a function of view angle and instrument.

3.3 Existing Sensors and Data Sets

The satellite data record goes back to 1972 in the case of Landsat. Valuable support data are provided in terms of time series and validation data sets from the high spatial resolution Landsat sensors (Multispectral Scanner (MSS) and Thematic Mapper (TM)) and the SPOT Haute Resolution Visible (HRV) sensors. The global record is best represented by the AVHRR data sets from NOAA's series of satellites extending back to 1981.

An overview of global data sets for land applications from the AVHRR is provided by Townshend (1994). These data sets vary significantly in spatial resolution (from 1 km to very coarse resolutions of 15 km and greater) and in the processing algorithms involved in their generation. The main lessons learned in the creation of global data sets from the AVHRR are discussed by Justice and Townshend (1994) and listed in Table 1 of that paper. The list includes two points pertinent to calibration and validation: globally comprehensive long-term, well-calibrated and characterized reflectance or radiance values of the sensor are the most important properties of remotely-sensed data sets for global change research; and the usefulness of global data sets is greatly inhibited by a lack of validation of the basic data sets and derived higher level products.

3.3.1 How Well Have the Requirements Been Satisfied?

With the exception of specific intensive field campaigns and certain Pathfinder data sets, calibration accuracies and continuity have not been pursued for Landsat-1 through -5 MSS data for almost a decade. The problem of cross-calibrating the Landsat MSS sensors and accounting for drift in sensor characteristics has been studied most extensively by Markham and Barker (1987). They found that postlaunch intercomparisons of the calibrations for the Landsat-2 to 5 MSS sensors agree to within 12 percent and that the calibrations tend to degrade as the sensors age.

An onboard internal calibrator system based on radiance lamps has been used to try to maintain the calibration of the Landsat-4 and -5 TM instruments. Data processing systems around the world have used these calibration lamp signals on a frame-by-frame basis to generate calibrated TM products. However, recent vicarious calibration results from White Sands indicate some drift in the Landsat-5 TM calibration that is not being captured by the onboard system and subsequent processing (Thome et al., 1993). Investigations involving the quantitative use of TM data over time will be significantly affected by this drift. Systematic vicarious calibration of the TM instrument should lead to sensor radiance accuracies on the order of 2 to 3 percent (Slater et al., 1987). In the absence of a reliable sensor calibration for TM data, some investigators have resorted to the use of empirical techniques that provide data calibration to surface reflectance, usually based on assumptions about the radiometric characteristics of pseudo-invariant reference targets. An advanced and promising form of empirical technique is that of radiometric rectification, which has the potential to approach 1 percent accuracy in surface reflectance units in the visible and near-infrared (Hall et al., 1991).

Radiometric calibration of the SPOT HRV instruments is based on a weighted blend of relative and absolute methods (Gellman et al., 1993) and is deemed operational. Calibration coefficients are provided to users with the image products. Uncertainties in these coefficients are estimated to be on the order of ±3 percent for SPOT-1 HRV-1, for example (Santer et al., 1992). Although user requirements in terms of calibration accuracy are not yet being met, valuable experience is being gained in the process.

The two shortwave channels of the NOAA AVHRR instruments have been radiometrically calibrated prior to launch but there have been no onboard systems to provide postlaunch calibration updates. Because of the widespread and continuous use of AVHRR data for regional and global studies, considerable effort has been devoted to arriving at postlaunch calibration updates (Teillet et al., 1990; Rao and Chen, 1994), with particular emphasis on the temporal behaviors of the NOAA-7, -9, and -11 AVHRRs. However, much of this effort has been experimental in nature and the transition to systematic calibration updates for operational use has been slow. In the interim, many operational users have adopted AVHRR calibration coefficients based on desert-site methodologies that are deemed to be better than the use of prelaunch coefficients (Teillet and Holben, 1994; Rao and Chen, 1994; Cihlar and Teillet, 1995). It is estimated that current sensor radiance calibration uncertainties for the AVHRR shortwave channels are on the order of 7 to 9 percent absolute. It is not known if these uncertainties are valid at the lower end of the radiometric dynamic range.

The recent AVHRR Pathfinder calibration activity represents the first concerted effort to establish a calibration on the part of some of the researchers involved (Rao and Chen, 1994). Another notable effort is that of the ISCCP, which recognized many years ago the importance of calibration to produce a good operational product and devoted the necessary time and effort accordingly (Brest and Rossow, 1992).

3.3.2 What Needs to be Done?

A program of retrospective vicarious calibration is in order for the Landsat MSS data record, making use of relevant test sites and any available ancillary data. Wherever possible, MSS and TM image pairs for the same locations and time should also be used. This program could be undertaken by an agency such as the EROS Data Center, Sioux Falls, South Dakota.

Operational providers of Landsat TM data should be advised of the calibration degradation which is not being captured by the onboard system. The White Sands calibration work should be consolidated and also enhanced by additional calibration information based on darker targets. The combination of vicarious calibration and empirical radiometric rectification techniques can be used to produce retrospective Landsat data sets in surface reflectance units.

The NOAA-14 AVHRR postlaunch calibration needs to be established on an urgent basis. The NOAA-11 AVHRR calibration record needs to be completed as soon as possible. Postlaunch calibrations for the NOAA-8, -10, and -12 AVHRR instruments have yet to be done and should be carried out as soon as possible. The postlaunch calibration characterizations for the NOAA-7, -9, and -11 need to be more widely disseminated internationally and also made available on the World Wide Web (WWW).

Global AVHRR data sets should benefit from state-of-the-art techniques for cloud screening, atmospheric and bidirectional corrections, and compositing. Emphasis needs to be placed on the global acquisition and availability of atmospheric water vapor content and aerosol parameters such as optical depth, which is more difficult to obtain over land than over oceans from satellite data sources

3.4 Future Data Sets From Existing Sensors and Follow-on Sensors From the Same Series

For the high spatial resolution sensors such as the Landsat TM and the SPOT HRV, the key lies in instituting systematic calibration updates based on vicarious methodologies such as the White Sands approaches. Attention also needs to be paid to the lower end of the dynamic range. Equally important is the necessity to make the calibration updates known and to provide ready and easy access to them so that operational processing can be done in the most accurate manner.

With respect to the AVHRR series of sensors, a more systematic and timely effort by the instrument provider is to be encouraged. A new airborne sensor system for congruent underflights would be extremely valuable and would provide better and cheaper results in the long run. Cross-calibrations between AVHRR and same-day imagery from other, better calibrated sensors should become more viable with the advent of many new satellite instruments over the next few years.

Consideration should also be given to the semicontinuous staffing and instrumentation of test sites. If the sites are well-chosen and properly used and if the funding is shared by many agencies, such an effort would benefit the radiometric calibration of most spaceborne sensors used for quantitative work. It would make it possible to provide near-real-time calibration to users who need it and would be a significant step in the direction of real-time calibration.

3.5 Anticipated Sensors and Data Sets

Over the course of the next few years, a significant number of satellite instruments will be launched with a primary mission to acquire land surface data. When it finally crystallizes, the next Landsat will likely carry an enhanced version of the very successful TM instrument. The SPOT-4 mission includes the VEGETATION sensor, a cross between the

SPOT HRV and NOAA AVHRR concepts, intended for broadscale monitoring of vegetation. The NOAA AVHRR series will be followed in due course by a new generation of operational meteorological instruments that will undoubtedly be used to acquire global land data sets. Scheduled for launch in 1998, the EOS AM-1 spacecraft will carry several instruments, most notably MODIS, (Running et al., 1994), which will contribute to land, ocean and atmosphere investigations. Shortly after the turn of the century, ADEOS-2 will carry the Global Imager (GLI), another broadscale sensor dedicated to land and ocean remote sensing.

While each of these forthcoming sensors will probably find a niche in terms of providing special land data sets that only they can provide, there will also be certain geophysical parameters that all of them will be called upon to provide despite significant differences in sensors characteristics. The most common denominators in this regard are likely to be surface reflectance and derived parameters such as vegetation indices. Self-consistent results will only be possible with well-organized calibration/validation efforts.

The panel discussions did not devote a lot of time to future sensors, but a few calibration/validation points were highlighted: careful prelaunch sensor characterization; stability of orbits, sensors, calibration sources and methods; use of the Sun as the ultimate reference source (experience to date indicates that on-board calibration using the Sun or Moon as a reference standard is essential to meet the strict calibration requirements); regular and systematic sensor intercomparisons; well-developed atmospheric correction algorithms; increased development of the validation component.

3.6 General Recommendations

The recommendations in this section are in addition to those made in earlier sections. Moreover, as with the deliberations of the land panel at the workshop, they pertain primarily to the calibration/validation of sensor data products as opposed to land surface parameters derived or modeled on the basis of the satellite data.

3.6.1 Calibration/Validation Activity Recommendations

It is recommended that prelaunch sensor characterization and postlaunch processing capability to level-1B radiance be the responsibility of the instrument provider. There is a concern that no provision is being made to provide level-1B data from Landsat-7.

Prelaunch round-robin measurement and methodology comparisons should be considered essential components for obtaining realistic uncertainty estimates. Calibration standards are not always handled the same way and error analyses can take varying approaches. Participants should include national standards laboratories, satellite instrument providers, and key experts such as those from the University of Arizona and the NASA/GSFC.

It is recommended that more systematic use be made of vicarious calibrations based on ground test sites. Considerable research is being done in the area of vicarious calibration but, with few exceptions, operational users are not benefiting from it.

It is recommended that a much greater effort be devoted to timely and ready access to documentation, data, metadata, and information on sensor characteristics and all relevant related calibration/validation. Instrument and data providers should also clearly identify individuals associated with various aspects of their operations so that users can readily contact appropriate personnel.

3.6.2 Data and Processing Recommendations

There should be less time elapsed between the development of state-of-the-art processing standards and techniques and their operational use in the processing of global data sets. CEOS is in a position to tackle this issue.

It is recommended that standing committees of international experts be established by the instrument providers to maintain calibration/validation for each particular instrument. Part of each committee's responsibilities would be to host periodic workshops to assess, document, and achieve consensus on state-of-the-art standards and techniques for improving the processing stream for land sensor data.

An international workshop should be convened by NOAA/NESDIS or by the Earth Resources Observation Systems (EROS) Data Center to recommend an improved algorithm string for the reprocessing of the existing AVHRR data set, including radiometric calibration, cloud screening, atmospheric and illumination corrections, directional and topographic corrections, and geometric resampling and remapping. This algorithm string should be used to reprocess existing global AVHRR data sets, and the products carefully evaluated with respect to the accuracy of retrieved surface reflectances.

Processing algorithms should be developed and applied in such a way as to make retrospective and future data sets as consistent as possible. Detailed comparisons and standardizations are needed between similar geophysical parameters derived from dissimilar measurement systems and concepts. An end-to-end systems analysis is recommended to ensure the consistency of accuracy requirements specifications from sensor data through to geophysical parameters and on to information products. The analysis should include elements such as radiometric and spectral calibration, atmospheric/illumination/ bidirectional/topographic corrections, geometric resampling and remapping, cloud screening, temporal compositing, etc.

Algorithm research and development are needed to ensure the proper and accurate integration of multisource remotely sensed and other types of data from radiometric, spectral, and spatial perspectives. Remote sensing instruments acquire imagery in very specific modes and geometries that have direct impact on the radiometric character and information content of geophysical parameters derived from Earth observation data.

A concerted effort is needed to make critical ancillary data more available and accessible. A case in point is the availability of digital terrain elevation data on a global basis with sufficient quality and accuracy. Efforts in this regard have escalated recently, but only after considerable pressure by the user community. Other data sets of wide interest and applicability in support of Earth observation, in general, and calibration/validation, in particular, include land cover, atmospheric parameters, and surface emissivity.

Sample data sets, accuracy assessment products, and algorithm benchmark/test data sets should be generated and made available to operational data providers and users to facilitate their product validation. Different processing systems seldom give identical products from the same source of sensor data. Based on past collaborations with respect to AVHRR processing, it is suggested that the Canada Centre for Remote Sensing and the EROS Data Center could undertake this type of work.

3.7 Concluding Remarks

From a technical standpoint, most of the recommendations in this paper are not new ones. They are known to be vital if new and historic satellite data are to provide a quantitative, long-term record of land surface parameters and global change. Scientific and technological knowledge and expertise have matured to the point where sensor calibration and data product validation are approaching the levels necessary to meet user requirements. Perhaps what has been missing is a general appreciation of the importance of precise radiometry and the resources needed to achieve it in platform, sensor, and processing design and development. The necessary resources are as much in the realm of collaborations and infrastructures as they are human and financial. An international body such as CEOS can exert considerable influence on the establishment of calibration/validation standards and methods for satellite sensor data and derived geophysical parameters. CEOS can bring to the global community the weight of conviction necessary to make adequate funds available to achieve these standards.

4. SENSORS GROUP SUMMARY REPORT

Participants in the Sensors Group included Bruce Guenther, Chair, NASA/Goddard Space Flight Center, Greenbelt, MD; Jim Butler, Co-Chair, NASA/Goddard Space Flight Center, Greenbelt, MD; Philip Ardanuy, Hughes STX, Greenbelt, MD; Ernest Hilsenrath, NASA/Goddard Space Flight Center, Greenbelt, MD; Nagaraja Rao, National Oceanic and Atmospheric Administration, Camp Springs, MD; Philip Slater, University of Arizona, Tucson, AZ; Michael Weinreb, NOAA/National Environmental Satellite, Data, and Information Service, Washington, DC; and Christopher Mutlow, Rutherford Appleton Laboratory, Oxfordshire, United Kingdom.

The Sensors Group recommendations are segmented into two areas: prelaunch in the laboratory and on-orbit within the sensor, and postlaunch using auxiliary data sets taken together with sensor data. The Sensors Group believes that the primary responsibility of a flight program is the development and distribution of the calibrated and Earth-located sensor data set.

4.1 Laboratory Preflight Calibration/Characterization

The Group is certain that funds used on characterization and calibration during the development of a sensor ultimately saves money by providing a data set that can be used by the science community with a minimum of resources necessary for validation of those data sets. The science data users frequently must commit significant resources to validate or recalibrate the data set before they can use it to produce geophysical data products, when the sensor is poorly calibrated.

Instrument calibration capabilities and facilities should be maintained over the long term. Present procedures rely primarily on instrument contractors for calibrating instruments. Experience, capabilities, and facilities often disappear after projects end. Therefore, new projects have to re-establish and relearn what has been lost. Loss of calibration measurement infrastructure is not cost-effective and will effect Mission to Planet Earth (MTPE) objectives for stable and accurate long-term measurements.

Although the National Institute of Standards and Technology (NIST) maintains absolute radiometric standards, and we assume they will continue to do so, it is the application of these standards to instrument calibrations that is important to the MTPE measurement objectives. A core calibration capability now exists within NASA, but could be dismantled with budget cutbacks. Continuation of this core capability is essential, and programmatic decisions should be made, as soon as possible, as to how this should be accomplished.

The Group feels that new scientific research frequently develops after the community studies and uses satellite data sets. This new scientific research often means the satellite data must be considered in a different way than was first anticipated. New uses of the satellite data often requires that the sensor calibration must be redone. Consequently, the sensor test data must be documented and preserved so the test data can be analyzed again later as the need arises. Actual use of the data usually calls into question the original analysis of the calibration data, so the prelaunch calibration will be revised later with these test data sets.

The calibration actually achieved for a sensor will be of higher quality and more useful to the community for developing geophysical data sets when the user community is involved early in the calibration process. Calibrations also are of a higher quality when people, in addition to those who build the sensor, are intimately involved in the calibration and validation data analysis process. Frequently the instrument builder has only enough resources to calibrate a sensor and interpret that data in the context that the instrument meets the performance requirements. The motivation is greater to show compliance with specifications than to provide a thorough accounting of the sensor performance for the sensor builder.

The ability to compare separate sensors requires that the calibration be derived in terms of physical standards and standard processes. When sensors are needed for global climate data sets, the sensor characterization and calibration must be well understood. Sensor performance is better understood when component and subsystem testing are performed and the calibration of the sensor system agrees with sensor sensitivity inferred from the subsystem and component level testing.

In instances where the calibration uncertainty requirements for a global climate data set exceed what is currently available in calibration for a sensor, then a research program is needed to develop improved calibration techniques. Calibration science must continue to evolve to support current and emerging measurement requirements, and success in developing these research ideas requires a firm funding commitment for laboratory calibration development. The laboratory calibration development must include support to contemporary standards and the application of those standards and to the development of the new standards which are required. Additionally, current calibration terminology indicates that sensor performance should be described in terms of uncertainties, rather than in terms of (estimated) accuracy. Further uncertainty or consistency are most reliably established through direct round-robin measurement comparisons to validate the error estimates.

The choice for a high quality calibration source is typically a source that has very low spectral, spatial and temporal contrast. Light sources for calibration standards are always analyzed with lamps which have a smooth continuum variation in the spectrum. Presently, the measurement technology is headed toward imaging systems, and the science of interest usually is derived from scenes with a high scene contrast. Imaging systems for which radiometrically accurate data sets are required are very complicated to interpret when a scene of high contrast is observed. This problem requires additional measurement and software science research to be adequately accomplished. This concern is applicable both to traditional systems such as MODIS and the emerging smaller, cheaper technologies under development.

4.2 On-board Calibrators and Vicarious Calibration Techniques

In the presence of an on-board calibration system, that system usually is available continuously to the sensor and is useful for tracking and trending sensor performance characteristics which other calibration techniques are not capable of doing. These individual calibration devices must be carefully designed and validated to allow best confidence in their use for developing climate data sets.

On-orbit lunar observations offer the prospect of high quality long-term reflectance tracking for individual sensors, and for cross-calibration among numerous sensors, used for climate data sets. Measurements obtained in this century can be compared to those obtained in coming centuries for comparison of Earth change studies. To accomplish the use of lunar observations, spacecraft must be capable of pointing their sensors at the Moon, and the reflectance characteristics of the lunar surface must be characterized over the surface for illumination phase angles and the 11-year lunar libration cycle.

Validation and vicarious calibration methods are needed throughout the lifetime of these missions. Each provide a needed and distinct capability which is required for the maximum usefulness of the derived calibrated data sets. Development of a consistent set of uncertainties across several sensors is important when these sensors are being used for vicarious calibrations. The uncertainties should be verified through a direct field comparison program.

When multiple sensors are used to accomplish a measurement objective, then overlap of the measurements is essential. In the case of radiance overlap, a few months is adequate. In the instance of the derived geophysical data sets, data overlap of more than 3 months is required. Overlap of sensors longer than 3 months for geophysical products is required when at least one of the sensors is poorly characterized. The only way to reduce overlap necessary for obtaining the geophysical products is by providing sensors to orbit that are better characterized than has been the case to date.

In some instances where overlap of observations from space are desired, the older sensor is carried on a spacecraft which has a different local time for equator crossing. In these instances, direct radiance comparisons between the two sensors is difficult. This Group recommends that further research be undertaken in the area of data analysis (assimilation) of the radiances to provide comparison data sets for sensor cross-calibration.

Two specific recommendations are provided to CEOS and GCOS. In the late 1980's the International Coordination Working Group agreed that for Earth observations platforms, each platform provider will create a calibration plan for each platform, and develop that platform consistent with that plan. In December 1994, the CEOS Calibration Validation Working Group recommended that instrument providers (which provide a calibration in the data header for "direct broadcast") provide an approach to broadcast current calibration coefficients. This Group recommends that for each sensor which users direct broadcast, that a current set of calibration coefficients be provided in a easily accessible medium, such as the WWW. This Group requests that CEOS and GCOS reaffirm these two recommendations and take an active role in their implementation by platform providers.

5. CONTEXT AND PERSPECTIVES

5.1 Data Continuity and Sensor Evolution

Mark Abbott, Oregon State University, College of Oceanic and Atmospheric Sciences, Corvallis, OR

Many critical processes in the Earth system vary on long time scales, and in many cases, these long-term changes are difficult to detect. These small, gradual changes require precise measurements over long times to assess their trends and quantify their magnitude. For example, small changes in total cloudiness can affect the atmospheric radiation balance which in turn can lead to either warming or cooling of the Earth. Essentially, we are examining small differences between large numbers, and we must have both accuracy and precision in our measurement systems. With both satellite and *in situ* observations, it has been extremely difficult to meet these measurement requirements over the many decades necessary to resolve low frequency variability in the Earth's environment. Subtle shifts in instrument calibration or performance and changes in processing software can be mistaken for natural variability in the Earth system. Thus, we must maintain the ability to carry forward the science information essential to reconstruct a consistent time series from multiple data sources.

To ensure continuity of data from different instruments (or from copies of the same instrument), there must be a careful program of calibration, sensor characterization, and intercomparison. Satellite sensors present particularly difficult challenges for the development of long-term, well calibrated time series. Such on-orbit tests are complex and often require a considerable amount of time and money. We have often relied on flying copies of sensors to simplify this process. Although instrument copies do not eliminate the need for calibration and characterization studies, such research is usually cheaper and produces more robust results.

In addition to the basic sensor instrument, the total observing system includes sensor calibration, data processing algorithms, and numerical models that use the final data products. Changes in observing characteristics (sensor performance, orbit crossing times, etc.) will introduce perturbations into the entire system. Such perturbations temporarily introduce gaps as the science community undertakes studies to characterize the new system and make appropriate changes to the processing algorithms and models. During this gap, there is essentially little scientific return from the sensor. The more complex the changes in the sensor system, the longer the gap will be. However, thorough preflight checks including sensor calibration and characterization will help reduce, but not eliminate, this time between launch and the delivery of scientifically useful data.

The AVHRR on the NOAA TIROS series of meteorological satellites serves as an example. This sensor was a follow-on to the AVHRR beginning in 1978. The AVHRR has slowly evolved over the past decade to meet operational requirements. These changes in sensor performance have been slow and orderly, reflecting NOAA's estimates of the costs of adapting the entire data collection, processing, and modeling system to the new data set versus the increased benefit to the operational mission. Thus any changes had to be justified in the context of the overall system performance. But even this orderly transition did not provide Earth scientists the calibration accuracies necessary to distinguish geophysical variability from sensor variability. The NASA/NOAA AVHRR Pathfinder activity is a major investment aimed at improving the quality of the time series, but the AVHRR itself remains uncalibrated.

If we consider our research from a business perspective, then changing instrument systems should be viewed as a cost that must be weighed with potential benefits in the form of increased science return to the overall system. This analysis must consider the impacts of sensor evolution on the end-to-end system performance, realizing that simply improving instrument performance does not automatically increase science return. Advances in sensor technology do not replace the need for careful calibration and sensor characterization that are essential to separate sensor variability from the often small but critical changes in long-term data sets of the Earth system.

5.2 Committee on Environment and Natural Resources

Robert Winokur, NOAA Assistant Administrator for Satellite and Information Services

The Task Force on Observations and Data Management is the cross-cutting entity of the CENR. It is composed of representatives from CENR Subcommittees and U.S. and other Government Agencies. The Task Force is charged by the CENR Chairman to provide a long range view of the observing system and data management requirements across the subcommittees and represent the U.S. internationally in these areas. The Chairman is Dr. Charles Kennel, NASA Associate Administrator for Mission to Planet Earth; the Vice-Chair is Mr. Robert Winokur, NOAA Assistant Administrator for Satellite and Information Services; and the CENR Task Force Secretariat is Dr. Robert Schiffer, NASA.

The Task Force is responsible for developing and overseeing implementation strategies for the U.S. components of an international global observing system. The system will provide global-to-regional-to-local observation and monitoring of the Earth's environment and natural resources, and management of the resultant and related data; advanced scientific understanding; help develop predictive and assessment capabilities, products, and services of broad value to society; and support CENR and related research and key national and international missions and initiatives.

The Task Force objectives are to coordinate the development of an inventory of observations and data, coordinate research observation and data systems requirements, identify gaps and overlaps, promote the development of a more comprehensive system of observations, promote development of a system to manage this extensive information, ensure data collected can be used to provide credible information, establish a closer relationship with the environmental modeling community, develop and test ways to improve the usefulness of observations and data, provide the U.S. secretariat for international global observing system programs, and promote the continuance and extension of the policy of full and open exchange of data for environmental research.

In summary, the Task Force provides a forum for coordination of observations and data management needs of the U.S. participants in environmental and natural resources research; a forum for coordination of U.S. contributions to, and requirements for, international observing and data management systems for this research; and a forum for active working relationships with international entities.

President Clinton established the National Science and Technology Council (NSTC), a cabinet-level committee for science, space, and technology throughout the Federal Government. NSTC established nine committees, including the CENR. CENR is chaired by Dr. D. James Baker, NOAA Administrator. The CENR Task Force, as a national focal point for GCOS, will be a mechanism for national GCOS coordination, and will provide the U.S. Secretariat for international global observing system programs (GCOS, Global Terrestrial Observing System (GTOS), Global Ocean Observing System (GOOS)).

With respect to joint polar-orbiting systems, in May 1994, Presidential Decision Directive called for convergence of the Department of Commerce/NOAA's polar-orbiting environmental satellites with the Department of Defense/Defense Meteorological Satellite Program (DOD/DMSP), and as appropriate, components of NASA's EOS. The Integrated Program Office was established October 1, 1994. NOAA is providing overall coordination and leading satellite operations, DOD is leading acquisition, and NASA is leading technology transition. NOAA is also working with European counterparts on a joint polar mission (EUMETSAT will fly the morning mission and NOAA will fly the afternoon mission). We need to consider implications of such programs to calibration/validation issues and plan ahead.

The NOAA/NASA Pathfinder Program concept was based on the question of "What can be done now to further global change research?" The Program evolved to take advantage of data presently archived, while awaiting remotely sensed data from future satellites. The Pathfinder document describes space-based systems such as AVHRR, TOVS, Geostationary Operational Environmental Satellite (GOES), and SSM/I; and data availability such as product interuse and product distribution.

There are a number of calibration/validation and data set continuity issues for operational satellites:

- Calibration of Operational Visible and Near-IR Sensors—Currently there is no on-board calibration, only preflight. Because sensors degrade in orbit, candidate on-board calibration techniques need to be evaluated. Figure 2 shows the apparent rate at which the reflectance factor of an Earth scene, taken as 100 percent on the day of launch of the spacecraft, would decrease in time because of the in-orbit degradation of the instrument. This information, when coupled with the absolute calibrations performed with congruent path aircraft/satellite measurements, make it possible to correct the AVHRR-generated environmental products for radiometer degradation.

- Intersatellite Calibration—Slight differences in instruments of similar type; changes in instruments to measure same variable; need overlap period to cross-calibrate.

- Satellite Orbital Drift—NOAA Polar Orbiting Environmental Satellite (POES) PM satellite: Equator crossing time becomes later during operational life (4 hours over 5 years).

- Product Compatibility—Different spatial and temporal resolutions; different sensors; need community consensus algorithms.

- Product Validation—Remote sensing provides indirect measurements; need ground-truth; use conventional observing where possible; field comparisons and special ground-based networks for some variables. Figure 3 illustrates the effects of the Mt. Pinatubo stratospheric aerosol cloud on satellite-derived SST measurements as compared to *in-situ* measurements.

- Convergence of U.S. Military and Civil Polar Satellites (see Figures 4 and 5)—Military provides good imagery; the civil sector provides good radiometry; we must ensure product viability.

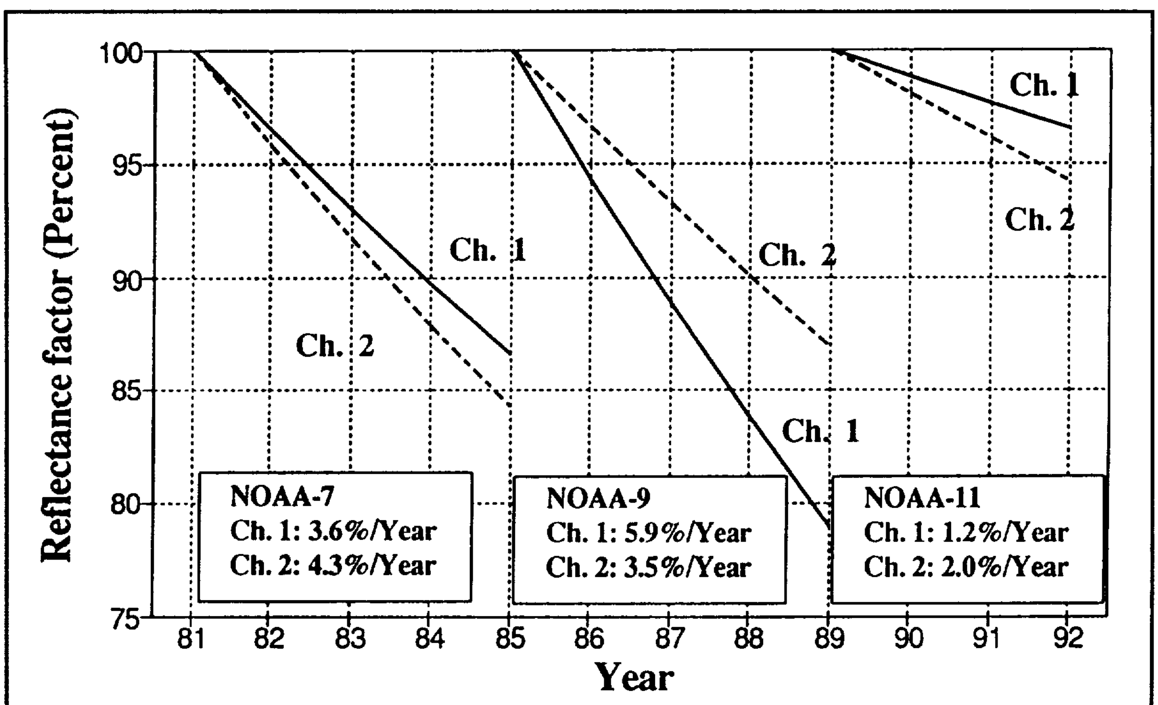

Figure 2. The relative degradation of the visible (channel 1) and near-infrared (channel 2) channels of the AVHRRs on NOAA-7, -9, and -11 spacecraft, determined using the southeastern Libyan desert as a time-invariant calibration target.

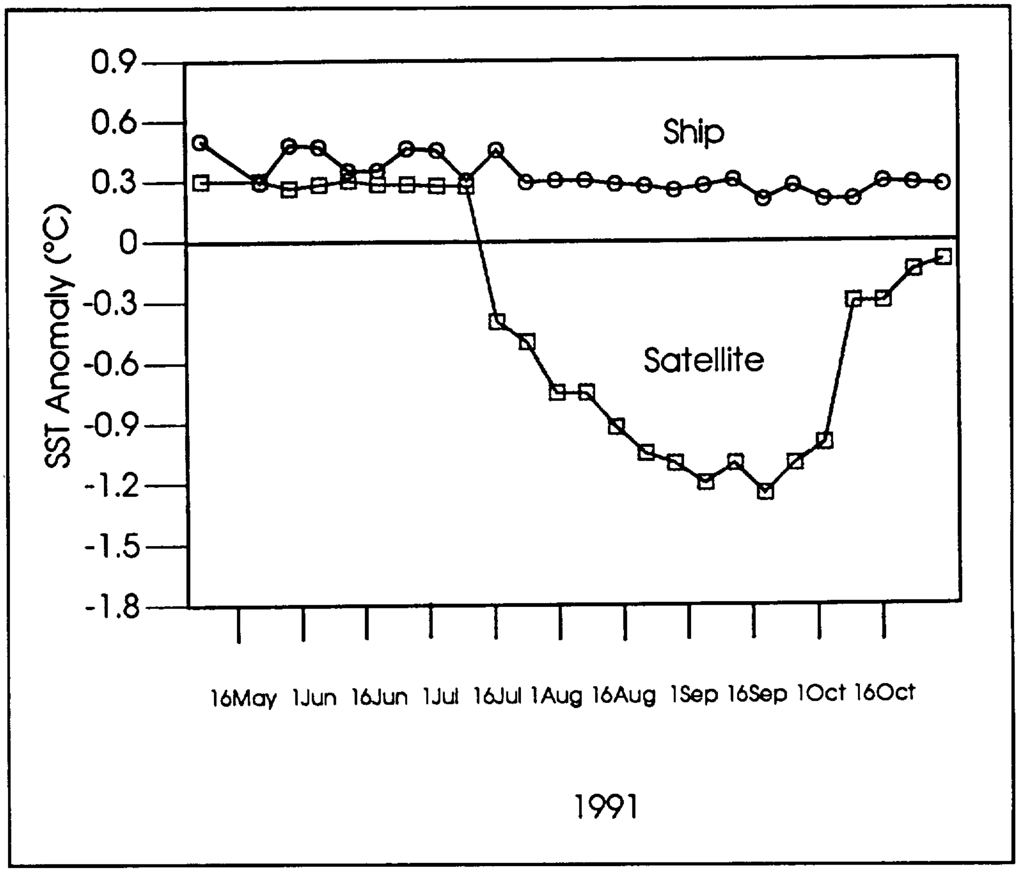

Figure 3. Tropical sea surface temperature (20°S to 20°N).

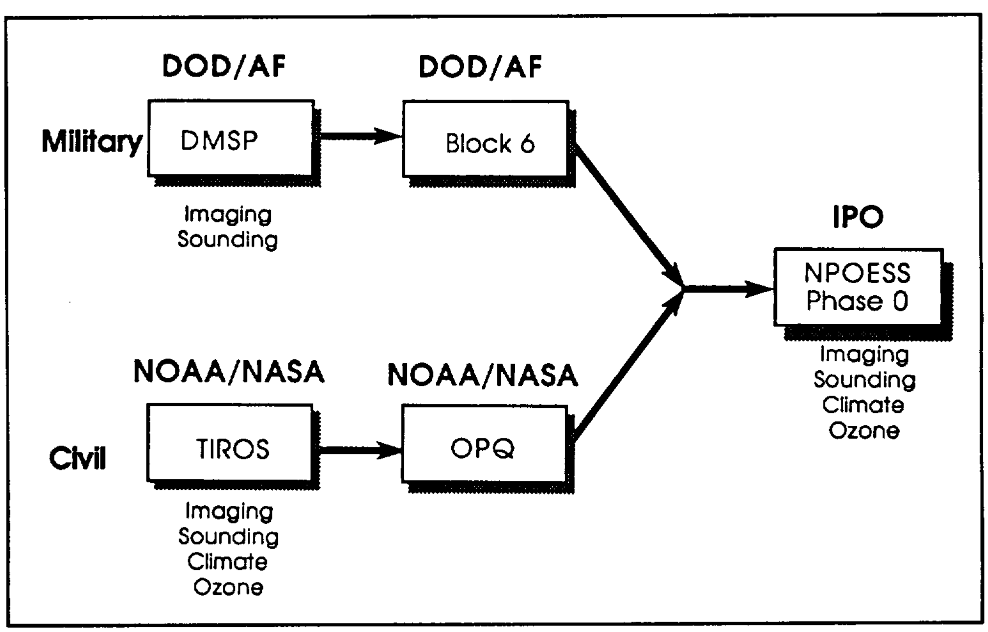

Figure 4. Meteorological satellite system evolution.

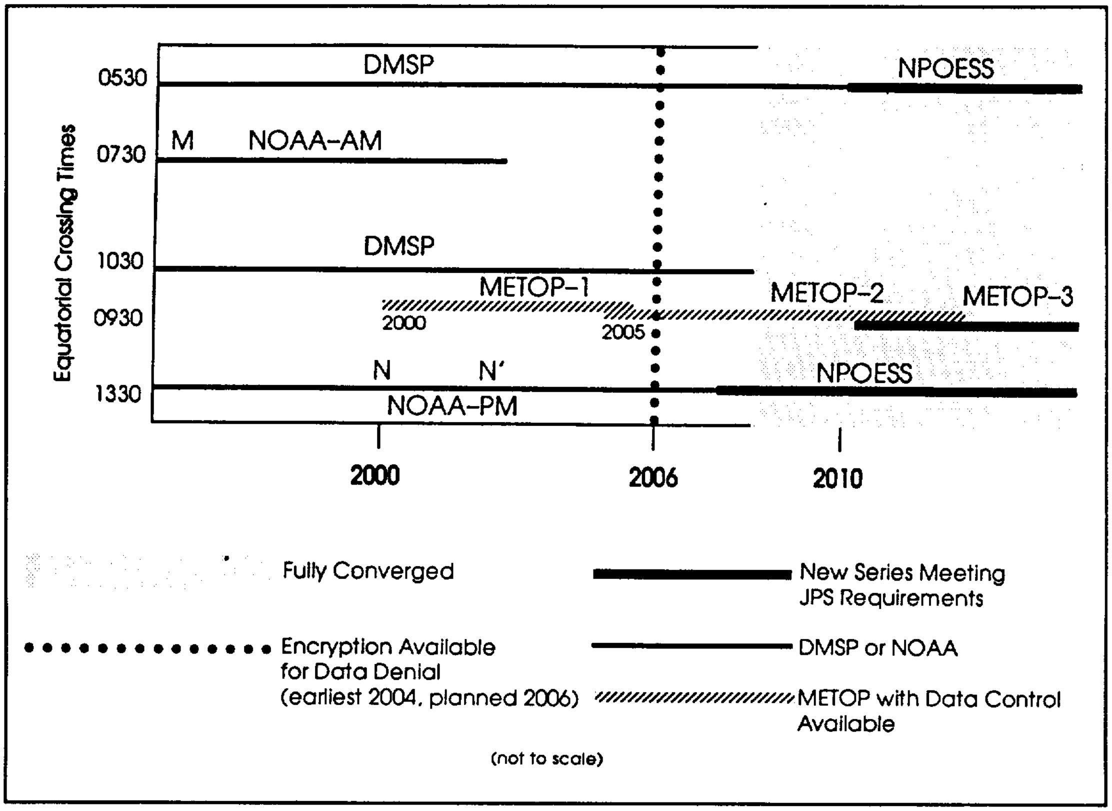

Figure 5. Convergence synchronization.

5.3 A Strategic Vision for High-Quality Earth Observations Data

Susan Till, Chair CEOS Working Group on Calibration and Validation, Canada Centre for Remote Sensing, Ottawa, Ontario, Canada

> *"Success in global change/climate and environmental research and monitoring requires a continuing commitment to the establishment, maintenance, validation, description, accessibility, and distribution of high-quality, long-term data sets, many of which rely on spaceborne observations.[1]"*

This report is a summary of the Strategic Plan of the CEOS Working Group on Calibration and Validation, prepared in September 1994 by the members of the Working Group and its subgroups. The Plan entitled "Leadership to Ensure High-Quality Earth Observation Data: A Strategic Vision" is updated on a regular basis and is available on the CEOS Infosys System at http://ceos.esrin.esa.it/CCEOSinfo.

5.3.1 The Committee on Earth Observation Satellites

Improvement of global Earth observation data for both operations and global climate research requires strong coordination and cooperation among all nations. The operators of civilian Earth observing systems from space have organized to form CEOS. CEOS, in response to this growing observation system, has two technical working groups that are having a major impact on assuring product compatibility. The Working Group on International Systems and Services was set up in 1995 to establish standard formats and networks for data compatibility (it combines the work of the former Working Group on Data and the former Working Group on International Network Services). The Working Group on

[1]The CEOS Satellite Data Exchange Principles in Support of Global Change Research, adopted by the CEOS Plenary.

Calibration and Validation, currently chaired by Canada, fosters technical coordination and cooperation for space and ground segments to document data quality.

CEOS members exchange technical information on the coordination of space and ground segments, and investigate the means for increasing data utility and cost-effectiveness for both operators and users. These means might include the standardization of data formats where appropriate, the increase in compatibility of data archives, and future directions/opportunities in Earth observations from space. Underlying all data and information are the issues of data quality and the mechanisms by which these data have been calibrated and validated.

CEOS has empowered its Working Group on Calibration and Validation (WGCV) to prepare a strategic plan that addresses the issues and needs of calibration and validation on a long-term basis. Here, the CEOS strategic vision is presented, the individual working groups are identified, and their mission and objectives introduced.

5.3.2 Calibration and Validation

Calibration of satellite sensors sets the stage on which the integrity of all data subsequently derived is based. Without such calibration, using primary or secondary standards and fundamental engineering and scientific methods, the data may be, at best, qualitative or may provide only an index of the geophysical parameter being derived.

Validation is usually a measure of how well the derived product from a given sensor describes a given geophysical parameter as measured by independent means. The most effective satellite sensors are those which have good prelaunch and postlaunch calibration, supported by rigorous surface and *in situ* programs to ensure the quality of the products on a continuing basis.

Only with both elements of calibration and validation present can changes in the product be determined to be either a change in the instrument or a change in the environment. This change in the environment may not necessarily be related to a change in the geophysical parameter calculated through appropriate algorithms. As an example, an infrared sensor measuring SST is very much affected by the atmosphere. Basic atmospheric corrections can be made, but a large explosive volcanic eruption can alter the infrared (and visible) signal for long periods. Meantime, the increased optical thickness of the atmosphere caused by the volcanic ash and aerosol clouds can lower the SST, and the only way to resolve the differences fully is through good calibration and validation.

Seasonal-to-interannual and decadal-to-centennial climate measurements require ever increasing lengths of data records. Satellite-derived data are of relatively short-term duration in an historical sense. The use of satellite information for climate purposes demands that satellite remote observations be of very high quality over long periods of time. Slight changes that may occur in either the calibration of a sensor or its validated product may generate erroneous data records unless the data are of very high quality. Because satellites are the only known technology that provides economical global coverage, this need for high-quality satellite climate data was recognized more than a decade ago. Thus, CEOS has been most concerned about data and its availability, and the quality of the data as determined by calibration and validation.

Calibration and validation considerations will affect the long-term credibility of any satellite data set. Not only must the immediate needs of understanding data quality be met, but all basic data sets used in the calibration and validation process must be maintained to ensure that improved products provide the needed information for understanding physical, chemical, and biological processes. The calibrated and validated data sets, when maintained properly, become an inherent part of the total data ensemble. In 10 to 20 years, new "pathfinder" data sets for global parameters will enhance the ability to separate sensor-, product-, and environment-introduced changes not recognized today.

5.3.3 Working Group on Calibration and Validation

A review of the satellite ensemble, either launched or to be launched in the next 10 to 20 years, shows that approximately 60 Earth observing satellites will be available to support measurements of about 35 basic geophysical parameters with hundreds of derived products.[2] Typically, each of these satellites carries from one to five or six different sensors. Each space agency has its own plan for calibration and validation of these sensors. Usually these plans are known within the respective agency and meet that agency's standards for calibration and validation, subject to resources. Selected scientists and principal investigators have specific knowledge of calibration and validation programs and may be a part of these programs. However, these calibration and validation plans are generally not well known internationally, but the international user community depends on understanding the quality of the data.

If two space sensors measure the same geophysical parameter, then ideally they product the same results. If these sensors are identical, then the same geophysical algorithm can be used. Any difference in the geophysical results then would be attributable to calibration. When this is the case, then specific coefficients must be generated for each of these sensors based on its prelaunch calibration. If the sensor has on-board calibration, the calibration can be maintained during the mission. If not, then short- and long-term sensor stability becomes a problem that can only be recognized through a rigorous validation program. However, as noted earlier, when validation is based on comparison with the measurement of a geophysical parameter, other environmental "noise" can contaminate the signal and give erroneous results, even when the sensor has remained stable.

Because CEOS recognized the need to both understand and quantify the current and future data derived from satellites, the committee established the WGCV in 1984. The work of this Group has taken on a more important role since early in 1990 when the full importance of high-quality satellite data and potential climate change observations came to the forefront. The need for data quality improvement was led internationally through such programs as those under the guidelines of the World Meteorological Organization (WMO), the Intergovernmental Oceanographic Commission (IOC), and the International Geosphere-Biosphere Program (IGBP). These programs include the GCOS, the GOOS, and the GTOS. All require high quality data over long periods of time. For CEOS to fulfill its mission, the work of WGCV was expanded (as shown in the Figure 6).

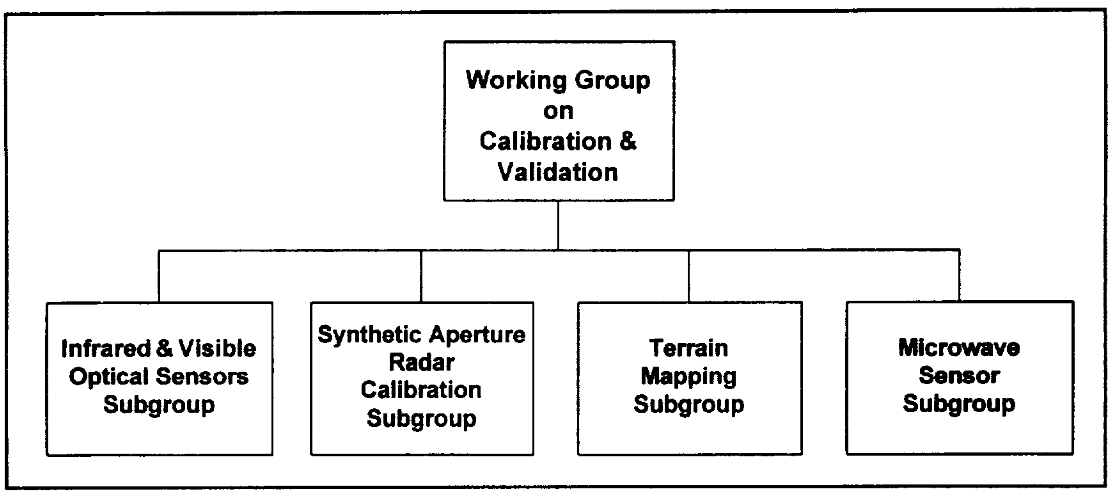

Figure 6. CEOS calibration and validation subsystem.

The Working Group on Calibration and Validation is developing a strategic plan to address the future of high-quality satellite data. The need for international cooperation has been shown to be both desirable and necessary. This need and necessity sets the stage for the vision of this group and its future.

The objectives of the WGCV are to enhance coordination and complementarity, to promote international cooperation, and to focus activities in the calibration and validation of Earth (cryosphere, hydrosphere, lithosphere, biosphere, and atmosphere) observations for the benefit of the CEOS members and the international user community. Meeting these objectives will include the promotion of:

---

[2]See for example the CEOS Dossier of Satellite Missions, Ground Segment and Data Products, and Global Environmental Programs, September 1993.

- Exchange of technical information and documentation
- Investigation of possibilities for technical coordination and cooperation for space and ground segments
- Coordination of calibration and validation campaigns and programs
- Optimizing and sharing of available facilities, expertise, and resources as appropriate

Specific objectives are:

- Sensor-Specific Calibration and Validation—To document and establish forums for the assessment and recommendation of techniques and standards for prelaunch and postlaunch characterizations and calibration.

- Geophysical Validation—To document and establish forums for the assessment and recommendation of techniques for validation of geophysical parameters derived from Earth observation satellites.

5.3.4 Specific Considerations

5.3.4.1 ***Calibration/Validation Dossier.*** One major activity of the WGCV involves preparation of a volume on calibration and validation for inclusion in the CEOS dossier. Major topics include an overview of calibration and validation procedures for each of the main sensor types currently in use, a list of major facilities in use, a list of calibration and validation campaigns and programs, and current and future needs. In 1993, the WGCV prepared a pilot draft, and recommended to the plenary that the complete calibration/validation volume be funded and completed. This recommendation was endorsed and the CEOS members and secretariat were tasked with pursuing funding mechanisms. With support from NASA, the preparation of the Dossier has now started and is to be completed in mid-1996.

This dossier will be a reliable source of reference for instrument designers needing to know who or what activity is able to offer specific services and calibration facilities. It will also help to provide data users with the capability of assessing the quality, validity, and usefulness of the data and to provide contacts for further help, if needed. The information exchange will thereby facilitate:

- coordination of calibration and validation campaigns and programs,
- investigations of possibilities for technical coordination and cooperation for space and ground segments,
- using and sharing available facilities, expertise, and resources as needed by the WGCV, CEOS, and the global community of users.

The dossier will be a dynamic document that maintains past, present, and future plans of the group. As new activities are completed, an evolution of the activities will move from future, to present, to past documentation. As a minimum, it will provide:

- a definition of terms to facilitate unambiguous communications
- a list of major calibration/validation facilities including details of each facility and the type of sensor data processed
- a record of past, present, and future calibration/validation campaigns and programs
- a summary of instrument calibration/validation activities with a history of each instrument on the past, present, and future missions.

5.3.4.2 ***WGCV Newsletter.*** One of the objectives of the Working Group is to enhance coordination through the exchange of technical information. To address this, the WGCV publishes a regular newsletter. The newsletter, CEOS Cal/Val, is issued after each meeting of the Working Group, approximately once every 9 months. It is produced and distributed by BNCS, UK (editor M.S. Hutchins). It contains news items on calibration and validation, agency reports, meeting reports, contact points for information on calibration and validation, and updates of calibration/validation coefficients, bibliographies, etc. To be placed on the mailing list, please contact M. Hutchins, UK fax 44 252 522 959.

Issue 4 was published in the spring of 1995 and includes articles on X-SAR data calibration (DLR), the Arctic lidar observatory for atmospheric research (NSC), and calibration activities at USGS. Issue 5 was published in September of 1995, with articles on microwave radiometric data calibration using geophysical auxiliary data (RAS) and postlaunch calibration for NOAA-7, -9, and -11 AVHRR channels (NOAA).

5.3.4.3 ***Terminology***. Noting that different agencies and disciplines have used similar terms but with different meanings, the WGCV is addressing the need for a common terminology in the area of calibration and validation. To date the WGCV has defined the primary terms as follows:

- Calibration—is the process of quantitatively defining the system response to known, controlled signal inputs.
- Validation—is the process of assessing by independent means the quality of the data products derived from the system outputs.

A list of other terms in common use, such as accuracy, precision, error, characterization, and verification, are included in the Appendix of the WGCV Strategic Plan (May 1995), for reference and discussion purposes. There terms are being reviewed by a subset of WGCV, and ISO and other standard references are being incorporated where appropriate.

5.3.4.4 ***Geophysical Validation***. The same geophysical product can be derived from different sensor types. Examples of this include SST, which could be derived from infrared or microwave region instruments, and land vegetation biomass, which could be derived from active microwave or passive optical sensors.

Ideally, the derived geophysical parameter estimates, from whatever source, should be identical. However, such parameters need to be validated rigorously, and the results derived from different remote sensing sources should be compared with each other and with other ancillary data and information.

5.3.4.5 ***Test Sites***. There are numerous test sites developed by agency and/or national programs that offer unique opportunities for expanded international roles. Some of the sites are instrumented and are in use only in times of specific campaigns. The WGCV studies and recommends structures and implementing mechanisms that will allow expanded use of specific test sites beyond those of national interest(s). Considerable work in this area has already been carried out by the WGCV Subgroups (see below).

5.3.4.6 ***Data Continuity***. Reliable data, with adequate quality to address the multitude of environmental problems and issues, are fundamental to the understanding of the Earth system and to the prediction of future events. Data collected in the past and present are used to understand oceanic, land, and atmospheric processes and the coupling between them. Data collected in the future will also be used for these purposes. Data continuity and the compatibility of past, present, and future data so that the observational record is free of artificial (nonphysical) changes, is, therefore, critical in the construction and use of data sets.

5.3.4.6.1 *Surface Verification Requirements*. As validation is a continuing phase of remote sensing, so also must be the availability and continuity of surface verification measurements. For operational satellite programs, the usual mode is to obtain data from a comparable operational surface program. This relation must be maintained. Satellite operators should make the surface operators aware of the need for their data and establish a system whereby both systems operate in support of each other. Activities should be maintained for all corroborating data sets.

5.3.4.7 ***Intercomparisons of Prelaunch Calibration***. Closely tied with the need to consider specialty groups to perform validation of identical geophysical parameters derived from dissimilar space sensors, is the problem of two or more similar sensor making similar, but not necessarily identical, measurements of the same geophysical parameters. In this case, it is critical that there be documentation of the methodologies by which sensors in orbit have been calibrated. No amount of validation can overcome poor or dissimilar calibration—it can confirm only that there are different results without understanding the cause.

Most Earth-observing space sensors are characterized by the determination of their absolute and relative spectral responses and the spatial uniformity of these responses. These characteristics are usually supplemented by the determination of the out-of-band or out-of-field response of the sensor. These prelaunch data are revised by the use of in-orbit calibration systems, which are rarely as comprehensive as the laboratory characterizations. (If they are active sensors, such as radars or altimeters, other characterizations may be more important.)

The documentation of prelaunch and postlaunch methodologies permits an analysis of the systematic differences that are associated with the calibration and use of that sensor. There are two methods of intercomparison. The first method requires that satellite sensors be calibrated using the same equipment. The second method requires the use of a transfer standard (i.e., instrumentation which compares the calibration equipment). Systematic errors in calibration often arise from the methodologies that are used in performing the calibration. Consequently, the CEOS should encourage the regular publication of such cross-calibration methodologies.

The CEOS has recommended the regular interchange of transfer standards. This activity permits the regular intercomparison of calibration results of different laboratories. It further sensitizes each laboratory and each agency to the methods that are employed by other organizations.

The WGCV also is concerned with postlaunch calibration and performance measurement of various sensors. This measurement often is conducted with ground-based or airborne instrumentation; it is easier to conduct comparisons of such instruments by bringing them to a common site than it is to bring spaceflight sensors together. This methodology may very well be an important characteristic of postflight comparison. However, it does not obviate the need for the transfer standards which are used to establish the proper calibration of the validating instruments.

5.3.4.8 Intercomparisons of Postlaunch Validation. Unlike the calibration of individual sensors, which is typically conducted by its developers, validation of the measurements of those sensors in terms of geophysical parameters is a more diffused activity, typically involving coordination of sensor data acquisition with associated *in-situ* measurement campaigns. Further, validation may be accomplished by relating at-sensor measurements with the results of independent field activities, or the results of ongoing, long-term ground monitoring, often conducted by researchers whose primary motivation is something other than validation of remotely sensed data. Finally, validation may be accomplished, in part, by the cross-calibration of sensors on different platforms.

Validation data are produced primarily by the scientific investigators most interested in an application of data from a given sensor. Such validation exercises are sponsored by the agencies responsible for supporting the related scientific investigations, and by agencies responsible for the dissemination of archived data sets. The "keepers" of validation data sets then, logically, will be those agencies. The WGCV must facilitate the dissemination of these data.

5.3.4.9 Linkage with Related Calibration and Validation Group. Linkages are being established with other international groups involved in related activities that are not currently under the umbrella of CEOS influence. A key mechanism envisioned for this linkage is the encouragement and use of informatic tools. These tools are best represented by the opportunity to use MOSAIC/WWW for the exchange of Earth observation-related data, whether satellite-derived or from other sources. Cooperation with the CEOS Working Group on International Systems and Services is critical to the success of linkages for calibration/validation purposes, maximizes the availability of data and information, and minimizes the efforts of participating groups/institutions.

WGCV and its subgroups are maintaining contact with other groups interested in calibration/validation, such as the (former) WGD subgroup on auxiliary data sets, the ESA ERS-1 FRINGE interferometry group, and the IOC Ocean Color group. In addition, the WGCV activities such as the specialist presentations are open to nonmember experts.

5.3.4.10 Outreach to Developing Countries. The WGCV will increase its outreach to calibration and validation experts from development countries, particularly regarding the use of test sites in the host developing country. Such interaction should be mutually beneficial. Researchers in CEOS agencies will benefit by gaining access to critical surface measurements and local scientific expertise. Researchers in developing countries will benefit by becoming involved in international programs and by gaining access to data sets for both scientific and training purposes. Both research groups will benefit from the coordination of campaigns.

5.3.5 WGCV Subgroups

The WGCV has four established subgroups:

- SAR Calibration Subgroup—Chaired by A. Freeman/Jet Propulsion Laboratory (JPL), this subgroup addresses specific calibration issues through technical working groups, workshops, discussion papers, and special journals.

- Infrared and Visible Optical Sensors (IVOS) Subgroup—Chaired by I. Barton/Commonwealth Scientific and Industrial Reserach Organization (CSIRO), this subgroup has met on a regular basis since its inception in 1992. The subgroup has undertaken three major documentation activities: 1) long-term calibration/validation sites, 2) onboard calibration systems for key satellite sensors, and 3) ground calibration facilities and laboratories. The subgroup's fifth meeting, in 1995, included a special session on the calibration/validation of ocean color sensors.

- Terrain Mapping Subgroup—Chaired by I. Dowman/University College-London on behalf of British National Space Center (BNSC), this subgroup has met on a regular basis since its inception in 1992. The subgroup is considering user requirements and available technologies, and is reviewing test sites. In particular, the subgroup has compiled a dossier of test sites by means of a site survey form, and is assembling data sets of the test sites that would be used to validate Digital Elevation Models (DEMs). The data on the test sites are to be distributed on exabyte or CD-ROM. The group is preparing a guide for evaluating terrain data derived from satellite sensors.

- Microwave Sensors Subgroup—Chaired by J. Shiue/GSFC, this subgroup addresses a range of systems issues, including those pertaining to passive sensors, scatterometers, and altimeters, and the issue of the potential RF interference to microwave radiometers from active sources. The objectives of the subgroups are listed below.

5.3.5.1 Synthetic Aperture Radar (SAR) Calibration Subgroup. The mission of the SAR Subgroup is to foster high-quality SAR imagery from airborne and spaceborne systems through precision calibration in radiometry, phase, and geometry. The objectives of the SAR Subgroup are to:

- Act as a forum for international technical interchange on the evolving methodologies, techniques, and equipment of SAR calibration.

- Determine standard definitions and calibration requirements for SAR imaging systems and to support changes in CEOS formats and user products as appropriate.

- Facilitate international cooperative programs in the calibration of SAR systems.

5.3.5.2 Infrared and Visible Optical Sensors (IVOS) Subgroup. The mission of the IVOS Subgroup is to ensure high quality calibration and validation of infrared and visible optical data from Earth observation satellites. The objectives of the IVOS Subgroup are to:

- Identify and agree on calibration and validation requirements and standard specifications for infrared and visible sensors, including on-board calibration systems.

- Promote international and national collaboration of all infrared and visible optical sensors and, thus, assist in the improved application of data from satellite sensors.

- Include all sensors (ground-based, airborne, and satellite) where there is a direct link to the calibration/validation of satellite sensors.

- Identify test sites and encourage continuing observations and intercomparisons of data from these sites.

- Encourage the timely and unencumbered release of data relating to calibration/validation activities, including details of prelaunch and in-orbit calibration parameters.

5.3.5.3 Terrain Mapping Subgroup. The mission of the Terrain Mapping Subgroup is to provide opportunities for validation and comparison of DEMs from satelliteborne sensors; to support surface campaigns by provision of test data; and to provide a forum for discussion of results and future developments. The objectives of the Terrain Mapping Subgroup are to:

- Provide test data sets of DEMs and other relevant material over a number of test sites covering different terrain, land cover, and climatic conditions.

- Provide guidelines for the use of the test data to allow evaluation and comparison of DEMs from different sensors and as a means of validating data from such sensors.

- Provide a forum for discussion of matters related to DEMs derived from satellite data.

5.3.5.4 Microwave Sensors Subgroup (MSSG). The mission of the MSSG shall encompass both active and passive microwave sensors. Active microwave sensors such as altimeters, scatterometers, and real-aperture radars are all included in the MSSG, with the exception of SAR. The microwave spectrum shall include decimeter to submillimeter wavelengths. The objectives of the MSSG are to:

- Promote national and international cooperation in the calibration and validation of microwave sensors, and to enhance coordination and complementarity in experiments and test facilities for the benefit of the CEOS members and the international user community.

- Promote accurate calibration and validation of microwave sensors through standardization of terminology and measurement practices.

- Provide a forum for discussion of current issues and for exchange of technical information.

6. CASE STUDIES AND CONSENSUS STATEMENTS

6.1 Lessons Learned From the SBUV/TOMS Analysis

Robert Hudson, Department of Meteorology, University of Maryland, College Park

6.1.1 Introduction

Hilsenrath et al. (1995) recently published an article in the Earth Observer on the calibration of BUV satellite ozone data. This article gives an excellent account of the more than 15-year history of the program at GSFC. My intention in this paper is to build on the work of Hilsenrath et al. and discuss the philosophical issues that the BUV processing has raised. I have broken these issues into 4 major categories: algorithm, data archiving, calibration, and validation.

6.1.2 Algorithms

An algorithm is a mathematical model used to invert the measured radiances to a geophysical quantity. In almost every case it is based on current knowledge of the system that is being investigated and contains assumptions and approximations. However, the objective of any research program is to improve the knowledge of that system. Thus in the course of the analysis of the radiances, assumptions will be shown to be incorrect, and better approximations will be formulated. No doubt, the EOS algorithms will change with time.

Time and time again the Ozone Processing Team (OPT) at GSFC had to go back to the original counts for the Total Ozone Mapping Spectrometer (TOMS) and SBUV as their knowledge of the instrument and the atmosphere changed. Some of these events are discussed in Hilsenrath et al. (this report, Section 7.3).

- Re-analysis of the preflight calibration leading to wavelength changes and new calibration constants.
- In-flight calibration of the diffuser plate.
- Photomultiplier nonlinearity.
- Solar zenith angle dependence. Others are the effect of Raman scattering, which fills in the Fraunhaufer Lines in the solar spectrum, leading to a small correction to the algorithm; and the sensitivity, or rather lack of sensitivity to tropospheric ozone.

6.1.3 Data Archiving

The improvements to the algorithm described above would have been of little use if the raw data (i.e., the measured counts) could not have been accessed. On-board processing has a nice ring to it, "the frontier of data management." But it also has its drawbacks, for, as discussed above, the algorithm in the spacecraft cannot be final. It should also be noted that one cannot reconstruct the original counts from the derived geophysical data, as these data are the "best fit," and one has lost the subtle but significant deviations from this fit. In the case of the BUV data, we would lose any knowledge of the effects due to aerosols.

Another reason for insisting that the instrument counts be the fundamental data to be archived, is the ability of the scientific community to come up with new ways to use the radiances. The following are examples for the TOMS data:

- Sulfur dioxide measurements from volcanic plumes.
- Tropospheric ozone determination for regions of biomass burning, especially in the tropics.
- The derivation of the ultraviolet irradiance at the Earth's surface directly from the measured TOMS radiances.

6.1.4 Calibration

If one is concerned only about the comparison of simultaneous data from different instruments, then the absolute accuracy of the instruments is of paramount importance. But as the concern of the EOS program is the derivation of data sets to observe global change, then the precision of the data becomes of paramount importance. As a potential user of the EOS data, it seems to me that much more effort has been devoted to the absolute accuracy, than to the precision, principally because of the need to cut instrument costs. Monitoring the precision of an instrument requires some form of on-board calibration, which implies an increase in complexity, size, and weight.

One lesson from the BUV instruments is that absolute calibration is not easy, especially if accuracies below 5 percent are required. Neither were the efforts to obtain a 1 percent precision. I would suggest that EOS examine carefully what accuracy and precision are actually required for each geophysical quantity, and then examine how the current instruments can meet these requirements. My personal bias is that EOS would do better by putting its money into precision rather than absolute accuracy.

6.1.5 Validation

As many of you know I do not like the use of the word "validation," it implies truth. In the TOMS/SBUV analysis, a deliberate attempt was made to steer away from this concept. The reason is that no one instrument can be expected to have the truth. Think of the normal scientific process. We never really believe a scientific experiment until the measurement has been performed by more than one group, and we really prefer to have the second measurement performed by an independent technique. Why is this? What does it gain? The answer is confidence. Confidence that unknown systematic errors have not contaminated the results. Unfortunately, I do not know of a way of quantifying this confidence.

When confronted in 1985 with the problem of an unknown diffuser plate degradation, the OPT avoided using the ground-based Dobson spectrometer network to calibrate the SBUV and TOMS instruments, developing instead internal calibration techniques. Thus in the end there were two independent sets of total ozone measurements which gave the same total ozone trends within the errors of the respective instruments.

I understand that the total ozone measurements could be considered a special case. Not many quantities can be measured independently from a ground-based network. And in many cases the quantity to be derived is the result of an analysis using a regression algorithm based on "ground truth." However, I contend that we have to rethink exactly what is meant by the word "validation," and wherever possible, we should work towards the principle of comparison of independent data sets.

6.2 Lessons Learned from ISCCP Radiometric Calibration and Monitoring

Christopher Brest, Science Systems and Applications, Inc. (SSAI), NASA Goddard Institute for Space Studies, 2880 Broadway New York, NY 10025

6.2.1 Introduction

This paper discusses procedures developed and lessons learned from efforts to calibrate ISCCP Stage B3 and Stage C data. Recently, ISCCP has begun processing Stage D data, which is a replacement and improvement on Stage C. Among the improvements are better calibration, based in part on the experience gained in developing the techniques discussed here. The methods described are those used to develop the initial calibration; the lessons learned from that experience are still valid. More complete descriptions of techniques, including references for many of the points made in this paper, can be found in Brest and Rossow (1992). The improved calibration methodology and results used to produce Stage D data will be described in a forthcoming paper.

6.2.1.1 ISCCP Stage B3 data. The ISCCP, the first project of the World Climate Research Program, began its operational data collection in July 1983. The project is designed to take advantage of the global satellite coverage provided by current and planned operational weather satellites, both geostationary and polar-orbiting. One objective is the collection of a uniform global radiance data set which can be analyzed to obtain a climatology of cloud properties to improve their parameterization in climate models (Schiffer and Rossow, 1983). Data have been collected (to date) from the imaging radiometers on the NOAA polar orbiters (NOAA-7, -8,-9, -10, -11, and -12) and from the geostationary satellites, GOES (GOES-5, -6 and -7), METEOSAT (MET-2, -3, -4, and -5) and GMS (GMS-1, -2, -3, and -4). Such a satellite data set was collected for 1 year for First GARP Global Experiment (FGGE), but the ISCCP data collection represents the first comprehensive, global, multiyear data collection.

The ISCCP Stage B3 data (Schiffer and Rossow, 1985; Rossow et al., 1987) are the raw image radiances from all the satellites that have been reduced in volume, navigated, radiometrically normalized, and placed in a common format. Three different calibration tables are available to convert the digital count values to radiances. The first calibration table (nominal) represents the best initial information available for each satellite, often the prelaunch values. The second calibration table (normalized) represents the normalization of the geostationary satellites to the current polar orbiter, and the normalization of succeeding polar orbiters to NOAA-7. Corrections for short-term changes or long-term drifts in the calibration are provided in the third (absolute) calibration table. Normalization and calibration corrections are made only for the 0.6 μm and 10 to 11 μm wavelength channels that are common to all satellites; for the NOAA AVHRR, these are referred to as Channel 1 and Channel 4, respectively. The ISCCP Stage B3 version of AVHRR data are identical to the Global Area Coverage (GAC) form, which has a nominal resolution of 4 km, except that they are truncated from 10 to 8 bits and are reduced by sampling the GAC pixels at a 24-km spacing.

These data form one of the two main products produced by ISCCP and are archived at the ISCCP Central Archive (Satellite Data Services Division, NOAA/NESDIS, Washington, DC 20233). The other is the Stage C cloud products (now being reprocessed as Stage D).

6.2.1.2 Calibration of Satellite Data. To provide quantitative measurements for Earth studies, accurate and comprehensive calibration of satellite radiometers is critically needed. The recent increase in the use of satellite data for climate studies calls for the retrieval of physical parameters from the measured radiances and, therefore, for absolute calibrations that are known over long time periods. However, even the use of classification or "index" type analyses of satellite data to monitor changes in climate requires that the relative stability of the satellite radiometers be known for long-term data sets. Moreover, plans to collect global satellite data over decadal periods, to monitor changes in surface conditions and in climate, require a calibration standard that can be transferred from one satellite to another in a series.

Although most instruments undergo a thorough calibration prior to their launch on a satellite, there appears to be no predictable relationship between these prelaunch calibrations and the postlaunch performance. Thus, comprehensive, well-documented postlaunch calibrations are needed. Since the solar channels used for imaging on most operational satellites do not have on-board calibration capabilities, an Earth-target approach to calibration is the only method available. The number of these problems which are to be avoided determines the amount of effort and expense associated with a field campaign to obtain such calibration data; making such measurements routinely and for more than one target is usually not attempted. This limits results to two-point calibrations (target plus space) and to one or two measurements over the lifetime of a particular satellite.

Use of Earth surface targets to monitor the relative calibration of satellite instruments over long time periods introduces a number of other factors that cause diurnal and seasonal changes in the radiation from the target as seen by the satellite: variations in viewing and illumination geometry, changes in the atmosphere, navigation (Earth-location of the individual image pixels) errors in heterogeneous areas, changes in the surface characteristics (such as soil moisture and vegetation), and cloud variations.

6.2.1.3 ISCCP Calibration Monitoring. The NOAA polar orbiter AVHRR data from Channels 1 and 4 play a crucial role in the project by serving as the radiometric calibration standard for all of the satellites. Although a thorough prelaunch calibration of all AVHRR channels is performed, only the infrared channels are monitored after launch using an on-board calibration target. At the beginning of the project, the absolute calibration of Channel 1 for the NOAA-7 AVHRR was not known.

To be able to use the Stage B3 radiances to determine a climatology of cloud properties, the radiances must be calibrated and the calibration maintained as a constant over the whole data set. In addition to the normalization of the geostationary satellite radiometers, this requires monitoring of the calibrations of the AVHRR over long periods. Since the polar orbiters are replaced periodically, the calibration standard must also be transferred from one satellite in the series to the next.

Lacking a comprehensive field measurement program to establish an absolute calibration, the prelaunch calibration of Channel 1 was checked by comparing the surface reflectances obtained for a variety of surface types to available values reported in the literature. In addition to uncertainties arising from surface and atmospheric effects that are not accounted for in the analysis, there is a degree of uncertainty in the reported literature values. Often description of the measurements is incomplete, regarding the type of instrument used, its calibration, spectral response, relevant characteristics of the surface measured, viewing and illumination geometry, and (where needed) the corrections for atmospheric effects, such as broken cloudiness or aerosols. These factors are responsible for a large range of reflectance values reported in the literature for specific surface types.

Despite these factors, there was sufficient agreement between visible reflectances obtained from NOAA-7 AVHRR with the published literature values to adopt the performance of the AVHRR Channel 1 on NOAA-7 in July 1983 as the standard reference for all visible radiance measurements in the entire ISCCP data set (Rossow et al., 1987). This calibration was later revised when better information was obtained.

The normalization of the geostationary satellite data to the polar orbiter is performed at the ISCCP Satellite Calibration Center in Lanion, France. The monitoring of the polar orbiters and the normalization of succeeding polar orbiters is done at the ISCCP Global Processing Center at NASA's Goddard Institute for Space Studies (GISS). The calibration of the visible channel of the polar orbiters and the lessons learned from it are the subject of this paper. For the reprocessing of the data to form the Stage D products, the calibration procedure has been modified and improved; however, that is not discussed here. The lessons learned from the original procedure are still valid.

6.2.2 Methodology

As will be discussed below the ISCCP calibration procedure differs from most in that instead of using one small selected site (e.g., a desert target), we use a wide variety of targets and the entire globe itself as a target. This brings up some interesting contrasts with other procedures which are being developed.

6.2.2.1 Surface Reflectance Retrieval. All of the B3 data are processed, representing approximately 20 million daytime image pixels per month per satellite. Data are corrected for seasonal variation in solar irradiance, corrected for ozone absorption, and for Rayleigh scattering as a function of illumination and viewing geometry. This step produces a "surface" reflectance (whether clear or cloudy) and reduces variability by removing some of the angular, seasonal, and latitudinal dependencies of visible radiances at the top of the atmosphere. The optical constants in the radiative models for visible radiance are adjusted to account for the spectral response of the NOAA-7 AVHRR Channel 1 and to simulate the observed spectral radiance as a function of viewing geometry. All optical properties are assumed to be homogeneous in a single AVHRR field of view.

A "surface" reflectance is retrieved for all radiance measurements by comparing them to a table calculated from the visible radiance model with no clouds. This procedure neglects the anisotropy of cloud, land, and water surfaces; however, collection of observations over an entire month at varying geometries provides a stable statistical measure of these reflectances.

6.2.2.2 Target Histograms. Reflectance frequency histograms are collected for nine surface/vegetation classes and 28 specific geographic targets. The data base used to sort each pixel by vegetation type is a global classification (with 1° resolution) compiled by and is based on the United Nations Educational, Scientific, and Cultural Organization (UNESCO) hierarchical classification. For this research 32 major vegetation types have been aggregated into nine vegetation (i.e., surface) "classes" based on characteristics of canopy structure and surface morphology. Previous work has demonstrated the importance of canopy structure in remote sensing of vegetated surfaces. Thus, the basic distinction is between tree-canopy and low-canopy height vegetation: grassland, shrubland, and tundra. Three tree-canopy vegetation classes are defined: deciduous, evergreen, and rain forest. Lastly, the distinct morphological surface classes of desert, snow/ice, and water are also defined.

These nine vegetation classes are further divided into geographic "targets," representing distinct regional and hemispheric occurrences of each class, to avoid problems of contamination of the class due to seasonal climate changes in cloud cover or snow cover, to derive a better (i.e., more specific) set of reflectance filters for the cloud detection step that avoids variations of a single vegetation type from region to region, and to separate seasonal illumination effects by monitoring the same surface types in the northern and southern hemisphere. Additionally, certain higher latitude areas are not available year round.

A total of 28 surface targets are chosen. Each target consists of the occurrence of the appropriate land cover type within a predefined latitude/longitude window. These targets are well distributed geographically, and comprise the bulk of the Earth's land areas.

Another set of "targets" is also defined to encompass larger areas (i.e., continents and hemispheres) with no breakdown by surface type. This target set includes the entire globe, all land surfaces, hemispheres, hemispheres subdivided according to land and water, and individual continents. Complete sets of histograms are also produced and examined for these targets, although they are not used directly in the calibration comparisons.

6.2.2.3 ***Reflectance Filters and Global Maps***. Because the CLEAR histograms are temporally stable, they can be used to define reflectance limits for each surface type and/or vegetation class, and each geographic target or region. The distances from the mode reflectances to the half-mode frequency values are doubled to define the filter limits. This procedure yields a more reasonable sample of clear-sky surface reflectances than a minimum brightness/maximum temperature approach which is too vulnerable to errors produced by spurious extreme values.

The reflectance filters are used to sort data into global maps with a latitude/longitude grid of one-half-degree resolution. Four maps are created: monthly mean surface reflectance, two biweekly surface reflectance maps representing the first and second half of the month, and a mean "total" (surface and clouds) reflectance map which represents an average of all the data available for the month. The monthly mean surface maps are used in a variety of comparisons. This method is used to monitor the calibration of each polar orbiter and, in a slightly modified form to normalize succeeding polar orbiters to the ISCCP standard.

6.2.3 ISCCP Calibration

6.2.3.1 ***Satellite Calibration Normalization***. The satellite normalization procedure will be discussed using NOAA-9 and NOAA-7 as an example. NOAA-7 served as the ISCCP normalization standard through January 1985. In February it was replaced by NOAA-9. To maintain a uniform calibration over the life of the project, it was necessary to normalize NOAA-9 to NOAA-7. For this purpose, 3 weeks of overlapping data were obtained.

The first step in assessing the difference between the calibrations of NOAA-7 and NOAA-9 AVHRR's was to process the 3 weeks of overlap data for both with the monitor software. Next the two global SURFACE reflectance maps were compared. The resulting slope of 0.896 indicated a substantial calibration difference. This difference is also evident in all of the histograms as a shift in distributions of reflectance and in the mean reflectance for all targets. The differences were greater at the brighter end of the scale, which suggests that the calibration difference includes a difference in sensitivity (gain).

At the time NOAA-9 was launched, NOAA-7 had drifted to a much later equator crossing time than the nominal 14:30 local standard time (LST). At the beginning of the ISCCP data collection period in July 1983, NOAA-7 had an equator crossing time of approximately 15:05 LST. By January 1985, the crossing time was 16:00 LST, almost a full hour later. NOAA-9 had a 14:20 LST crossing time in January 1985 which also drifted and reached approximately 16:07 LST by November 1988. The 1 hour and 40-minute difference in equatorial crossing times created a significant difference in solar geometry between the two satellites during the 3-week overlap period. The difference in μ_0 (cosine of solar zenith angle) between modal values for the Earth was 0.34.

Because of the significant difference in μ_0 between NOAA-7 and NOAA-9, we developed corrections for varying solar zenith angles by analyzing 1 year of NOAA-7 data (July 1983-84) for eight targets representing the major vegetation and surface types. Two dimensional histograms of surface visible reflectances (obtained by using the monitor filters to remove cloud effects) versus solar zenith angle were collected for each month of data and composited into an annual aggregate.

Reference μ_0 values were defined for each target by the average of the two modal values from NOAA-7 and NOAA-9. In a modified version of the monitor code, the reflectance of each pixel was corrected based on the difference between μ_0 and the reference value and the surface type slope correction factor. This correction was added (subtracted) from the original reflectance value to create the μ_0-corrected reflectance. The data were corrected to the reference μ_0 values for each target to minimize error introduced by the procedure.

The reflectance filters for NOAA-7 and NOAA-9 during the overlap period were derived in the same way as the monitor filters using modes and half-mode values. Many of the NOAA-7 filters were similar to the previous AVHRR monitor filters, but a few were significantly different because they included some snow cover or persistent cloud cover. The broader filters were advantageous for two reasons. First, because we have a limited sample (only 3 weeks) of data to work with, taken during Northern hemisphere winter when some of our targets were snow covered, the new procedure gives us a larger population to work with by defining the reflectance filters for the specific time period rather than for a whole year. Secondly, this approach gives us more data in the brighter range of reflectances, between the Sahara and snow/ice reflectances, which should be beneficial to our linear regression analysis. The filters were defined separately for each satellite based on the modal populations. This gives two sets of independent, but analogous filters.

The normalization was an iterative procedure in which gains and offsets were applied to the NOAA-9 data. Then the regression of the global surface maps, the target histograms, and the target mean reflectances were compared to the NOAA-7 data. The chosen values were a compromise selection based on the best overall results for these various comparisons.

6.2.3.2 Satellite Calibration Monitoring. Each satellite must be monitored for calibration drift over its lifetime. We monitor the monthly mean reflectance for each of the clear-sky targets and the entire globe. Analysis of the time record is complicated by the seasonal variation of most types of targets, and the drift in orbit to later in the day. Using NOAA-9 as an example, it was readily apparent that, despite these complicating factors, the calibration drifted significantly. Therefore, we developed a correction for the observed sensor drift. Given the monotonic decrease in reflectance, we chose to fit a straight line to the data, using the monitor output from February 1985 through November 1988. Because of the seasonal variations of many targets, it is necessary to use at least several complete seasonal cycles to derive a proper correction to the sensor degradation trend. We fit trend lines to the reflectance for a variety of targets, as well as the whole Earth. The slopes derived for each target were normalized by dividing by the target's mean reflectance.

Because of the global nature of ISCCP, we chose the fit to the global map as the best compromise for all targets. The trend line for the sensor degradation was derived from global mean reflectance values obtained from monthly equal-area clear-sky reflectance maps. The accuracy of our trend line is confirmed by the close agreement between the resulting calibration and an aircraft absolute calibration obtained in November 1988.

6.2.3.3 Absolute Calibration. The results of an intercomparison of several absolute and relative calibration methods for NOAA-9, combined with the absolute measurements obtained from simultaneous and coincident aircraft measurements (from the NASA ER-2 collected in October 1986), provides an absolute calibration for AVHRR Channel 1 (Whitlock et al., 1990). This value is much more accurate than that originally adopted for ISCCP, which was based on the *qualitative* agreement of measured vegetation reflectances with literature values. Consequently, this new absolute calibration is applied to all ISCCP visible data before an analysis using a radiative transfer model to obtain cloud and surface properties.

The correction to the original visible calibration recorded on all ISCCP (B3) radiance data tapes is to multiply by 1.2. This illustrates a key attribute of the ISCCP data since the original count values are reported on the data tapes, together with the original calibration supplied by the satellite operator, and any corrections performed by us to normalize to the ISCCP reference standard, users cannot only re-examine these calibration adjustments, but also exploit any new information obtained after the tapes were produced. Thus, we easily incorporated this absolute calibration into the cloud climatology analysis by introducing the correction at the beginning of the analysis software.

6.2.4 Discussion

6.2.4.1 Comments on Method. General requirements for a viable satellite calibration procedure are: a calibration standard which is well-defined and stable over time, a comparison to the calibration reference standard over most of the dynamic range of the instrument, an examination of the linearity of the instrument response, a check for shifts in spectral response, sufficient temporal resolution, sensitivity studies, and cross comparisons. Specific features of our procedure address each of these criteria.

6.2.4.2 Calibration Standard that is Well-Defined and Stable Over Time. When using a natural surface as the calibration standard, this can be a difficult goal to achieve. Even with the use of a single good target, such as White Sands, New Mexico, problems are encountered. Examples are changes in the sand dunes and their shadows caused by the winds, and high water tables in some portions that cause variable soil moisture. Desert vegetation can also vary seasonally and from year to year depending on rainfall. The ISCCP calibration program uses multiple targets and is based on the fundamental assumption that the global aggregate of regional variations of surface visible reflectance is not changing with time. Of course, on-going human modifications of the surface and climate are expected to cause some systematic changes in regional surface albedo; however, these changes are not expected to be very large, particularly at 0.6 µm, over periods of 5 to 10 years. Since we did not have a routine calibration program for the satellite radiometers at the start of ISCCP, the best available method to monitor calibration was the use of the Earth's surface. The contrast in the results between NOAA-7 and NOAA-9 provides *post facto* support for this assumption by showing both how constant the global reflectance can be and how rapidly an instrument can change. We also did not have very accurate data on the reflectivity of various surfaces, so this analysis provides a calibration relative to a somewhat poorly defined standard. As discussed previously, when a better absolute calibration is available, it can easily be applied to the entire data set.

A key feature of our method is that we do not rely on the constancy of the surface at any one location, but rather on the constancy of the global aggregate of targets. In effect, we use over 200,000 individual targets (although all the water locations do not provide independent information) to check for calibration changes. We explicitly monitor not only the statistical variations (means and distributions) of a large number of surface types and various-sized geographical aggregations, but also the changes in each 50 x 50 km^2 map grid cell covering the whole globe.

This approach supplies both a massive statistical weight and a sensitivity to different kinds of instrument changes. The former ensures that only a change in the instrument would produce a systematic shift of all the measured quantities, whereas the latter enables the detection of changes in instrument linearity or spectral responses. Although we have not emphasized this aspect of the results and have calculated only linear shifts in calibration, we have not observed any changes in the character of the regressions that would suggest either linearity changes or spectral response changes. The response of the instrument is very linear; moreover, the shifts in target mean reflectances do not indicate significant differences between water, vegetation, and deserts, which would appear with large changes in spectral response.

6.2.4.3 Comparison to a Standard Over Most of the Dynamic Range of Instrument. It is critical that the calibration be performed over as much of the range of the instrument's response as possible. For example, calibration at only the low end of the instrument's response then requires extrapolation to the remainder of the instrument's response, and with unknown error. In the ISCCP program, the use of a wide variety of surface types ensures that the measurements cover a large portion of the instrument's dynamic range. The clear-sky radiances of some desert areas are almost 50 percent of the solar insolation. The use of reflectances (radiance divided by cosine of solar zenith angle) makes the land ice sheets the "brightest" objects, even though their clear radiances are only about 20 to 30 percent of the solar constant. This means that any discrepancies in radiance measurements will amplify discrepancies in reflectances for these locations. Our corrected calibrations produce very good results for these targets, too.

6.2.4.4 Examination of Linearity of Instrument Response. Employment of a single bright target and an assumption about the dark end (space counts) does not allow for assessment of the linearity of the instrument's response. Again, using a large number of targets offers significant advantages.

6.2.4.5 Check for Shifts in Spectral Response. The recent activity in calibration has focused on response of instruments, and little work has been done on the possibility of spectral shifts from sensor aging. A method that uses multiple targets, with different spectral response characteristics (e.g., vegetation, snow, or sand), could be used to look for such changes.

6.2.4.6 Sufficient Temporal Resolution. For an aircraft-based absolute calibration program, a reasonable goal would be to aim for four flights per year. This would probably ensure two or three flights and one or two usable calibrations per year. This would not only be a significant improvement over previous calibration activities, but it should be sufficient for a viable calibration program. Whenever there is a change in satellites, the flight schedule should be altered to collect data from both satellites during an overlap period. For a relative procedure, based on the ISCCP experience, 2-week data aggregation is sufficient.

6.2.4.7 Sensitivity Studies. This is an important step in the development of a calibration scheme. It is necessary to accurately define a threshold at which to act. For example, the ISCCP synthetic sensitivity study showed that the method was probably not able to detect a calibration shift smaller than 1 to 2 percent (absolute) reliably, especially if the shift was due to a degradation of sensor sensitivity. Therefore, it was initially decided to ignore any indications of calibration change smaller than 2 percent in the ISCCP results.

6.2.4.8 Cross Comparisons. There are several calibration experiments currently being conducted, each with its own assumptions and limitations, and therefore it is important to compare results. For example, Staylor (1990) has monitored the average of measured visible radiances (converted to narrowband albedo using an empirical bidirectional model for deserts) over the Libyan desert obtained from NOAA-6, NOAA-7, and NOAA-9. Despite the different treatments of angle dependence (ISCCP neglect of solar zenith angle dependence in the monitoring procedure causes the small seasonal oscillation in the results), the agreement is excellent: calculated trends are the same to a precision better than 1 percent. Such comparisons are important to establish confidence in results.

The resulting calibration measurements represent the best global compromise for all target types and reflectance ranges. By normalizing the entire ISCCP radiance data set to this standard and maintaining constancy over the whole time period, we ensure the best approximation to a uniform radiometric standard that is obtainable.

6.2.5 Recommendations

Based on what was learned from the calibration program for ISCCP, there are a number of recommendations to be made.

6.2.5.1 Immediate actions that should be taken are:

- Obtain a consensus on calibration of current (and recent) satellites.
- Provide wide dissemination of results on a frequent basis, either in the form of a newsletter, an online data base, or both.

6.2.5.2 Actions that should be taken in the near future are:

- Formally adopt a particular methodology and set of results as best representing the calibration of particular satellites.
- Make several prelaunch calibrations with the last one being as near to launch as practical. This would provide a better basis for a prelaunch calibration than just one measurement.
- Plan 1 month of overlapping data collection operations when a satellite is being replaced by the next in the series.
- Have an aircraft campaign during this time period that calibrates both satellites.
- The data producer should conduct routine statistical monitoring of data as they are being produced. This not only serves as a quality control mechanism, but it can also indicate calibration changes.

6.2.5.3 Activities that must be started now, and that will pay off in the long-term future are:

- Technological development and implementation of onboard calibration for shortwave channels.
- Develop, test, and implement operational (i.e., near real-time) calibration procedures.

6.3 A High Resolution Global Sea-Surface Temperature Analysis

Richard Reynolds, National Centers for Environmental Prediction, NOAA/NWS, Camp Springs, MD

The SST analyses used at the National Centers for Environmental Prediction (NCEP), formally the National Meteorological Center, use both *in-situ* and satellite data. The *in situ* SST data are obtained from two different sources. The data source from 1990 to present is based on real-time data from the NCEP file of all ship and buoy observations available to NCEP on the Global Telecommunication System (GTS) within 10 hours of observation time. Prior to 1990, the data were obtained from the Comprehensive Ocean-Atmosphere Data Set (COADS; Slutz et al., 1985). COADS adds additional delayed data to the GTS data. After a wait of several years, the procedure can roughly double the number of *in-situ* observations. The distribution of observations partially depends on ship traffic which is most dense in the midlatitude Northern Hemisphere. The deployment of drifting and moored buoys has partially been designed to fill in some areas with little ship data. This process has been most successful in the tropical Pacific and Southern Hemisphere. However, it should be noted that there are areas, such as the tropical Atlantic, that have almost no buoy SST observations.

The satellite observations are infrared measurements from the AVHRR on the U.S. NOAA polar-orbiting satellites. These data are produced operationally by NOAA's NESDIS. (This operational function is now being done by the U.S. Navy.) The satellite SST retrieval algorithms are "tuned" by regression against quality-controlled drifting buoy data using the multichannel SST technique of McClain et al. (1985) and Walton (1988). This procedure converts the satellite measurement of the "skin" SST (roughly a millimeter in depth) to a buoy "bulk" SST (roughly 0.5 m). The tuning is done when a new satellite becomes operational or when verification with the buoy data shows increasing errors. The algorithms are computed globally and are not a function of position or time. Although the AVHRR cannot retrieve SSTs in cloud-covered regions, the spatial coverage of satellite data is much more uniform than the coverage for the *in-situ* data.

In-situ and satellite observations are sparse near the ice edge. To supplement these data, sea ice information is used on a 2° grid. If a grid box is ice-covered (concentration of 50 percent or greater), an SST value is generated with a value of -1.8°C. The freezing point temperature of sea water with a salinity of 33 to 34 psu is -1.8°C. This range of salinity is typical near the ice edge in the open ocean.

The superior coverage and greater density of satellite SST data would tend to overwhelm the *in-situ* data in most conventional analyses. This would only be a problem if the satellite data have biases on large time and space scales. These biases have occurred in the operational satellite data set. The most severe cases occurred following the March 1982 eruptions of El Chichón (Reynolds et al., 1989) and the June 1991 eruptions of Mount Pinatubo (Reynolds, 1993). The stratospheric aerosols from these eruptions resulted in strong negative biases in the satellite algorithms.

To illustrate the effect of one of these events, the average weekly anomaly from *in situ*, daytime, and nighttime satellite observations was computed between 20°S and 20°N during the period with strong stratospheric aerosols from Mount Pinatubo (see Reynolds, 1993). The results (see Figure 7) show that the SST anomalies were all tightly grouped during May and June 1991. After this period, the *in-situ* anomaly remained relatively constant while the day and night satellite anomalies became more negative. The nighttime anomalies reached a minimum during September; the daytime retrievals reached a minimum during August. The difference between the *in situ* and satellite anomalies shows that the satellite observations had average negative biases with magnitudes greater than 1°C in the tropics in August and September 1991. An attempt was made on October 3, 1991 to correct the nighttime algorithm. However, as shown in the figure, this correction was only partially effective. As discussed in Reynolds (1993), this correction lead to other satellite biases in the Southern Hemisphere midlatitudes. The aerosols and the associated tropical biases gradually became weaker until the biases became negligible in April 1992.

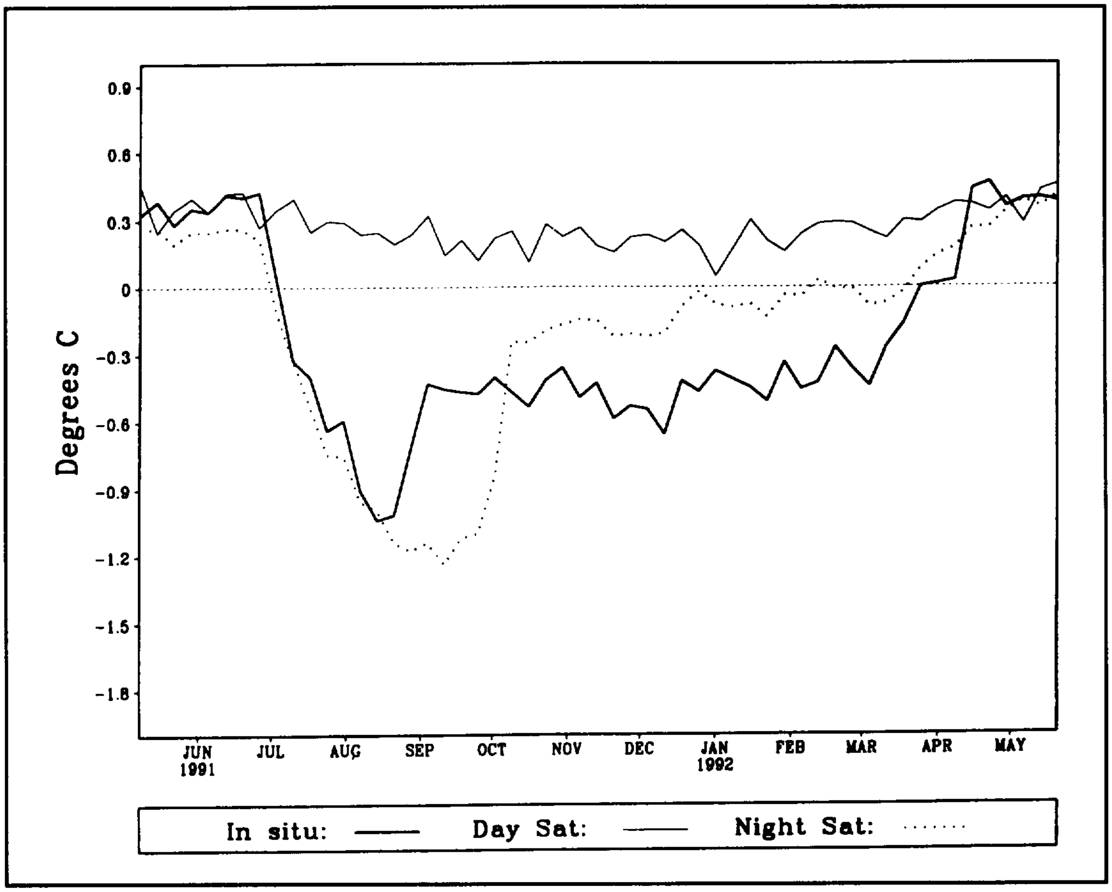

Figure 7. Weekly in-situ, daytime and nighttime satellite SST anomalies obtained from observations. The anomalies are averaged between 20°S and 20°N from May 1981 to May 1992.

The NCEP analysis of Reynolds (1988) and Reynolds and Marsico (1993) used Poisson's equation to remove any satellite biases relative to the *in-situ* data before combining the two types of data. This analysis, henceforth called the blend, was produced monthly from January 1982 to December 1994 on a 2° grid with an effective spatial resolution of 6°. The analysis used the *in situ*, satellite, and sea-ice SSTs discussed above. In this procedure the analysis resolution was degraded to a resolution that could be supported by the *in-situ* data.

To improve this resolution, an optimum interpolation (OI) analysis was developed. The OI is done weekly on a 1° grid and uses the same data that were used by the blend. The OI method (see Reynolds and Smith, 1994) assumes that the data do not contain spatial biases. To correct for the satellite biases, a preliminary step using the blended method provides a smooth correction on a 12° grid for each week. The satellite data are adjusted by this correction and used in the OI along with the *in-situ* data. In the OI, error statistics are assigned to each type of data (ship, buoy, etc.). Because the *in-situ* data errors are not smaller than the satellite data errors and because the satellite distribution is so much better than the *in-situ* distribution, the satellite data overwhelm the *in-situ* data in the OI. However, the *in-situ* data are critically important in correcting the satellite data for biases before these data are used in the OI.

The OI has now been computed from November 1981 to present. November 1981 was selected as the starting point because that is the date when the AVHRR data first became operational. For comparison, the OI has also been computed without the preliminary satellite bias correction. This analysis will be referred to as OI-UC where UC stands for uncorrected.

To verify the accuracy of the differences among the blend and the two versions of the OI, the monthly SST anomalies from the analyses are compared with independent data. These data are the monthly averaged SST anomalies from the Tropical Ocean Global Atmospheric (TOGA) Tropical Atmosphere Ocean (TAO) equatorial current moorings (e.g., see McPhaden, 1993). Three locations have been selected with the longest records: 110°W, 140°W, and 165°E. The monthly root mean square (RMS) difference between buoys and each of the three analyses (blend, OI, and OI-UC) are computed for the period from January 1982 through 93 and for each month. The results are summarized in Table 1 for a high aerosol year (1991) and for the entire period. In all cases, the OI is superior to both the OI-UC and the blend. The OI is superior to the blend because of its better resolution. Because the spatial gradients are greater in the eastern than in the western Pacific, analysis differences between the blend and the OI are greater at 110°W and 140°W than at 165°E. In years without strong satellite biases, the OI and OI-UC analyses behave similarly. However, the large biases during periods such as 1991, cause the degradation of OI-UC analysis relative to both the OI and the blend.

Table 1. Monthly RMS Differences Between Buoys and Analyses for 1991 and for the 144-Month Period: January 1982 to December 1993 (the buoys are all located on the equator).

Longitude	Analysis	RMS (°C) Difference 1991	RMS (°C) Difference 1982-93
110°W	OI	0.19	0.38
110°W	OI-UC	0.87	0.47
110°W	Blend	0.40	0.86
140°W	OI	0.15	0.39
140°W	OI-UC	0.79	0.43
140°W	Blend	0.52	0.73
165°E	OI	0.15	0.24
165°E	OI-UC	0.71	0.39
165°E	Blend	0.18	0.28

An example of a time series of the buoy and three analyzed SSTs are shown at the equator and 110°W in Figure 8 for the period January 1988 to December 1991. This period was selected to show an anomalous negative period in 1988 and an anomalous positive period in 1991 with the satellite biases shown in the Figure 7. The figure shows that the OI is in better agreement with the buoy than the blend, particularly in the anomalously negative periods of 1988. This occurs because the spatial gradients in eastern tropical Pacific tend to be stronger during periods of lower SST. The OI-UC behaves similarly to the OI during most of the record. However, the negative satellite biases overwhelm the OI-UC analysis and produce buoy to analysis differences of almost 2°C in August and September 1991.

The OI analysis with the satellite bias correction yields high quality global SST fields. The satellite data provides high resolution weekly and monthly global analyses in areas with little or no *in-situ* data. These fields have been used in the NCEP atmospheric reanalysis and for NCEP coupled model verification. In addition, the OI fields have also been used to improve SST analyses from 1950 to 1981 when satellite data were not available. In this method, spatial patterns from empirical orthogonal functions (EOFs) are obtained from the OI fields. The dominant EOF modes (which correspond to the largest variance) are used as basis functions and are fit, in a least squares sense, to the *in-situ* data to determine the time dependence of each mode. A complete field of SST is then reconstructed from these spatial and temporal modes as described in Smith et al., (1995).

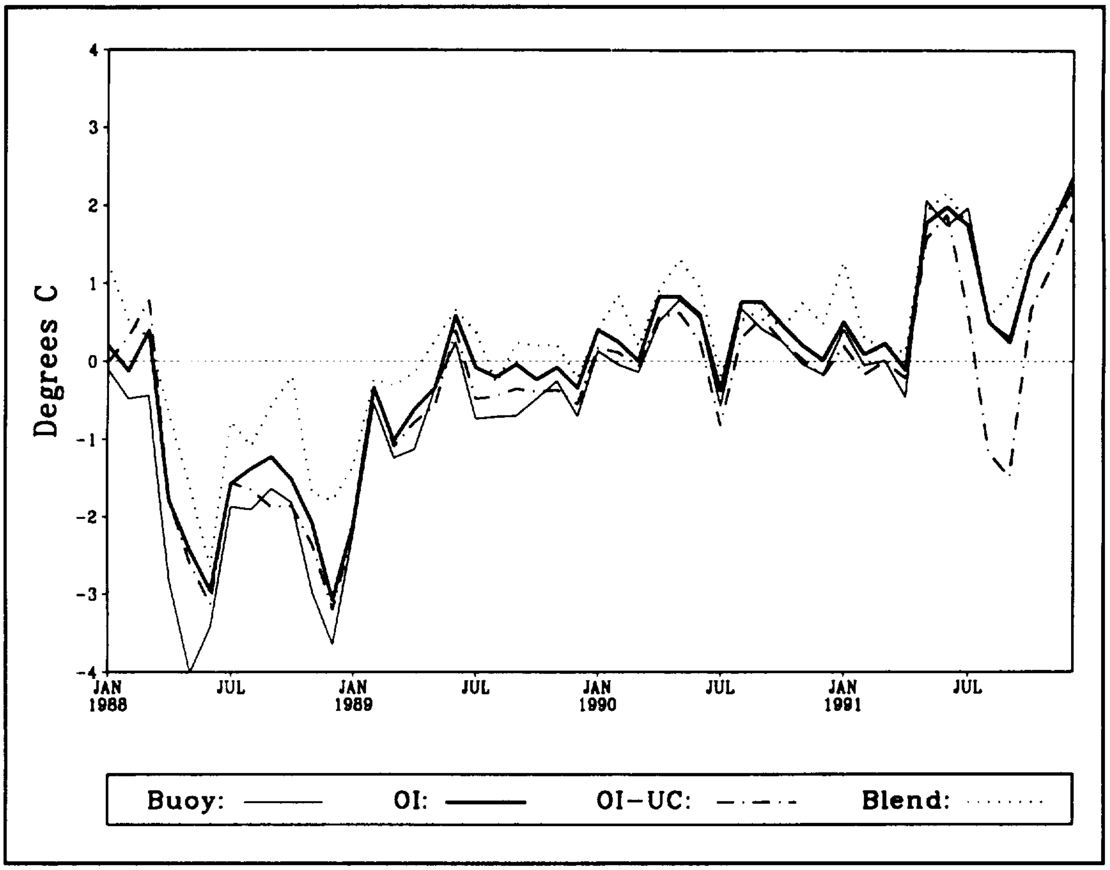

Figure 8. Monthly mean SST anomaly for the blend, the OI, the OI-UC (the OI with uncorrected satellite data) and the TOGA TAO buoy at the equator and 110°W. The time period is January 1988 to December 1991.

6.4 Data Requirements for Global Land Surface Modeling

Forrest Hall and Piers Sellers, NASA/Goddard Space Flight Center, Greenbelt, MD

A complex and coupled set of energy and constituent exchanges takes place between the land surface ecosystem and the physical climate system (Figure 9). Every class of ecosystem models (Water-Energy-Carbon, Carbon and Biogeochemistry, and Ecosystem Structure and Function) is essential for the study for global change and constrained by similar deficiencies in scaling and data needs.

- Scaling: Scaling issues are derived from using results from small-scale studies, combined with very simple aggregation assumptions, to describe regional-scale processes and surface-atmosphere exchanges. Are the remote sensing algorithms and process models developed at homogeneous plot levels valid for larger-scale regions where surface and boundary layer heterogeneity may be large?

- Data Needs: There are very few reliable, consistent, accessible, large-scale data sets in existence for the purposes of initialization and validation of these models on regional and global scales.

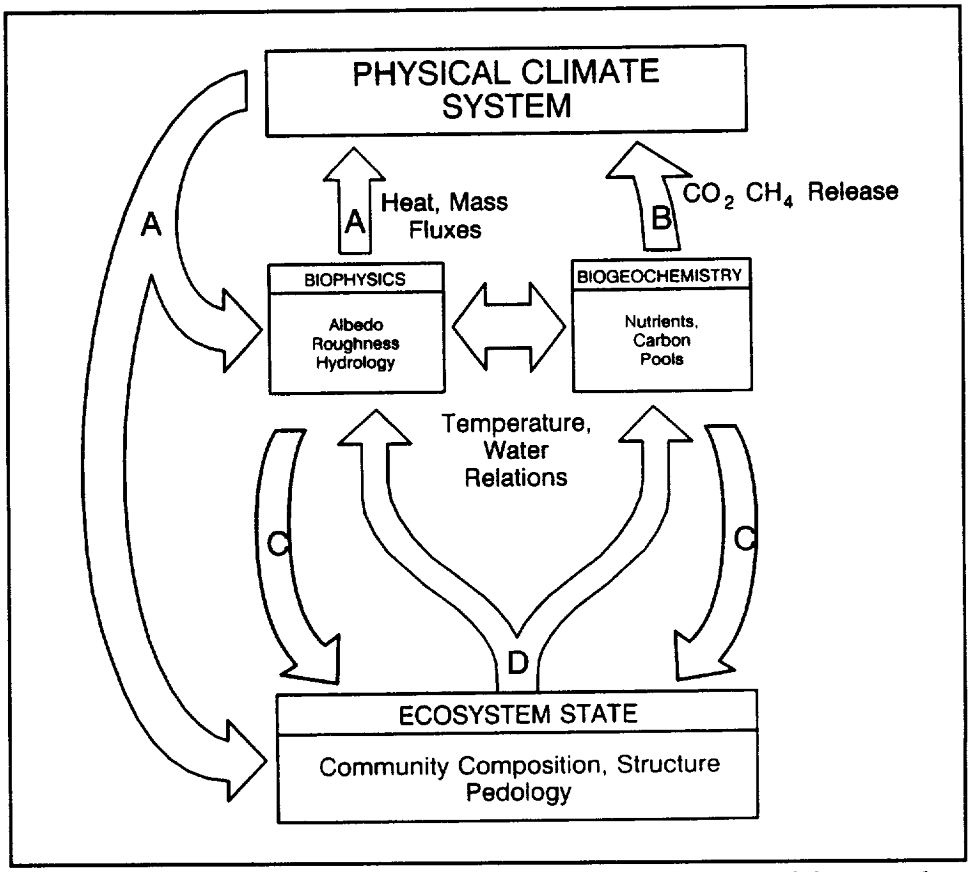

Figure 9. Important interactions between the land surface ecosystem and the atmosphere with respect to global change. A: Influence of changes in the physical climate system on biophysical processes. These effects may feed back to the atmosphere through changes in energy, heat, water, and CO_2 exchange. B:Changes in nutrient cycling rates; release of CO_2 and CH_4 from soil carbon pool back to the atmosphere. C: Changes in biogeochemical processes and water and nutrient resources influence community composition and structure. D: Change in species composition causes changes in surface biophysical characteristics and biogeochemical process rates.

It is hoped that a combination of large-scale field experiments and a stream of satellite data products can be used to deal with these two issues. Research work conducted within and parallel to the field experiments can lead to improved algorithms, which would then be used to generate better regional and global data sets. Satellite data algorithms that deal with land-surface studies have been developed under the aegis of the responsible government agencies.

6.4.1 Introduction

Early in ISLSCP, an algorithm workshop was held to determine the state of the art in satellite data processing algorithms. Sellers et al. (1990) concluded that:

- Algorithms were available to calculate many of the important surface and atmospheric state variables required by the modelers.
- Few of the algorithms had been thoroughly evaluated with regard to accuracy and precision.
- There was a lot of room for improvement in the algorithms in terms of calibration, geometric correction, and cloud screening procedures.

- Few of the algorithms had been tested sufficiently or were innately robust enough for operational use. Based on this finding, a series of field experiments were planned to develop and test satellite algorithms and ecosystem/atmosphere process models, as described in the next section.

6.4.2 Field Experiments

The International Satellite Land Surface Climatology Project (ISLSCP) field experiments and parallel activities performed by the World Climate Research Program (WCRP), the International Geosphere-Biosphere Program (IGBP) and other organizations were designed to address the two fundamental issues described above. Several experiments, including FIFE, BOREAS, HAPEX-Mobilhy, and others found the following:

- Process Models: The problems of scaling soil-vegetation-atmosphere models from local scales up to several kilometers do not appear to be as severe as originally feared. This point was re-emphasized at a joint ISLSCP-Biological Aspects of the Hydrological Cycle (BAHC; an element of IGBP) workshop held in Tucson, Arizona in March 1994 which focused specifically on scaling. A number of studies presented at the workshop indicated that the radiative transfer, and mass and heat transport models used to describe processes on the scale of individual plants or small plots could be used to calculate large-scale (10 to 50 km) surface-atmosphere fluxes to acceptable accuracies using relatively simple spatial-aggregation techniques. In some of these studies, explicit checks were made on the accuracy of these methods using a variety of surface and airborne instruments to cover the scale range from a few centimeters out to several kilometers.

- Satellite Data Algorithms: The field experiments sponsored by ISLSCP and other organizations involved the collection of integrated data sets which allowed end-to-end evaluation of procedures for calculating surface state parameters from exo-atmospheric radiances. It was found that several components of the surface radiation budget (insolation, downward photosynthetically active radiation (PAR), reflected shortwave) could be estimated from sensors on geostationary platforms to good accuracy and that useful estimates of downward longwave and net radiation could also be calculated. Satellite data were used to calculate surface biophysical parameters, including the fraction of photosynthetically active radiation absorbed by the green portion of the vegetation canopy (FPAR), unstressed stomatal conductance and photosynthetic capacity. These parameters have been used in simulation models to calculate the surface-atmosphere fluxes of carbon and water. Significantly, the remote sensing methodologies, the parameters, and the models themselves have been shown to be largely scale-invariant. This indicates that the local-scale models tested o the field experiment scale could be combined with large-scale satellite data sets to produce continental-scale fields of energy and mass (H_2O) and CO_2) fluxes.

The field experiments succeeded in dealing with the two major issues that framed their design. The next task was to take the lessons learned from the experiments and apply them to improve models and to generate better data sets. It can be argued that the modeling community directly benefited from the work: several offline models and at least three operational GCM's currently utilize formulations that are based on field experiment results. However, it is also clear that the experimental results were only occasionally used to help generate improved large-scale data sets from satellite data.

6.4.3 The Need for Global Data Sets

It was concluded at the 1992 ISLSCP workshop that the communities working on model development and on past and planned field experiments had their activities in hand. However, it was made clear that the availability and/or accessibility of global data sets for land-atmosphere models was unsatisfactory. Each modeling group reaffirmed the need for global data sets for initialization/boundary conditions, forcing and validation (see Figure 10).

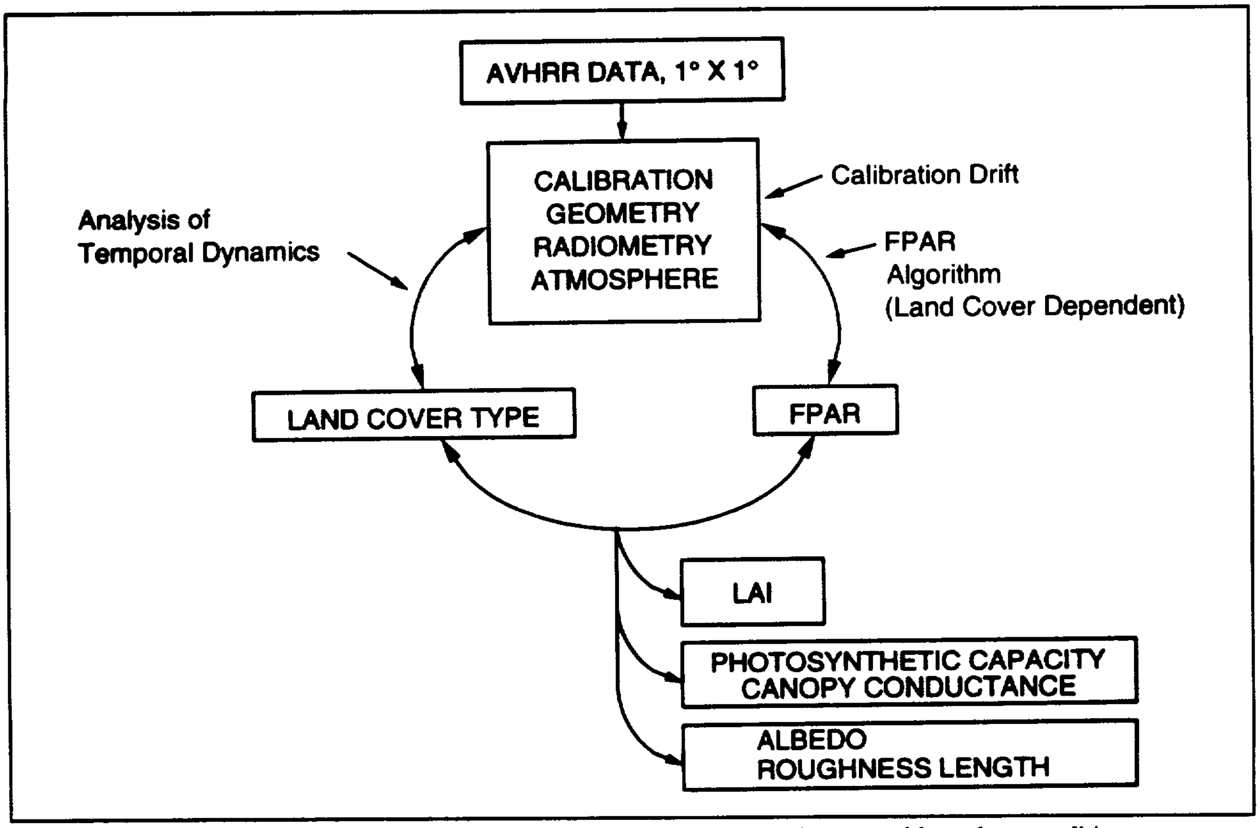

Figure 10. Land models rely on validated global data sets for energy exchange and boundary conditions.

The stated intention was to thoroughly test the surface models independent of atmospheric models, which cannot be relied on to provide realistic forcings, so as to highlight the components of the land models that need attention. Figure 10 shows the roles of the different data required to do this task. These are summarized below:

- Surface Boundary Conditions. Land cover type and associated biophysical attributes, including FPAR, leaf area index, roughness length, albedo, etc., are all necessary to specify the state and activity of vegetation in the models. Soils, snow cover, and ice data are needed by hydrological submodels.

- Forcings: Near-surface meteorological conditions (temperature, humidity, windspeed), radiation fluxes and precipitation are needed by almost all land-atmosphere models. Many of the energy-water-carbon models require that the diurnal cycle be resolved in these data and that the precipitation forcing be divided into convective and large-scale fractions.

- Fluxes: The energy-water-carbon and biogeochemistry models calculate the land-atmosphere exchanges of energy, water, carbon, and trace constituents and changes in equivalent storage quantities within the vegetation-soil system. For example, land surface parameterizations in GCMs typically produce time-series of evapotranspiration, soil moisture, snow and ice storage, and runoff. With the exception of some satellite-derived surface radiation budget data sets, there are no truly global data sets available that can be used to continuously validate the output from these models; the communities have to make do with temporally and spatially sparse surface-atmosphere flux data sets, which are mainly derived from field experiment data, and a few runoff records.

Within this framework, each modeling group prepared its own prioritized list of data sets. When these were analyzed and compared, it was found that there was a large overlap in the stated requirements. Table 2 lists the consolidated data needs as prioritized across the three working groups at the workshop. At the time, these high priority data sets were perceived to be unavailable or inaccessible to the modeling community. Specifically:

- Operational meteorological agencies generate streams of four-dimensional data assimilation (4DDA) products, including near-surface meteorology, radiation fluxes, soil moisture fields, etc., but the required information specified by the working groups was expensive and difficult to extract from the product archives,

- Very few satellite-based products were actually available, and

- Other data sets based on surface survey work (soils, topography, runoff) were available but required considerable further analysis or reduction to make them directly useful to the modelers.

With regard to data accessibility, it was thought that with some effort the situation could be greatly improved. In most cases, such as the 4DDA products, soils information, topography, etc., it was thought to be more a questions of institutions deciding to take on the job and committing resources to see it through rather than the solution of difficult technical problems.

The situation with respect to data availability was different: archives of satellite data certainly existed in the form of instrument counts, exo-atmospheric radiances, or in some cases atmospherically-corrected surface radiances. In only a handful of cases, for example the International Satellite Cloud Climatology Project (ISCCP) cloud products and the Earth Radiation Budget Experiment (ERBE) surface (clear-sky) albedo products, were there global fields of surface or atmospheric parameters. For some of the satellite-based products specified in Table 1 (vegetation, incident PAR, insolation), the raw satellite data existed but the processing had not been carried through to the production of global data sets of physical or biophysical parameters. However, most of the necessary tools and materials for undertaking such a project were available at the time of the workshop: the data existed, many of the algorithms had been developed and tested using field experiment data, and the required data product list was defined. What was required was an initiative to bring all of these things and the appropriate scientific expertise together to actually produce the global data sets. It was repeatedly pointed out that huge resources had been expended by agencies to design and launch satellite instruments, collect and archive the observations and conduct the necessary investigations to understand and use the data. The final step, applying recent scientific experience to produce global data sets of useful and usable parameters, was a clear priority and would be relatively cheap to execute, but had been done in only a few cases. These general assessments formed the basis for some specific recommendations.

6.4.4 Workshop Recommendations

Three initiatives were put forward by the workshop. These cover the immediate generation of global data sets, the improvement of methodologies and algorithms for follow-on data sets, and the improvement of communications between different elements of the Land Science community. These are discussed in turn below.

Table 2. Consolidated, Prioritized Data Needs Across the Science Areas: Water-Energy-Carbon and Biogeochemistry; and Ecological Structure and Function

Data Field	Domain	Spatial Resolution	Temporal Resolution	Source/Methodology	Action
Vegetation (cover type, phenology, disturbance, LAI, FPAR, etc.)	Regional and Global	50 x 50 km to 1 x 1 km	Monthly	• AVHRR • Landsat, SPOT	• Use an existing AVHRR product for now • Support 1 x 1 km land surface data set effort • Revitalize efforts to correct data and apply algorithms to define biophysical parameters
Near-Surface Meteorology	Global	50 x 50 km	Diurnal Cycle, Monthly Means	NMC, ECMWF, JMA, 4DDA, and Observations	• Initiate work to process 4DDA products into usable data sets
Precipitation	Global	100 x 100 km	Monthly Means and Selected Days	WCRP-GPCP, Operational Meteorological Agencies, Surface Data, Thermal IR, 4DDA	• Implement NMC workshop to analyze surface network data • Check that the above is linked to WCRP-GPCP • Provide resources for gridding data if necessary
Radiation Fluxes (S↓, S↑, L↓, L↑, PAR↓)	Global	250 x 250 km to 50 x 50 km	Diurnal Cycle, Monthly Means	GOES, METEOSAT, ERBE, AVHRR, TOMS, ISCCP, ESA, NASA Analyses	• Define interested communities, dialog with ISCCP • Check regressions using Pathfinder data • Validate against long-term data
Soil Physics: Texture, Depth, Porosity Chemistry: Minerology, PII	Global	100 x 100 km to 1 x 1 km	Once	FAO Product and Supporting Material; New Initiatives, Notably IGBP	• Assign soil physics parameters to the FAO soil descriptor fields for now • Support new initiative and encourage early deliveries
Topography	Global	10 x 10 km to 1 km or better	Once	USGS, DMA, ERS-1	• Support efforts to release all data from DMA • Check across data sets for consistency
Runoff	Regional to Global	Catchment Grid Formats 50 x 50 km	Monthly	World Runoff Data Center in Germany	• Strong encouragement to GRDC in Germany, enlist WMO support • Encourage continuous updating of the data set; gridding and averaged products
Snow and Ice	Regional to Global	25 x 25 km	Monthly	NOAA, NASA, Russian and Canadian Agencies; SSM/I and Surface Observations	• Apply existing techniques • Develop and apply improved algorithms and international communications links • Investigate use of SAR

Initiative I. Immediate Generation of High Priority Global Data Sets

The original 1992 workshop recommendation is restated more-or-less verbatim here. "It is proposed that some essential global data sets could be put together within 2 years (i.e., by the summer of 1994) and released to the community. Existing or planned data management systems should be involved in this effort from the beginning. The data sets are listed in order of priority and are shown schematically in Figure 11.

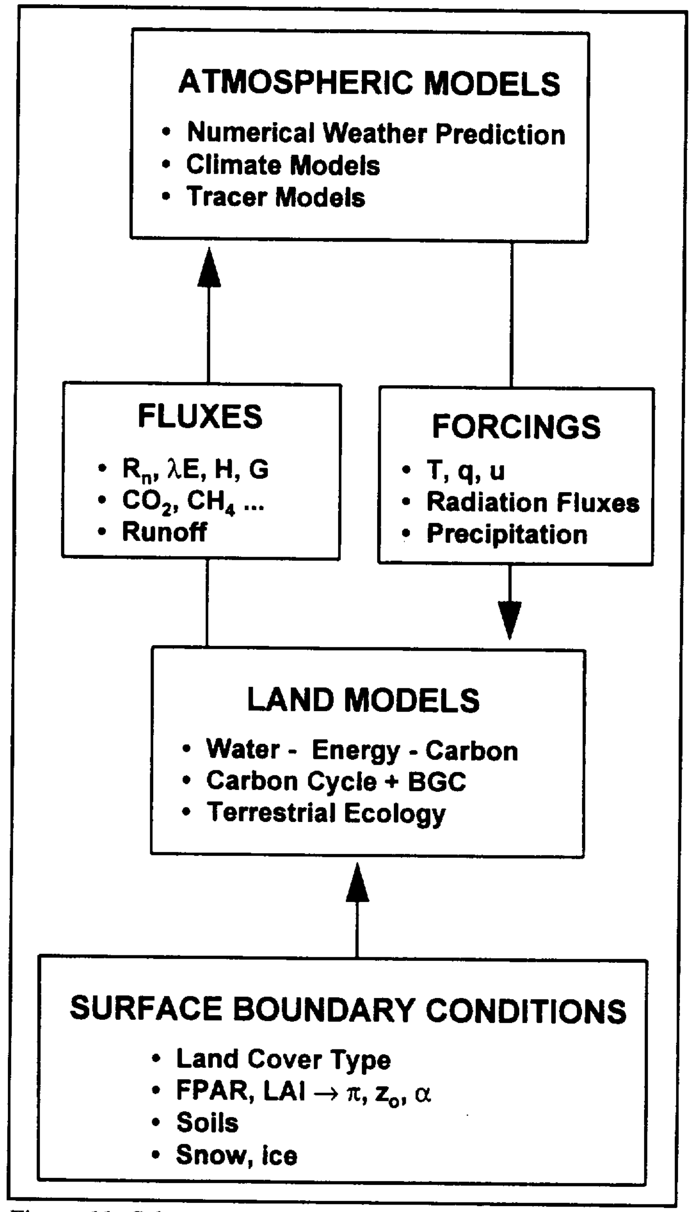

Figure 11. Schematic relationships among global data sets.

The workshop made the following recommendations for the four areas of vegetation, hydrometeorology, radiation, and soils.

- Vegetation. Global, monthly data sets of vegetation-related parameters should be generated at good spatial resolution, 100 x 100 km or better is preferred, and monthly time resolution. The available AVHRR data should be used as the basis of this effort and algorithms applied to calculate fields of cover type, phenology, FPAR, and leaf area index.

- Hydrometeorology. Near-surface meteorological data sets should be extracted from the four-dimensional data assimilation (4DDA) streams generated by operational meteorological agencies. Specifically, near-surface temperature, humidity, wind vector, surface temperature, soil moisture content, radiation components, and precipitation should all be saved. Temporal resolution should be sufficient to resolve the diurnal cycle (preferably 4 or more reports per day).

A number of institutions hold archives of rainfall data. A gridded product (100 x 100 km or better) is required with monthly time resolution and some information, direct or indirect, on the proportion of convective to large-scale precipitation. Runoff data are stored at the Global Runoff Data Center (GRDC) in Germany—these data should be processed to yield mm/day numbers (monthly means) for selected large catchments. This data subset would be of more direct use to modelers.

Snow and ice data are collated by NOAA and NASA in the U.S. and also by Canadian and Russian operational agencies, largely from analyses of optical satellite data and *in situ* observations. The temporal resolution of the data should be sufficient to resolve weather-related changes in snow extent.

- Radiation. There is a strong desire to have many components of the surface radiation budget available at resolutions down to 50 x 50 km, although it is clear the community could do good work with coarser resolution (250 x 250 km) products. Again, the temporal resolution should be sufficient to resolve the diurnal cycle. ISCCP holds data sets on insolation and longwave fluxes on a 250- x 250-km grid. The continuing ERBE work provides surface albedo estimates and net surface shortwave radiation fluxes on the same scale.

- Soils, Soil Moisture, and Topography. Global soils data sets with quantitative, even if only best-guess, soils physics and soils chemistry information are needed. Soil texture, depth, porosity, mineralogy, and pH fields are required by some water-energy-vegetation and most biogeochemistry modelers. A data set could be quickly generated based on the Food and Agricultural Organization (FAO) global 1° x 1° data base and supporting or related information.

  Soil moisture information is very useful for validating all classes of models. It was recommended that the soil moisture remote sensing community be tasked with producing some global or regional products from existing sources, such as *in situ* observations and spaceborne microwave sensors (e.g., SSM/I), even if this turns out to give only qualitative spatial and temporal patterns of soil moisture climatology, rather than precise information at a single point under ideal retrieval conditions. (These patterns would be very useful for checking 4DDA fields and other soil moisture estimates.)

Good topographic data sets are available but not easily accessible. Every effort should be made to extract the best available product from the U.S. Geological Survey (USGS) and/or the Defense Mapping Agency (DMA).

This recommendation framed ISLSCP Initiative I. After the 1992 workshop, an ad hoc ISLSCP Science Steering Committee supported by the staff at NASA/GSFC worked to put together a mutually consistent collection of data sets. It should be noted that, as requested, all the data were reformatted to a common 1° x 1° grid and cover the same period from 1987 to 1988.

The Initiative I CD-ROMs should be an invaluable resource for initializing, forcing and validating all three classes of land model (see Figure 2). One example, the IGBP, may use the CD-ROMs as a baseline initial condition and meteorological forcing data set for a global carbon model intercomparison exercise. Another example, the data on the CD-ROMs will be used to force offline versions of land surface parameterizations (LSPs) to calculate more realistic global fields of hydrological variables including evapotranspiration, soil moisture, runoff, etc. This last project is sponsored by GEWEX-ISLSCP and IGBP-BAHC. Besides these and similar applications, the data set will provide a strong starting position for global studies that will help the Land Science community prepare for the Earth Observing System data stream.

Initiative II. Improved, Follow-On Data Sets

The data sets specified in Initiative I were generated over a 2-year period (i.e., with existing data and the available robust and simple algorithms. The resulting products go some way towards satisfying the immediate needs of the modelers and will exercise every aspect of the data-algorithm-modeler pipeline as well as (hopefully) a data system or two en route. However, it is clear that great improvements could be made over this first data release, mainly in the areas of temporal coverage, algorithm improvement, and validation. ISLSCP Initiative II has the aim of releasing an improved set of global data in 1997 which should cover the period from 1986 to 1995.

Initiative III. Improved Communications Within the Land Science Community.

The workshop highlighted the extent to which related research thrusts can become separated from each other even when it is obvious that there are strong mutual scientific interests at stake. It was recognized that top-down coordination by management could provide only part of the answer. It is equally important to provide regular forums where the different communities discuss their areas of overlap on a scientist-to-scientist basis. It was observed that many recent workshops had drifted into within-discipline discussions (e.g., wish-list writing, experiment design, etc.) with little time to focus on the so-called bottleneck issues (e.g., implementation of algorithms to produce global data sets, incorporation of late-developing model needs into experiment design, etc.). Clearly, these cross-cutting issues need explicit attention.

6.4.5 Concluding Remarks

In June of 1992, an interdisciplinary Earth Science workshop was convened in Columbia, Maryland to assess recent progress in land-atmosphere research, specifically in the areas of models, satellite data algorithms, and field experiments. At the workshop, representatives of the land-atmosphere modeling community stated that they had a need for global data sets to prescribe boundary conditions, initialize state variables, and provide near-surface meteorological and radiative forcings for their models. The data sets collated on these CD-ROMs represent a first attempt to meet this need.

The data sets on the CD-ROMs are grouped under the following headings: Vegetation; Hydrology and Soils; Snow, Ice, and Oceans; Radiation and Clouds; and Near-Surface Meteorology.

All data sets cover the period from 1987 to 1988, and all but a few are spatially continuous over the Earth's land surface. All have been mapped to a common 1° x 1° equal-angle grid. The temporal frequency for most of the data sets is monthly. A few of the near-surface meteorological parameters are available both at 6-hourly values and as monthly means.

The Initiative I data sets should provide modelers with many of the fields required to prescribe boundary conditions, and to initialize and force a wide range of land-biosphere-atmosphere models. All of the data have been processed to the same spatial resolution (1° x 1°), using the same land mask, and steps have been taken to ensure spatial and temporal continuity of the data. The data sets cover the period from 1987 to 1988 at 1-monthly time resolution for most of the seasonally-varying quantities and at 6-hourly resolution for the near-surface meteorological and radiative forcings.

ISLSCP Initiative II aims to improve on this effort by covering a longer time period (1986 to 1995), at higher spatial resolution (0.5° x 0.5°), using superior data sources and algorithms where possible. In addition, GEWEX-ISLSCP and other organizations, for example IGBP-BAHC, are pursuing approaches for collating validation data sets to check the Initiative II data set at a few times and places embedded within these global data sets.

6.5 Ocean Color Multisensor Calibration Meeting

David Herring, MODIS Administrative Support Team ; Science Systems and Applications, Inc., Lanham, MD

During the third week in February, some 80 oceanographers, engineers, computer scientists, and program and project managers within the international Earth science community held a 3-day meeting at the University of Miami to discuss the future strategy for handling global ocean color remote sensing data from multiple platforms. Co-chairs for the meeting were Robert Frouin, MODIS Co-Program Scientist, and Wayne Esaias, MODIS Ocean Discipline Group Leader.

At the meeting, participants created the framework for reaching their goal as presented by NASA Headquarters: to develop a plan and approach for conducting coordinated cross-calibration and validation of ocean color satellite sensors. NASA is specifically interested in establishing data system requirements necessary for the combined use of satellite ocean color products from SeaWiFS, MODIS, Ocean Color and Temperature Sensor (OCTS), GLI, Medium Resolution Imaging Spectrometer (MERIS), POLDER, and other ocean color sensors, to address the needs for decadal-scale observations within the NASA MTPE and international Global Change research community. The meeting participants' objective is to submit a report by early May 1995 to NASA MTPE that addresses the following:

- Scientific and agency needs and objectives,
- Radiometric calibration requirements and approach,
- Global geophysical product algorithm validation,
- Multisensor data comparison and merging procedures,
- Data and data system requirements for multisensor data,
- International and national coordination, and
- Five-year budget requirements.

On Wednesday, February 22, the international contingent met in an all-day plenary session to share status reports on their respective projects, and to set the stage for addressing the issues listed above. On Thursday, meeting participants were divided into five groups to discuss discipline-specific concerns. On Friday, the attendees reconvened for a Final Plenary Session to report on each groups' conclusions and/or recommendations.

6.5.1 Radiometric Calibration and Characterization of Sensors

Chuck McClain, SeaWiFS Project Scientist, reported on Group 1's recommendation to build on the framework developed under the joint SeaWiFS-MODIS calibration and validation program. McClain suggested that a U.S. Ocean Color Intercalibration Executive Committee be formed to oversee this program. McClain also urged the ocean color community to continue developing measurement protocols—laboratory and field.

Regarding laboratory calibration efforts, McClain stated that the ocean color community should strive to expand the scope of the SeaWiFS Calibration Round-Robin beyond the present radiometric source round-robins which have been hosted at San Diego State University Center for Hydro-Optics and Remote Sensing (although round-robins of this nature should be continued). For instance, the NIST could initiate training workshops to facilitate these additional round-robins. McClain recommended that the ocean color community develop a prelaunch sensor characterization standard that describes the key parameters and tests that should be performed and documented.

Regarding postlaunch on-board calibration and stability, McClain suggested that the community could develop a cumulative description of the solar calibration, internal lamp calibration, calibration pulse, dark current, and sensor engineering data collection schemes for the present suite of ocean color sensors. He endorsed the support that the EOS Project is giving Hugh Kieffer the lunar measurement program. McClain said the community should obtain, if possible, witness filter samples for all U.S. ocean color instruments and maintain samples in a vacuum environment. NASA Headquarters should also support efforts, such as field studies, to evaluate and correct sensor anomalies such as stray light and bright target recovery.

Regarding vicarious calibration, McClain stated that the community should support additional calibration mooring sites, and other vicarious calibration programs at both high latitudes and high altitudes. Atmospheric optical measurements near calibration mooring sites should be supported. Additionally, U.S. initialization cruises for every ocean color mission launch should be supported. A common atmospheric correction scheme for all applicable ocean color sensors should be implemented by the community. McClain recognized that international agreements for data exchange must be established whereas few currently exist. Ultimately, a plan must be developed and supported by the community for evaluating and comparing onboard and vicarious calibration information and associated uncertainty budgets.

McClain stated that NASA should support an ocean color calibration data archive for prelaunch and postlaunch satellite calibration, characterization, and sensor engineering data. Additionally, match-up data and calibration round-robin data should be archived.

6.5.2 Global Geophysical Product Validation

Wayne Esaias and Frank Muller-Karger, University of South Florida, presented a summary of Group 2's deliberations. Esaias defined validation as "the process of defining the spatial and temporal error fields and regional limits for a given biological/geophysical product throughout the mission." Esaias stated that comparison of satellite-derived values with real *in situ* values is the basis for determining the accuracy of a data product in extended ranges. Error fields can also be interpreted in terms of the spatial and temporal statistics of the geophysical variables of interest. Every mission has a very basic, minimal validation program, but none are global in scope. Some validation programs are tuned to specific regions and sensors of interest. Esaias said the international sharing of validation data (*in situ* and ancillary) increases the spatial and temporal coverage by about a factor of 10 over individual projects, and enables cross-comparison of data products. Esaias added that the community should use the Sun photometer network to help with vicarious calibration.

Esaias said there is a need to identify key regions that may have inadequate *in-situ* validation activities ongoing. Global survey cruises and focused field expeditions can be used to collect uninterrupted time series data on important variables. Esaias pointed out some key areas of concern: extreme optical environments, high latitude bio-optical moorings, increased sampling of the Southern Ocean, the tropical Atlantic (and northwest Africa to characterize its dust), and major river plumes. He recommended that the GSFC Wallops Flight Facility be used to obtain aerosol optical depth data.

Validation data collection efforts should emphasize normalized water-leaving radiance, chlorophyll-a, aerosol optical depth, diffuse attenuation coefficient (k), SST, productivity, coccolith, suspended sediments, and fluorescence. Esaias stressed that there is a need to implement quality control measures in collecting validation data.

Muller-Karger stated that there is a need for data consistency, and standardized, simplified collection methods. He recommended establishing an ocean color working group with an international forum to define methods and protocols for compiling and distributing data products. Muller-Karger also recommended augmenting each international partner's validation data base several fold so that the effort becomes global in scope. In short, access to data should be easy and completely open to international partners.

Muller-Karger suggested developing a global automated observation network for physical oceanography and ocean biogeochemistry. The intent is to establish a near real-time validation data base that facilitates fine tuning of data products. He pointed out that the basic technology for such a network does exist, but needs more development (for example, biofouling is a concern). Muller-Karger encouraged the community to take advantage of ships of opportunity wherever possible—consideration should be given to using commercial vessels that frequent shipping lanes and fishery areas, and operational assets such as NOAA and U.S. Navy vessels. He recommended that ocean color community delegates conduct a feasibility study to identify vessels and shipping lanes. Additionally, technology and protocols must be refined to ensure data quality, while an international framework is established to ease sampling restrictions in foreign waters.

6.5.3 Multisensor Product Comparisons and Merging Data

Janet Campbell, University of New Hampshire, asserted that the only methods for monitoring global oceanic (or terrestrial) primary production require observations from space. Therefore, she said, our long-range goal is to produce a continuous time series of bio-optical and geophysical variables derived from ocean color satellite data. This data base will enable oceanographers to monitor changes in coastal and open ocean biological production that might occur as a direct or indirect result of climate change and human population growth.

The time series will begin in 1996 with SeaWiFS and OCTS data and, subsequently, data from the MERIS, MODIS, and GLI sensors will also be incorporated. Campbell pointed out that all of these sensors draw from the common heritage of the Coastal Zone Color Scanner (CZCS), and thus, there is a basis for merging data. Because of their high degree of compatibility, data from SeaWiFS and OCTS will be the easiest to merge. Changes in spatial and spectral resolution will make the task more challenging as the later sensors come on line.

It is unclear whether data from other sensors (POLDER, MOS Priroda, and others) will be merged because these sensors employ techniques or have other differences that may render them incompatible. Group 3 proposed creating a data set that would allow oceanographers to determine whether the data from these sensors can be merged directly, versus providing important comparative information.

The purpose of the time series is to monitor environmental change. Thus, the variables chosen include (to begin with): a CZCS-like pigment concentration (derived from CZCS bands) that will enable us to begin the time series in 1978 with CZCS data, chlorophyll-a, diffuse attenuation coefficient, and aerosol optical depths. Other variables (e.g., primary productivity, coccolithophore concentration, etc.) will be added as these become operational products at a later date.

Campbell does not recommend the production of time series for variables (e.g., water leaving radiances, epsilons, etc.) solely for the purpose of interpreting the higher-level derived variables. These data sets will exist within the project. However, the need to make adjustments or corrections to make data sets compatible is anticipated. No doubt, data from the earlier satellites (SeaWiFS and OCTS) will have to be "corrected" to make them compatible with later sensors. To this end, she recommends the creation of a Test Data Set (or Diagnostic Data Set) that will contain the information necessary to figure out how to accomplish the adjustments. This information (calibration constants, sensor gains, raw digital counts, algorithm parameters, etc.) is readily available and accessible during the initial processing of the data, but is highly inaccessible after the data are processed. Campbell recommends that a diagnostic data set be created by each sensor project at the time the initial data are processed.

The diagnostic data set will contain data and ancillary information for individual pixels located at a fixed spatial grid. The grid-point spacing will be relatively large in open ocean areas, but will get finer near shore. The total number of grid points will be on the order of 30,000 globally. Thus, the diagnostic data volume will be manageable, not overly burdensome, but extremely valuable in later years as oceanographers work out the details of how to merge the data from multiple sensors over a 15-year time period.

6.5.4 Data System Requirements

Gene Feldman, SeaWiFS Data Processing Manager, stated that recommendations on a data system could not be spelled out until processing requirements from Groups 1, 2, and 3 are better defined. Feldman suggested that he should establish a home page on the WWW for the ocean color community. The page could help clearly identify who the members of that community are, and the members' respective missions. Additionally, a Mosaic-like browser could be implemented for every ocean color project affiliated with distributed data servers running SEABASS, or a similar system in which *in situ* data are available for public use.

Feldman recommended putting together an actual data set package from a field program (e.g., the Southern Ocean JGOFS October 1996 through April 1997 campaign) that includes current ship data, and data from buoys, moorings, and other sources. Metadata for each data set should also be preserved and made available. Additionally, while JGOFS is in operation, spacecraft ocean color missions also in operation should provide access to coincident satellite data in near real-time to the JGOFS researchers and data users.

Feldman also recommended designating a group/program/investigator to develop an ongoing effort to support the collection, formatting, cataloging, and distribution of ocean color, *in situ*, and/or field support data. Feldman said that this designate may provide links to a highly distributed system, or coalesce the data into a DAAC. Feldman added that a standard grid, such as the ISCCP 1-km nested grid, must be adopted.

6.5.5 National and International Coordination

Robert Frouin stated that oceanographers need long-term data sets to study interannual phenomena, such as global and coastal change, carbon cycling, and spatial scaling. Currently, there are six satellite sensors capable of global coverage that are scheduled to launch within this decade—a great opportunity to start building a long-term data base. To take advantage of this opportunity, the ocean color community must develop composite calibration/validation data sets; pool metadata on sensor characteristics; compare and assess algorithms for product extraction; arrange for Level 1 data exchange after launch; and evaluate and distribute final products.

Frouin reiterated the need to establish a science working group on ocean color (such as an extended Japanese-U.S. Working Group on Ocean Color (JUWOC)), or to make recommendations and take the lead on strategic planning. International space agencies, government bodies, and CEOS should take the lead in setting policies, establishing multilateral agreements, writing memoranda of understanding, or whatever is needed for sharing data among the international community. Additionally, the United States should establish an interagency work group on ocean color to facilitate the coordination of programs, funding, etc.

Frouin recommended using pilot projects to lay the foundations for international collaboration. These pilot projects could help force the community to resolve problems relating to multisensor calibration and validation, data formatting and exchange issues, sharing agreements, etc.

7. CONTRIBUTED PAPERS

7.1 Calibration and Validation for Space-Based Global Environmental Studies: A Programmatic Perspective[3]

Jack Kaye, Science Division (YS), Office of Mission to Planet Earth, NASA Headquarters, Washington, DC

Disclaimer: the manuscript which follows is the author's personal opinion and is not the official position of NASA or any of its component organizations!

The usefulness of space-based data for long-term studies of the Earth requires a well-thought-out program for prelaunch and postlaunch instrument calibration. Such calibration must have sufficient accuracy that long-term changes in the measured parameter can be separated from apparent changes due to differences in instrument calibration or on-orbit degradation. In the past, calibration activities have been focused on a single instrument or a single series of instruments coming from one agency's or nation's program. However, as space-based Earth observation becomes a more international discipline, the focus on calibration issues must shift more towards ensuring consistency in calibration between like sensors coming from different agencies and nations.

This change in nature of the field leads to a significant challenge to the agencies which build and/or fly space-based instruments for Earth observations for which there is strong scientific interest in looking across the data sets from several instruments. As noted above, the rate of change of an environmental parameter may be sufficiently small that typical interinstrument calibration differences and/or on-orbit instrument degradation may be of comparable magnitudes. In some cases, temporal overlap between instruments may mitigate some of the interinstrument calibration issues. If gaps develop in a measurement series, as is likely for those parameters which are measured as part of "research" rather than "operational programs" then the interinstrument calibration issue rises in importance. Comparison to ground- or balloon-based measurements may provide a way to establish an effective calibration, but this procedure can introduce additional uncertainty due to differences in space and ground- and/or balloon-based measurements.

The difficulties associated with ensuring consistency in calibration are complicated because calibration requires a long-term investment in people and facilities. If questions about calibration of an instrument come up at some future point, it is important that the information from the original calibration, and, where possible, some of the calibrating tools and standards, still be available. Further, (and this is an assertion which I cannot fully justify at this point) much of the information learned as part of the calibration process resides mainly in the heads of people who have been involved previously. Thus, if calibration facilities and groups come and go, the ability to reconsider previous calibration processes is likely to be compromised.

This need to retain facilities and expertise makes for a particular challenge for a measurement program which is interagency and international in scope. Most agencies recognize their responsibility to provide for prelaunch and postlaunch calibration for their space projects. If calibration is always done on an individual project basis, however, the possibility of interinstrument biases developing is dramatically increased unless clear attention is paid to the use of traceable standards and metrologic procedures. Exchange of standards and/or diagnostic tools can help remove some of this danger, but this places additional demands on agencies other than that directly responsible for the instrument of immediate concern. Thus, for example, if one agency has been flying an instrument for measuring a certain parameter, and another agency will be flying its instrument to measure that parameter for the first time, it would be highly beneficial for there to be a close working arrangement between the calibration efforts of the previous and new agency teams.

It is easy to rephrase the above challenge as a financial one to space agencies—to optimally ensure consistent calibration across a series of measurements of a given parameter, agencies must provide financial support for calibration efforts which may not be focusing on their own instrument for extended periods of time. It is my assertion that this proves to

[3]Disclaimer: The manuscript which follows is the author's personal opinion and is not the official position of NASA or any of its component organizations.

be a problem for agencies, especially when funding gets tight during tough times. Then, it is likely that "protecting one's own" will take priority over "supporting one's partners." When that happens, it is easy to imagine that calibration teams will be broken up, people will move onto other things, and detailed experience that might be needed some time in the future is much less likely to be available.

While much of my attention has been paid to prelaunch calibration issues, it is clear that the maintenance of calibration after launch is critical as well, especially for measurements which are to be part of a series. If an instrument degrades on orbit in a way that is not accurately corrected for, then it is difficult to understand the origin of a calibration difference between the prior and new instrument; it could be a calibration problem in the new instrument or it could be uncorrected degradation of the previous instrument. Thus, it is especially important that calibration/validation efforts be maintained to support existing instruments, and that this support continue throughout the lifetime of an instrument and sufficiently long into the period of operation of the next instrument so that issues of prior instrument degradation vs. new instrument calibration offset can be resolved.

Coordination of the calibration effort across multiple agencies requires that up-front thought be paid to this potential pitfall. It is quite possible that the financial burdens associated with ensuring calibration consistency are not large, but they are likely not insignificant either. Redundant facilities operated simultaneously may not be needed, but support for personnel with expertise and hands-on experience is necessary. Given that most international cooperation is done on a no-exchange-of-funds basis, it becomes incumbent upon one agency or nation to maintain calibration support for its scientists, even when working on something that may be looked at as largely supporting the flight project of a partner agency.

As an example of this, it is worthwhile to consider the situation of space-based ozone monitoring. The U.S., largely through the efforts of NASA and NOAA, has led the measurement of ozone using ultraviolet wavelengths. To date, six instruments using the BUV technique have flown: the TOMS instruments on Nimbus-7 and Meteor-3, the SBUV instrument on Nimbus-7, and the SBUV/2 instruments on NOAA-9, -11, and -14. In Aorukm 1995, the European Space Agency (ESA) launched the Global Ozone Monitoring Experiment (GOME) on its ERS-2 satellite. GOME uses a modified BUV technique called Differential Optical Absorption Spectroscopy (DOAS), so consistency in calibration with the U.S. instruments was desired.

A strong calibration effort has been carried out at the NASA/GSFC, making use of facilities at the NIST. Activities at GSFC have included the recent development of a BRDF facility at GSFC. The GSFC team has participated in calibration procedures for all of the U.S. instruments and the ESA GOME instrument.

Most significantly, NASA has operated a Shuttle SBUV (SSBUV) program to help ensure the calibration accuracy of the SBUV/2 instruments, and this program was applied to the TOMS instrument data. There have been seven flights of the SSBUV instrument going back to 1989. SSBUV provides the best method for assessing the calibration of the SBUV/2 instruments. Simultaneous solar measurements will allow for correction for SBUV/2 instrument degradation which cannot otherwise be quantitatively assessed. Coincident ozone measurements can provide the most accurate correction of SBUV/2 ozone measurements, including an accurate check on the correction for degradation of the SBUV/2's diffuser plate (note on NOAA-9 the on-board calibration system did not work, so that the SSBUV correction is critical).

However, budget pressure within NASA led to the planned phaseout of the SSBUV program after the launch of STS-72 on November 30, 1995. This flight provided underflight information for three new space-based ozone instruments: the NOAA-14 SBUV/2 (launched December 1994), the ESA GOME (launched April 1995), and the NASA Earth Probe (EP) TOMS instrument (previously scheduled for launch in July 1995, although the recent failure of the Pegasus XL launch vehicle has left the launch date undetermined at this time). In addition, the STS-72 flight may also be compared with the Chilean Ozone Layer Monitoring Experiment (OLME) instrument, scheduled for launch on a Ukrainian rocket in the summer of 1995.

Thus, in spite of the plans for several other launches of UV-based ozone measuring instruments in the coming decade (3 additional SBUV/2 instruments, 3 additional TOMS instruments, the Dutch-German SCIAMACHY instrument on ENVISAT in 1998, the EUMETSAT OMI instrument planned on METOP in 2001, and the Japanese ODUS instrument planned for launch on EOS CHEM in 2002), we are looking at a reduction in existing infrastructure (both ground- and space-based) for interinstrument and on-orbit calibration.

It is important that we know what lessons can be learned from the ozone issue, for which the international nature of the measurement program is increasing more rapidly than for many other environmental parameters. If possible, it would be worthwhile to protect some of the existing capability, especially the laboratory and ground-based calibration efforts, which could potentially be maintained at much lower costs than the space-based component associated with SSBUV.

7.2 The Characterization and Validation of Data From Satellite Instruments

J. Brownscombe, P. D. Curtis, and R. W. Sauders

The calibration, characterization, and validation of operational meteorological satellite instruments is becoming increasingly important now that their value for climate monitoring has been recognized—the impact of human activities on climate continues to be a topic of global concern and there is a compelling case for data monitoring programs to try to narrow some of the uncertainties associated with long-term climate change. Over the last 30 years, satellites have provided data routinely for weather forecasting and the quality, quantity, and coverage of satellite measurements has steadily improved over the years and will remain a vital source of data for operational meteorology and climate in the future. Long time series of global data are already available for some parameters; however, it is a major task to confirm that these data are consistent and that any trends—even in the more sensitive indicators of climate change—have not been masked by instrumental changes, changes in measurement technique, or absolute calibration.

7.2.1 The Requirements for Characterization and Validation of Data from Operational Satellite Instruments

The more stringent requirements for the characterization and validation of data from operational satellite instruments posed by their use for climate-related observation have been the subject of a number of studies worldwide. These requirements have been partially addressed for existing instruments and data sets by a combination of in-orbit intercomparison of instruments in a series to transfer accurate calibrations and the use of specific external targets of opportunity over a period of time for calibration in the case of instruments without on-board calibration facilities. The "Pathfinder" data sets, which are being produced by NOAA from the AVHRR imager and the TOVS instruments, are examples of the application of these techniques. A combination of very careful prelaunch calibrations and long-term in-orbit comparisons between instruments has also been undertaken by the U.K. Meteorological Office for the Stratospheric Sounding Units (SSUs) which are flown as part of the TOVS sounding package.

In recent reports for the Commission of the European Communities on the Potential for Enhancing the use of EUMETSAT data and, more generally, in the first draft of the GCOS Space Plan, the critical importance of accurate knowledge of instrument calibration and its variation with time have been emphasized. When measurements are used for climate monitoring purposes, it is very important that all instrumental biases be removed to avoid step changes between satellites and/or other instruments measuring the same parameter and that any drifts in calibrations be identified. It is difficult or impossible to obtain information on biases for the current generation of atmospheric sounding instruments.

In the future, much more attention will need to be paid to the accuracy of initial calibration of instruments in custom-designed facilities capable of reproducing the in-orbit environment of the individual instruments. It is important to maintain consistency of calibration and characterization both between members of a series and between current and new generations of instruments—to this end, long-term intercalibration and testing of instruments in common facilities is desirable. These facilities should maintain traceable and well-documented links to established standards to guarantee the long-term consistency. In addition, programs of in-orbit intercomparison of different members of the same series of instruments and for intercomparing different instruments play an important role in ensuring long-term consistent data sets.

The current operational meteorological instruments are passive radiometers looking at visible, infrared, or microwave radiation with the quantitative information largely from thermal IR and microwave channels. To understand the instruments, we need to know not only the radiometric performance, but also to understand the physical processes relating the observations from space to the parameters of interest. This includes validation of the radiative transfer models by aircraft, ground-based, and laboratory measurements of both the atmosphere and the surface.

Some specific examples of techniques used for accurate prelaunch calibration and characterization, for postlaunch validation, and for verifying the radiative transfer models are described in the following sections in the context of both numerical weather prediction and climatological requirements.

7.2.2 The UKMO Instrument Characterization Facilities

The U.K. Meteorological Office's ground-based test facilities are being used to carry out tests on the next generation of operational passive microwave instruments. The instruments are placed in a 3-m thermal vacuum chamber which is evacuated to a pressure of better than 10^{-5} torr to simulate the space environment. There are a variety of tests on the satellite instruments which have to be performed. First, the thermal behavior of instruments can be studied by cycling the instrument between the minimum and maximum temperatures expected in orbit. Various heaters and coolers allow an instrument to be maintained at any desired temperature between 263 and 323 K to within 0.1 K. Second, the radiometric performance and calibration accuracy of a total power microwave radiometer can be determined by viewing blackbody targets placed in both the Earth-viewing direction and the space-viewing direction. The temperature of the targets is known to an accuracy of better than 0.1 K, and the microwave emissivities have been measured to be >99.999 percent.

The noise figures for each channel are determined by viewing the Earth target, and a check on any nonlinearity of the detectors is obtained by varying the Earth target temperature between 80 K and 330 K. The difference between the measured target temperature and the radiometric temperature measured by the radiometer gives an estimate of the absolute accuracy of the instrument calibration. In addition, this test facility has the capability to move the position of the Earth target during tests allowing any scan dependence in the calibration to be measured. This unique feature, together with the prove ability in temperature control of an instrument and the targets, has resulted in much interest being shown by the builders of operational microwave instruments.

The high frequency part of the AMSU-B has been fully characterized in this facility. The AMSU will fly on the NOAA operational polar orbiters NOAA-KLM from 1996. Results have been obtained for the AMSU-B engineering model and all three flight models. The engineering model of the lower frequency component of AMSU (AMSU-A1) is also due to be tested in the facility during the summer of 1995. Plans are also underway to test the MIMR and the MHS both planned to fly on the EUMETSAT METOP platform early next decade. Airborne instruments can also be tested and, for instance, the NASA microwave imaging radiometer was fully characterized during 1994.

The test carried out on the AMSU-B flight models showed their radiometric performance was within specification over the predicted operating temperature range of the instruments. For instance, the noise figures for the five AMSU-B channels are all below 1 K (with the exception of one channel on the first flight model which is 1.1 K). The departure from linearity for all the channels was always less than 0.1 K which is well within the specification of 0.3 K. The absolute calibration was shown, as expected, to have a small scan dependence for cold targets (a change of 0.4 K from nadir to edge of scan) and this was fully characterized and a correction for the in-orbit data formulated. The results of the tests will be made available to the user community to allow them to make better use of the in-orbit AMSU data.

7.2.3 Airborne Instrumentation for Validation of Measurements and Models

The Meteorological Research Flight (MRF) of the U.K. Meteorological Office operates a modified C-130 aircraft for use as an airborne platform for atmospheric research. It has an endurance of over 11 hours and can operate from 20 m to 10 km altitude. The aircraft is equipped with a host of instruments for measuring all significant atmospheric variables. In support of remote sensing studies, a suite of radiometers are installed or under development. There are two microwave radiometers on the aircraft in support of the in-orbit data validation and model development for the AMSU program. A two-channel microwave radiometer at 89 and 157 GHz can view both the nadir and zenith and has been used to test the radiative transfer models and effects of clouds and surfaces at AMSU-B frequencies. A lower frequency two-channel radiometer at 23 and 50 GHz has just been completed for measurements at AMSU-A frequencies.

Clear air measurements at 89 and 157 GHz have been made through a variety of different types of atmosphere from over the Baltic Sea in winter to the tropical Atlantic. This data set has enabled a comparison with several microwave radiative transfer models to be carried out over a wide range of conditions and an optimum model for operational applications to be selected. Also measurements through water cloud have been made to develop techniques for remotely retrieving cloud liquid water path (LWP) from the AMSU-B radiance measurements. Accuracies of 50 g/sq.m in retrieved LWP were found by comparison with *in situ* measurements. Measurements of a variety of different surface types including snow and ice are now underway at all four microwave frequencies to help recognize and characterize surface types from the AMSU data.

For visible and infrared wavelengths, a 16-channel filter wheel radiometer is mounted in a wing pod which can view both the nadir, zenith, and two calibration blackbodies. The filters cover relatively wide spectral bands, and so to improve the spectral resolution, an interferometer is being developed with much higher spectral resolution (0.7 cm^{-1}) to cover the thermal infrared region of the spectrum. This Airborne Research Interferometer Evaluation System (ARIES) will be used to develop improved models and retrieval algorithms for future satellite interferometers (e.g., Improved Atmospheric Sounding Interferometer (IASI)). It will also be used for studies of clouds and surfaces at higher spectral resolution and for remote measurements of minor atmospheric constituents and aerosols. Together with the microwave radiometers on the C-130, the inclusion of ARIES will allow a complete airborne equivalent to the proposed next generation of space atmospheric sounders (i.e., a combination of an infrared interferometer and AMSU). Current plans are for ARIES to fly for the first time during 1996.

7.2.4 Future Development of Requirements

Numerical Weather Prediction (NWP) models of necessity use data from a wide range of both conventional and remote sensing instruments. Increasingly sophisticated data assimilation schemes are being used. It is expected that by the end of this decade, such models will be able to use fully 4-dimensional data assimilation and be able to assimilate direct radiances measured by satellite instruments rather than requiring the separate generation of products such as temperature profiles before assimilation.

To utilize any type of measurement in the most effective way, the NWP models require information on the error characteristics of the measurements so that they can be given the appropriate weight when they are assimilated. Bias errors between different measurements of essentially the same physical parameter are particularly damaging. Thus the importance of knowing the accurate radiometric performance of an instrument and the uncertainties in the spectroscopic models will be even greater in the future for effective use in NWP.

For climate uses, the most stringent requirements will flow from the desire to detect long-term changes in atmospheric and surface variables. This will put an increasing premium on the accurate identification of changes of instrument calibration with time and will demand increasingly effective use of prelaunch and postlaunch calibration and intercalibration facilities. An important consequence will be the need to maintain expertise on detailed instrument performance and its measurement over long periods of time. Only by safeguarding this knowledge will it be possible to continue long-term measurements with known error characteristics which will allow the production of historical single instrument data sets for comparing and contrasting to determine as accurately as possible long-term climatological changes. Although assimilation is a powerful technique for using the dynamical equations to constrain and, to an extent, quality control input data it is not able to maintain long-term consistency from a changing mix of observations. Hence, it is also important to maintain long-term single instrument data sets for the accurate detection of long-term changes in climate.

7.3 Calibration of BUV Satellite Ozone Data—An Example for Detecting Environmental Trends

Ernest Hilsenrath, NASA/GSFC, Greenbelt, MD; P.K. Bhartia, Laboratory for Atmospheres, NASA/GSFC, Greenbelt, MD; and Richard Cebula, Hughes STX Corporation, Greenbelt, MD

The environmental and economic implications of ozone depletion were just emerging when the Nimbus-7 satellite was launched in 1978. This marked the beginning of ozone monitoring with observations from the SBUV and TOMS instruments. Since that time, a National Plan for Stratospheric Monitoring (1989) was put into place. Its keystone was the continuation of ozone observations on the NOAA polar orbiting satellite series with the SBUV/2 instrument. In parallel to this plan, NASA is flying a series of improved TOMS instruments on the U.S. Earth Probe and Russian and Japanese environmental satellites. Both the NOAA and NASA series of instruments will operate into the next century. A version of TOMS, supplied by the Japanese, may also fly on the EOS Chemistry platform. An advanced instrument using SBUV principles (called the GOME)), will be launched by the ESA in 1995 on the ERS-2 satellite. This instrument will measure several important atmospheric constituents including ozone.

A combination of satellite- and ground-based observations has now verified that global ozone depletion has occurred. Measurements of photochemically active gases in the stratosphere and model calculations indicate that this depletion is most likely the result of anthropogenic gases reaching the stratosphere (WMO, 1991). Because atmospheric ozone blocks harmful ultraviolet radiation from reaching the Earth's surface, a global policy was established to phase out certain gases, such as chlorofluorocarbons, known to be harmful to the ozone layer. With this phase out, ozone is predicted to recover to levels existing prior to 1980 about the year 2030 (WMO, 1991). Nevertheless, it remains clear that global-scale ozone monitoring should be continued to verify model predictions and to monitor for possible unpredicted events such as the Antarctic ozone hole and the recent accelerated ozone depletion in the Northern midlatitudes attributed to Mount Pinatubo aerosols (Gleason et al., 1993).

Global total ozone has decreased by approximately 2 to 3 percent (Stolarski et al., 1992; Bojkov et al., 1990) while upper stratospheric ozone has decreased by about 5 to 8 percent (Hood et al., 1993) since 1978. These trends, derived from both satellite- and ground-based data, are consistent with model predictions (WMO, 1991). Detecting and verifying these trends has been a major challenge to the scientific community. As part of an international effort, the OPT at the GSFC has been committed for over 15 years to refining and validating ozone data from SBUV, SBUV/2, and TOMS (collectively called BUV) instruments flying on the NASA, NOAA, and Russian satellites. This task is being realized through algorithm improvements and comprehensive studies of the various instruments' prelaunch and postlaunch calibrations. The purpose of this article is to review how the BUV instrument performance has been assessed and to describe some of the tools employed to correct prelaunch and postlaunch calibrations. An example will be given that demonstrates the suitability of corrected and combined ozone data sets from two satellites for climatological and trend studies.

7.3.1 Satellite Instrument and Data

The SBUV instruments which view in the nadir, measure both column amounts of ozone and ozone profiles from roughly 25 to 55 km with an instantaneous field of view (IFOV) of 200 km square. The TOMS instruments measure only total column ozone. Employing a cross-track scanner TOMS can map, with 50-km square resolution at the nadir, the entire daylight portion of the globe in about 24 hours. SBUV and TOMS instruments operate on similar principles and employ a common algorithm based on measurements of the Earth's geometrical albedo in the wavelength range where ozone absorbs ultraviolet radiation (250 to 340 nm). Measurement of the geometrical albedo (proportional to the ratio of Earth's backscattered radiance to the incoming solar irradiance) has a distinct advantage among remote sensors and has been critical for accurately detecting long-term changes. In principle, the albedo measurement allows a canceling of nearly all time-dependent instrument changes that are common to both the radiance and irradiance measurements. For BUV measurements, the most critical component that does not cancel is the reflectivity of a diffuser which is deployed only for the solar irradiance measurements. The accuracy of the ozone measurement depends on the accuracy of the prelaunch albedo calibration and the uncertainty of the solar diffuser time-dependent changes. There are other noncancelling parameters, such as instrument linearity, that must also be tracked carefully over time. The BUV instruments and their operating periods studied by the OPT are listed in Table 3.

The NOAA-14 SBUV/2, Earth Probe TOMS, and ERS-2 GOME are scheduled to be launched over the period from December 1994 to June 1995. The data from these new instruments will represent the continuation of the long-term record started by Nimbus-7. Therefore, the challenge is expanded not only to track precisely each instrument's performance over time, but also to assure consistency among measurements, since each instrument will likely have different calibration biases and a unique observing geometry.

Table 3. BUV Satellite Instruments Studied by OPT

Instrument	Period
BUV (Nimbus-4)	1970-1975
SBUV (Nimbus -7)	1979-1990
SBUV/2 (NOAA-9)	1984-present
SBUV/2 (NOAA-11)	1989-present
SSBUV (Shuttle)	7 flights since 1989
TOMS (Nimbus-7)	1979-1992
TOMS (Meteor-3)	1991-1994

7.3.2 Prelaunch Calibration

Consistent prelaunch calibration of an instrument of one type is essential for traceability when successive instruments of one type are flown. However, accurate absolute calibrations are essential for understanding differences among instruments which employ different observational techniques. Calibration of BUV instruments requires standards and measurements provided by NIST. The primary standard for irradiance is the well-established FEL lamp, a 1000-watt tungsten filament lamp with a quartz envelope. These lamps have been reliable and have long demonstrated consistency and stability of the order of 0.5 percent. FEL lamps have a two-fold purpose in BUV albedo calibrations. First, they are used for the irradiance calibration and second, they illuminate a laboratory diffuser target for the radiance calibration. The target's BRDF must be known to compute its radiance. In the past, NIST has provided these BRDF measurements, but recently they have not been able to meet all the NASA and NOAA requirements for BRDF measurements. As a result, NASA has established a facility at GSFC to measure BRDF of targets used for the BUV and other satellite-borne environmental observing instruments.

Recently an integrating sphere has been successfully employed to calibrate the radiance sensitivity of BUV instruments (Heath et al., 1993). The technique involves transferring the irradiance of a standard FEL lamp to the sphere. The sphere radiance is computed from its irradiance through a geometrical expression. The integrating sphere has now been used to calibrate the SBUV/2, SSBUV, TOMS, and GOME instruments. These instruments have also been calibrated using targets whose BRDFs have been measured in the GSFC facility. Comparison of the integrating sphere and diffuser calibrations has shown agreement of about 1 percent (Janz et al., 1994) for SSBUV, SBUV/2, and GOME. For TOMS, the agreement ranged from 1 to 3 percent. The NOAA-14 SBUV/2, which was launched December 30, 1994, and the Earth Probe TOMS and ERS-2 GOME launched in 1995 will, therefore, have had common and consistent prelaunch calibrations. Overlap of these instruments with the presently operating instruments (Table 3) is not guaranteed. However, a Shuttle SBUV (SSBUV) instrument flight will provide a link from the old to the new instruments. The SSBUV will be discussed below.

7.3.3 Postlaunch Calibration

An accurate description of an instrument's time-dependent characteristics is fundamental to detecting trends in most of the Earth's environmental parameters from space. This critical issue has been recognized by the EOS community, and, therefore, instrument stability has been a major consideration in specifying performance on the EOS platforms. Figure 12 illustrates the total output change experienced by the Nimbus-7 TOMS instrument during 14 years in orbit. The changes shown here include changes to both the solar diffuser's reflectance and drift in the instrument optics throughput. For most of this period, the decrease in output is dominated by diffuser degradation. The Nimbus-7 SBUV/TOMS instrument did not include an on-board calibration system. Therefore, indirect methods were developed to assess diffuser change based on the observation that the diffuser degraded approximately exponentially with solar exposure (Cebula et al., 1988). Although the model worked well for the first 6 years of the mission, by the mid-1980's it was clear that the model was underpredicting the amount of diffuser degradation based on comparisons of the satellite data with ground-based ozone monitoring stations. This underestimate was probably due to diffuser degradation occurring even when the diffuser was not exposed to the Sun.

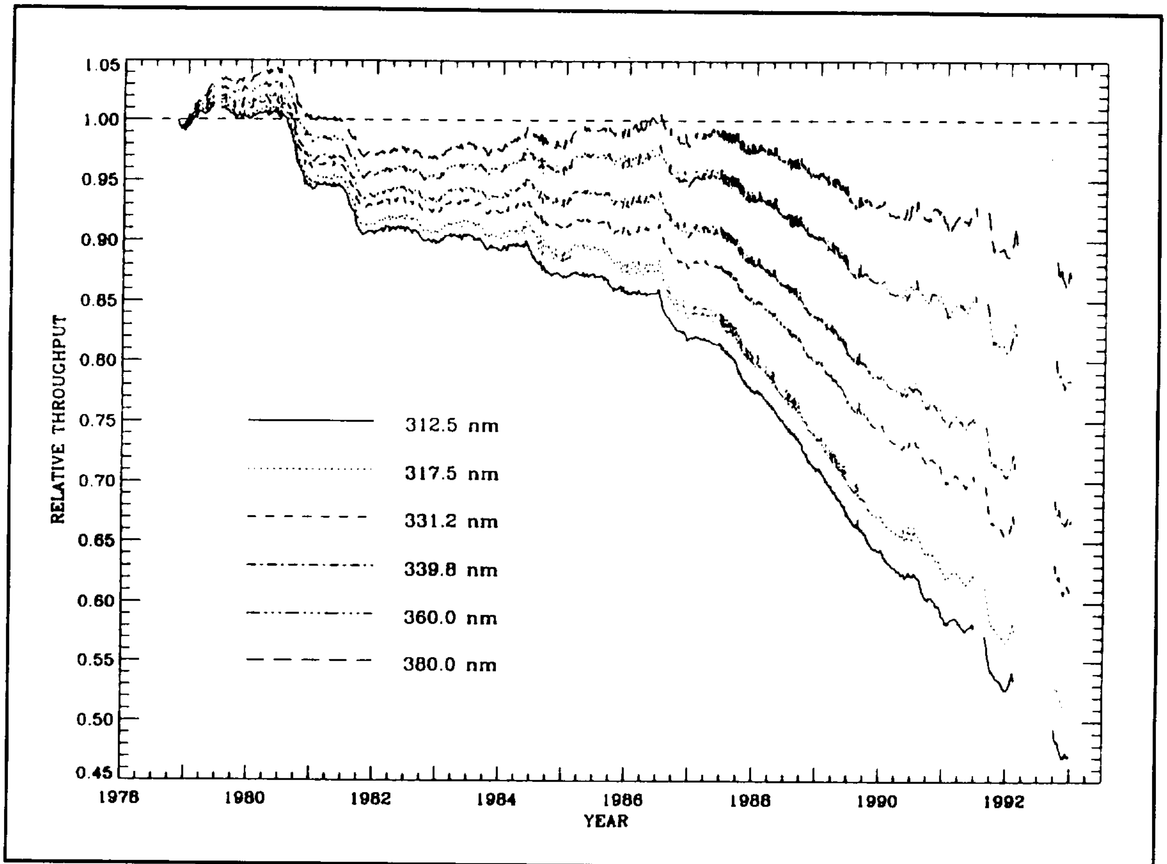

Figure 12. TOMS total instrument throughput degradation over time relative to first day. The shorter wavelengths degrade the most.

Figure 13 summarizes diffuser degradation experienced by several BUV type instruments. For each instrument the diffuser degradation rate increases with decreasing wavelength. However, the exact rate (percent change in diffuser reflectance per unit solar exposure) and the spectral dependence varies greatly from one instrument to another. These degradations are thought to result from polymerization of contaminants due to exposure to the solar UV and, therefore could vary from one instrument and spacecraft to the next. Like the Nimbus-7 SBUV and TOMS, both the Nimbus-4 BUV (not shown) and the Meteor-3 TOMS instruments also experienced diffuser degradation that was approximately exponential with solar exposure. In contrast, however, the diffuser degradation experienced by the NOAA-11 SBUV/2 instrument over the period from 1988 to 1993 is approximately linear with solar exposure for the small amount of diffuser degradation observed during this period. These observations indicate only a few of the complexities encountered in characterizing diffuser temporal and spectral behavior.

To overcome this problem, several techniques have been employed by the OPT to characterize and correct diffuser degradation. The following is a review of three of these techniques: (1) spectral discrimination method, (2) on-board systems, and (3) intersatellite comparisons. Other techniques, not discussed here, take advantage of steady scenes such as the Antarctic and Greenland ice caps and clear-sky ocean reflectivity.

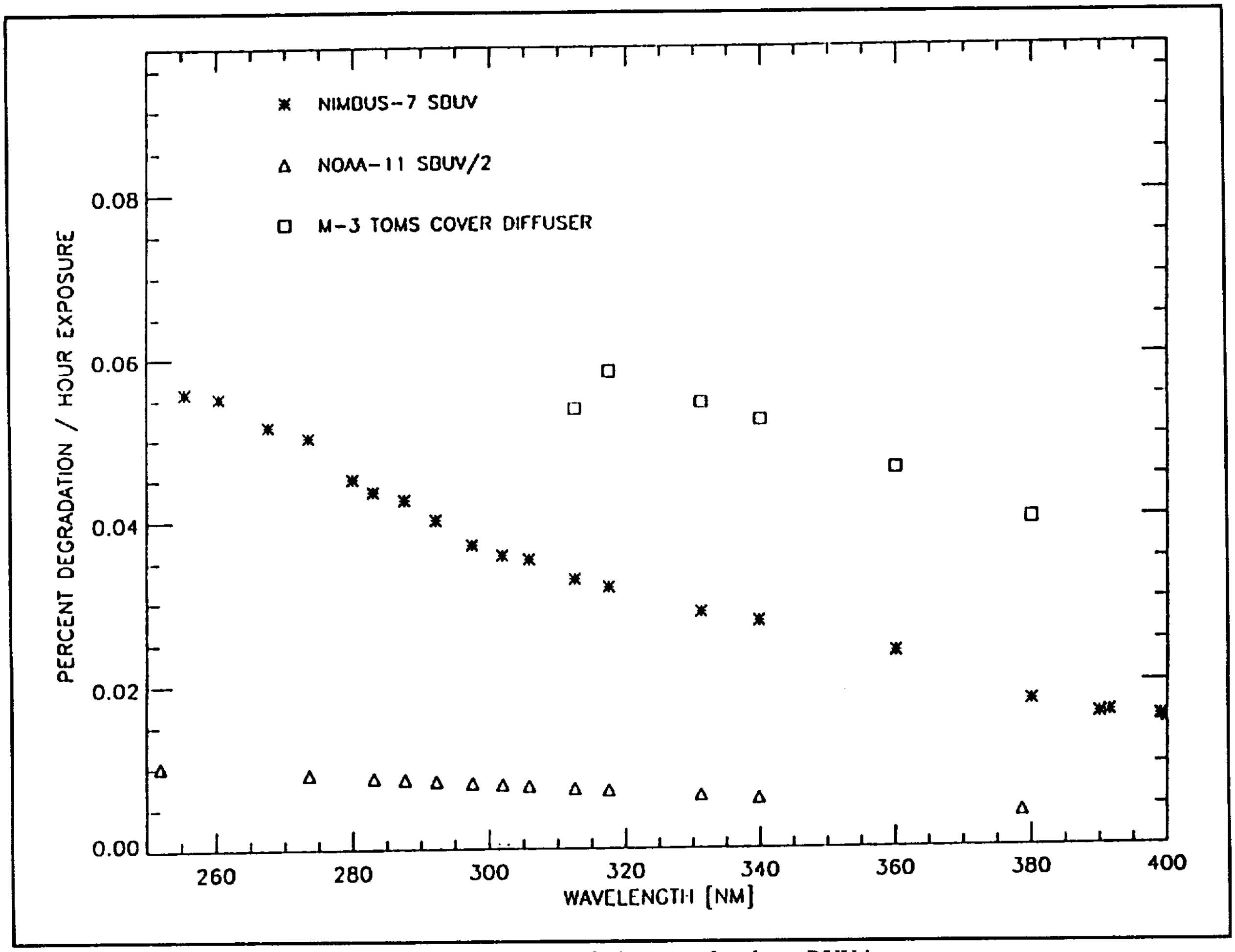

Figure 13. Solar diffuser degradation rate for three BUV instruments.

7.3.4 Spectral Discrimination Method

As an alternative to adjusting the Nimbus-7 satellite SBUV/TOMS data to the ground-based data, it was necessary to develop internal calibration methods to more accurately describe the diffuser degradation in the absence of a direct measurement. One of these methods is called the "Spectral Discrimination Method." It takes advantage of the fact that the observed albedo response to diffuser degradation is roughly linear with wavelength in the 305 to 400 nm region (Figure 12), while the effect of a real decrease in atmospheric ozone would produce an exponential change with wavelength of the measured albedos. Given a sufficient number of wavelengths, these two spectral features can be easily separated. A simplified version of this method that used only selected pairs of wavelengths (called the Pair Justification Method) was used to correct total ozone measurements from the SBUV/TOMS instrument (Herman et al., 1991). This method could be improved further by using continuous spectral data. Although the SBUV has a limited amount of spectral data, the GOME instrument has been designed to take full advantage of this technique.

Unfortunately, this method is limited to wavelengths longer than 305 nm. At shorter wavelengths, ozone changes produce linear rather than exponential change in radiance with wavelength that cannot be readily separated from diffuser degradation. For shorter wavelengths, a different technique was developed, inspired by the so-called Langley Plot method, long used for the calibration of ground-based ozone instruments. This method compares the ozone values derived from measurements made at different solar zenith angles. Ozone can be observed under these conditions when the polar region is in perpetual sunlight and the same latitude band is observed twice (once during the ascending portion of the orbit and again during the descending portion). Since the solar zenith angles for these two sets of measurements are different, the Langley Plot method can be applied to these data. The actual application of this method to BUV measurements is complex (Bhartia et al., 1995), and subject to uncertainty from the diurnal variation of ozone in the upper stratosphere. Nevertheless, this method worked quite well based on ground and satellite intercomparisons. The uncertainty in estimating the 274-nm albedo is about ±3 percent. This translates to uncertainty in ozone at 2 mb of about ±5 percent over the 12-year life of the SBUV instrument, during which the diffuser itself suffered a degradation of about 40 percent at this wavelength.

7.3.5 On-Board Calibration Systems

To overcome the difficulty experienced with the Nimbus-7 diffuser degradation, the NOAA SBUV/2 instruments carry an on-board system to measure diffuser reflectivity change. This system consists of a low-pressure-discharge mercury lamp, which is viewed by the instrument directly and then is viewed by the instrument via the diffuser. This system failed on the NOAA-9 instrument but worked satisfactorily on NOAA-11. The diffuser reflectivity is determined from weekly measurements of six strong lines of the mercury lamp. Figure 14 illustrates the diffuser reflectivity change at the 253.7-nm line of the lamp (noise in the data during the first year was caused by some imprecision in the grating drive). With this and similar data at other mercury lines, a linear fit over time and then over wavelength was derived. This yielded time-dependent correction coefficients at all wavelengths covered by the SBUV/2 measurements through January 1993. (Since that time NOAA has employed an extrapolation of those coefficients for operational processing). Other time-dependent factors including temperature, gain, and instrument goniometry (which was sensitive to the NOAA satellite drifting orbit) were all corrected for in the most recent reprocessing of the NOAA-11 SBUV/2 data (Hilsenrath et al., 1995).

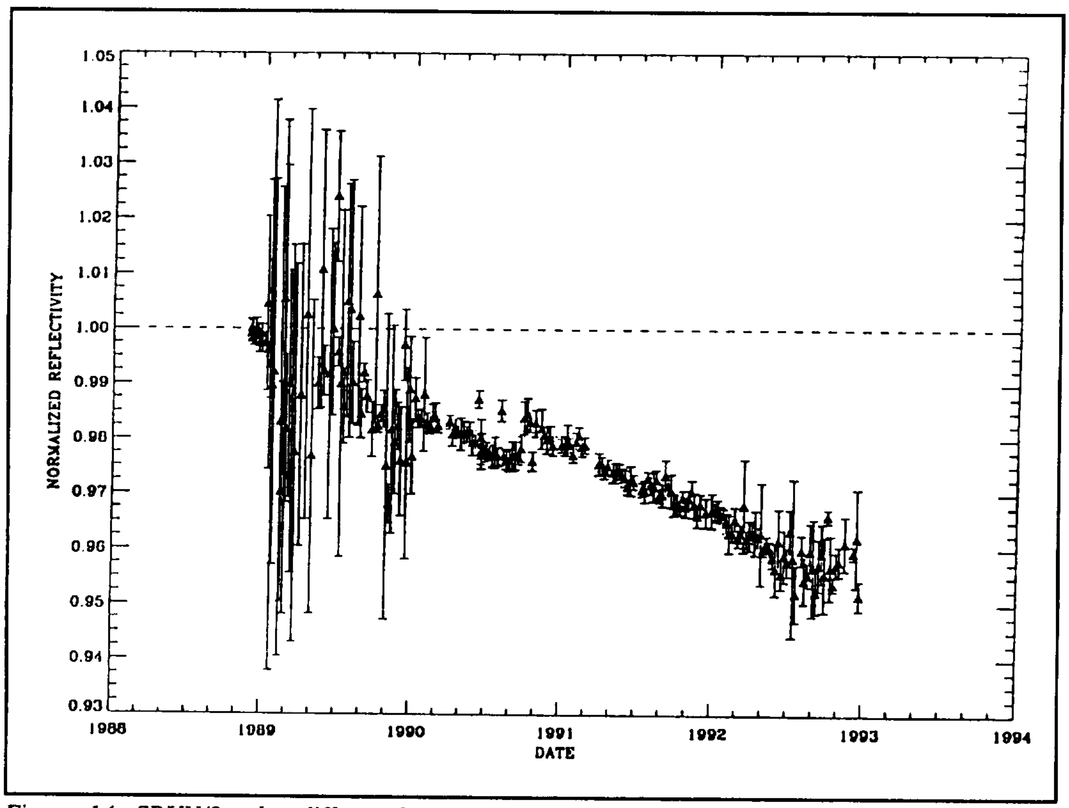

Figure 14. SBUV/2 solar diffuser degradation as measured using the on-board system. Noise appearing in the first year was due to grating drive errors.

The Meteor and subsequent TOMS instruments incorporate further refinements to minimize and measure diffuser degradation. These instrument diffusers consist of three surfaces in a carousel configuration. The first surface is a cover and exposed to space most of the time. The second surface is the working diffuser and exposed to the Sun about 1 hour per year. Finally, the third surface is used as a reference and only exposed less than one-third hour per year (Jaross et al., 1994). The upcoming TOMS, to be flown on the U.S. Earth Probe and on the Japanese ADEOS-1, will incorporate both the carousel and an on-board system to measure diffuser reflectivity.

7.3.6 Intersatellite Comparison

Another method to evaluate and calibrate an aging satellite instruments' sensitivity is to compare its data with data taken from a freshly launched instrument. In accordance with this concept, the National Plan (1989) called for regular flights of an SBUV instrument (called the SSBUV) on the Space Shuttle as additional assurance that the drift in the NOAA SBUV/2 instruments would be accurately corrected. The strategy for employing SSBUV to check the calibration of the NOAA instruments was described by Frederick et al. (1990). Accurate flight-to-flight SSBUV calibrations are critical to achieving this goal and are checked regularly. Calibration precision has been tracked with a precision of 1 percent (Cebula et al., 1989; Hilsenrath et al., 1993). SSBUV has flown seven times since 1989 and has conducted near simultaneous observations with all of the instruments listed in Table 3 except for the Nimbus-4 BUV.

The NOAA-11 SBUV/2 data were corrected using the on-board system discussed above as a first step. Its calibration was then further checked using nearly coincident SSBUV observations in the following way. From a coincident data set (approximately 60 coincidences per mission), an Albedo Normalization Factor (ANF) was computed from the difference in the measured albedos between the two instruments with a further correction factor to account for the fact that the measurements from the two instruments are not taken under identical conditions. The correction to measured albedo difference is calculated from a radiative transfer code which accurately predicts the albedos for a given solar zenith angle, surface reflectivity, wavelength (ozone cross section), and ozone amount (measured from one of the two instruments). The ANF values were computed for three SSBUV flights covering the period 1989 to 1991 and then used to correct the NOAA-11 SBUV/2 prelaunch calibration and to verify the time-dependent corrections derived from the on-board systems (Hilsenrath et al., 1995). Table 4 summarizes those results and lists percent ozone corrections to the NOAA-11 operational (uncorrected) data set for the period January 1989 to April 1993. In principle, an ANF correction could be used between any number of BUV instruments in orbit after selecting one as the standard and then defining coincidence criteria between the two.

Table 4. Percent Corrections for NOAA-11 SBUV/2 Data

	Prelaunch	Time Dependent (per year)	Total (after 4 years)
Total Ozone	1	0.4	2.5
Ozone at 15 mb	0	0.8	3.2
Ozone at 1 mb	5	3.0	16.0

The NOAA-11 SBUV/2 data, reprocessed with an improved algorithm and the calibration corrections discussed here, have been reported by Planet et al., (1994) to describe the effect of Mount Pinatubo aerosols on the recent trends in Northern Hemisphere ozone global amounts. SSBUV data, using a refinement of the process described above, are now being employed to verify the calibration of Nimbus-7 and Meteor-3 TOMS with promising results. However, correction of the TOMS data is much more complex because of large scene variability. Larger scene variations occur with TOMS comparisons because of its higher spatial resolution and because observations are made at wavelengths (>300 nm) that are reflected from the Earth's surface.

7.3.7 Results

With these corrections, the NOAA-11 SBUV/2 data can now be compared to overlapping data from the Nimbus-7 SBUV. These data, which overlap for about 18 months starting in 1989, are compared in Figures 15 and 16. Figure 15 illustrates the difference in total ozone between the two instruments as a function of latitude and time. The differences are on the order of 1 percent except at high-latitude winters where they differ by as much as 4 percent. This probably results from the fact that the measurements from the two instruments are made at different solar zenith angles at the same latitude. This solar zenith angle effect probably results from residual calibration errors in one or both of the instruments. Figure 16 illustrates a time series of ozone centered at 3 mb, over Southern Hemisphere midlatitudes, for both SBUV and SBUV/2. This is the altitude most sensitive to ozone destruction by anthropogenic chlorine via homogeneous chemical catalytic processes. The corrected SBUV/2 data match the SBUV data fairly well. The annual cycles and the long-term trend in ozone superimposed on the 11-year solar cycle are also clear. The differences in ozone are about 6 percent which translates into an albedo difference of only 3 percent for the corresponding wavelength. This is a remarkable result considering the history of the two data sets. Several studies are now underway to combine these data sets to derive ozone trends as a function of latitude and altitude for the period from 1979 to 1993.

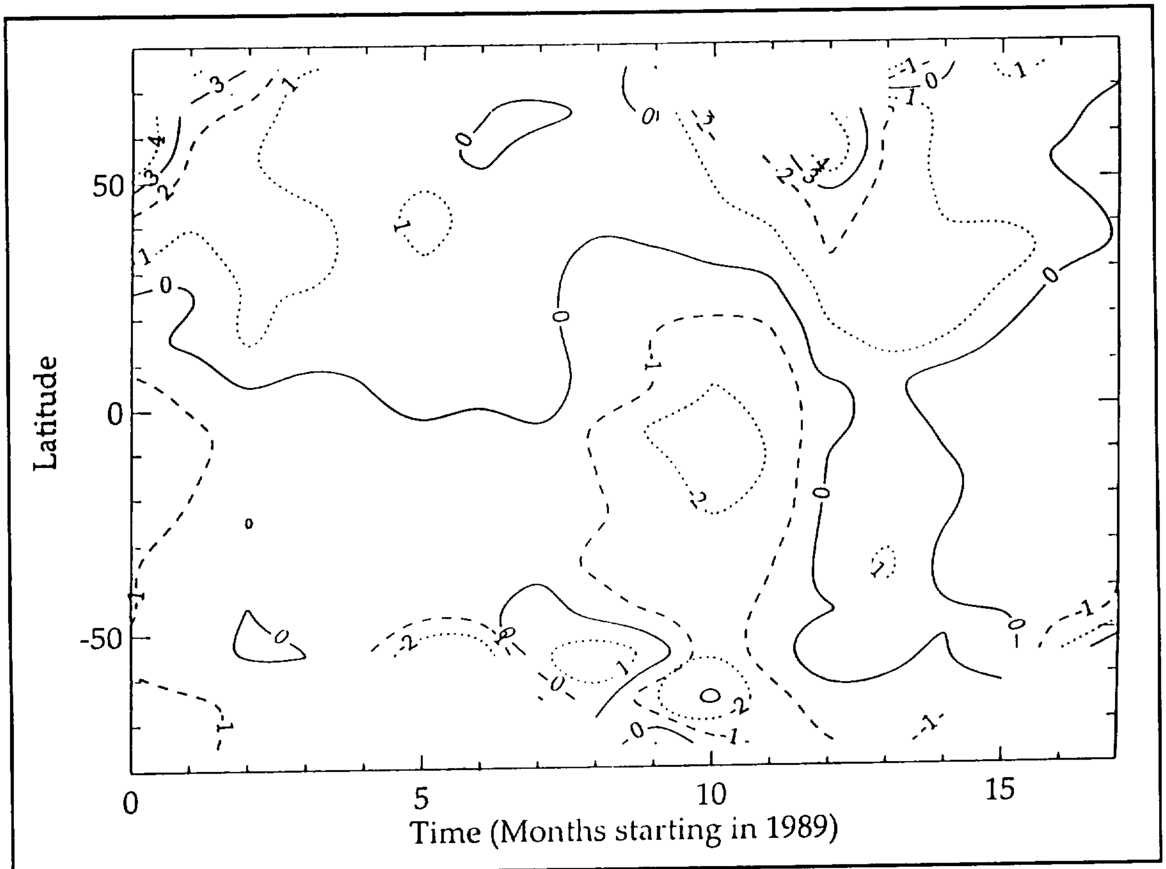

Figure 15. Comparison of SBUV and SBUV/2 total ozone for the period when the two measurements overlap as a function of latitude and time. Contours are ((SBUV - SBUV/2)/SBUV) x 100.

7.3.8 Conclusion

We have discussed some of the tools that the OPT has used over the years to create a trend-quality ozone data set from the BUV instruments. Development of these techniques has required close collaboration between instrument and algorithm scientists and data users. In addition, a long-term project with stable personnel has been critical to the success of these efforts. We have discovered that creating high-quality satellite data sets is a highly iterative process. One starts with an incomplete understanding of how the Earth's atmosphere and surface affect the radiances measured by the satellite instrument. Uncertainties in both prelaunch and postlaunch calibration of the instrument add to this uncertainty. For climate research, where one is dealing with small changes embedded in large short-term variability, these uncertainties make it difficult to separate between "true" change and those that are caused by small systematic errors in calibration and the algorithm. However, as the length of the data record builds and multiple instruments of the same type are launched, identification of various sources of errors becomes increasingly precise.

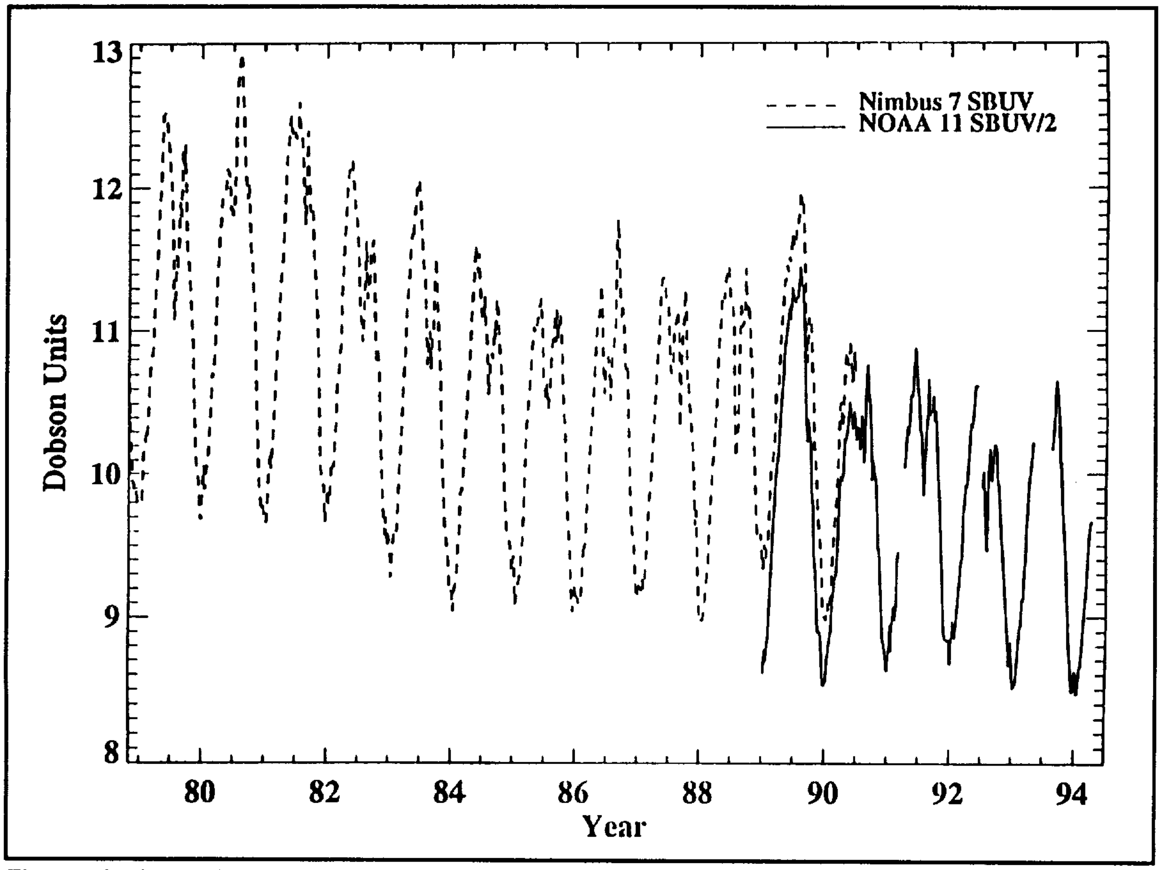

Figure 16. Time series of ozone in a layer 2-4 mb zonally averaged over 30°S to 50°S. Nimbus-7 SBUV data are dashed and NOAA-11 SBUV/2 data are solid lines.

7.4 Calibration and Validation of the ATSR Series of Instruments—A Model for the Calibration and Validation of Global Change Data Sets?

Christopher Mutlow, Space Science Department, Rutherford Appleton Laboratory, Didcot, Oxfordshire, OX11 0QX, United Kingdom

The ATSRs are a series of imaging radiometers that have been expressly design for the purpose of climate change detection; the ATSR-1 instrument is flying on the ESA ERS-1 platform, the ATSR-2 has just been launched on ERS-2, and the Advanced ATSR (AATSR) will fly on ENVISAT-1 later in the decade.

All three instruments feature channels at 1.6, 3.7, 10.8, and 12 μm which, together with the instruments innovative feature of "along track scanning," are designed to provide accurate observations of SSTs for the purposes of climate change/global warming detection. The ATSR-2 and AATSR instruments also include new channels for vegetation monitoring at 0.55, 0.67, and 0.85 μm. Data acquired by these channels will be used to monitor the changes in the distribution of vegetation cover which are expected to be the result of climate change.

The detection of climate change requires a very sensitive observing system capable of detecting small changes that occur over long periods of time against a highly variable background signal. This goal is difficult to achieve, and requires the instrumentation used to be well-calibrated and extremely stable over time if a reliable attribution to climate change/global warming is to be made (Allen et al., 1995a,b). Therefore, as the monitoring of climate change is the major objective of the ATSR instrument program, considerable care has been put into the ATSR instruments to achieve this through the following:

- Careful design of the instruments and their in-flight calibration systems.

- Careful prelaunch characterization and calibration of the instrument by a team independent of the instrument contract and with the involvement of scientists who expect to use the data from the instrument.

- Postlaunch validation of the data using high-quality ground- and aircraft-based instrumentation.

- Cross-calibration of the instruments using the same preflight calibration facility and through in-flight comparisons of the instruments (where possible).

- Through in-flight consistency checks on in-flight calibrators.

The procedures followed are detailed in the following sections.

7.4.1 The ATSR Instrument Program

7.4.1.1 Design of the ATSR Instruments. All of the ATSR instruments have been designed as high-quality radiometers, with short- and long-term stability as a major design goal and with a reliable, accurate, precise and stable on-board calibration system for both the infrared and visible channels. The main design features of the ATSR instrument are:

- Along-track scanning to provide improved atmospheric corrections.

- Low-noise signal channels with the good infrared performance being achieved through the use of a closed cycle cooler to actively cool the detector.

- In-flight calibration of the infrared channels through the use of two blackbody calibration targets whose temperatures span the expected range of sea-surface scenes.

- In-flight calibration of the visible channels using the Sun and a Russian Opal diffuser.

- Careful design of the instrument structure, electronics, and mechanisms to achieve stable thermal and electrical environment.

Full details of the instrument can be found in Edwards et al., 1990. An example of the careful attention that is required if an instrument is to be used for climate change detection is the care taken in the design of the ATSR's on-board calibration systems. Mullard Space Science Laboratory, who designed the system, paid considerable attention to the design, construction, and finish of the blackbodies. The targets were designed to have very high emissivity and good around-orbit temperature stability to minimize the posthoc corrections that need to be made to the data to take into account these imperfections and temperature changes. This has paid off as the black bodies change their temperature by less than 0.5 K around an orbit. Also, considerable effort was expended in ensuring the long-term stability of the temperature monitoring system used in the calibration targets. During the development program, Mullard Space Science Laboratory (MSSL) discovered that after temperature cycling the platinum resistance thermometers (PRTs) used to monitor the ATSR black bodies showed changes in their calibration. Further investigation showed that after repeated temperature cycling, each PRT's characteristics became stable and repeatable. Therefore, each of the PRT's used in the ATSR black bodies was subjected to a preconditioning and characterization process before being used in the flight instrument. MSSL also retains a life-test black body which is being used to monitor the expected in-flight degradation of the ATSR blackbodies.

Similarly, considerable care has also been taken in the design of the visible calibrator for the ATSR-2 and AATSR visible channels. A Russian Opal is used as a diffuser to view the Sun through a piece of high-quality optical glass which protects the Opal from the effects of radiation damage (proton and gamma ray and UV). The glass also has the effect of eliminating the fluorescence from the Opal as it significantly reduces the UV incident on the diffuser. Radiation damage experiments performed on samples of Opal show that the presence of the optical glass between the diffuser and the Sun significantly increase the long-term stability of the system by minimizing the damage over the 5-year expected lifetime of each ATSR instrument.

7.4.1.2 ***Preflight Characterization and Calibration of the ATSR Instruments.*** As well as the considerations given to the design of the ATSR instruments, considerable attention has also been given to the preflight testing of each instrument. To ensure that this characterization and calibration process reveals the true performance of the instrument, both the ATSR-1 and -2 instruments have been calibrated by a team who is largely independent of the team/contractor who built them to ensure probity. The contractor delivers the instrument to the calibration facility after performing a predelivery test to ensure the correct function of the instrument calibration team. Prior to this testing, the instrument subsystems have been calibrated at component level and subsystem level (i.e., measurement of the spectral response of each optical component within the focal plane assembly (FPA) followed by all-up spectral calibration of the FPA).

The ATSR ground test program includes:

- Spectral calibration of the detectors and other spectrally defining components compared with the all-up spectral response of the FPA.

- Static and dynamic field of view tests to characterize the shape, responsivity, and co-alignment of the instrument channels.
  - Stability of field of view and co-alignment at the beginning-of-life and end-of-life instrument temperatures.

- Detailed radiometric calibration of the channels including:
  - A full range of temperatures from LN2 to 320 K.
  - Tests around the full ATSR scan in both the forward and nadir views.
  - Stability of radiometric performance at beginning-of-life and end-of-life instrument temperatures.
  - Stability of radiometric performance during simulated orbital temperature cycles.

To ensure that complete characterization and calibration data were collected, the progress of the test program was reviewed weekly by a calibration review team, which included potential users of the ATSR data set. These reviews took place before moving on to the next test in the sequence to ensure that adequate data had been collected. Furthermore, the characterization and calibration program for ATSR-1 and -2 was performed using the same calibration equipment, and where possible, the same calibration team (it is intended to do the same for AATSR).

It is also interesting to note that in testing the ATSR-1 instrument, its internal calibration sources were so good that we were able to detect temperature errors in the temperature monitors in the external blackbody targets caused by self-heating of the sensors used. For ATSR-2, the sensors in the external targets were changed to properly conditioned rhodium-iron thermometers to get around this problem. The ATSR-2 channels have been calibrated prelaunch using an integrating sphere and calibration lamp.

7.4.1.3 ***In-flight Consistency Checks and Validation.*** The consistency of the ATSR instrument's calibration has been checked several times during the mission, and in particular, after launch and just prior to launch of ATSR-2 on ERS-2.

This is achieved by switching the role of the two blackbodies, the hot one becomes the cold reference and the cold the hot reference. As the two targets are switched, by heating one and turning off the power to the other, the radiances from each target are monitored. At the crossover point, where the two targets are at the same temperature, the same radiance should be observed if the two targets have remained constant (or at least have deteriorated in the same way). So far the checks we have performed show that the two blackbodies are still the same to within 50 mK. This gives us confidence that they have remained stable over time since each have been operating at very different temperatures, the cold one at -10°C and the hot at +30°C, for nearly 4 years.

7.4.1.4 ***Cross-Calibration.*** With ATSR-1 and -2 now being in orbit and operating simultaneously, we have a unique opportunity to intercompare the two instruments for a period of up to 1 year. This will allow us to cross-compare the SSTs from the instruments and understand the reasons for any offsets and biases between the two data sets—while both instruments are still operational. We are very lucky in this regard, most instruments are operating alone either because they are an on/off or because they are the replacement for an operational instrument that has just failed. In this situation, there is likely to be little or no overlap of operations, and sometimes even a fairly long gap in the record of data due to the time taken to launch the replacement instrument. For climate research, it is essential to have this continuity of data sets if we are to achieve the goal of collecting consistent long-term data sets. With ATSR, we have a unique opportunity to collect a data set that can be traced back to the same ground standard sources, over three missions and a 10- to 15-year period (if all goes well). Lets hope our experiences over the next few months show that this case has been worthwhile.

The ATSR instruments are, and will continue to be, cross-compared with data from AVHRR, but the different orbit times makes this more tricky as we need to understand diurnal changes to resolve some of the mismatches. This may be possible by using the geostationary satellites as transfer standards, but this work is ongoing through collaborations with NOAA and other agencies.

7.4.2 Concluding Comments

The ATSR program team has worked hard to develop an instrument together with preflight and in-flight that can collect consistent and well-characterized data sets required for climate change detection. Within the next few months, the work on the recently launched ATSR-2 instrument will show whether the project team has achieved their goal.

However, we believe that some of the strategies we have employed in the development of the ATSR program have wider application to the calibration of other climate change data sets and could be modified to enable more direct intercomparison of data from different instruments. One particular aspect of this is ensuring the traceability in calibration of each instrument to international standards.

7.5 Sensor Requirements for the Radiometric Consistency of Long-Term, Global, Land Data Sets

Philip Slater, Remote Sensing Group, Optical Sciences Center, University of Arizona, Tucson, AZ

Three topics are addressed that relate to the production of consistent, long-term data sets from one, or a series of nominally identical satellite sensors. First, the need is identified for detailed sensitivity analyses to determine how temporal radiometric stability and uncertainty in absolute calibration are related to monitoring changes of geophysical quantities. This will establish the stability and calibration uncertainty required for an individual sensor and also for a sensor series. Second, the importance is discussed of developing consistent performance specifications for a sensor, through an end-to-end system analysis. Third, some examples are provided of the uncertainties with which cross-comparisons can be made between sensors, on the same platform or in the same series. These are shown to be of adequately low uncertainty, at the less-than-1 percent level in most cases, for them to ensure adequate consistency in the production of long-term data sets. However, this conclusion does emphasize the need to reserve a period for nearly simultaneous, overlapping ground coverage for sequential sensors in a series.

7.5.1 Sensitivity Analyses

The increasing number and diversity of Earth observing programs, over the past few years, have stimulated the development and refinement of algorithms for the conversion of remotely sensed data into geophysical products. Among the input data that have to be included in most of these algorithms is the calibration of the sensor. This is because geophysical products are usually related to the spectral radiance or reflectance of the surface, rather than the number of digital counts associated with each sampled area.

Armed with these algorithms, we are now better prepared quantitatively to defend the requirements for sensor calibration. (Until now this has been largely the domain of system engineers, who have usually specified the best they thought achievable.) The task now is for the sensor's science team members to exercise their algorithms to determine how temporal radiometric instability and calibration uncertainty affect the accuracy of monitoring global change and the quality of the geophysical products they plan to provide. Note that stability is often more important in studies of global change than the uncertainty of the absolute calibration. However, the periodic absolute calibration of a sensor provides a means to check the sensors stability and to provide the means for retrieving surface radiance, reflectance, or temperature.

Analyses of this kind, based on the sensitivity of the algorithms used to derive geophysical products, should provide strong scientific justification for specifications of temporal stability and absolute calibration. There is no *a priori* reason to believe that such analyses will dictate more rigorous specifications than are currently required by system engineers. In fact some specifications may be relaxed, which could result in a considerable savings in cost: changes to the specifications of temporal stability or calibration uncertainty of a factor of two may correspond to a cost saving of a factor of ten!

7.5.2 Consistent Specifications for Sensor Performance

The specification of the performance of a sampled-image, multispectral sensor is a difficult undertaking. The need for accurate absolute calibration is determined by the requirements to retrieve spectral reflectances and temperatures and to record changes in these quantities unambiguously. There are several conditions under which calibration accuracy can be defined, unfortunately some of these are generally not well understood or appreciated. These conditions are as follows:

- Preflight, when the sensor views a uniform-radiance field that usually just fills its aperture and field of view.
- In-flight when the aperture and field of view are uniformly irradiated over part of their areas.

The results from such calibrations are then used to retrieve reflectances and temperatures under the following conditions:

- An inhomogeneous target, undersampled by the sensor, the recorded image containing radiometric artifacts due to Modulation Transfer Function (MTF), aliasing, stray light, out-of-band response, etc. all of which are present in different proportions to the preflight or in-flight calibration conditions.

- As in previous bullet, but with the image resampled to conform to a given map projection or simply to register to other multispectral bands acquired nearly simultaneously by the sensor.

It is clear that, in general, the radiometric uncertainties associated with 1 and 2 above bear little relation to those under the operational conditions associated with 3 and 4. The only way to minimize the differences is to pay extraordinary care to specifying appropriate sensor performance and image handling methods. It is futile to require an accurate preflight laboratory calibration if the sensor, when in operation on orbit, introduces radiometric uncertainties that are greatly in excess of the preflight calibration values. However, this mistake is made repeatedly: high MTFs are specified at Nyquist that introduce radiometric artifacts, narrow filters are specified that cannot be adequately blocked for out-of-band response, stray-light specifications are written that prohibit accurate radiometry in the vicinity of contrasty images of interest, and the image has to be resampled to be useful. In such cases, the overall radiometric uncertainty under operational conditions can be many times that of the standard calibration procedure, making the effort and money devoted to high accuracy calibration a waste of time and resources.

This paper stresses four points related to the production of long-term, global, land data sets:

- High temporal stability, the sine qua non for consistent data sets and highly accurate absolute radiometric calibration.

- Realistic specifications for sensor performance, particularly absolute calibration, because of the time and cost involved in achieving accurate calibration.

- Consistency between the requirement for absolute calibration, as obtained conventionally by the sensor's response to uniform radiance fields, and the sensor s characteristics such as signal-to-noise ratio (SNR) and digitization.

- Consistency between the requirement for absolute calibration and the radiometric errors introduced in the operational image-recording mode of the sensor, such as aliasing, stray light, and spectral out-of-band rejection.

An example of the second point is that of the spectral filter out-of-band specification for a sensor presently under construction. One specification calls for <5 percent over the range 0.2 to 20 μm, the other calls for the out-of-band response to be <0.01 percent of the peak transmittance. The <5 percent specification is inconsistent with the radiometric stability required of the sensor to meet the 3 percent and 1 percent calibration accuracy requirements in the solar-reflective and thermal-infrared ranges, respectively. However, no detector is available to cover this range for a scanning system. To demand the design of filters to meet this irrelevant specification is unnecessarily exacting and costly. The <0.01 percent specification is unnecessarily restrictive and, therefore, inconsistent with the radiometric stability needed to support the required calibration accuracy. This specification will have a substantial impact on the cost of the filters for the silicon-detector range.

The third point concerns the consistency of specifications like stray light and out-of-band response with absolute calibration accuracy. For an imaging sensor whose IFOV has been selected spatially to resolve targets of special interest with an absolute radiometric accuracy of 3 percent or 1 percent, according to spectral interval, it is inappropriate to specify a radiometric error of 6 percent for a pixel at a radius of 4.5 pixels from a high radiance area. (This specification was also associated with the sensor mentioned above.)

More subtle problems relate to radiometric artifacts introduced by undersampled multispectral imagers and digital-image resampling procedures. As mentioned earlier, these problems are not accounted for when radiometric accuracies are being specified, just as they are ignored later when the images from the sensors are analyzed. This shortcoming in our understanding and appreciation of the radiometric errors introduced by sampled imaging systems is in urgent need of correction. At present we are probably significantly over specifying absolute calibration accuracy, and spending a great deal of time and money for the privilege.

It is strongly recommended that a detailed study be conducted to review the calibration requirements for future sensors as complete systems: from the target, through the atmosphere, to the image plane for both small and large targets, and then through sampling, registration, and resampling to map projections. The results of such a study should define sensor performance specifications based on the size of the smallest target that needs to be sensed with the required radiometric accuracy. A result of such a study might show that, if the requirement is for the sensor to obtain 3 percent radiometry over a 1-km-square area, then the IFOV should be about 200-m square and the image should then be resampled at 1-km spatial resolution. Such a condition is not generally well-appreciated in the scientific and engineering communities.

7.5.3 In-Flight Cross Comparisons of Sensor Calibrations

The in-flight cross comparison of sensor calibrations is potentially of enormous benefit to those trying to analyze the data quantitatively from more than one sensor on a platform or from several sensors in a series. For a uniform site, a preliminary study indicates that satisfactory accuracies can be obtained both between sensors on the same platform and for consecutive sensors in a series.

Cross comparisons were simulated in the solar-reflective range between MODIS, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), the Multiangle Imaging Spectroradiometer (MISR), and Landsat using 6°S. Midlatitude summer and winter atmospheres with visibilities of 5 km and 100 km were used with a solar zenith angle of 45° and a nadir view angle. An approximation to the spectral reflectance of Lunar Lake and Railroad playa, both in Nevada, was assumed. An elevation of 6,000 ft. was used and a homogeneous target assumed.

With overlapping bands for wavelengths greater than about 500 nm, comparisons between bands from the above sensors were stable to within about 1 percent for all four atmospheric conditions. If verified, these results would aid in reconciling any differences noted between the in-flight calibration of the above sensors and thus improve the consistency of similar data products from these sensors. An accurate comparison can be achieved between sequential, nominally identical sensors in a series, provided they can be maneuvered to be within about 10 minutes of each other in the same inclination orbit for a few weeks. Over extended wavelength ranges, comparisons between different MODIS bands, not in absorption regions, showed a similar 1 percent stability. This could prove to be of value as a simple way to check relative band-to-band calibration in flight—of critical importance for data products produced by ratioing methods.

Considerable work needs to be done to verify these early results and estimate the range of conditions under which they are valid. In particular, the above results are very sensitive to site nonuniformity, and the search for uniform sites is therefore very important. Note that the uniformity of a site should be over a 3 x 3 pixel area for comparing different sensors, because of the registration problem. For comparing bands within the same sensor, the area should be consistent with the band-to-band registration, generally a 1.5 x 1.5 pixel area should suffice.

7.5.4 Concluding Remarks

This paper emphasized the importance of four points for the efficient production of long-term, global data sets:

- High temporal stability, the sine qua non for consistent data sets and highly accurate absolute radiometric calibration.
- Realistic specifications for sensor performance, particularly absolute calibration, because of the time and cost involved in achieving accurate calibration.
- Consistency between the requirement for absolute calibration, as obtained conventionally by the sensor's response to uniform radiance fields, and the sensor s characteristics such as SNR and digitization.
- Consistency between the requirement for absolute calibration and the radiometric errors introduced in the operational image-recording mode of the sensor, such as aliasing, stray light and spectral out-of-band rejection.

In addition, results of a preliminary study were reported, which showed that in-flight cross comparisons can be conducted with an accuracy approaching 1 percent. The limiting constraint is the uniformity of the target site.

It is recommended that further work be conducted in these topics to improve the consistency and accuracy of future long-term, global data sets.

7.6 Perspectives on Proper and Useful Calibration/Validation in Terrestrial Remote Sensing

Philip Teillet, Canada Centre for Remote Sensing, Ottawa, Ontario, Canada

Interest in the calibration/validation aspects of terrestrial remote sensing has been on the rise and significant resources are being devoted to relevant areas of research and development. However, while capabilities in calibration/validation have continued to improve, sensor calibration remains a difficult field and data product validation has received relatively less attention. Thus, the challenge to provide operational data products with proper calibration/validation has only partially been met. Rather than consider the specifics of any given method, this paper looks at calibration/validation from a larger perspective and addresses some of the technical issues that have more to do with infrastructure and information aspects. The discussion is divided into three topics: calibration/validation activity, data and processing, and users and products. In each case, several issues are highlighted to emphasize areas that warrant more attention or to document some of the lessons learned in calibration/validation.

7.6.1 Introduction

The international Committee on Earth Observation Satellites (CEOS) Working Group on Calibration/Validation (WGCV) defines calibration and validation as follows: calibration is the process of quantitatively defining the system response to known, controlled signal inputs; validation is the process of assessing by independent means the quality of the data products derived from the system outputs. These definitions refer specifically to sensor calibration and data product validation. However, calibration/validation, or calibration/validation for simplicity, as a combined expression has also become synonymous in the context of remote sensing with the suite of processing algorithms that convert raw data into accurate and useful geophysical quantities that are verified to be self-consistent. Thus, any radiometric and spectrometric data processing prior to analysis tends to be lumped into the calibration/validation category. Although it need not be so, geometric resampling and remapping are usually considered to be in a separate processing category, even though there is a fundamentally direct interplay between radiometry and geometry.

For the purposes of this paper, calibration and validation are considered to be sensor calibration and data product validation, as defined by the CEOS WGCV. However, most of the discussion dwells on larger issues related to making remotely sensed data accurate and useful in the more general sense of the calibration/validation expression. Thus, in the process of going from raw sensor data to calibrated sensor data, and then to validated geophysical data and biogeophysical information (as shown in Figure 17), there are elements of calibration/validation activity (Table 5), data and processing (Table 6), and users and products (Table 7). In each case, there is a whole series of elements and activities that are almost disciplines unto themselves. The lists in Tables 5 through 7 are not intended to be comprehensive, but rather to convey the complexity of calibration/validation and the significant level of effort required. The objective of this paper is to highlight some of the elements and activities in these lists that warrant more attention than they have received to date. The paper is somewhat disjointed as a result, but it is written in the spirit of documenting some of the issues and lessons learned in calibration/validation.

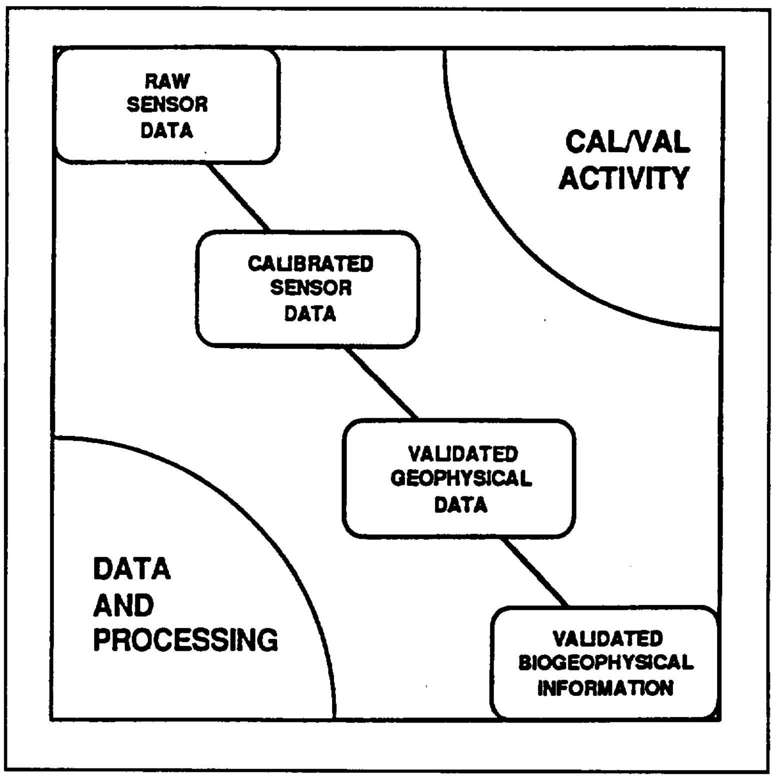

Figure 17. Conceptual data and information flow related to the calibration/validation of Earth observation data. In all cases, it is to be understood that there are users and products.

7.6.2 Calibration/Validation Activity

Most of the items listed in Table 5 are self-explanatory and the overall list represents infrastructures of resources and activities that are far from trivial. For the development and use of any given instrument, the challenge is to make optimum use of these calibration/validation infrastructures and ensure access to useful results. Three areas from the list in Table 4 are discussed in the following sections.

Table 5. List of Some of the Principal Elements of Calibration/Validation Activity in the Remote Sensing Context.

Calibration/Validation Activity
• Calibration standards and laboratories • Standard measurement protocols • End-to-end calibration/validation plans and error analyses • Prelaunch measurement intercomparisons • Onboard calibration systems • Postlaunch vicarious techniques • Integrated multiple methodologies • Test and confidence sites • Coordinated field campaigns • Synergy from crossover data of test sites • International cooperation • Access to data and information

7.6.2.1 Integrated Multiple Methodologies. The need for several independent calibration methods has been documented by Slater (1988). When there are multiple approaches to sensor calibration, there also needs to be a plan for the use of these techniques over the lifetime of the sensor (Guenther, personal communication), and an algorithm for the weighted integration of the results from the various techniques into a single set or sequence of calibration coefficients for operational use (Slater, personal communication). There has been very little experience with this type of integration process. For postlaunch calibration of the NOAA AVHRR, the most common approach has been to give equal weight to different calibration results (Che and Price, 1992) or to use results from one or two methods only (Brest and Rossow, 1992; Teillet and Holben, 1994; Cihlar and Teillet, 1995) for consistency in the absence of a detailed evaluation of the various approaches. Postlaunch radiometric calibration of the SPOT HRV, instruments is based on a weighted blend of relative and absolute methods (Gellman et al., 1993) and is deemed operational, although the details of the recipe are not known.

7.6.2.2 Synergy From Cross-Over Data of Test Sites. In the case of vicarious calibration of optical sensors based on ground-level test sites and measurements, the reliability of calibration methods depends on the accuracy and precision of the field instrumentation and atmospheric models used. The ground target must also be well- characterized in spatial, radiometric, spectral, and temporal domains. However, the procedures for acquiring the necessary field data are usually constrained and suboptimum because of the difficulty in rapidly characterizing large surface areas and the costs involved in a systematic program of field deployment.

Because of the importance of calibration test sites, any additional information on their characteristics is worth consideration, including remotely sensed data acquired at other wavelengths such as in the thermal infrared or in the microwave portions of the electromagnetic spectrum. For example, multitemporal ERS-1 SAR data have been obtained for three optical calibration sites by Teillet et al. (1995a). The three sites are White Sands, New Mexico, and the Lunar Lake and Railroad Valley playas in Nevada. The study reports on an initial examination of multitemporal SAR image data sets generated for the three test sites and focuses on the significant pattern changes observed in the scenes, largely due to surface roughness, soil moisture, and run-off. Such cross-over data sets can contribute to a baseline understanding of ground targets and provide insight into the usefulness of such targets for in-flight calibration of optical sensors.

7.6.2.3 International Cooperation. The topic of international cooperation could be discussed at length, but only a couple of examples will be noted here to make the point. With respect to NOAA AVHRR data sets and multitemporal composites, the EROS Data Center and the Canada Centre for Remote Sensing have been able to standardize their radiometric processing streams to generate data sets for the North American Continent (Morasse and DIorio, 1992). Similarly, processing standards were established by collaborators of the International Geosphere Biosphere Programme to help spearhead the realization of the global 1-km AVHRR data sets (Townshend et al., 1994). The CEOS WGCV is in a position to play a useful role in the expansion of such international cooperation.

7.6.3 Data and Processing

The role of data characteristics and processing algorithms in the calibration/validation of remotely sensed data has probably been underestimated. There are numerous facets that influence the quality and usefulness of information derived from Earth observation imagery (Table 6). Some of these are mentioned briefly in the sections that follow.

Table 6. List of Some of the Principal Elements of Data and Processing Related to the Calibration/Validation of Remotely Sensed Data.

DATA AND PROCESSING
• Sensor characteristics and updates • Ancillary data • Spectroradiometric correction algorithms • Resampling and remapping algorithms • Proper data integration and continuity • Biogeophysical models • Data representation and storage impacts • Algorithm benchmark data sets • Access to data and information

7.6.3.1 ***Sensor Characteristics and Updates***. The characteristics of sensors change with time after launch and, therefore, it is important to monitor these changes and provide updated information to data processors and users, in addition to building as stable an instrument as possible in the first place. Monitoring, characterizing, and documenting responsivity changes in the shortwave channels of the NOAA AVHRRs has proven to be a difficult technical and organizational challenge for the relevant scientific community (Teillet et al., 1990; Teillet and Holben, 1994; Rao and Chen, 1994; Cihlar and Teillet, 1995). The study of postlaunch spectral band changes and their impacts has received relatively little attention to date (Teillet, 1990; Suits et al., 1988; Flittner and Slater, 1991).

7.6.3.2 ***Ancillary Data***. The availability and use of key ancillary data also play a critical role, particularly in surface reflectance algorithms. For example, the lack of readily available atmospheric aerosol and water vapor information has necessitated the use of various approximations in the atmospheric correction of Landsat TM and AVHRR imagery (Teillet and Fedosejevs, 1995; Teillet, 1992), among others. The inclusion of digital terrain elevation data in radiometric processing has also been relatively slow to develop (Teillet and Staenz, 1992; Running et al., 1994).

7.6.3.3 ***Data Representation and Storage Impacts***. There is considerable room for improvement in the careful choice and clear documentation of data products in terms of their representation. A case in point is the widely used NDVI. NDVI can depend significantly on whether it is a function of digital counts, or radiance, reflectance, and on whether it is defined as a function of the data themselves or their scaled representation used for storage and manipulation on computer systems (Teillet et al., 1995b). The study by Teillet et al. (1995b) examined the issues for several vegetation indices based on spectral bands for SPOT HRV, Landsat TM, NOAA AVHRR, EOS, MODIS, and Envisat Medium Resolution Imaging Spectrometer (MERIS).

7.6.3.4 ***Proper Data Integration and Continuity***. Algorithm research and development are needed to ensure the proper and accurate integration of multisource remotely sensed and other types of data from radiometric, spectral, and spatial perspectives. Remote sensing instruments acquire imagery in very specific modes and geometries that have direct impact on the radiometric character and information content of geophysical parameters derived from Earth observation data.

A practical example of data continuity is that of vegetation monitoring by means of NDVI. Use of NDVI is likely to span the lifetime of multiple sensors of a given type and also encompasses several different sensor types. Thus, one faces the important and difficult task of ensuring that the same vegetation information, in terms of NDVI in this case, can be computed from all of these sensor systems. Study of the impact of radiometric, spectral, and spatial sensor characteristics on NDVI has only begun recently (Qi et al., 1994; Teillet et al., 1994, 1995b; Guyot and Gu, 1994).

7.6.4 Users and Products

Because calibration and validation are very much science- and technology-oriented, there is often a tendency to lose sight of the users and products that exist to varying degrees at each step of the process. This applies to both calibration/validation information itself and remotely sensed data and information products that have been calibrated and validated. Table 7 lists some of the principal elements in this regard and a few of these points are discussed briefly in the following sections.

Table 7. A List of Some of the Principal Elements Concerned with Users and Products Related to the Calibration/Validation of Remotely Sensed Data.

USERS AND PRODUCTS
• User requirements • Data and information products • Accuracy assessment products • Figure-of-merit products • Sample data sets • Metadata • Access to data and information

7.6.4.1 Accuracy Assessment and Figure-of-Merit Products. Consistent with the data integration and continuity issues is the concept that users should be provided with indicators of product reliability. Accuracy is not uniformly distributed in remote sensing products. An example is the variation in spatial resolution from pixel to pixel in a maximum NDVI multitemporal composite. There are reliability and responsibility issues that increasingly arise as Earth observation data become an integral part of decision making information systems.

7.6.4.2 Access to Data and Information. A much greater effort needs to be devoted to timely and ready access to documentation, data, metadata, and information on sensor characteristics and all relevant related calibration/validation. While considerable resources are finally being dedicated to calibration, operational users are not yet benefitting accordingly.

7.6.5 Concluding Remarks

From the point of view of calibration/validation, it may be said that the data flow in remote sensing progresses from raw sensor data to calibrated sensor data, and then to validated geophysical data from which validated biogeophysical information is derived. This process encompasses the expertise, facilities and activities of many institutions, from the instrument and data providers along the chain to the end users. The process also involves many technical considerations that need to be addressed if self-consistent and useful data and information are to be obtained. This paper documents a few of these technical considerations from the perspectives of calibration/validation activity, data and processing, and users and products. The most common thread is a need for more timely and ready access to calibration/validation data and information.

7.7 The NOAA/NASA Advanced Very High Resolution Radiometer Calibration Activity

C.R. Nagaraja Rao, Satellite Research Laboratory, NOAA/NESDIS Office of Research and Applications, Washington, DC 20233

The NOAA/NASA Pathfinder program has for one of its main objectives the reprocessing of long-term records of geophysical products generated with sensors on board NOAA's Polar-orbiting Operational Environmental Satellites (POES) and the Geostationary Operational Environmental Satellites (GOES) (Ohring and Dodge 1992). The aim is to provide researchers with data sets of acceptable continuity and high quality for the Pathfinder period (1981 to present) which can be used in climate and global change research. Some of the data sets initially identified as suitable candidates for the Pathfinder activity were those obtained with the Advanced Very High Resolution Radiometer (AVHRR) and TIROS Operational Vertical Sounder (TOVS) suite of instruments on board the polar orbiting satellites.

At the time of inception of the Pathfinder program in the early 1990's, the importance of the proper assessment of the calibration of the various satellite-borne sensors, and of their postlaunch performance was recognized and emphasized to ensure the continuity and quality of the geophysical products that would be generated. In addition, it was also realized that establishment of calibration linkages among similar sensors—such as the AVHRR on different NOAA spacecraft—or among different sensors used to generate similar geophysical products was crucial to render the quality and useability of the Pathfinder data sets optimum. These sensor-related requirements in the context of establishing long-term climate-related Pathfinder data sets are very close to the basic requirements (e.g., Becker et al., 1988) of the design and execution of programs such as the Global Climate Observing System (GCOS) in general. In addition, the activities of the AVHRR Pathfinder Calibration Working Group, to be described later, will be seen as highly germane to the objectives of the calibration/validation workshop jointly sponsored by the Calibration/Validation working group of the Committee on Earth Observation Satellites (CEOS) and GCOS.

7.7.1 The AVHRR Pathfinder Calibration Working Group

The AVHRR Pathfinder Calibration Working Group was formed in the spring of 1991, with its membership made up of representatives of the AVHRR Pathfinder Atmosphere, Land, and Ocean Science Working Groups, and other experts in the area of satellite sensor calibration to address and resolve sensor-related issues mentioned in the previous section. The working group decided to assess initially the postlaunch performance of the AVHRRs on NOAA-7, -9, and -11 spacecraft—the afternoon satellites—for programmatic reasons, and identified the following as the principal elements of its charter:

- Evaluation of the in-orbit degradation of the visible(channel 1: ≈0.58-0.68μm) and near-infrared (channel 2; ≈0.72-1.1μm) channels of the AVHRR;

- Establishment of calibration linkages for channels 1 and 2 of the AVHRRs on board the NOAA-7, -9, and -11 spacecraft;

- Development of user-friendly procedures to correct for the nonlinearities in the response of the Mercury-Cadmium-Telluride (Hg-Cd-Te) sensors in the AVHRR thermal infrared channels (channel 4: ≈10.3-11.3μm; channel 5: ≈11.5-12.5μm); and

- Dissemination of calibration information among the user community in an effective, timely manner.

7.7.2 Visible and Near-Infrared Channels of the AVHRR

The AVHRR visible (channel 1) and near-infrared (channel 2) channels degrade in orbit, initially because of the outgassing from the radiometer components, and rocket exhaust, and subsequently because of continuous exposure to the harsh space environment. Any assessment of the in-orbit degradation is rendered difficult by the absence of on board calibration devices. Therefore, vicarious calibrations based on: (a) the use of terrestrial targets which can be considered radiometrically stable in time, (b) model simulations of the upwelling radiance at the top of the atmosphere, with measured inputs for any, some, or all of the model parameters, or (c) matching the performance of the sensor under study with the performance of a different sensor whose calibration is known are generally resorted to. In the present study, the calibration working group decided to assess the relative in-orbit degradation of the three AVHRRs using the southeastern part of the Libyan desert (21°N to 23°N; 28°E to 29°E) as a radiometrically stable target; further, the AVHRRs on NOAA-7 and -11 spacecraft would be normalized to the AVHRR on NOAA-9, after the relative degradation of the two channels of the latter (NOAA-9 AVHRR) had been rendered absolute by anchoring the same to the absolute calibrations based on congruent path aircraft/satellite radiance measurements made over White Sands, New Mexico during September/October 1986. Intersatellite calibration linkages thus established would be used in the reprocessing of the AVHRR atmosphere, land, and ocean data sets. Details of the ISCCP Stage B3 data used in the evaluation of the in-orbit degradation of the two channels, the quality control criteria adopted, the assumptions made regarding the sensor degradation and the reflectional properties of the Libyan desert calibration target, and of the methodology of deriving the in-orbit degradation and establishing intersatellite calibration linkages are found in Rao et al. (1993) and in Rao and Chen (1995).

We have shown in Table 8 the formulae for the calculation of the radiances measured in AVHRR channels 1 and 2 after the sensor degradation has been duly accounted for; here, C_{10} is the AVHRR signal in 10-bit counts, and d is elapsed time in days since launch. We have attempted validation of these formulae by comparing the predicted trend in "slope" expressed in units of W/(m^2 sr μm count)—the entire expression appearing in front of (C_{10} - number) in the formulae—for the AVHRR on NOAA-11 with independent, but rather infrequent, determinations of the same from congruent path aircraft/satellite radiance measurements (Abel et al., 1993); and by demonstrating that application of these formulae to the reprocessing of reflectance measurements made over different Earth scenes results in the elimination of spurious trends in the records of geophysical products obtained over the operational life of the AVHRR, and of the discontinuities in the record during the transition from one satellite to the next.

Table 8. Formulae for the Calculation of Calibrated Radiances

NOAA-7 was launched on June 23, 1981; NOAA-9 on December 12, 1984; and NOAA-11 on September 24, 1988. The two sets of formulae for NOAA-9 yield the same radiances.

Spacecraft	Radiance (W/m^{2}srμm)
NOAA-7	
Channel 1	$0.5753 \times \exp(1.01 \times 10^{-4} \times d) \times C_{10} - 36)$
Channel 2	$0.3914 \times \exp(1.20 \times 10^{-4} \times d) \times C_{10} - 37)$
NOAA-9 (Set A)	
Channel 1	$0.5465 \times \exp[1.66 \times 10^{-4} \times (d - 65)] \times (C_{10} - 37)$
Channel 2	$0.3832 \times \exp[0.98 \times 10^{-4} \times (d - 65)] \times (C_{10} - 39.6)$
NOAA-9 (Set B)	
Channel 1	$0.5406 \times \exp(1.66 \times 10^{-4} \times d) \times (C_{10} - 37)$
Channel 2	$0.3808 \times \exp(0.98 \times 10^{-4} \times d) \times (C_{10} - 39.6)$
NOAA-11	
Channel 1	$0.5496 \times \exp(0.33 \times 10^{-4} \times d) \times (C_{10} - 40)$
Channel 2	$0.3680 \times \exp(0.55 \times 10^{-4} \times d) \times (C_{10} - 40)$

7.7.3 The Thermal Infrared Channels

The nonlinearity corrections in the thermal infrared channels of the AVHRR are typically in the range from about -1°C to +2°C, depending on the Earth scene temperature. The AVHRR Calibration Working Group focussed its attention on the development of a simple analytic expression for the sensor nonlinearities over the entire range of Earth scene temperatures encountered in orbit. The resulting formula for the calculation of the corrected radiances measured in either channel 4 or 5 is based on the "nonzero radiance of space" concept. Thus, the measured radiance [mW/(m^2 sr cm^{-1})], which has been corrected for sensor nonlinearity, is given by $N = a + b(N_{LIN}) + c(N_{LIN})^2$ where N_{LIN} is a pseudolinear radiance calculated using the concept of nonzero radiance of space, and the instrument responses when it views the internal calibration target and space. The corrected radiance is then used to calculate the brightness temperature. The coefficients a, b, and c, and the method of calculating N_{LIN} from the sensor response when it views space and the on-board internal calibration target are furnished to the user for the different AVHRRs. The novel feature of the above formula is that it is independent of the operating temperature of the AVHRR—a feature that is absent in earlier formulations of the nonlinear corrections, either in the radiance or temperature representations (e.g., Weinreb et al., 1990). Details of the method, including the rationale behind the concept of "nonzero radiance of space" are found in Rao et al. (1993).

7.7.4 Applications/Validation

We used only about 15 percent of the data for the Libyan desert calibration target in the establishment of intersatellite calibration linkages for the AVHRR channels 1 and 2; thus we felt it may be scientifically meaningful to show here the results of applying the radiance formulae shown in Table 8 to the entire set of data. The results are seen in Figure 18 where we have shown the time series of surface albedo spanning the period 1981 to 1991 obtained with the prelaunch and postlaunch calibrations. The spurious downward trends and the discontinuities in the time series when the prelaunch calibrations, which do not account for sensor degradation, are used and removed with the application of the postlaunch calibration. Similar improvements in the time series of surface albedos, and of the NDVI have been noticed for different locations over the globe by different investigators. In particular, the radiance formulae have found application in: (a) the reprocessing of the AVHRR Pathfinder land data set for the period 1981 to present; (b) NOAA's Clouds from AVHRR (CLAVR) studies; (c) the validation of NOAA's aerosol product over the oceans; and (d) the establishment of correspondence between the surface reflectance of several sites in Spain measured by the AVHRR on NOAA-11 and channel 3 of the TM on Landsat-5; appropriate literature citations are found in Rao and Chen (1995). The nonlinearity corrections for the thermal infrared channels have been applied in the reprocessing of the AVHRR Pathfinder Ocean data product. The same technique is presently in use in the operational determination of SST with the AVHRR on NOAA-14 spacecraft.

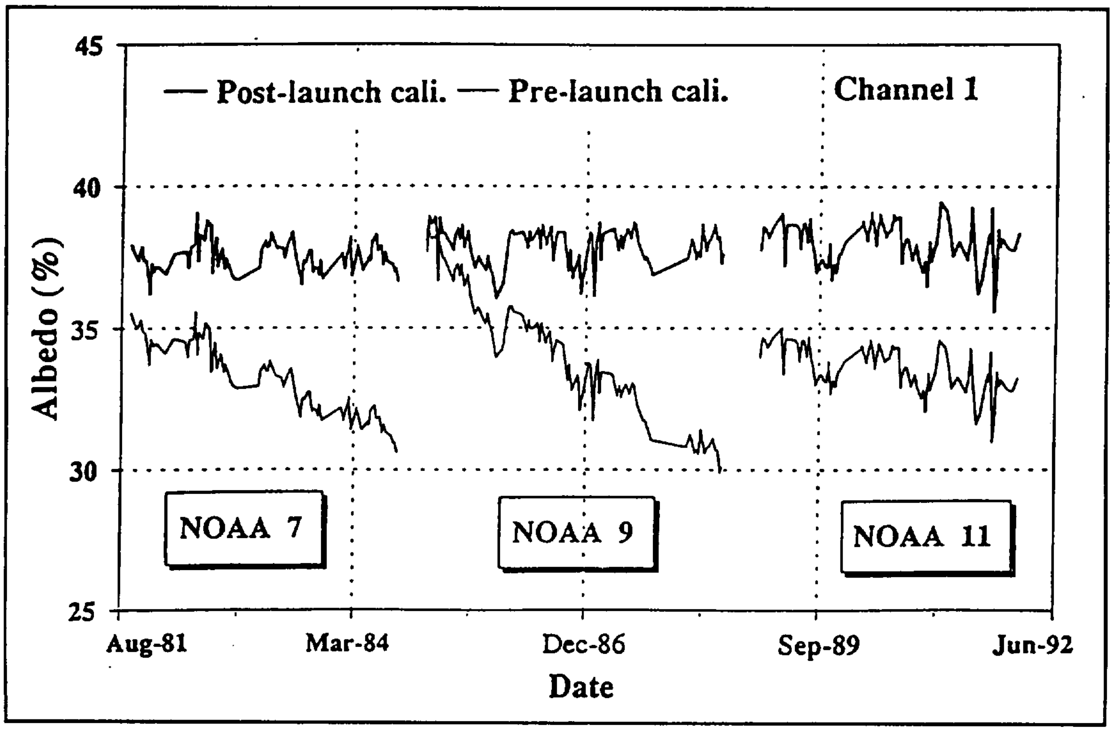

Figure 18. Albedo in AVHRR channel 1 of the southeastern Libyan desert.

7.7.5 Conclusion

Some of the lessons learned during the AVHRR Pathfinder calibration activity may prove to be helpful in addressing calibration-related issues in the future. We feel that it is desirable to undertake a detailed study of the relative merits of the various vicarious techniques that have been proposed for the calibration of AVHRR channels 1 and 2, with particular attention paid to the identification of any feature(s) of the technique that may cause the results to differ from those of others. Techniques that use terrestrial targets which are assumed to be radiometrically stable will benefit if multiple targets, with reflectances which cover a significant part of the dynamic range of the sensor, are identified and used by the calibration community. The sensitivity of postlaunch calibrations based on model simulations to the uncertainties in model inputs such as atmospheric aerosols, ozone, water vapor, and surface reflectional properties should be examined thoroughly. Since absolute calibrations based on congruent path aircraft/satellite radiance measurements are invaluable as anchor points to establish calibration trends, they should be performed more frequently than they have been to date, and a dedicated program should be established for the same with requisite resources on a long-term basis. Comprehensive field campaigns, involving ground-based measurements of model input parameters, congruent path aircraft/satellite radiance measurements, and a suite of similar satellite sensors, or of different sensors directed towards the measurement of the same geophysical parameter should be planned and conducted, with agencies such as CEOS participating in the coordination of the same. The feasibility of rugged on board calibration devices for the visible and near-infrared channels of AVHRR-type sensors should also be investigated. Some of these issues are presently being addressed by the Meteorological Satellite Instrumentation/Calibration Group at NOAA's NESDIS Office of Research and Applications.

7.7.6 Acknowledgment

The work reported here was supported by the Information Management component (Manager: Dr. P. Topoly) of the NOAA Office of Global Programs. Drs. J.T. Sullivan and C.C.Walton, both of the NOAA/NESDIS Satellite Research Laboratory, and J.Chen (QSS., Inc., Washington, D.C.) have made valuable contributions to the work reported here.

7.8 A 16-Year Record of Global Temperatures From Eight Microwave Sounding Units

Roy Spencer, NASA/Marshall Space Flight Center; John R. Christy, The University of Alabama in Huntsville

The TIROS-N MSUs have been imaging deep layer temperatures since 1979. We have intercalibrated a total of eight MSU's flying on successive satellite platforms into a 16-plus-year record of global deep layer temperature variations at various spatial and temporal averaging scales (e.g., Figure 19). While overlaps in operational coverage of the Earth by successively launched satellites permit intercalibration to within about 0.01°C for monthly global temperature anomalies, absolute accuracy of the brightness temperature (T_b) measurements themselves might not be much better than 1°C. Deep layer temperatures are monitored for the lower troposphere (from MSU channel 2 at 53.74 GHz, Spencer and Christy, 1992a,b) and the lower stratosphere (from channel 4 at 57.95 GHz, Spencer and Christy, 1993). Independent validation with 10 years of radiosonde data has also been performed (*ibid*). MSU channel 3 (at 54.96 GHz) cannot be used for long-term trend calculations due to calibration drift in at least two of the MSU's, but can be used for computation of dynamical quantities associated with spatial gradients in the data, and interannual variability for many of the nondrift years. Transition to the AMSU's, scheduled for first launch in 1996, will require the formation of multichannel (and possibly multiview angle) averaging kernels that match those currently utilized by us for the MSU's. The various sources of intersatellite differences in deep layer temperature measurements are discussed in the following sections.

7.8.1 Intercalibration Issues

Intercalibration of the first eight MSU's into a continuous record of temperature variability involves many decisions which must be made. Changes in any of these decisions will alter the final results. Due to ground software changes impacting the NESDIS historical MSU calibrated data record, we perform our own calibration of the MSU data (Spencer et al., 1990). While we would prefer that intercalibration of two successive copies of the same sensor (such as the MSU) be as simple as finding a single additive offset, our analysis of the MSU data suggests that there are several sources of intersatellite differences which must be accounted for. We have found that latitude- and season-dependent offsets alone have been a convenient way to intercalibrate. These offsets are computed empirically, based upon the differences that separate satellites measure when both satellites are operating. The overlap periods during which these offsets are computed have ranged from 3 months to several years (Figure 20). Physical origins of these differences can be deduced with some confidence from their space/time characteristics. They can be categorized generally as diurnal, sensor, and sampling effects.

7.8.2 Diurnal Effects

Even if the MSU's were all identical in their response, the TIROS-N satellites are launched into two very different (from a temperature perspective) Sun-synchronous orbits at 7:30 a.m. and p.m. (morning satellites), and 2:30 a.m. and p.m. (afternoon satellites). Because of NOAA data archival and satellite tracking policies, the MSU archive includes virtually no overlaps between two satellites at the same orbit time, requiring intercalibration between 2:30 and 7:30 satellites. Because of this, significant diurnal temperature effects must be removed from the data to obtain a record of global temperatures suitable for climate studies. The diurnal effects vary with latitude, geography, season, and altitude, and generally range between 0.1°C and 1.0°C for MSU channels 2, 3, and 4.

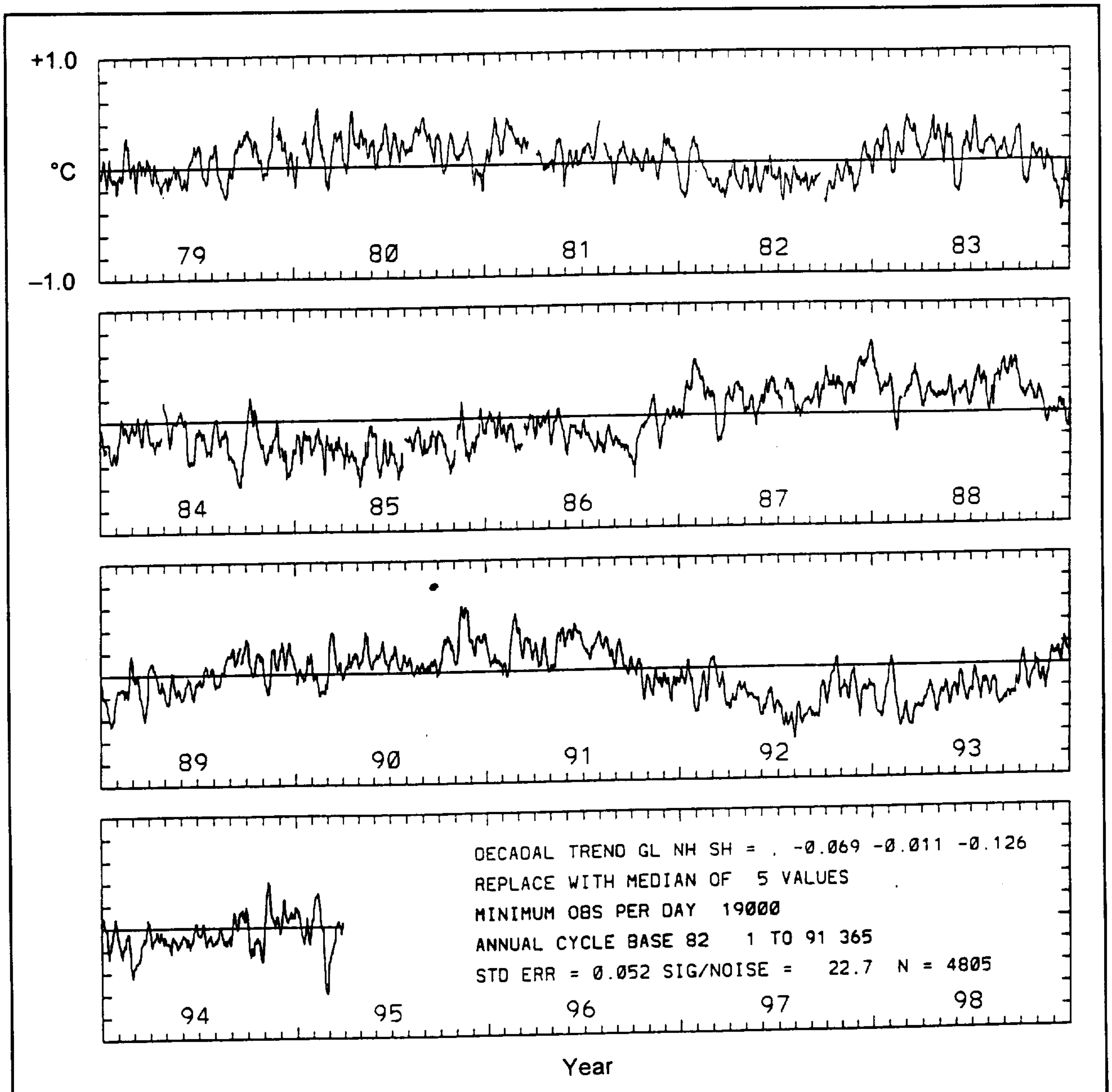

Figure 19. Intercalibrated record of daily global averaged temperature anomalies for the lower troposphere from MSU channel 2R.

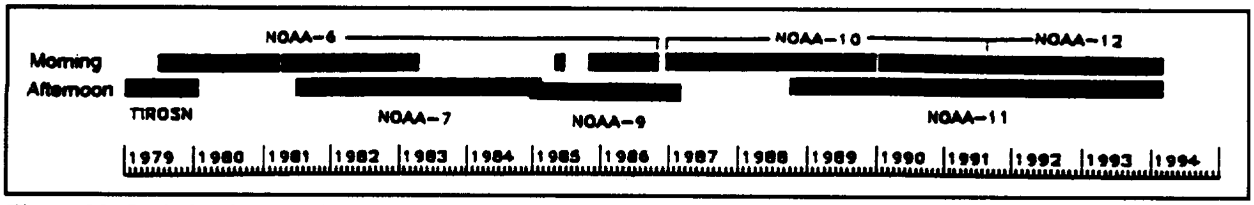

Figure 20. Periods of record for the MSU's between 1979 and 1994. Nominal satellite times are 7:30 a.m. and p.m. for the morning satellites and 2:30 a.m. and p.m. for the afternoon satellites.

In general, we have not tried to separate the sensor differences from the diurnal differences. However, some estimates of the magnitude of the diurnal effects can be made by computing 12-hour differences from the ascending and descending portions of orbits for the same satellite. By examining several 2:30 (or 7:30) satellites from different years, we have determined that these 0.1° to 1.0°C 12-hour differences are repeatable to hundredths of a degree. This gives some confidence in the size of the diurnal effects. The most accurate documentation of the average diurnal effects requires compositing of all 16-plus years of data. Recent (not yet published) computations by Dr. W. D. Braswell suggest that the MSU's channel 2, 3, and 4 diurnal influence over the ocean is approximated quite well by Lindzen's classical tidal theory (Chapman and Lindzen, 1970), which quantifies the zonally symmetric temperature effects of diurnally varying solar heating of ozone and water vapor. Deviations from these traditional (ozone and vapor) tidal effects seem to be dominated by cloud solar absorption effects in the deep convective zones of the tropics. Over land, warmth of the tropical upper troposphere and cooling of the lower stratosphere occurs during daytime convective activity over Amazonia, the Congo Basin, and the Maritime Continent. These areas often experience diurnal (12-hour) differences of 0.5° to 1.0°C. However, the largest diurnal effects occur where the lower tropospheric channel (channel 2) has significant direct influence from the land surface, especially when that surface is subject to high solar inclination combined with low moisture content or high elevation.

The diurnal effects in the differences between the 2:30 and 7:30 satellite data show up as one or more harmonics of the annual cycle when a time series of the T_b difference between the two satellites is plotted (Figure 21). We usually work with daily time series of 2.5° zonal T_b averages. We arbitrarily assume a base period during the NOAA-6 and NOAA-7 overlap period, with all other satellites ultimately being referenced back to NOAA-6. With the addition of each new satellite having at least 1 year of data, any residual harmonics of the annual cycle are removed (without detrending the data) from the time series of daily differences between the new satellite and the previous satellite. In the case of NOAA-11, there was a large orbital drift to late in the afternoon which introduces a spurious trend in its T_b differences with NOAA-10 and NOAA-12. This trend was removed before inclusion of the NOAA-11 data (i.e., the NOAA-11 trend during orbital drift was forced to match the NOAA-10 and NOAA-12 trends). Additional details about the MSU intercalibration can be found in Christy et al. (1995).

Clearly, any temperature monitoring effort for climate purposes must either make all of the observations at the same solar time, year after year, or must account for these large diurnal effects.

7.8.3 Sensor Effects

There are many possible sources of subtle differences between successive copies of the same MSU instrument. While the sensor fabricator might try to make two sensors identical, only lengthy intercomparisons on-orbit can ultimately determine the success of such efforts. After approximate adjustments for diurnal effects, we see MSU's differ in their calibrated response by as much as 1°C, but usually by less than this. Possible sources of these sensor differences include variations in antenna patterns; changes in channel response linearity (which was originally measured at JPL); warm target Platinum Resistance Thermometer (PRT) biases (unlikely in most cases); temperature gradients across the warm calibration targets; differing emissivities of the calibration targets; different amounts of stray radiation received in the sidelobes of the antenna pattern from the spacecraft, to name a few of the more obvious possibilities. There are also T_b asymmetries between the left and right sides of the MSU swath, although we have determined that some portion of this effect is real, due to the difference in local solar times across the 2,000+ km wide swath.

When we intercalibrate the MSU's, these sensor biases are removed together with the diurnal biases. Even though we have not yet completely quantified the diurnal signals which would then allow a sensor intercalibration, preliminary indications are that the effects are comparable in magnitude.

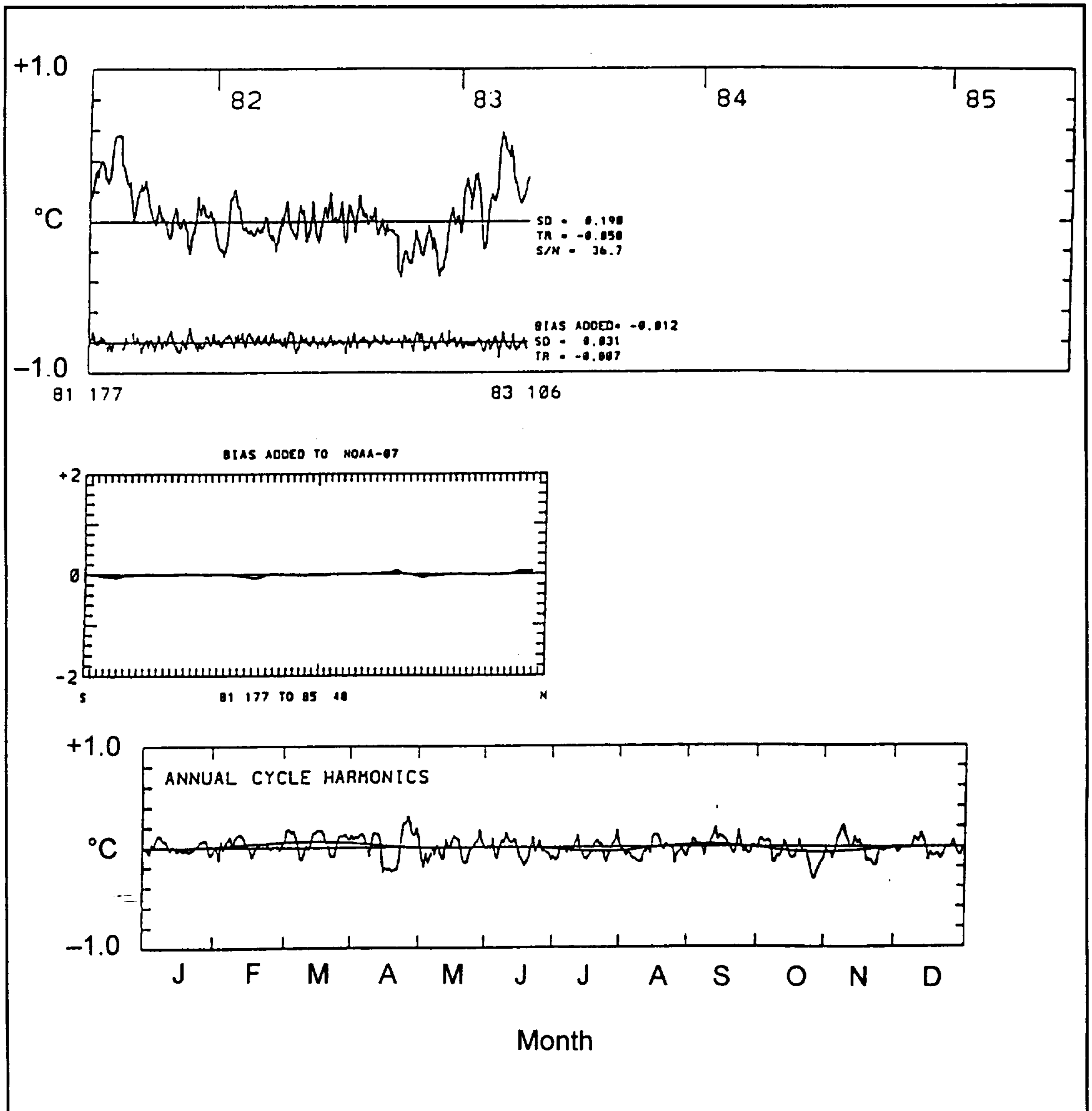

Figure 21. Example intercalibration information for the overlap between NOAA-6 and NOAA-7 MSU's: Daily, globally-averaged sums and differences of channel 2R Tb after final intercalibration (top); the average latitude-dependent biases to force NOAA-7 Tb to match NOAA-6 Tb (middle); and the annual harmonics removed from the satellite difference time series (bottom).

7.8.4 Sampling Effects

When computing intersatellite differences, there are many decisions which must be made. Many of the sampling-related issues deal with the fact that the MSU is a cross-track through-nadir scanner which has limb-darkening effects in its T_b measurements. To provide useful gridpoint temperature products at time resolutions much better than 1 month, this limb-darkening effect must be removed. If zonal averages are required, then one can avoid limb corrections and just average data from the center 50 percent of the swath where the weighting functions do not deviate too much from the nadir weighting function. Still another possibility for larger averaging areas or times is a multiview angle retrieval (Spencer and Christy, 1992b, called channel "2R") in which we exploit the closely spaced weighting functions from the different scan angles within a single scan to deconvolve to a much sharper averaging kernel peaking in the lower troposphere. Because this retrieval destroys any horizontal temperature gradient information across the swath, this method is only appropriate for long averaging times and/or large averaging space scales.

All sampling decisions impact the intercalibration between satellites. For instance, if one wants to intercalibrate just MSU channel 2 between two satellites, several decisions must be made. Should the 11 footprints of data be treated separately or together? Should the two sides of the swath be considered symmetric, allowing combination of the same view angles from either side, or should they be treated separately? What should the time averaging scales be: daily, pentad, monthly, seasonal? What should the space averaging scales be: 1°, 2.5°, zonal, hemispheric, global? Should we screen out high elevation or precipitation-contaminated data, and with what thresholds? For satellites whose orbits drifted from the nominal 2:30 and 7:30 observation times, how should these off-time data be handled? While we have answered all of these questions to build the data sets that we provide to the research community, it is possible that there are somewhat better choices for any of our decisions.

The possibility of precipitation-size ice causing low T_b biases deserves special mention. We screen local minima in channel 2 T_b data which are due to large, deep thunderstorm complexes. Experiments were carried out in which successively more of the MSU data were screened out and we examined their effects on computed interannual variability in monthly temperature anomalies. We found that our current threshold might allow small regional biases in interannual variability where there are large interannual changes in deep convection (e.g., with ENSO) but that hemispheric and global averages were essentially unaffected (less than 0.01°C variations).

7.8.5 Radiosonde Validation

Extensive, independent validation with 10 years of U.S. radiosonde data has been carried out (see references). Here, we show a time series of 63-radiosonde station averages (U.S., Caribbean, and West Pacific stations) monthly channel 2R (lower tropospheric retrieval) temperature anomalies for the MSU and radiosonde separately (Figure 22). The bottom time series in Figure 22 is the difference between radiosonde-calculated and MSU-observed channel 2R, illustrating the fact that the MSU decadal trend is about 0.01°C cooler than that computed from the radiosondes. This is a composite result at just the 63 locations of the radiosonde stations, while the global grid includes about 10,000 gridpoints. This result provides considerable confidence in the merging procedures and the validity of the global temperature trends computed form the MSU data.

7.8.6 Cross-Calibration to the MSU

The first AMSU will be launched sometime after late 1996. In general, its channel frequency selection is different from the MSU, although it has considerably more channels. Because these channels' weighting functions overlap heavily, it will be possible to construct nearly the same MSU weighting functions from the AMSU channels using the oxygen absorption theory combined with a radiative transfer model. The disadvantage of this cross-calibration approach, however, is the higher probability that any one of several AMSU channels will fail versus a single MSU channel failing. The alternative would be the use of the AMSU channel which provides the weighting function that most closely approximates the corresponding MSU channel. Unfortunately, this might well introduce different diurnal and seasonal effects that would have to be quantified and corrected for. Obviously, accuracy of the oxygen absorption theory will determine the accuracy of the results of the weighting function-matching portion of the cross-calibration method, and empirical adjustments will likely be required after sufficient amounts of AMSU and MSU data are collected.

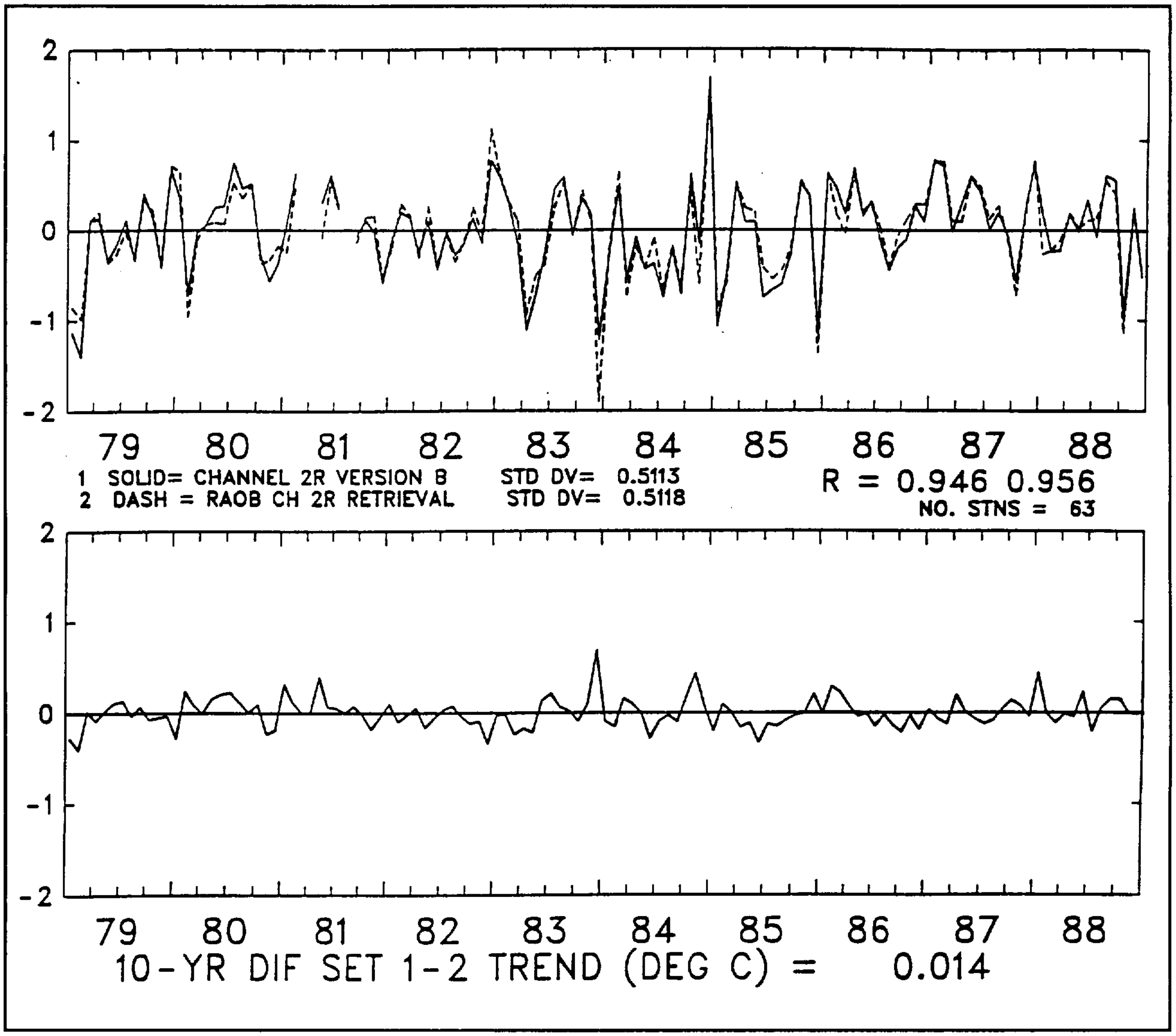

Figure 22. 10-year time series of monthly anomalies in 63-station averaged radiosonde-computed (dashed) and MSU observed (solid channel 2R T_b (top); difference time series (bottom).

7.9 Lessons Learned Calibrating and Validating Nimbus-7 ERB

Dr. H. Lee Kyle, NASA/Goddard Space Flight Center; Dr. Philip Ardanuy, Hughes STX Corporation, 7701 Greenbelt Road, Suite 400, Greenbelt, MD 20770

The Nimbus-7 ERBE instrument, and its calibration and measurement program have been described in scientific journals (Hickey and Karoli, 1974; Jacobowitz et al., 1984; Hoyt et al., 1993; and Kyle et al., 1984, 1985, 1995a,b,c) with some additional details given in reference publications (Soule, 1983; Kyle et al., 1990a, 1993a, 1994a,b). The scientific results, published in numerous articles, are reviewed by Kyle et al. (1993b). This article briefly reviews the major postlaunch calibration challenges and, from them, compiles some lessons learned to guide future experiments. Almost all of these questions are discussed separately in the above-mentioned articles. However, they have never been compiled into a single list.

The Nimbus-7 ERB experiment made useful measurements for 15 years (November 16, 1978 through December 12, 1993), with major data gaps in 1992 and 1993. The ERB instrument had 22 sensors separated into three instrument packages. There were 10 solar sensors to measure the total solar irradiance and its broadband spectral components; four wide-field-of-view (WFOV) sensors measured the emitted and reflected terrestrial irradiances in three spectral ranges (0.2 to >50 μm; 0.2 to 3.8 μm, and 0.7 to 2.8 μm); eight sensors in a scanner measured the reflected and emitted terrestrial radiances in two spectral regions (0.2 to 4.8 μm and 4.5 to 50 μm). The scanner failed after June 20, 1980, while due to budget problems, the WFOV measurements were only calibrated through October 1987 (9 years). The scanner measurements were stable to about 1 percent with no noticeable drift over their 20-month lifetime (Jacobowitz et al., 1984). The WFOV measurements were, in the mean, stable to about 0.5 percent over their 9-year calibrated history. There was no noticeable drift in the calibrated shortwave reflected radiation, but the longwave, emitted radiation had a slow downward drift amounting to a total of 0.5 percent over the 9 years (Kyle et al., 1995c). The total solar irradiance measurements were stable, in the mean, to a few parts in 10,000 over the 15-year measurement period (Hoyt et al., 1992). The referenced articles describe the calibration efforts required to achieve these results. Due to various issues, including cost, the solar spectral measurements have never been adequately calibrated (Kyle et al., 1993a).

7.9.1 The Scanner: Channels 13-22

The scanner was biaxial; each scan pattern combined a forward and back scan with a scan to either the left or right side. This scan pattern allowed a straightforward comparison between the scanner measurements and those of the WFOV sensors. There were four shortwave sensors (13-18) and four longwave sensors (19-22), so grouped as to increase the spatial coverage of each scan. At nadir, the terrestrial footprint of each sensor was a square with a side of about 90 km.

The shortwave scan channels were calibrated in the laboratory by viewing a diffuse target, using several methods (Jacobowitz et al., 1984). The information obtained from these preflight tests was then used for the initial calibration of the in-flight measurements. Additional postlaunch checks on the shortwave scan channels were made by comparing the observed brightness levels of nearly isotropic surfaces (i.e., cloud-free snow) to published ground-based observations, comparing the mean shortwave flux obtained from the scan channel radiances to the flux obtained from the difference between the WFOV total-spectral channel 12 and the integrated longwave scan channels over a 2-week period, and comparison of channel sensitivities obtained from views of an on-board diffuse target illuminated by the Sun with prelaunch values. These postflight calibration checks all indicated an increase in the shortwave channel sensitivity since launch which varied with each sensor (see Jacobowitz et al., 1984, Table 5). Required sensitivity adjustments were made, but the cause of the sensitivity shifts were never determined. Additionally, it was discovered that one of the shortwave scan channels (18) became extremely noisy by the end of 1978 (Jacobowitz et al., 1984). As a result, only the measurements of the first three channels were used in subsequent studies.

The sensitivity and offset for the longwave scanning channels were periodically checked in flight by observing a calibration blackbody and cold space. According to Jacobowitz et al. (1984), the deviations of the in-flight sensitivities and offsets remained constant to within ±1 percent of their initial values. A later review of the longwave scan channel calibration revealed that a slightly incorrect temperature coefficient was used to determine the temperature variations of the longwave sensors. While the determined coefficients of several Platinum Temperature Monitors (PTMs) made at the same time were on record, no record could be found for the ones actually flown. The single coefficient first used belonged to a PTM used in the laboratory but it was not close to the mean of the other recorded values. The mean of the recorded values was, therefore, used to evaluate the temperatures for the final calibration of the longwave flight measurements (Kyle et al., 1985, 1990a).

The purpose of the longwave sensors was to measure the Earth-emitted longwave radiation at wavelengths greater than 4.5 μm. The longwave filter consisted of deposited layers on diamond substrate. The response function of this filter (Figure 23) is far from that and this causes problems in estimating the true total longwave Earth-emitted radiation. One of the ERB team members, Larry Stowe, directed a study which derived a correction curve (Figure 24). Using the response function shown in Figure 23, filtered and unfiltered radiances were computed for 105 atmospheric cases compiled in the Meteorological Satellite Laboratory Report #10 (August 1962). The radiance at five zenith angles (0, 20, 45, 60, 785) is given in 77 spectral intervals between 0 cm^{-1} and 2,500 cm^{-1} in this report for each of the 105 atmospheric cases. The energy outside this spectral interval was estimated by assuming blackbody emission at the equivalent blackbody temperature derived from the unfiltered radiances for the 77 spectral intervals.

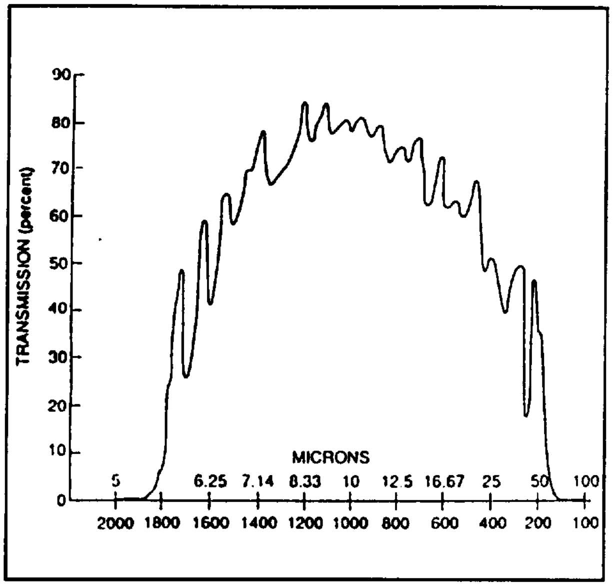

Figure 23. The transmittance function of the longwave scanning sensors 19 to 22.

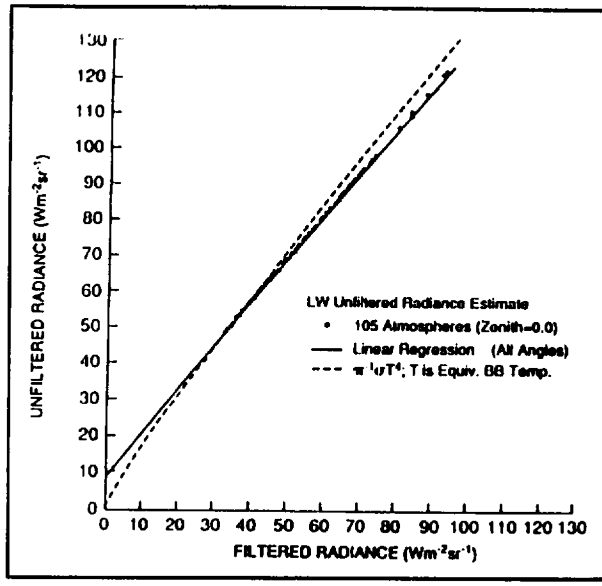

Figure 24. Plot of filtered vs. unfiltered radiances for longwave sensors 19 to 22.

The final results varied only slightly with solar zenith angle so a mean (all angle) result was used to "unfilter" the flight measurements. The unfiltering algorithm uses the regression line down to 30 $Wm^{-2}sr^{-1}$; below this the dashed, blackbody, line is followed towards zero. This correction was reported in the ERB Science Team minutes, but it has not previously been reported in open literature.

In addition to the above problems, the operation of the scanner interfered with the operation of a short-lived companion experiment on the Nimbus-7 satellite. Because of this, operation of the scanner was often suspended during part of the orbit or even for days at a time during the winter and early spring of 1979 (Kyle et al., 1990a). There was a noticeable decrease in the noise in the ERB total solar irradiance sensor (channel 10c) after the final turnoff of the scanner (Hoyt et al., 1992).

7.9.2 The WFOV Sensors

The WFOV sensors channels (11-14) were calibrated before launch, but possessed no internal calibration facility on the satellite. The in-flight calibration monitoring depended on comparison with each other, with the scanner measurements, on periodic direct looks at the Sun, and on analysis of the data to identify measurement irregularities.

The four sensors covered three spectral ranges. They were all type N3 thermopile sensors. Channels 11 and 12 were redundant channels with no spectral filters and a response range (<0.2 to >50 μm). Channel 11 had black- painted baffles and was, at first, generally kept covered to reduce possible degradation. Channel 12 had polished aluminum baffles similar to those on an earlier ERB instrument that flew on the Nimbus-6 satellite. Channel 13 was covered with two concentric Suprasil W domes (Figure 25) which passed shortwave radiation in the range (0.2 to 3.8 μm). Channel 14 had a RG695 insert between its two Suprasil W domes which restricted its pass band to (0.7 to 2.8 μm). These sensors saw the entire visible disk of the globe plus a small ring of cold space.

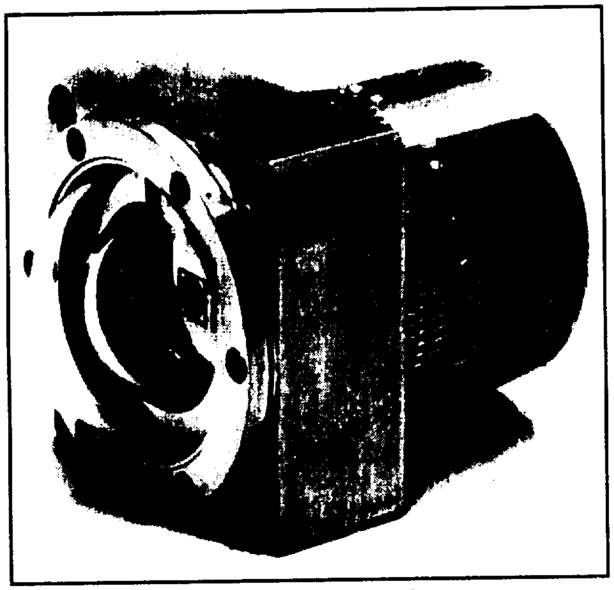

Figure 25. Suprasil W dome.

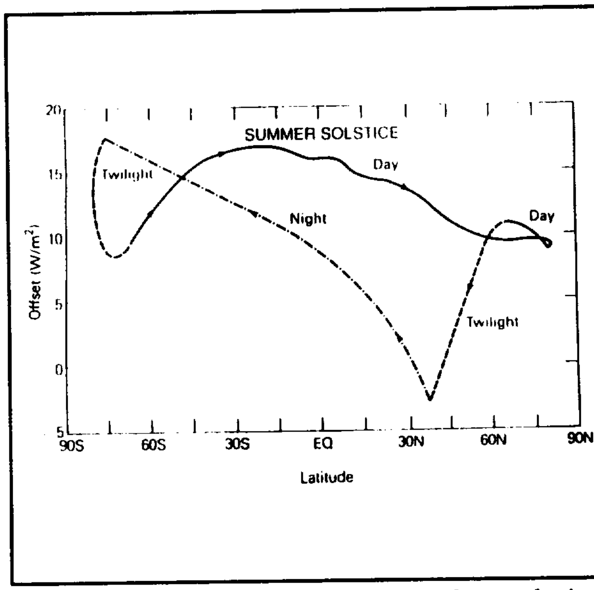

Figure 26. Channel 13 estimated thermal perturbations over the course of one 104-minute orbit near June 22.

Numerous calibration problems were discovered after the sensors were in orbit. First, the total channels 11 and 12 were calibrated on the ground using a blackbody which filled their entire field of view. In space, the polished baffles on channel 12 allowed it to radiate more energy to cold space than had been anticipated. A corrected sensitivity coefficient was derived using improved laboratory measurements with a cold ring about the target blackbody and checking this by calculations based on the period Sun looks and by comparison with the scanner measurements. Before the launch of the Nimbus-7, NASA discouraged the use of black paint because it tended to deteriorate and change the sensor characteristics. However, the Chemglaze Z306 black used on channel 11 and some of the other ERB sensors stood up very well. It was cured by baking in a vacuum before launch (Kyle et al., 1984, 1995b).

Secondly, after launch channel 12 also required a bias adjustment. Channel 12 and the scanner measurements were compared and an adjustment was made to bring the mean channel 12 nighttime measurements into agreement with the longwave scanner value.

Thirdly, channels 13 and 14 were found to be sensitive to the solar-driven heating and cooling cycle that occurred during each 104-minute orbit. They were also thermally affected by the duty cycles of the ERB instruments, and to a lesser extent, the cycles of nearby experiments. The magnitude of these thermal perturbations are illustrated for channel 13 in Figure 26. Only the daytime shortwave measurements were used in Earth radiation budget studies. The nighttime measurements and additional laboratory measurements (Maschhoff et al., 1984) were used to develop models to remove these thermal perturbations. Such heating and cooling cycles were significant in channels 11 and 12. It is hypothesized that most of the heating and cooling signals in channels 13 and 14 arose from the heating and cooling in the covering domes (Kyle et al., 1995a,b).

Fourthly, all the WFOV sensors were contaminated by both scattered and direct sunlight for solar zenith angles between the terminator and the Earth's shadow. This caused large gaps in the important longwave terrestrial flux measurements at mid- and high latitudes. An interpolation procedure was developed to remove the contaminating solar signal (Kyle et al., 1995b).

Fifthly, all the WFOV sensors degraded with time and the degradation was spectrally dependent. In addition, the degradation on channels 13 and 14 was angularly dependent with the degradation largest in the forward-facing portion of the covering domes. In channels 11 and 12, the degradation was monitored by the periodic Sun looks. Figure 27 shows the sensitivity changes in channel 12 derived from the Sun looks. This refers only to the shortwave radiation. The longwave changes were determined from a time series analysis of the channel 12 global mean nighttime observations. The procedure was validated by also calculating the shortwave degradation in channel 12. This checked with the results of the solar monitoring program. In practice, sensitive changes were determined twice a month and interpolation used to obtain a smooth variation. Table 9 summarizes the longwave and shortwave degradation in channel 12 by listing the degradation factors once a year. There was no detectible degradation in channel 11, which was kept covered a good deal of the time. It is hypothesized that most of the degradation arose from changes in the black paint which covered the sensor chips.

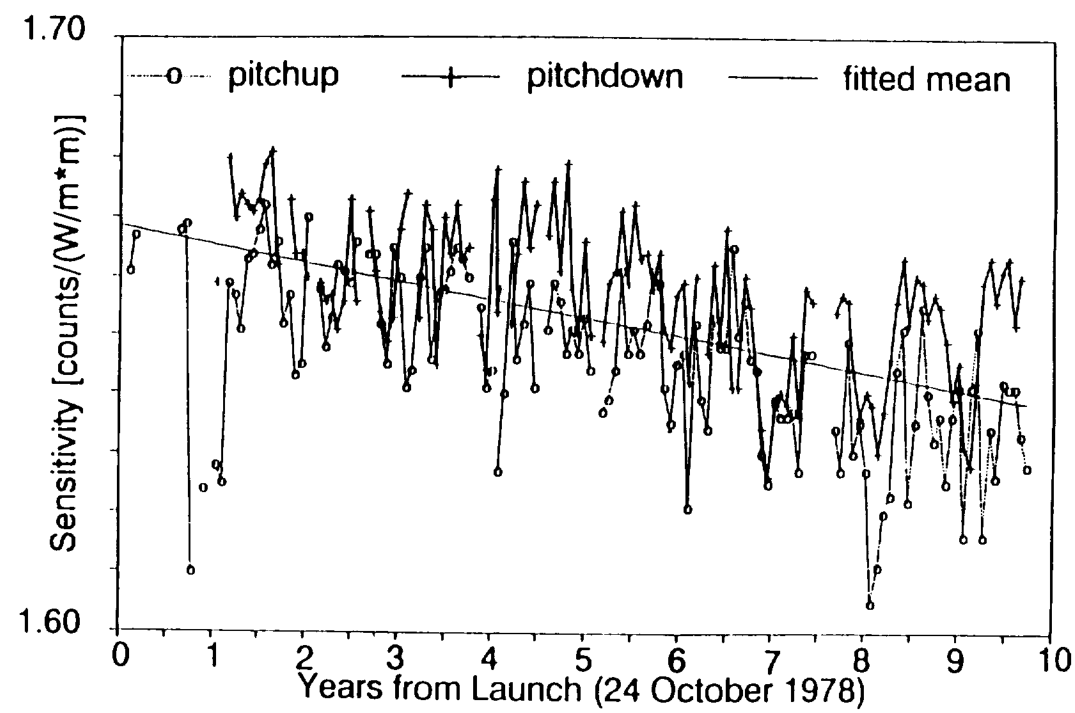

Figure 27. Channel 12 shortwave sensitivity changes determined by solar observations at sunrise (pitchup) and sundown (pitchdown).

The solar observations demonstrated the angular dependence of the degradation in channels 13 and 14. This is shown in Figure 28 for channel 13; the degradation is much greater on the forward-facing (sunrise) portion of the covering dome. The solar looks monitored only a small portion at the front and back of the domes. The definitive darkening of the Earth-facing portion of the domes was monitored by comparison with the channel 12 observations (Figure 29). The comparison procedure is described in detail in Kyle et al. (1994a, 1995c). The degradation was considerably larger in channel 13 than in channel 14 because of the spectral dependence of the degradation. This spectral dependence is indicated in Figure 30. Sensors similar to those on the ERB, were flown on the Long Duration Exposure Facility (LDEF) satellite. This satellite was recovered after about 6 years in orbit and the sensors were analyzed for changes (Hickey et al., 1992). In channel 13 and 14, the chief degradation seems to have occurred due to films forming on the covering domes from material outgassing from the experiment and from the rest of the satellite. In the presence of solar ultraviolet radiation and perhaps atmospheric atomic oxygen, this film darkens and decreases the transmissivity of the dome. The darkening is greatest on the forward-face of the domes which see the Sun directly at satellite sunrise, once each 104 minutes. The back of the dome sees the Sun at satellite sunset but not as clearly. During periods of solar excitation, the number of ionized and neutral oxygen atoms at satellite altitude (950 km) increases sharply. At such times, the film is partly cleaned from the ram direction, or the front of the domes (Predmore et al., 1982). The time-dependent degradation of channels 13 and 14 is summarized in Table 10. Because of the wavelength, dependence of the degradation channel 13 is affected much more than channel 14. A correction algorithm, based on a comparison of channel 13 and scanner measurements, was developed to correct for the asymmetric darkening of the channel 13 domes (Kyle et al., 1984). No correction was made for the smaller asymmetry problem in the near-infrared channel 14, nor was any correction developed for the wavelength-dependent darkening in channel 13.

Table 9. Channel 12 Longwave (T_{LW}) and Shortwave (T_{SW}) Degradation Factors

Year	Date	T_{LW}	ΔT_{LW}	T_{SW}	ΔT_{SW}
	October 24, 1978	1.00		1.00	
0	November 16-30, 1978[a]	0.9961	0.0039	0.9997	0.0003
1	October 16-31, 1979	0.9801	0.0160	0.9975	0.0022
2	October 16-31, 1980	0.9780	0.0021	0.9953	0.0022
3	October 16-31, 1981	0.9765	0.0015	0.9938	0.0015
4	October 16-31, 1982	0.9755	0.0010	0.9930	0.0008
5	October 16-31, 1983	0.9752	0.0003	0.9926	0.0004
6	October 16-31, 1984	0.9752	0.0000	0.9926	0.0000
7	October 16-31, 1985	0.9752	0.0000	0.9898	0.0028
8	October 16-31, 1986	0.9752	0.0000	0.9868	0.0030
9	October 16-31, 1987	0.9752	0.0000	0.9848	0.0020

ΔT_i is the difference $(T_{i-1}) - T_i$). At launch, October 24, 1978, it is assumed that T=1.
[a]November 16, 1978 was the first measurement day. The listed degradation that had occurred by this time is an estimate.

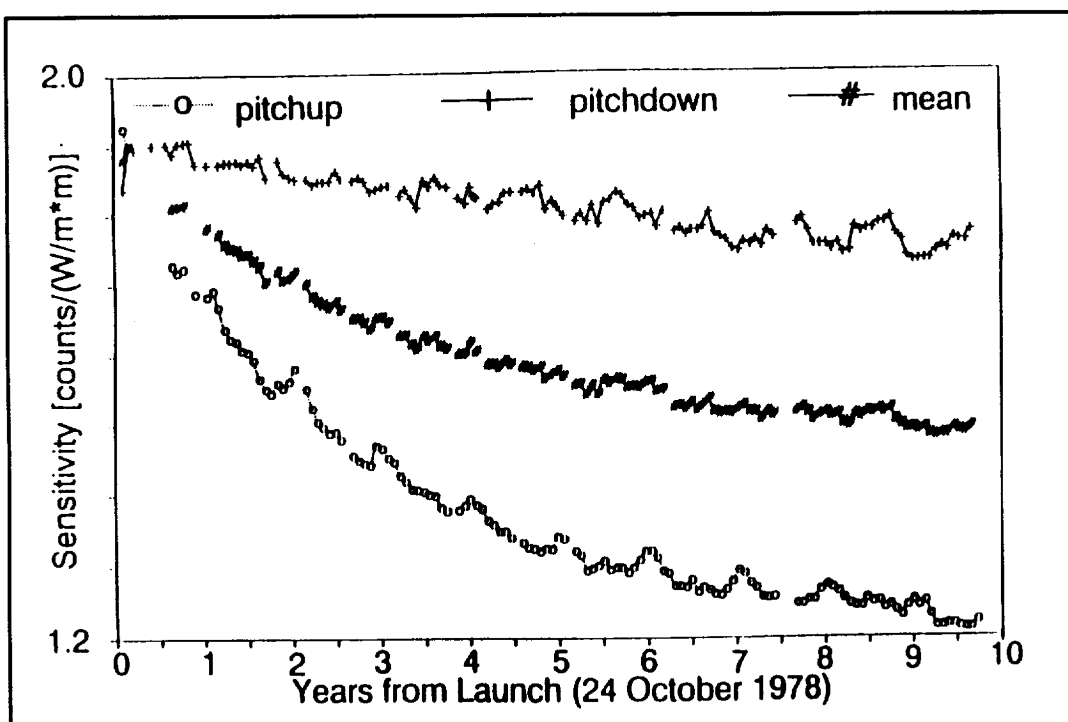

Figure 28. Channel 13 shortwave sensitivity changes observed at sunrise (pitchup) and sundown (pitchdown).

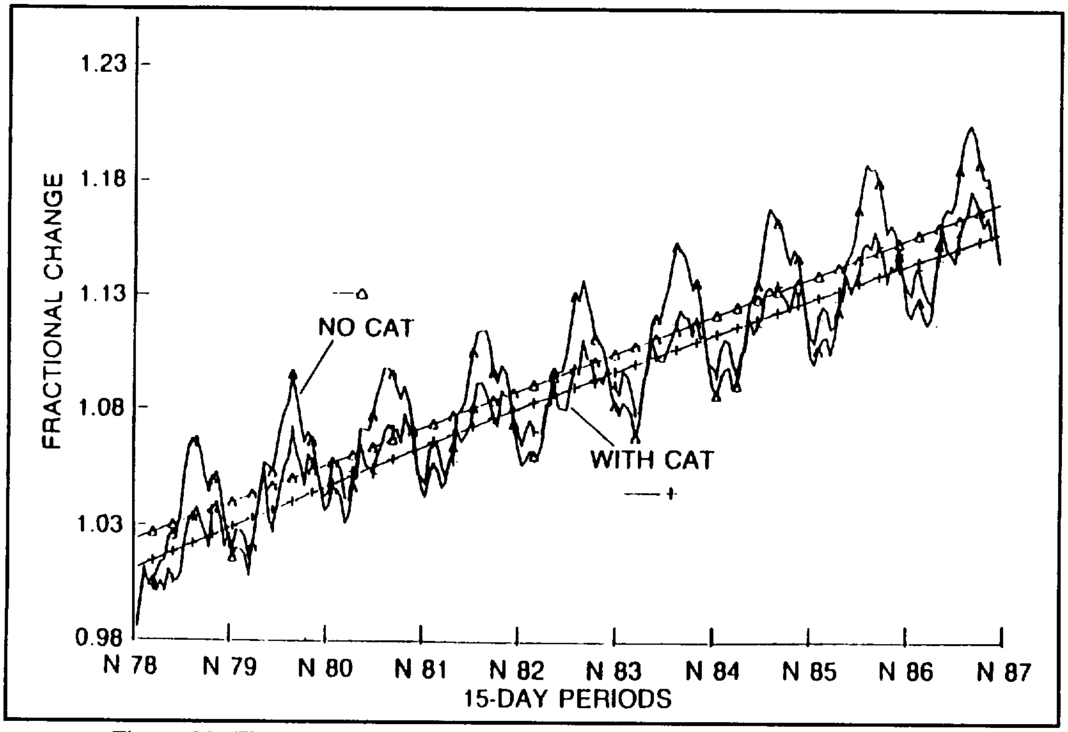

Figure 29. The mean channel 13 darkening from comparison with channel 12.

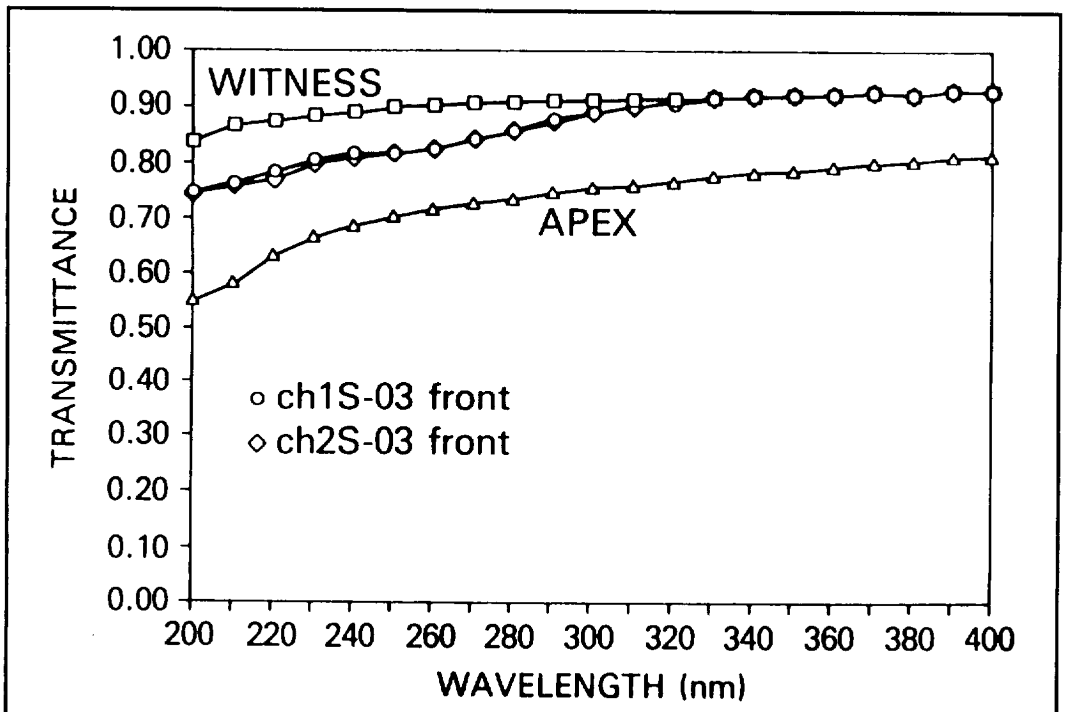

Figure 30. Measured wavelength-dependent darkening on ERB-like windows recovered from the LDEF satellite. The witness window was not flown. The apex windows received more contamination than did the middle curve.

Table 10. Channels 13 and 14 Mean Degradation Factors

Year	Date[a]	T_{13}	ΔT_{13}[b]	T_{14}	ΔT_{14}
	October 24, 1978	1.00		1.00	
0	November 16-30, 1978[c]	0.9961	0.0039	1.00	0.00
1	October 16-31, 1979	0.9703	0.0258	0.9878	0.0122
2	October 16-31, 1980	0.9518	0.0185	0.9768	0.0110
3	October 16-31, 1981	0.9362	0.0156	0.9680	0.0088
4	October 16-31, 1982	0.9241	0.0121	0.9633	0.0047
5	October 16-31, 1983	0.9123	0.0118	0.9587	0.0046
6	October 16-31, 1984	0.9007	0.0116	0.9542	0.0045
7	October 16-31, 1985	0.8895	0.0112	0.9496	0.0046
8	October 16-31, 1986	0.8786	0.0109	0.9451	0.0045
9	October 16-31, 1987	0.8672	0.0114	0.9406	0.0045

Notes:
[a]Degradation factors were determined for every half-month period. Here one value per year is shown.
[b]$\Delta T_i = T_{i-1} - T_i$ or degradation rate.
[c]The degradation for this period is an estimate since the measurements started on November 16, 1978.

As a final validation (Kyle et al., 1990b), the ERB WFOV measurements were compared with follow-on, independent measurements from the Earth Radiation Budget Experiment (ERBE, Barkstrom et al., 1989). In a number of the comparisons, the final time- and space-averaged products were used. Thus, differences in the calibration, temporal and spatial sampling, and procedures to invert the measurements to the top of the atmosphere were all involved. Data for the 4 months (April, July and October 1985, and January 1986) were treated. Comparisons were made with both the ERBE scanner and WFOV measurements. For the global mean longwave radiation, the ERB and ERBE WFOV measurements agreed equally well with the ERBE scanner measurements. For the reflected shortwave radiation, the ERB WFOV measurements were close to the ERBE scanner values, but both were larger than the ERBE WFOV measurements.

7.9.3 The Solar Measurements

The 10 solar sensors (channels 1-10c) observed the Sun briefly once per orbit at satellite sunrise. Channels 1 and 2 were similar to channel 13 with a pass band of (0.2 to 3.8 μm) while channel 3 was similar to channel 12 (<0.2 to >50 μm). Channel 10c was a self-calibrating cavity radiometer with a pass band (<0.2 to >50 μm). Only channel 10c (c for cavity) had both the stability and sensitivity to track variations in the total solar irradiance over a 15-year period. The solar variability is shown in Figure 31 where monthly means of the channel 10c measurements are compared with those of the Wolf sunspot number, which is often used to track solar activity. The short-term stability of channel 10c was ±0.35 Wm^{-2}. The long-term stability was at least this good and may have been as good as ±0.1 Wm^{-2} for most of its lifetime.

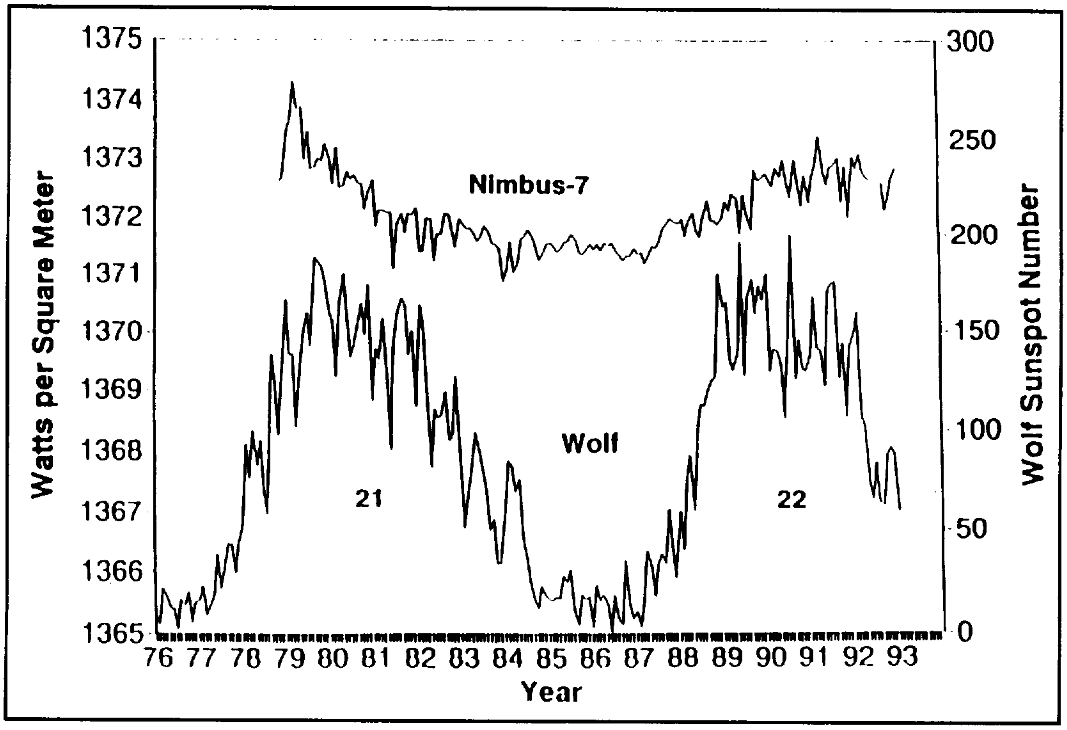

Figure 31. Channel 10c measurements through January 1993 compared with the Wolf sunspot number.

This level of stability for channel 10c was achieved by a continuous calibration program from the initial prelaunch calibration through the end of the measurement program. A major review occurred in 1989 to 1991 and resulted in improvements in most of the calibration coefficients and parameters. These included new procedures for calculating the sensor's off-axis angle, zero offset, temperature sensitivity coefficient, and the Earth-Sun distance. All of the measurements were reprocessed using the new procedures (Hoyt et al., 1992). The recalculation of the sensor's off-axis angle was the most important improvement. This removed discontinuities in the solar irradiance of the order of 0.06 percent. It was found that the device which recorded the solar telescope's off-axis angle slipped by 0.5x twice during the experiment. As a final validation, the channel 10c measurements were compared to the independent solar measurements taken by the Solar Maximum Mission (SMM) satellite (Willson and Hudson, 1988, 1991). Both measurement sets showed the same irradiance variations and trends, but the ERB measurements were in the mean about 4.4 Wm^{-2} higher than the SMM values. This bias shift was caused by uncertainties in the absolute calibration of the two sensors.

In-flight calibration adjustments were not developed for ERB channel 1-9. The measurements are shown in Figure 32. The degradation history of channels 1-3 can be accurately tracked by comparison with the channel 10c measurements. However, the spectral channels 4-9 cannot accurately be corrected by the same method since the variation in the solar irradiance has a spectral dependence with the largest variations occurring at the shortest wavelengths. The original spectral pass bands of channels 4-9 in microns are, respectively, 0.536 to 2.8, 0.698 to 2.8, 0.398 to 0.508, 0.344 to 0.460, 0.300 to 0.410, and 0.275 to 0.360. The results from the LDEF experiment (Hickey et al., 1992) suggest that fogging films formed on the Suprasil W windows caused the major problems. For the first several years channel 1 was kept covered most of the time, but it was opened briefly about once every 12 days. This is indicated by the dots. Some degradation occurred during but it was small compared to that of its twin sensor (channel 2). In 1984, channel 1 was uncovered and left open, except for a break in 1986, until the fall of 1988, when it was again covered except for brief measurement periods. It was reopened in spring of 1990 during an excited Sun period. A rapid cleansing occurred at this time. Additional analyses of these channels can be found in Kyle et al., 1993a and 1994b. Some useful information could probably be obtained from ERB-type spectral channels if each had two or three backup channels which would normally be kept shuttered and opened only for brief comparison measurements. With careful analysis, some information about the spectral variability of the Sun during excited Sun periods could probably be derived from the

Nimbus-7 ERB measurements. Detailed examination shows that the true solar signals which appear in channel 10c when large sunspot groups are present, also appear in channel 1-9. However, care would have to be taken in determining the relative change in two or more of the channels.

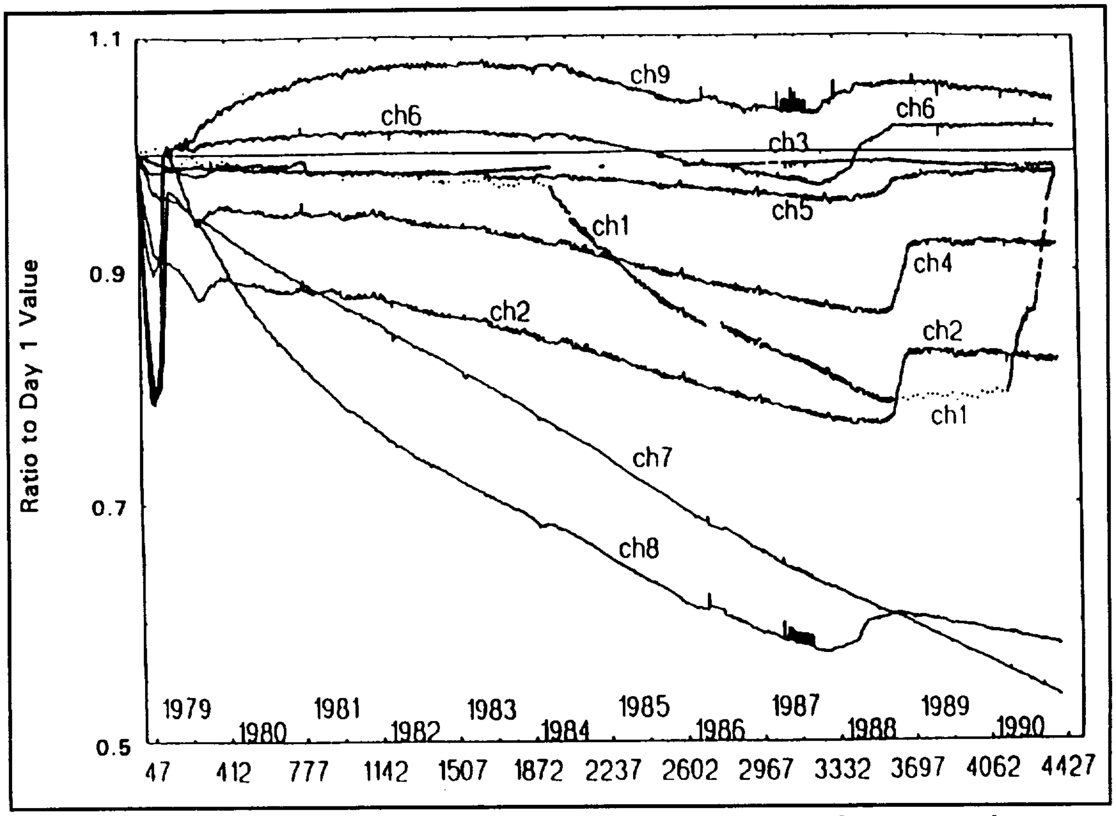

Figure 32. The measurements from solar channels 1 to 9 have not been corrected.

7.9.4 Lessons Learned

- In addition to the prelaunch calibration program, an extensive postlaunch characterization effort was required to properly calibrate about 15 years of total solar irradiances and 9 years of Earth-emitted longwave and reflected shortwave radiation recorded by the Nimbus-7 ERB experiment.

- Only after this comprehensive on-orbit characterization and subsequent reprocessing was the 9-year ERB data set optimized for global change research. Improvements permitted quantitative analysis of the relationship between clouds and radiation, isolation of the El Niño/Southern Oscillation signal, detection of the effects of the explosive volcanic eruption of El Chichón, and analysis of other important topics.

- As far as practicable, the effect of the sensor's space environment on the measurements should be studied both before and after launch. This should include heating and cooling due both to the satellite diurnal cycle and to operational instrument and satellite duty cycles. The effect of both mechanical vibrations and electromagnetic interference should be also included. Complete shielding is impossible, but power modeling can correct some of the residual problems. Either a flight or laboratory model of the instrument should be kept available for additional laboratory tests during the measurement programs to assist ongoing characterization research.

- On-board calibration facilities should be as good as possible. In addition to internal calibration sources, external sources such as the Sun, the Earth, and deep space can often be used as important calibration standards for large FOV instruments.

- When practical, sensors and on-board calibrators should be flown in groups of 2 to 4. One for standard measurements, and spares kept normally covered but opened periodically to check for degradation.

- Sensor degradation is frequently spectrally dependent, and means to detect and correct for this dependency should be devised if possible.

- Time series analysis is a powerful calibration/validation aid if the measurements continue over a sufficiently long time.

- Follow-on programs designed to extend time series of critical measurements should overlap in time to allow cross-calibration and normalization of the measurement series.

7.9.5 Acknowledgements

Many persons worked for a long time on the Nimbus-7 ERB calibration problems. The authors particularly wish to recognize the efforts of John Hickey of the Eppley Laboratory, Inc.; Herbert Jacobowitz and Larry Stowe of NOAA/NESDIS; Douglas Hoyt, Lanning Penn, Richard Hucek, and Brenda Vallette of RDC; Brian Groveman of CSC; and Robert Maschhoff of Gulton Industries.

7.10 Ensuring Continuity From ERBE to ScaRab

Robert Kandel, Laboratoire de Météorologie Dynamique du CNRS, Ecole Polytechnique, Palaiseau, France

Because energy exchange between the climate system and the cosmos is overwhelmingly dominated by radiation, accurate determination of ERB components, of their geographical distribution, and of their variation on various time scales, is needed for a correct understanding of climate system energetics (Hartmann et al., 1986; Ramanathan, 1987). It is particularly important to measure cloud radiative forcing, its seasonal changes, and its variations in the course of different climate anomalies. However, we must make sure that the interannual changes that we detect are indeed geophysical in nature, and not simply artifacts arising from calibration errors or sampling biases.

ERB determinations depend on radiation measurements made from space. There are several levels of difficulty here (cf. e.g., Stowe, 1988; Kandel, 1990), and in considering them we follow the main lines of ERBE data processing:

- The first is assuring flat spectral response over broad shortwave (SW) and longwave (LW) bands, and accurate absolute calibration. In practice the instruments do not have a perfectly flat spectral response (Lee and Barkstrom, 1991; Monge et al., 1991). The correction for spectral filtering depends both on the instrument properties and on the spectrum of the scene being viewed (Smith et al., 1986). In what may be called a "primary data processing" stage (ephemeris and count conversion in ERBE), individual scene properties are ignored. Raw data counts for each channel are converted into so-called "filtered" radiances, using both ground calibration results and data from views of on-board calibration sources. The filtered radiances may be considered to be purely instrument products, their accuracy expressed in physical units ($Wm^{-2}sr^{-1}$) depending on the ground and on-board calibrations.

- To take the scene spectrum into account, one needs information about the nature of the scene. In ERBE, one calculates a first approximation to the "unfiltered" broadband SW and LW radiances, using a spectral correction based only on instrument spectral response and using "average" SW/LW terrestrial spectra without consideration of the actual scene being viewed. The pixel location on the Earth is determined in the primary data processing, so that the underlying scene type or "geotype" is known apart from varying snow/ice cover. Viewing geometry is also determined. However further information is needed regarding cloud cover. This is the essential role of the scene/cloud identification (SCI), considered as part of the "inversion" subsystem in ERBE processing (Smith et al., 1986; Wielicki and Green, 1989). The ERBE SCI uses first-order approximate values of the unfiltered SW and LW radiances together with the instantaneous viewing geometry, and a set of angular models and very rudimentary diurnal variation models, to determine whether the scene is clear, partly or mostly cloudy, or completely overcast. This SCI then determines which set of geometry- and scene-dependent spectral correction coefficients are to be applied to the filtered total (TW), SW, and LW radiances to provide a better estimate of the unfiltered LW and SW radiances. It is extremely important to note that this spectral correction procedure depends both on geophysics and on the characterization of the instrument.

- Converting measured radiances into fluxes requires taking into account the anisotropy of the reflection and emission of radiation by the Earth-atmosphere system. The radiance-to-flux conversion involves strong angular corrections for the reflected SW, depending on the nature of the scene and on the geometry of observation. The SCI already used in the spectral correction, plays an essential role here, determining which angular correction model to use. When a limited and therefore necessarily inaccurate set of angular distribution functions (bidirectional reflectance distribution functions for the SW) is used, restricted angular sampling will lead to systematic errors (bias) in the results.

- Estimates of daily, decadal, and monthly mean ERB quantities require assumptions regarding the variation of these quantities between actual observation times (Brooks et al., 1986). Up to now ERB measurements have been made from satellites in low orbits (generally but not always polar), providing only a few measurements per 24-hour period. Inadequate time sampling and incorrect assumptions will bias the results.

7.10.1 The Spectral Correction Problem

It is important to recognize that, despite or because of the ERBE algorithm structure, complete separation of the different aspects of ERB determination is impossible. In particular, the separations between the primary (count conversion) and secondary ("inversion," i.e., spectral correction, SCI, and radiance-to-flux conversion) stages of data processing are not absolute. In fact, the spectral correction procedures, formally located by ERBE in the "inversion" subsystem of SDP, constitute an interface between primary and secondary data processing in which both instrument and Earth scene properties come into play. Daytime LW radiance (and flux) determinations depend on the SW channel calibration and on the blackbody calibration of the TW and LW channels.

Analysis of the ERBE 3-channel (SW, TW, LW) filtered radiances has revealed systematic differences between the daytime unfiltered LW radiances from the scanners operating on the NOAA-9 and -10 satellites as compared to the Earth Radiation Budget Satellite (ERBS) (Thomas et al., 1995). These differences affect the ERBE determination of LW diurnal variations, and can be traced to small discrepancies in the SW calibration and spectral characterization of the 3 different ERBE scanners. The response differences are misinterpreted by the ERBE spectral correction procedure which applies coefficients calculated assuming perfect calibration and identical spectral response in the three instruments. Moreover, the daytime LW estimates are obtained using what is essentially a weighted spectral subtraction LW = 3D (TW - r = 99SW) in the three-channel ERBE spectral correction procedure, whereas nighttime estimates are independent of the (zero) SW measurements. Thus the differences arising from instrument calibration problems are transformed by the ERBE spectral correction procedure into time-dependent biases which impact the ERBE "time-space averaging" system.

This analysis was possible because of the existence of the ERBE LW channel providing strongly filtered but purely longwave radiances. The method can be adapted to the case of ScaRaB (Kandel et al., 1994) or CERES (Smith et al., 1994), for which the third channel is a relatively narrow infrared window channel. The important property is that it is sensitive to LW radiation alone, and that statistically it is a good predictor of the broadband LW radiance. This has made it possible to check the ScaRaB SW calibration independently of on-board lamp sources, using the nighttime TW/IR correlation to provide a basis for estimating LW(IR) by day, and forcing weighting factor r to have a value that makes the spectral subtraction agree on the average with the window estimate. However, this procedure implicitly makes the assumption that the relative SW/LW response in the TW channel is stable and known.

7.10.2 The Problem of Unpublished Coefficients

Part of our spectral correction analysis was guesswork, because the ERBE spectral correction coefficients are not publicly available, except for a few samples found in the ERBE Science Team Meeting Minutes. Of course, these coefficients are specific to the presumed spectral response of the ERBE channels, and of no use for other instruments. However, a major difficulty in assuring a ScaRaB secondary data processing system compatible with ERBE has been the unavailability in the open literature of other coefficients used in the SCI. We recall that the ERBE "scene identification" uses a maximum likelihood estimate (MLE; Wielicki and Green, 1989) operating on the broadband SW and LW radiances estimated with the preliminary spectral correction, and comparing these with SW and LW radiances predicted for the same illumination and viewing geometry on the basis of the statistical angular models constructed from a limited set of Nimbus-7/ERB data.

In our preparation of the "ERBE-like" inversion algorithms for ScaRaB, we were able to use the published ERBE angular models (Suttles et al., 1988). However, certain parameters were not available. These include the *a priori* values of regional mean albedos and LW radiant exitances and cloud cover probabilities (some 120 coefficients), and the values used for taking account of LW diurnal variation of the cloud-free portions of land and desert scenes, in the MLE scene identification (cf. Wielicki and Green, 1989). For the ScaRaB ERBE-like processing, we have assigned "reasonable" values to these parameters, based in part on analysis of ERBE results carried out at LMD (Ahmed, 1994). The question is whether ScaRaB products, based on an ERBE-like SCI using values which are not identical to those used in ERBE itself, can be considered to be a continuation of the ERBE data set. We (Viollier et al., 1995) have tested this using the April 1985 ERBE data which contain both the input data to the scene identification and inversion processing subsystems, and the corresponding output products using the original ERBE inversion; the input data have been submitted to the ScaRaB ERBE-like inversion, and the output products obtained have then been compared with the ERBE output products. We find scene identification in agreement at the 80- to 90-percent level (Table 11) and even though there is disagreement, effects on the resulting global mean fluxes are very small (Table 12). The one significant difficulty which does appear has to do with the evaluation of clear-sky fluxes. The regional means of clear-sky fluxes are based on very small samples. The conclusion then is that even though all-sky mean fluxes are relatively insensitive to the details of the SCI and inversion, the mean clear-sky fluxes can depend on these fairly strongly. Inasmuch as the clear-sky flux is essential to determining the cloud radiative forcing, it remains important to improve the identification of clear scenes and the rejection of cloud contamination.

Table 11. Rate of Agreement (in percent) of Scene (cloud) Identification, Comparing ERBE-Like ScaRaB Processing with Original ERBE Processing, Applied to April 1985 ERBE Data

	Night			*Day*		
MLE Parameters Used	Total	Land	Desert	Total	Land	Desert
No Regional Adjustment	88.3	66.3	13.5	87.6	87.3	87.1
LW Regional Adjustment	88.6	76.5	86.6	86.7	81.2	83.7
LW and Albedo Regional Adjustment	Same	Same	Same	86.8	81.1	85.0

7.10.3 Time Averaging Problems

In working with ScaRaB data, we must have an "ERBE-compatible" processing system, although we hope to adapt certain aspects of this to take advantage of the four ScaRaB channels and to take into account the particular time sampling of the Meteor-3 orbit. However, modifications of the different subsystems—SCI, radiance-to-flux conversion, diurnal interpolation and averaging—must consider consistency issues. Angular models intervene both in SCI and in diurnal interpolation of SW measurements, and in the formal radiance-to-flux conversion. By contrast, although the formal structure of the ERBE "inversion" subsystem explicitly links angular correction and scene identification, we believe that a better LW angular correction can be made using the ScaRaB IR window channel in addition to the broadband LW estimate, without explicit scene identification (Stubenrauch et al. 1993). With regard to the SW domain, the diurnal interpolation/extrapolation procedure (DIEP) depends on the SW angular models adopted, and if one modifies the procedure, one must consider whether the modification is consistent with the ERBE (or other) angular models. At the same time, analyses of diurnal variations require taking into account the generally limited angular sampling properties of any individual satellite/scanner combination, whether in low or in geostationary orbit.

Table 12. Global Mean Fluxes (Wm^{-2}) and Flux Differences, Comparing the Original ERBE Products and Results Obtaining Our ERBE-Like ScaRaB Processing, All Averaged Over April 1985. The Difference in Global Mean Flux is Given by D, and s is the Standard Deviation of the Corresponding Regional Differences

	All-Sky		*Clear-Sky*	
April 1985 ERBE/NOAA-9	LW	SW	LW	SW
ERBE Product (reference)	234.94	100.75	268.03	55.39
ScaRaB Simulation (sampling 1 ERBE pixel/2), using the ERBE Scene Identification	234.97 Δ = +0.03 σ = 2.07	100.57 Δ = -0.18 σ = 2.57	269.58 Δ = +1.55 σ = 5.63	56.62 Δ = +1.23 σ = 6.03
Complete ScaRaB Simulation (sampling 1 pixel/2), but without Regional Adjustment	235.00 Δ = +0.06 σ = 2.10	100.51 Δ = -0.24 σ = 2.95	276.95 Δ = +8.92 σ = 9.89	54.69 Δ = -0.70 σ = 7.09
Complete ScaRaB Simulation (sampling 1 pixel/2) with Regional Adjustment	235.01 Δ = +0.07 σ = 2.10	100.45 Δ = -0.30 σ = 2.96	273.56 Δ = +5.53 σ = 6.13	55.80 Δ = +0.41 σ = 6.98

7.10.4 Concluding Remarks

Although it may be tempting to represent the overall problem in a flow chart with different tasks clearly put into separate boxes, in reality things do not work that way. Certainly a data processing system works in a series of steps, but interactions must be considered. Careful attention must be given to the interactions between the areas of instrument calibration and spectral characterization, and the secondary data processing both regarding spectral correction and diurnal interpolation. One must examine how purely instrumental errors propagate, sometimes canceling, sometimes being amplified, in the algorithms designed to take the presumably noninstrumental problems into account. In some cases, diagnostics in one step of the process must feed back into an earlier step, so as to get the best possible result. The spectral correction problem is at a critical interface, and the results found regarding the spurious ERBE LW diurnal cycle generated, via the ERBE spectral correction, by a SW calibration error, illustrate the importance of monitoring carefully the spectral characteristics of the broadband instruments. Because of the importance of the 5-year ERBE data set for climate studies, the daytime LW errors in the NOAA-9 and NOAA-10 scanner products must be corrected. Our validated correction algorithm is easy to apply.

Even for relatively "simple" experiments such as ERBE or ScaRaB, involving basically only processing of 3 channels of data, it appears that large sets of models (angular, time interpolation) and coefficients are needed to take into account the sampling issues in the processing. Therefore, even if the instruments are perfectly calibrated so that the filtered radiances indeed represent "absolute truth," the final output products depend on the algorithms, models and coefficients used. Some of these have been published in the open literature and so form the basis for checking the "ERBE-compatibility" of later products. However, others have not been made public. In some cases, this is understandable in that not all the coefficients may be "publishable" in the usual sense; they may not be well- established scientifically, but must be chosen more or less arbitrarily to have an operating processing system. It may be that the best approach is to make sensitivity studies, and to show (as we have done in Viollier et al., 1995) that the ERBE algorithms are robust and not unduly sensitive to the specific values of the coefficients in the "toolbox". Indeed, if the contrary turned out to be true, one might have serious doubts as to the accuracy of the products, unless it could be shown that the values adopted were true, and that would certainly deserve publication. Nevertheless the fact that some of the essential processing parameters remain confidential means that the corresponding products are not completely documented, and that their reproducibility cannot be independently assessed. This may not have raised any problems when practically all the missions were NASA missions, but in a world in which other agencies are making significant contributions to observations of the changing geosystem, there certainly should be some mechanism for exchanging information on these "unpublishable" but nonetheless essential data.

Existing and near-term instrument data can be used to improve the ERB data products, in particular in the area of the SCI and in that of LW diurnal interpolation. Although a single satellite only provides two LW radiation samples per 24 hours in low and middle latitudes, existing geostationary satellite IR window data in the ISCCP B2 or B3 format can be used to provide excellent correction of diurnal variation bias. This is certainly the case for "contemporary" B2/3 data; moreover when such data are used, the complete time variation can be well represented with 3-hour resolution. Even if there is only a single ERB instrument in low orbit, the time sampling biases can be substantially eliminated. Considering that the time sampling of global ERB data will continue to be incomplete over most of the decade, this is an important conclusion. These results can only be global if global data is available. Thus it is essential that coverage of the Indian ocean sector either by Insat or by GOMS be maintained, and that the data be available.

Because a single satellite provides only one SW radiation sample per 24 hours in low and middle latitudes, and because the reflected SW radiation can be strongly anisotropic in ways that depend on the surface and cloud cover, only partial and imperfect correction of diurnal variation bias can be expected. Data from existing geostationary satellites can contribute to an improved SW DIEP, using existing SW BDRFs and directional models. However significant improvements will depend on progress with the SW angular models and on the availability of broadband SW data from geostationaries (as in the GERB project for second-generation Meteosat). Without waiting for these, it still appears desirable to implement an improved SW DIEP making use of ISCCP data, so as to have at least a basis for assessing diurnal variation bias in the time averaging process. Of course, this also requires maintenance and availability of global coverage by geostationaries. Substantial progress will depend on the development of improved SW angular models.

The development of improved SW angular models and of improved scene/cloud identification must be carried forward in a consistent fashion. However, without waiting for the next generation of angular models, it is important to refine the clear-scene identification procedure because of its importance for the determination of cloud radiative forcing.

Finally, from the scientific point of view, the ultimate aim of Earth Radiation Budget research is a better understanding of how radiation processes govern climate. We would like to have a reliable estimate of climate-cloud-radiation feedback that does not depend on actually observing the climate change. We certainly will need more than the present concept of cloud radiative forcing, perhaps refining it to take into account different cloud types, or replacing it by some other parameter. We need to know more about the three-dimensional structure of the cloudy atmosphere, and this requires as a first priority the further development of passive and active microwave capabilities for sounding clouds. The Earth Radiation Budget as such, even considering its SW and LW components separately, and considering also the cloud radiative forcing components, remains an "integrator" of atmospheric state. Although it is not a "signature" of any single atmospheric process, it provides essential constraints to climate modelers.

7.10.5 Acknowledgments

This work was supported in part by CNES (the French national space agency), and by the European Space Agency. I thank my many colleagues (including ERBE Science Team members) who contributed advice, ideas, help, support and criticism which made this work possible. However the responsibility for any misstatements or errors in this paper is my own.

7.11 Calibration/Validation: GOES-I/M for Global Change

Dennis Chesters, NASA/Goddard Space Flight Center, Greenbelt, MD

7.11.1 Radiometric Calibration of GOES-I/M

Radiometric calibration of the GOES-I/M sensors is carried out by the vendor, ITT at Fort Wayne, Indiana. NOAA (Mike Weinreb) is responsible for operational calibration. The GOES-I/M satellites have design lives of 5 years each, providing at least 10 years of continuous radiometric observations from nearly fixed station.

Visible calibration of the Imager and Sounder uses an integrating sphere that is also used for the AVHRR instrument on the NOAA polar orbiters. Traceability to absolute standards is unknown to me. There is no on-board source, active or passive, for visible calibration in orbit. Relative response is specified as S/N of 150/1 from 0.5 percent to 100 percent albedo, and the instruments achieve approximately 250/1 S/N in prelaunch testing. The 8 individual silicon detectors are significantly more stable than the previous GOES' generation of Photomultiplier Tubes (PMTs), but they still require histogram normalization to compensate for faint striping. NOAA expects to adjust the relative visible response a few times per year.

Prelaunch calibration to determine nonlinearity in the IR detectors is performed using the vendor-built external target in thermal-vac. On-orbit infrared calibration of the Imager and Sounder uses an internal blackbody and frequent looks to space. Neither target is traceable to absolute standards. Absolute accuracy is specified to be 1 K, but that appears to be unachievable. Determination of the nonlinearity is uncertain to within uncertainties in the test set-up (approximately 1 percent). Space is used for a DC-restore once every 36.6 seconds. The internal target is maintained near 320 K, and used to determine gain approximately once every 15 minutes. Gain is determined between frames on the Imager, during frames on the Sounder. Trends in gain are monitored by NOAA. Detector-to-detector IR striping is noticeable in the noisier channels—a fraction of the single-sample noise. NOAA plans to model gain-changes between detectors and during the diurnal cycle to reduce relative error.

Regular de-icing of the detector windows on the Sounder is required to maintain acceptable performance. On the Sounder, the contaminant is unknown, with a spectrum dominated by absorption in the 9.7 micron channel. On the Imager, water-icing of detector windows was observed in prelaunch testing, but it does not occur on-orbit after the set-temperature for the window was raised.

Both instruments suffer from angle-dependent emissivity of the glass-overcoated scan mirror. The SiO_2 reststrahlung effect peaks around 8 microns, and is most noticeable in the middle infrared (6.7 to 12 microns for the Imager, 6 to 14 microns for the Sounder), where the apparent brightness of outer space varies along the east-west scan. Fortunately, the effects are predictable and measurable as a function of position and channel, and the angle-dependence is being corrected in real-time software before being broadcast to the users. The effect is only a few percent for the ±5-degree oscillation of the scan mirror about its normal degree position. The calibration look to the internal blackbody is also performed at 45 degrees. The use of the east or west-most scan positions for the space-look DC-restore is significant, and NOAA looks to the side farthest from the Sun.

Both instruments suffer from stray light and erratic calibration near subsatellite midnight, when sunlight streams into the face of the telescope. NOAA may "flywheel through" this period, rather than use obviously corrupted gain factors. NOAA performs on-orbit validation of the sensitivity and stability of the GOES sensors using the same radiometric under flights made for AVHRR.

7.11.2 Suggested Improvements and Scenarios

- The vendor's visible and infrared calibration targets should have regular validation using secondary standards traceable to NIST. MIT/LL has proposed to maintain the IR secondary standard, but they are not funded to do so, since there is no NOAA requirement for traceability.

- GOES-EAST and GOES-WEST should routinely scan the same regions under the same solar illumination during the duration of the years of operations to cross-characterize the two platforms.

- Each GOES-I/M Imager and Sounder on each satellite should routinely scan the same region of the Earth simultaneously to cross-characterize the instruments on the same platform.

- GOES-EAST and GOES-WEST should routinely scan the same regions being observed by the EOS platforms to cross-characterize all similar radiometers.

7.11.3 GOES for Global Change

The ability to determine a trend in diurnal or interannual cloud amounts becomes significant with better absolute and relative stability. NOAA's Global Change program should set the calibration/validation requirements for GOES, and a small but reliable budget established to fund calibration research and development for such a stable platform.

Can GOES achieve 2 percent uncertainty in the prelaunch absolute calibration of the visible channels? Can relative changes between detectors and on the optical surfaces be controlled? Can 1 K absolute calibration be actually achieved in the infrared?

7.11.4 Opportunity for Continuous Observations

The unique opportunity presented by GOES is the station it occupies for monitoring the changes in regional cloudiness in the moisture basins (marine stratus, Amazon, Gulf of Mexico, the central United States). A decade of continuous observation can help to discover the phase and magnitude of shifts in the diurnal cycle during ENSO events and as a result of deforestation. The observed diurnal shifts will be required to correct cloud estimations from Sun-synchronous polar orbiters.

7.12 Calibration and Validation of Spaceborne Passive Microwave Instruments for Global Change Monitoring and Studies

James C. Shiue, NASA Goddard Space Flight Center, Microwave Sensors Branch, Code 975, Greenbelt, MD; 301-286-6716, Fax 301-286-0294, email: jcshiue@meneg.gsfc.nasa.gov

The calibration and validation of spaceborne instruments for the purpose of monitoring global change and Earth science studies are of key importance and must be carefully planned and implemented to assure data quality and continuity. The dynamic range of the observable "signals" from the expected global change are rather small in magnitude, the rates of change are slow, and a meaningful time series of data set may take more than a decade to obtain. In such a long period, the data set may have to be obtained by different instruments, flying on different satellites, and on different orbits.

In some extreme cases, the same data may have to be obtained by instruments using different technologies, because of the rapid changing technology. A case in point is the stratospheric temperature, which used to be measured by the SSU which was based on infrared technology. It will be "replaced" by the upcoming AMSU-A, which is based on microwave technology.

Faced with the possibility of such diverse conditions, it is of utmost importance that the instrument and the whole process of measurement must be well characterized, so that data from different periods can be compared one to another with confidence.

The term "characterization" of an instrument here is used in its broadest sense; it includes the calibration part, which means quantitatively defining the instrument systems' response to known, controlled input signals. However, there are other aspects of the measurement process that must also be characterized. Some of them are particularly relevant for microwave sensors. These include a complete description of the spatial distribution of the radiation being entering the antenna aperture, and the way the samples of data are averaged and how the error were corrected. Different algorithms used for the correction and data reduction may result in subtle difference in the final data product. Essential elements are as follows.

7.12.1 Major Elements for Characterization of a Microwave Radiometer

7.12.1.1 ***Antenna Characteristics—Spatial resolution (IFOV) and beam efficiency.*** The footprint size or IFOV, is traditionally defined as the half-power full-beam width (HPBW) of the antenna's main-lobe radiation pattern. Only about half of the energy received by a microwave radiometer is inside the IFOV. Therefore, the knowledge of the side-lobes are also of importance, especially when the radiometer is viewing at a boundary line where the contrast of brightness temperatures between one region and another is rather high; a good example is the coastal regions between land and the ocean. The term beam efficiency, which is usually defined as the percentage of energy of a given polarization contained inside of a "main beam" is an important figure of merit defining the spatial character of the antenna.

7.12.1.2 ***Polarization***. The type of polarization (e.g., vertical or horizontal, or right or left circular), and its purity must be fully characterized. Polarization purity is important for measurements where there are large differences in brightness temperature between the vertical and horizontal polarizations. This is true for sea-surface measurements, because the emissivities are quite different between the vertical and horizontal polarizations, especially at large incidence angles. For example, cross-polarization (CP), defined as the percentage of the orthogonal polarization energy leaked into the main polarization, must be adequately corrected to yield the needed accuracy for SST measurement.

Similarly, microwave temperature sounders designed for measuring the stratospheric temperatures may depend on the purity of a single circular polarization, and the cross-polarization (usually known as the axial ratio) must be fully characterized.

7.12.1.3 Dependency of Antenna Characteristics on Beam Position. Most microwave sensors are of scanning type (i.e., its antenna beams can be scanned from one position to another). Depending on the method of scanning, the characteristics of the antenna beam may change as it scans. For example, an electronically scanned (phased array) antenna will change its beam width as it scans. Furthermore, the side-lobe shape, and even the reflection coefficients of the antenna may change during the scanning. These must be fully characterized as well.

7.12.1.4 Scanning Method. Scanning may modify the effective footprint size or the effective field of view (EFOV), therefore the method of scanning must be characterized. When the scanning is effected by "continuously slewing" an antenna mechanically (at a constant or variable angular velocity), the EFOV is then determined by the length of the integration time. In this case the EFOV is a convolution of the antenna IFOV, and its scan motion and the details of the scan motion is needed to fully characterize the resultant EFOV.

7.12.1.5 Beam-Pointing Errors and Co-registration. Pointing errors (control, knowledge, and stability), relative to spacecraft and spacecraft's own altitude errors must be fully characterized to interpret the data properly. The pointing errors not only affect the location of the beams, it can also change the incidence angle, which affect the surface emissivity (e.g., the SST measurement). Furthermore, if there are multiple beams, the relative position errors must also be known. If the beams are not physically all pointing at the same (coincidence) direction and is relying on some timing techniques to "reorganize" the EFOVs in the data package so that data from different frequency channels will be co-registered in clusters (for multifrequency comparison), they must be fully characterized in the data description.

7.12.1.6 Blockage by Other Sensors or Spacecraft Structure. The presence of objects (other instruments or spacecraft structures) in the near field zone of the antenna may affect its measurements. This is particularly important when the antenna is viewing the "cold space" for calibration. The effect of a "warm" object in the near field may modify the effective cosmic background brightness temperature (nominally 2.7 K) to a different value. If the presence of such objects in the vicinity of the antenna beam cannot be avoided, its effects must be fully evaluated.

7.12.2 Radiometer Characteristics

In addition to the spatial characteristics which are mostly related to the antenna radiation patterns, the characteristics of the radiometer (electronics) itself must be characterized. These include the method of calibration, its residue errors expected, and spectral characteristics of the passing bands.

7.12.2.1 Calibration. The method of calibration and its residue errors (systematic errors) from the prelaunch calibrations must be fully documented. It is also important to characterize the radiometer's in-orbit calibration, (external, internal, noise-injection, modulation type, or total power, etc.) and the degree of "averaging" being done to the calibration measurements. The stability of the calibration system must also be evaluated.

Frequently, an on-orbit, two-point (one hot and one cold point) external calibration system is used for microwave radiometers in which two targets with known brightness temperatures are measured from time to time. The hot point is usually a black-body target of unity emissivity and the cold point is usually the cosmic background microwave emission. With great care in design and implementation, this type of calibration system yields the best stability and smallest systematic error. For a linear system, the two-point external calibration removes all systematic errors in the time interval between the calibrations. This type of calibration should be used whenever possible.

7.12.2.2 ***Sampling Intervals***. Sample intervals both along the in-scan direction and along the down-track directions must be specified together with the data package; the antenna beam patterns are needed to allow the data users more flexibility in further corrections and resampling.

7.12.2.3 ***Bandpass (Spectral) Characteristics***. Band-center frequency, bandwidth, and their stability must be characterized. Moreover, the bandpass transmission characteristics (e.g., in-band "ripple" and out-of-band roll-off characteristics) must also be precisely described. The bandpass characteristics are of particular importance for microwave temperature sounders which are generally narrowband, and several channels may be crowded together near the peak of a resonant line so that the transmission characteristics of each channel must be fully characterized for the retrieval purpose. In addition, any potential Radio Frequency Interference (RFI), either from other instruments aboard the same host spacecraft or from certain areas of the Earth surface must be evaluated to ensure data quality.

7.12.2.4 ***Temperature Sensitivity***. Temperature sensitivity (alternatively known as the Noise Equivalent Delta-T, or NEdt)) is a measure of the noise character of the instrument (random part of the error). It is generally expressed in (experimental) standard deviation of the brightness temperature (or radiance). However, since it is a function of the length of the integration time, the integration time (or its equivalent EFOV) must be specified when quoting the numerical value of the temperature sensitivity of a radiometer. Furthermore, the temperature sensitivity should include all of the random errors of the radiometer throughout the complete chain of the radiometer system (e.g., the noise of the amplifiers and detectors, digitization errors (if any), cross-talk (leakage between channels in a multiplexed data system), and the noise introduced in the calibration, etc.).

7.12.2.5 ***Nonlinearity and Dynamic Range***. For radiometer systems where some slight nonlinearity cannot be avoided, the nonlinearity (within the dynamic range) must be fully characterized by prelaunch calibrations in a simulated operating environment; the test data can then be used in conjunction with the on-board calibration to correct for the nonlinearity.

7.12.3 Validation

It is best to calibrate each microwave radiometer so that its output is precisely related to the input (radiance or brightness temperature). However, some of the accumulated biases would be very difficult to completely remove. Therefore, the postlaunch validation is an important part of the complete calibration/validation process.

The validation method can include: (1) comparing the spaceborne measurement with known geophysical parameters obtained with *in-situ* measurements, the so-called "ground truth," (2) comparing the spaceborne data with aircraft measurements with the same or similar instruments, and (3) comparing one satellite instrument to another (with the same or similar instrument) from a different platform. In general, items (1) and (2) require well-designed field campaigns or ground truth networks.

Comparing two spaceborne instruments is a convenient technique. To maintain data uniformity needed to monitor climate change signals, it is mandatory to perform such intercomparisons (cross-calibration) between different instruments launched in time sequences. It would be most beneficial if there were at least a few months of overlap in time between two successive satellite missions and that the two orbits intersect so that a given area can be measured by both instrument at the same time.

The SSM/I on the DMSP satellite and the MSU on the TIROS and NOAA series satellites are good examples of a series of the same or similar microwave instruments. It is important that in the future when the AMSU-A (which has many similar channels as the MSU) is launched on NOAA-K (tentative launch date: mid-1996) to replace the MSU, a comparison of the AMSU-A data be made against that of the MSU. Such cross-calibration is of utmost importance for long-term climate monitoring and studies. However, such cross-calibrations will result only with careful planning and coordination.

7.12.4 Recommendations for Improvements in Calibration and Validation

During the last 2 decades, passive microwave (radiometers) sensors have become an increasingly important part of the spaceborne suite of sensors for quantitative Earth observation purpose. As the demand for high accuracy (small residue errors) and stability continues, more attention must be placed on careful characterization (including instrument calibration) of the microwave instruments. This timely workshop is a reflection of the fact that the importance of calibration is being recognized by the top agency management.

7.12.4.1 A Different and Better Instrument Development Methodology. Currently, most of the instrument calibration task is an integral part of the hardware procurement package, and the engineering firm which develops the instrument is responsible for the instrument calibration. While this mode of operation has its advantages in that the finished product is calibrated and ready for acceptance, the disadvantage is that often the calibration work has to be squeezed through between tight assembly line time schedules and even tighter budget constraints. Consequently, only "routine" calibration methods are used, because the "assembly line" environment is not conducive to methodical calibration and any new innovations, which may carry some risk.

A better mode of instrument development would be to establish a separate calibration task, as part of a government laboratory, to conduct the calibration task completely independent of the instrument developer, and to include it as part of the acceptance process for the instrument. This would completely alleviate any potential conflict of interest of the instrument developer who is often under pressure to keep within the budgetary and schedule constraints. It would also provide opportunities for the government laboratory to accumulate more experience and expertise in calibration technologies.

It is believed that if the calibration task is assigned to a government laboratory, the cost of instrument development would actually be lowered, because only a single calibration facility is needed for a series of similar instruments, regardless of where a given instrument is being developed. Currently each vendor must set up a calibration facility which is a costly investment. In addition, the government laboratory should also conduct research and development to explore new calibration technologies preparing them for future needs.

7.12.4.2 Support Needed for Developing New Calibration Technology. A number of ideas to explore new techniques for improving microwave radiometer calibration methodology have been proposed in the past years, but have yet to find any funding support. They include: (1) improving design of on-board warm load, (2) developing an on-board cold reference target to replace the cosmic background radiation as the cold reference, (3) developing techniques for evaluating the effect of near-field object and interference, (4) a radiometric method for beam efficiency measurement, (5) a new method for nonlinearity evaluation, and (6) monitoring long-term drift of on-orbit calibration. It is our hope that this workshop will stimulate more interests and may be some concrete actions that will result in new directions for instrument development and better calibration and validation.

7.13 Precipitation Radar (PR) Data and Their Calibration and Validation in the Tropical Rainfall Measuring Mission (TRMM)

K. Okamoto, Communications Research Laboratory, Ministry of Posts and Telecommunications

7.13.1 TRMM Precipitation Radar (PR) Data

TRMM is a joint project between the U.S. and Japan to measure the rainfall rate around the tropical and subtropical zone. The TRMM observatory is scheduled to be launched in August 1997, and the mission lifetime is considered to be 3 years. The TRMM precipitation radar (PR) will be the first spaceborne radar of its kind. This lack of heritage requires TRMM scientists to develop the TRMM PR algorithms and to consider calibration and validation methods from first principles. However, as the observation continuation of the tropical or the subtropical rainfall rate is quite important and is expected from the international climate program such as GEWEX, the mission to follow on TRMM is being studied in both countries.

Table 13 summarizes the standard products of TRMM PR. The products derived from radar data are classified into several processing levels (i.e., Level 1B, 1C, Level 2A, and Level 3A products). These products are also classified into qualitative and quantitative data.

Table 13. Summary of the Standard products of TRMM PR

Level	Qualitative Data	Quantitative Data
Level 1B	Rain/No Rain	Total Received Power, Noise Level, Storm Height
Level 1C		Profiled Zm
Level 2A	Bright Band (Yes/No) Rain Classifier (Stratiform, Convective Type)	Bright Band Height Storm Height $\sigma^\circ R$, Averaged σ°_{NR} (Land/Ocean) Profiled R (IFOV Basis) Path Averaged R (IFOV Basis)
Level 3A		Averaged Rain Parameters Over Space-Time Domain

Before extracting quantitative information, there are a number of qualitative descriptors of the storm. The most obvious question is whether rain is present along the radar beam. When a detectable signal is present, there are several issues that can be addressed. One of the most important is whether a "bright band" is present, and, if so, the height at which it appears. The bright band height is particularly important for the microwave radiometer TMI that will be installed on TRMM with the PR. TRMM will rely primarily upon the PR to determine the rain type classification that includes stratiform, convective, and so-called warm rain. The most detailed information available from the radar consists of the range-profiled rain rates R over each resolution cell (IFOV). Less ambitious goals include the estimation of path-averaged R. Other quantitative information includes surface cross-section data σ°_R when rain is present in the IFOV, a library of averaged surface cross-section data σ°_{NR} over both land and ocean when rain is absent, and averaged rain parameters over the space-time domain of 5° x 5° x 1 month.

7.13.2 Calibration Plan of TRMM PR

In this section, we discuss the overall plan of the PR calibration with an emphasis on the calibration plan after launch. Here we define "calibration" and "validation" as follows: the term "calibration" comprises those activities based solely upon an engineering consideration and performed routinely by the National Space Development Agency of Japan (NASDA) or implemented into the Level 1 processing. On the other hand, "validation" consists of those activities which are used to ensure the quality of the radar products. The validation plan will be discussed in a separate section of this paper.

7.13.2.1 ***Modeling of PR Parameter Variations***. Accurate calibration of the PR is important to establish the clear interface condition between Level-1 and higher level algorithms, thereby assuring accurate and stable rain products. To perform the PR calibration, variation and drift of the PR system parameters have been modeled to have "intermediate-term" and "long-term" components. The former is caused by the temperature change inside the PR and roughly has a period of one revolution of the satellite (about 91 minutes). Thus, the correction for this term can be done by monitoring the temperatures. The latter may occur due to gradual degradation of system performance (gain, loss, etc.) and/or failure of some active array elements. Since this term may involve changes in antenna characteristics and monitoring sensors, calibration using an external reference target is required.

7.13.2.2 ***Prelaunch Activities***. In the PR protoflight model development phase, a large volume of data will be obtained to establish the data base for the postlaunch PR calibration. The data set will include the temperature dependence of various parameters (e.g., gain, loss, phase) in the PR and the measured antenna pattern. Also PR sensitivity will be verified by using the Active Radar Calibrator (ARC) of the Communications Research Laboratory (CRL). These activities will serve as the baseline for the postlaunch PR calibration and validation.

7.13.2.3 Postlaunch Activities.

7.13.2.3.1 Internal Calibration. The internal calibration has been developed using a detailed PR system model which describes the temperature dependence of all system parameters related to the conversion process from the count value to the radar-received power or to the radar-reflectivity factor. The error analysis presented at the PR system Critical Design Review (CDR) indicates that an error of less than 1 dB can be achieved in the estimation of the radar reflectivity factor. The internal calibration using the temperature telemetry will be implemented in the Level 1 processing algorithms.

7.13.2.3.2 External Calibration and Internal Loop Calibration. External calibration of the PR will be performed using an ARC placed at a ground calibration site in Japan. The ARC will have three functions: radar transponder, radar receiver, and beacon transmitter. To reduce the errors caused by the uncertainty of PR antenna beam pointing, a special oversample antenna scan will be used in the PR external calibration mode. ARC echo levels obtained from the multiple beam directions allow a precise estimation of PR antenna pointing and "peak" ARC echo level corresponding to the PR antenna beam center position. One problem in the ARC calibration is that an ARC calibration can provide the calibration factor only at a specific angle bin. The internal loop calibration is performed by invoking the PR internal calibration mode, which is intended to measure the overall input/output (I/O) characteristics of the PR receiver IF and data processing units. Since this requires the interruption of science observations, this calibration would be performed in series with the external calibration using the ARC. The external calibration and the internal loop calibration will be performed every 2 to 4 weeks by NASDA. The results will be accumulated to monitor the long-term trend of the PR system parameters. Updates of PR calibration factors will be based upon a statistical analysis of the trend data.

7.13.2.3.3 Antenna Pattern Measurement. Postlaunch measurement of the PR antenna pattern are important to assess the in-orbit performance of the PR. The measurement of the along-track pattern is relatively easy because the time trend of the ARC received power and of the PR received power in the normal ARC calibration can be used to generate the transmit pattern and 2-way pattern, respectively. Similarly, the PR receive pattern can be obtained using the ARC beacon mode. On the other hand, measurement of cross-track (antenna scan plane) pattern is much more difficult in spite of the importance to assess the overall amplitude/phase stability of 128 active array elements. To make this measurement possible, a special spacecraft attitude (90° yaw maneuver) will be employed, in which the time trends of the received powers in the ARC calibration now provide the antenna scan plane pattern. The cross-track pattern measurement using the 90° yaw maneuver will be conducted at least once in the initial checkout period and about once per year after that (to be reviewed).

7.13.3 Validation Plan

Three kinds of validation methods for TRMM PR data are now considered. The first and most fundamental method is the direct comparison between the PR measurements and the data gathered by gound truth and validation (GV) activities. Of greatest interest are well-calibrated ground-based radar data (Z-value is the most straightforward physical value), rain rates measured by the rain gauge networks, and raindrop size distributions measured by the disdrometers. The data which are closest in distance and time to the TRMM overpass are the most useful for the validation of the TRMM PR IFOV data products while the monthly averaged rain rate at some GV sites are good validation for the monthly averaged PR products.

A second type of validation is the use of airborne radar data. Airborne measurements are useful especially over the ocean because of the difficulty in acquiring data by other means. Both NASA/JPL and CRL have airborne rain radars that operate at the same frequency as the TRMM PR.

The third method is to utilize the well-known scattering coefficient data of natural targets such as the backscattering coefficients of ocean surface or the Amazon rain forest. In addition, the existing and planned satelliteborne scatterometer data (such as TOPEX data at 13.6 GHz, and ADEOS/NSCAT at 14 GHz) will also be useful. In the following section, a more detailed introduction to the TRMM PR validation plan is given.

7.13.3.1 ***Ground Truth and Validation Data Prepared by the TRMM GV Activities.*** The TRMM PR data requests that the GV groups of both the U.S. and Japan develop at least one well-calibrated ground-based radar (error of less than ±1 dB), and rain gauge and disdrometer networks for both the instantaneous and statistical validation. Possible candidates are the NEXRAD at Melbourne, Florida and the dual-polarized Doppler radar at Darwin, Australia. In Japan, the southernmost radar of JMA at the Ishigaki-jima Island, can perform one 4-minute volume scan every hour on the hour, and will provide very good validation data over the ocean. Radar-AMeDAS composite maps from JMA (at the altitude of 1.5 km) will be useful. As the PR sampling rate is high at latitudes near 35°, field campaigns in Japan near the latitude of 34° to 35°, such as the MU radar site near Kyoto may be useful.

7.13.3.2 ***Airborne Radar.*** An important component of the validation plan for the TRMM PR is the use of airborne rain radars which will underfly the TRMM PR. At present, there are two such airborne radar systems that are ideally suited for the validation experiments. They are the Airborne Rain Mapping Radar (ARMAR) from NASA/JPL and the Multiparameter Radar (CAMPR) from Japan/CRL. These radars were developed with the same operation frequency and downward-looking geometry as the TRMM PR. These radars provide several unique data sets for the validation of the TRMM PR.

- The direct reflectivity measurements from the surface and rain cells can be compared with the TRMM PR results. For regions that are not raining, the surface cross sections as measured by the TRMM PR can be compared with the airborne radar results after appropriate spatial averaging. This will give additional calibration information for the TRMM PR. The airborne radar reflectivity measurements from raining regions can also be compared with the TRMM PR since the radars operate at the same frequency and, with appropriate spatial averaging, should show similar reflectivity and attenuation characteristics as the results from the PR at least for nadir-pointing observations. This step can provide information for the validation of the TRMM PR Level 1 data products.

- In areas where remote and conventional ground or shipborne *in-situ* measurements are not available, the data from these airborne rain radars can serve as a validation data source for rain retrieval process. Specific examples of these areas include the oceans and remote land regions. Since both of these airborne radars provide additional measurements beyond the basic radar reflectivities (such as Doppler measurements, radiometer brightness temperatures, multipolarization measurements, etc.) more sophisticated algorithms can be applied to the airborne data sets to retrieve rain rates at higher accuracy. The rain rates can then be used to validate the TRMM PR-derived rain rates in these remote areas that do not have other *in-situ* measurements. Of course, this process would require that the results from the airborne rain radars be cross-calibrated in some ways, and be validated in experiments over areas with well-calibrated rain measurement instrumentation prior to the TRMM PR validation phase.

- Since the airborne radar results are typically obtained with spatial resolutions higher than those of the TRMM PR, the results can also be used to evaluate the effects of nonuniform beam filling on the TRMM PR rain retrieval process at least for nadir-pointing observations. The rain retrieved at the higher resolutions from the airborne radar results can be averaged and then compared with the rainfall retrieved from the PR. Any systematic error obtained from these intercomparisons can be used to adjust the TRMM PR rain products.

7.13.3.3 ***Natural Target.*** Within the planned TRMM satellite coverage, the TRMM PR will make frequent observations over a number of natural targets whose radar cross-sections are spatially homogeneous and temporally stable. Because of these key characteristics, such natural targets can be used to validate the PR rain products. In this section, we will describe those natural targets currently being identified for PR validation.

7.13.3.3.1 *Ocean Surface at 10° Incidence.* It has been well-established through theoretical modeling, airborne experimentation, and SEASAT scatterometer data (Schroeder et al., 1982) that the rain-free normalized radar cross-section of the ocean surface is insensitive to the sea state and the wind condition at incidence angles of approximately 10°. At the PR frequency of 13.8 GHz, the experimental results obtained by the NASA/JPL ARMAR during the TOGA Coupled Ocean-Atmosphere Response Experiment (COARE) field campaign show a statistical mean of 8.5 dB and a standard deviation of ±0.5 dB over a 2-month period. Since the TRMM PR will make continuous observations over a

±17° scan throughout the entire mission, the rain-free ocean backscatter measurements collected at 10° incidence by the PR can be used to monitor the stability of the radar system and to determine the radar calibration constant.

7.13.3.3.2 Amazon Rain Forest. Because it is homogeneous and its normalized radar cross-section is relatively insensitive to the change in incidence angle, the Amazon rain forest has frequently been used as a primary external calibration target for many spaceborne radar missions. Using the SEASAT scatterometer data set, Birrer et al. have demonstrated that the normalized 14.6 GHZ radar cross-sections of the rain forest at incidence angles between 20° and 70° are quite uniform (with standard deviations of ~±0.5 dB) at a specific local time. Since the Amazon rain forest will be visited frequently by the TRMM satellite, it can be used as an external target to determine the long-term drift of the TRMM radar system parameters.

7.14 Long-Term Tropospheric CO and CH_4 Measurements: MOPITT/AM-1 and Beyond

Jinxue Wang, Paul Bailey, and John Gille, National Center for Atmospheric Research (NCAR), Atmospheric Chemistry Division (ACD), PO Box 3000, Boulder, CO

MOPITT, a gas correlation radiometer scheduled to be launched on the EOS/AM-1 platform in 1998, will provide column amounts of CO and CH_4 and profiles of CO in the troposphere (lower, middle, and upper troposphere). In this paper we will discuss our plan for the development of a long-term data set of tropospheric CO and CH_4 from MOPITT/AM-1, possible MOPITT follow-ons, and other EOS instruments. We will examine the availability of potential CO and CH_4 measurements from EOS instruments and its impact on the feasibility and quality of a long-term CO and CH_4 data set.

7.14.1 Introduction

One of the important objectives of the global change study is to assess the impact of human activities on the composition of the atmosphere. The troposphere, which contains about 85 percent of the total mass of the atmosphere, is a region of great chemical diversity and activity. It contains many chemically and radiatively important trace gases produced by natural and anthropogenic processes at the surface and within the troposphere. CO is a key gas regulating the oxidizing power of the troposphere. It was reported that the concentration of tropospheric CO has been increasing at a level of 1 percent/yr mainly due to increased human activities (Khalil and Ramussen, 1988). The rate of increase and the full range of the effects of the increased concentration of CO is not fully understood at the present time, but it is believed that CO is photochemically active and plays a major role in controlling the concentration of OH radicals in the troposphere. Increased CO may deplete tropospheric OH radicals, thereby reducing the yearly removal of many natural and anthropogenic trace species. In particular, this effect may add to the increase of CH_4, which in turn could further reduce OH concentration. CH_4 is a very effective greenhouse gas, and it may be second only to CO_2 in causing global warming. CH_4 also affects the oxidizing capacity of the atmosphere by removing tropospheric OH radicals and creating O_3, and it affects the ozone layer in the stratosphere by contributing water vapor and removing chlorine atoms.

Up to now there have been only limited measurements of CO and CH_4 by some ground stations, balloon flights, and the Measurement of Air Pollution from Satellites (MAPS) experiment (Reichle, 1986, 1990) on the Space Shuttle with very different techniques and instruments. It is extremely difficult to assemble a high quality long-term global tropospheric CO and CH_4 data set with those measurements only. MOPITT will provide the first continuous global data set of tropospheric CO and CH_4 to construct a long-term record of tropospheric CO and CH_4 measurements to be used in global tropospheric chemistry and global change study.

7.14.2 Experience with Long-Term O_3 Data from SBUVand SBUV/2 Measurements

The experience with the analysis of O_3 data from SBUV and SBUV/2 measurements from 1978 to 1994 by Bill Randel at the National Center for Atmospheric Research (NCAR)/ACD will be useful for us to construct a long-term tropospheric CO and CH_4 data set. A brief description of the technique used and the long-term O_3 data set is included here. More details can be found in an NCAR technical report (Randel and Wu, 1995).

In the analysis, data from Nimbus-7 SBUV (covering November 1978 to June 1990) and NOAA-11 SBUV/2 (January 1989 to April 1994) are combined into a continuous time series. The SBUV data are adjusted to match the SBUV/2 measurements, based on comparisons during the overlap period when both instruments were operational (January 1989 to June 1990). The quality of this long-term ozone data set has been studied by extensive comparisons with other satellite ozone data, such as TOMS data for column ozone and LIMS, SAGE II, Halogen Occultation Experiment (HALOE), and Microwave Limb Sounder (MLS) for ozone profiles. They showed that there is a consistent difference or bias of about 5 to 10 percent between SBUV and SBUV/2 data set and data sets from other satellite instruments. Actually, there is a difference of about 5 to 10 percent between any of the two data sets involved in the intercomparisons.

If the findings with the SBUV and SBUV/2 long-term O_3 data set can serve as an indication of the quality of a long-term data set from similar/different techniques and instruments, it will be difficult , if not impossible, to construct long-term tropospheric CO and CH_4 data set with data from different instruments with an accuracy better than 5 percent. Biases of 5 percent might be acceptable for CO measurement, even though undesirable, as the expected accuracy of MOPITT CO measurements is about 10 percent, but the 5 percent biases are not acceptable for column CH_4 measurements. The concentration of CH_4 varies over only a ±5 percent range, mainly due to interhemispheric differences. To investigate variations and their causes, a precision of <1 percent is required. The requirement of MOPITT column CH_4 measurement is 1 percent precision. The findings with the SBUV-SBUV/2 long-term O_3 data also indicate: (1) it is important to have overlapping periods of operation between different instruments for data validation and possible adjustments; (2) using the same or similar instruments consistently for the measurement of the same geophysical parameter will enhance the quality of the long-term data set; (3) it is important to have other satellite measurements for the validation of a long-term data set; and (4) it is useful to have ground-based instruments that can be calibrated periodically for satellite validation.

7.14.3 Long-Term Tropospheric CO and CH_4 Data Set with MOPITT/AM-1 and Follow-ons

In this section we will discuss our plan and strategy to construct a long-term (~10 to 15 years) tropospheric CO and CH_4 data set based on the assumption that there will be follow-ons to MOPITT/AM-1, that is there will be at least two more MOPITT in addition to MOPITT/AM-1 to provide CO and CH_4 measurements for a period of ~15 years. As discussed before, using consistent techniques and instruments will provide the best opportunity to construct a high quality data set for long-term trend detection.

MOPITT is an 8-channel IR radiometer using both pressure-modulation and length-modulation radiometry (Drummond, 1992). Characteristics of each channel are summarized in Table 14. More details about MOPITT and the data processing algorithm can be found in the MOPITT Algorithm Theoretical Basis Document (ATBD) (MOPITT ATBD, 1995). The objective of MOPITT CO measurement is to obtain profiles with a 3- to 4-km vertical resolution and an accuracy of 10 percent throughout the troposphere. A CO total column amount measurement will also be made with a 10 percent accuracy. For CH_4, the objective is to measure the column in the troposphere to a precision of better than 1 percent. Based on signal-to-noise and better chances of finding clear field-of-view, the horizontal resolution of MOPITT is selected to be 22 km. The column amounts of CO and CH_4 will only be available on the sunlit side of the orbit because solar absorption measurements are needed for the measurements. There is not enough thermal contrast between the lower atmosphere and the Earth surface to derive CO and CH_4 amounts in the boundary layer from thermal emission/absorption measurements alone. Since the atmospheric CO and CH_4 only change about 0.5 to 1 percent per year, and the total change will be less than 5 percent over the life of the MOPITT/AM-1 mission, careful prelaunch and postlaunch instrument calibration/characterization, data validation and intercomparisons are needed to ensure data quality. Our strategy to construct a long-term tropospheric CO and CH_4 data set is summarized below:

- Carry out careful prelaunch instrument model and retrieval algorithm validation. Conduct detailed prelaunch instrument characterization and calibration. A MOPITT Algorithm Test Radiometer (MATR) is under development at NCAR/ACD. It will be flown on an aircraft for the verification and test of correlation radiometer models and retrieval algorithms.

- Monitor the performance of the MOPITT engineering model. Conduct periodic recalibration of the engineering model to obtain data for the assessment of in-orbit instrument behavior.

- Perform frequent in-orbit instrument radiometric calibration.

- Have significant overlap periods between MOPITT/AM-1 and the follow-on MOPITT so that intercomparisons can be conducted.

- Process data sets from the different versions of MOPITT with the same data processing algorithm. If there is modification to the algorithm at anytime, reprocessing of prior data must be conducted with the improved algorithm.

- Conduct comprehensive data validation with CO and CH_4 measurements from ground-based instruments, airborne instruments, and other satellite instruments. We plan to use ground-based CO measurement network similar to that used during the validation of MAPS CO measurements. There is also discussion about the development of an airborne version of MOPITT (called MOPITTA) to underfly EOS/AM-1 and measure tropospheric CO and CH_4 for satellite data validation. MOPITTA will be re-calibrated periodically in the same calibration facility and with the same calibration targets as the spaceborne MOPITT.

Table 14. MOPITT Channel Characteristics

Characteristics \ Channel	1	2	3	4	5	6	7	8
Gas Species	CO	CO	CO	CH_4	CO	CO	CO	CH_4
Nominal Gas Pressure (KPa)	20	20	7.5	80	80	80	3.8	80
Mid-Wavenumber (cm-1)	2166	4285	2166	4430	2166	4285	2166	4430
Wavenumber Range (cm-1)	52	40	52	139	52	40	52	139
Mid-Wavelength (μm)	4.617	2.334	4.617	2.258	4.617	2.334	4.617	2.258
Wavelength Range (μm)	0.111	0.022	0.111	0.071	0.111	0.022	0.111	0.071
Modulator Type	LMC	LMC	PMC	LMC	LMC	LMC	PMC	LMC
Note: LMC stands for length modulation cell, and PMC stands for pressure modulation cell.								

7.14.4 Long-Term Tropospheric CO and CH_4 Data Set with MOPITT/AM-1 and Other EOS Instruments

If there are no follow-ons to MOPITT/AM-1, we will have to rely on other EOS instruments to extend the 5 to 6 years of CO and CH_4 data from MOPITT/AM-1 to 10 to 15 years. Several other EOS instruments could potentially provide CO and CH_4 measurements, but over different altitude ranges of the troposphere and with different projected accuracies. Based on the current baseline EOS mission profile, the instruments, time scale and overlap, and altitude ranges of CO and CH_4 measurements are summarized in Table 15. MOPITT is to be launched in June 1998 with a nominal in-orbit life of 5 years, with a goal of 6 years. CO profiles and column amounts of CO and CH_4 will be available from 1998 to 2003/2004.

EOS PM-1 will be launched in 2000. AIRS is the only instrument on PM-1 that might be able to provide column CO data as a research product, not a standard Level 2 product. The CO column is to be derived from AIRS measurement in the thermal band of CO at 4.7 μm. It was claimed that an accuracy of 10 percent might be achievable, based on retrieval simulation with the assumption of the knowledge of CO profile shape (Strow, 1994). Without assuming CO profile shape, there is not enough information from AIRS measurements in the CO thermal band at 4.7 μm to derive boundary CO amounts due to the very small thermal contrast between surface and the boundary layer. In many situations, the CO profile shape might be quite different from the assumed CO profile used in the retrieval. Therefore one must be cautious in using CO column amounts derived in this manner. Since there are about 2 years of overlap between MOPITT and AIRS, there will be opportunities for intercomparisons of CO column amounts from MOPITT and AIRS.

Table 15. EOS Instruments with CO and CH_4 Measurement Capability

Platform	Instrument Measurement	Time Scale (98 '99 '00 '01 '02 '03 '04 '05 '06 '07 '08 '09 '10 '11)	Projected Accuracy	Altitude Range
AM-1	MOPITT			
	CO, profile	◄── 98 – '03 ──►	10%	0 - 15 km
	CO, column	◄── 98 – '03 ──►	10%	
	CH_4, profile			
	CH_4, column	◄── 98 – '03 ──►	1%	
PM-1	AIRS			
	CO, profile			
	CO, column	◄── '00 – '05 ──►	10%(1)	
	CH_4, profile			
	CH_4, column			
CHEM-1	HIRDLS			
	CO, profile			
	CO, column			
	CH_4, profile	◄── '02 – '07 ──►	1-10%	10 -70 km(2)
	CH_4, column			
	TES			
	CO, profile			
	CO, column	◄── '02 – '07 ──►	?	
	CH_4, profile			
	CH_4, column	◄── '02 – '07 ──►	?	
PM-2	AIRS			
	CO, profile			
	CO, column	◄── '06 – '11 ──►	10%(1)	
	CH_4, profile			
	CH_4, column			

(1) CO column is potential research product of AIRS. The accuracy of 10% is based on current simulation studies (*Strow*, 1994).

(2) Depending on the cloud coverage and aerosol loading in the upper troposphere, the HIRDLS may extend to below 10 km under clear sky and low aerosol situation, it also may not be able to extend to 10 km under cloudy sky or high aerosol conditions.

EOS CHEM-1 is scheduled for launch in December 2002. Two instruments, the High Resolution Dynamics Limb Sounder (HIRDLS) and TES, will provide some CO and CH_4 measurements. HIRDLS, a high-resolution limb sounder, will be able to provide CH_4 profile data for the upper troposphere, depending on whether there are clouds in the upper troposphere. At this time, HIRDLS CH_4 measurements are projected to be about 1 to 10 percent accurate (Gille and Barnett, 1992). TES, a Fourier transform spectrometer, could potentially provide column amounts of CO and CH_4. Accuracy of TES CO and CH_4 column amount retrieval are unknown at this time. Similar to AIRS, the CO and CH_4 column amounts are to be derived from TES measurements in the thermal bands of CO and CH_4, and there are questions whether the boundary layer CO and CH_4 can be derived from thermal measurements alone due to the lack of thermal contrast between surface and the boundary layer. If the boundary layer CO and CH_4 amounts could not be derived from thermal measurements, then the CO and CH_4 column amounts derived from AIRS and TES are those above the boundary layer. There will be about a 6-month overlap between HIRDLS/TES and MOPITT and about a 3-year overlap between HIRDLS/TES and AIRS. Intercomparisons among those 4 instruments can be conducted for about 6 months, and intercomparisons among HIRDLS, TES, and AIRS can be conducted for about 3 years. Since MOPITT is the only instrument that will provide CO and CH_4 amounts from solar absorption measurements, it is important to validate the

CO and CH_4 column amounts data from AIRS and TES against MOPITT measurements, and the 6-month overlap might not be enough. Ideally, there should be more than a 1-year overlap to have data from all seasons. The situation would be much better if there is a MOPITT follow-on on the AM-2 platform.

Based on the above discussion, MOPITT is the only instrument that will provide CO profile data for lower, middle, and upper troposphere from 1998 to 2003/2004. HIRDLS will be able to provide upper troposphere CH_4 data from 2003 to 2007/2008. Therefore, there will be no global measurement of CO profiles beyond 2003/2004, and no information on the upper troposphere CH_4 beyond 2007/2008. MOPITT will provide column amounts of CO and CH_4, including the boundary layer, from 1998 to 2003/2004. Beyond 2003/2004, AIRS and TES might be able to provide column amounts of CO and CH_4 above the boundary layer. Therefore, there will be no data on CO and CH_4 in the boundary layer beyond 2003/2004. There will be no measurement of CH_4 column amount beyond 2007/2008. Beside problems with data availability, careful intercomparisons between CH4 instruments (MOPITT, AIRS, HIRDLS, TES) are essential to construct a long-term data set of tropospheric CO and CH_4. The best scenario is that a long-term data set of tropospheric CO and CH_4 can be constructed from measurements by those CH_4 instruments. Whether the long-term data set constructed this way will be good enough for trend detection is unclear at this time. It seems that MOPITT follow-ons are needed in order to construct a high quality long-term tropospheric CO and CH_4 data set for tropospheric chemistry and global change study.

7.15 Calibration and Validation Requirements for a Global Change Ocean Color Data Set

Curtiss Davis, Code 7212, Naval Research Laboratory, Washington, DC

The goal is to make a decades long data set of ocean optical properties and phytoplankton biomass and productivity to assess the impact of global change on the ocean food web and to assess the effect of changes in energy absorption, CO_2 uptake or DMS production by the marine food web which may dampen or accelerate global change. The observable changes will be in the range of a few percent per decade, and hence to observe those features requires a data set that is reasonably accurate and highly consistent. However, the global data set must be pieced together from data from a series of international satellites each with a unique design and with its own calibration and validation procedures. To achieve the required combined data set requires excellent calibration and validation, reference to a common standard and standardized processing to produce one or more products for the decades long time period which can be used to assess global change.

To assess the precision of a measurement requires that the errors associated with each step in the measurement process be propagated to the final product. In addition, an accurate product requires reference to a known standard, and the errors associated with the knowledge of that standard must be added to the errors associated with the measurement and the calculation of a product from that measurement. For example, the present plan is to validate SeaWiFS data in terms of measured radiances. This measurement can be traced to NIST standards from the preflight calibration, and checked with Moon and solar diffuser measurements during flight. The goal of ±5 percent for the measured radiances during the 90-day checkout period is difficult but achievable. Maintaining that accuracy for the duration of the 5-year lifetime is more difficult, and depends primarily on the Sun and Moon measurements and on vicarious calibrations (comparison with atmospherically corrected ground measurements).

An additional source of error is added when it is necessary to use data from more than one sensor. There are the errors associated with the calibration of each sensor, and with the comparison of the calibration scales used for those calibrations. For example, many of the National Standards Laboratories recently did a comparison of their radiance and irradiance standards. The range was ±2 percent for the visible and considerably larger in the UV and IR. The sources used for the calibration of SeaWiFS and other ocean color sensors are secondary or working standards derived from those primary standards and carry the error of the primary standards and additional errors associated with the transfer of calibration from the primary standards to the working standards. To minimize or avoid these additional errors, instruments whose data are to be used for long-term data sets should be calibrated against the same standards. The ideal scale is the Sun as the information about the ocean is contained in the absorption and scattering of sunlight by the constituents in the ocean and the water itself and using the Sun avoids errors that are associated with transferring calibrations from laboratory sources which have a different spectral shape. However, calibration using the Sun requires the use of a diffuser or screen to reduce the signal to the levels comparable to ocean radiances. The biggest concern with this approach is the degradation of the diffuser material.

A primary ocean color product is phytoplankton biomass as estimated by the concentration of chlorophyll. The satellite product is usually validated against shipboard chlorophyll measurements. The errors are much larger than those for the measured radiances, because of the errors associated with the chlorophyll model, and errors associated with the validation data. Absorption per unit of chlorophyll is a function of the type of phytoplankton, its vertical distribution in the water column, and its light and nutrient history. The satellite measurement is an absorption-based measurement, and thus the satellite estimate of chlorophyll has errors associated with these variables. These estimates are less accurate than the radiance measurements and improvement will only come as models improve to account for most of the absorption per unit of chlorophyll variability.

At the next higher level, the goal of satellite ocean color measurements is an estimate of phytoplankton primary production as a measure of the role of the marine food web in the global CO_2 cycle. This estimate includes the errors associated with the measurement of ocean color, and the models that produce the estimate of production from those measurements. There is no agreed-upon algorithm for this product at this time, and present models disagree with each other, and with at-sea production measurements. For sake of argument, let us assume that the error associated with the model for the estimation of primary production from ocean color data is ±30 percent. The best ocean color satellite instruments should give ocean color estimates with an accuracy of ±5 percent so the error associated with the model is by far the largest error associated with the production estimate. Also, the discrepancy between models is much larger than the changes that are expected in the next decade due to global change.

Comparison of data from different satellite instruments should be at the radiance level. A long-term radiance data set could be used for the determination of phytoplankton variability by applying the same model to the radiance data. Many models can be tested using a large, long time series data base of *in-situ* measurements as a reference for assessing the success of the models. The models may not be perfect, but if it is applied uniformly, they should show trends over time.

The alternate approach is to combine the best estimates of chlorophyll or primary production from each of the satellites, each using their own models. The likelihood of seeing change as a function of different models is greater than the errors associated with the radiance estimates. If this approach is used, it would have to be validated against a large long time series of sea measurements, and each satellite data set adjusted to give the best estimate of that data set. Otherwise, there would not be a consistent data set for trend analysis.

A key point for developing a long-term global ocean color data base is that all of the sensors be calibrated against the same radiance standard, or standards that are intercompared and shown to agree to within a percent or two. *In situ* instruments that will be used to collect validation data sets should be calibrated against the same standards. The SeaWiFS and MODIS projects have initiated such an intercalibration program, and are working to include international partners who are building ocean color sensors for collecting validation data. This is a good approach that should be continued and expanded. Ideally the standard should be the Sun. This may require better diffusers, or at a minimum, a better characterization of the present diffusers, and standardization of the methods for using them.

Large long time series reference data sets are essential for the validation of radiance data and calculated products such as chlorophyll and primary production. Again, there must be a concerted effort to compare techniques, evaluate blind standards, and otherwise show that all of the data in the international data base are comparable. The JGOFS program is developing a standardized data base for chlorophyll and primary production measurements which will be useful for this purpose.

8. DESCRIPTION OF THE WORKSHOP

8.1 Organization

The CENR Task Force on Observations and Data Management CEOS/GCOS Global Change Calibration/Validation Workshop was held May 10-12, 1995 in Arlington, Virginia.

The CENR Secretariat hosted a meeting-of-experts from the international scientific community to develop recommendations for calibration and validation of global change data sets taken from instrument series and across generations of instruments and technologies. 49 scientists from 9 countries attended. The U.S., Canada, United Kingdom, France, Germany, Japan, Switzerland, Russia, and Kenya were represented.

The workshop began with the presentation of case studies describing experience in developing long-term geophysical data sets such as ozone trends, ISCCP, blended SST, and land climate.

Splinter sessions (by "traditional disciplines") extended and developed key components from case studies

- Atmosphere
- Ocean
- Land
- Sensors

Findings and recommendations (lessons learned) were presented by each discipline in the plenary session to the steering committee for integration.

8.2 List of Participants

Name and Country	Affiliation	Email Address
Ardanuy, Philip USA	Hughes STX Corporation	pardanuy@ccmail.stx.com
Arkin, Phillip USA	National Meteorological Center	parkin@sgi17.wwb.noaa.gov
Barnett, John UNITED KINGDOM	Oxford University	j.barnett@physics.ox.ac.uk
Brest, Christopher USA	NASA	clclb@nasagiss.giss.nasa.gov
Butler, Jim USA	NASA/GSFC	butler@highwire.gsfc.nasa.gov
Chesters, Dennis USA	NASA/GSFC	chesters@climate.gsfc.nasa.gov
Chu, William USA	NASA/Langley Research Center	chu@arbv1.larc.nasa.gov
Cihlar, Josef CANADA	Canada Centre for Remote Sensing	cihlar@ccrs.emr.ca
Croom, David UNITED KINGDOM	Rutherford Appleton Laboratory	dlc1@rl.ac.uk

Name and Country	Affiliation	Email Address
Davis, Curtiss USA	Naval Research Laboratory	None
Guenther, Bruce USA	NASA/GSFC	guenther@highwire.gsfc.nasa.gov
Hall, Forrest USA	NASA/GSFC	fghall@ltpsun.gsfc.nasa.gov
Halpern, David USA	Jet Propulsion Laboratory	halpern@pacific.jpl.nasa.gov
Haruyama, Yukio JAPAN	National Space Development Agency	None
Hilsenrath, Ernest USA	NASA/GSFC	hilsenrath@ssbuv.gsfc.nasa.gov
Hinsman, Donald SWITZERLAND	World Meteorological Organization	donald.hinsman@itu.ch
Hudson, Robert USA	University of Maryland College Park	hudson@atmos.umd.edu
Kandel, Robert FRANCE	Ecole Polytechnique	kandel@lmdx04.polytechnique.fr
Karpov, Alexandr RUSSIA	Russian Federal Service for	None
Kaye, Jack USA	NASA/Mission to Planet Earth	jkaye.mtpe.hq.nasa.gov
Kinyodah, Gideon KENYA	Kenya Meteorological Department	None
Malay, John USA	NOAA/NESDIS	None
Mandt, Gregory USA	NOAA/NESDIS	gmandt@nesdis.noaa.gov
Moodie, Linda USA	NOAA/NESDIS	lmoodie@nesdis.noaa.gov
Morel, Pierre USA	NASA/HQ	None
Mutlow, Christopher UNITED KINGDOM	Rutherford Appleton Laboratory	chris@atsrsc.ag.rl.ac.uk
Okamoto, Ke'ichi JAPAN	Communications Research Laboratory	okamoto@crl.go.jp
Rao, Nagaraja USA	NOAA	None

Name and Country	Affiliation	Email Address
Rao, P-K USA	NOAA/NESDIS, U.S. Department of Commerce	krao@orbit.nesdis.noaa.gov
Rasool, Ichtiaque FRANCE	IGBP-DIS	rasool@biogeodis.jussieu.fr
Ratier, Alain FRANCE	CNES	a.ratier@cst.cnes.fr
Reeves, Robert USA	NOAA/GOOS	breeves@noaa.gov
Reynolds, Richard USA	NOAA	wd01rr@sgi11.wwb.noaa.gov
Rockwitz, Klaus-Dieter GERMANY	DARA	None
Ryder, Peter UNITED KINGDOM	Meteorological Office	None
Schiffer, Robert USA	NASA/HQ	None
Shiue, James USA	NASA/GSFC	jcshiue@meneg.gsfc.nasa.gov
Slater, Philip USA	University of Arizona	pslater@kona.opt-sci.arizona.edu
Spence, Tom SWITZERLAND	Joint Planning Office	jpo@gcos.wmo.ch
Spencer, Roy USA	NASA/Marshall Space Flight Center	roy.spencer@msfc.nasa.gov
Teillet, Philippe CANADA	Canada Centre for Remote Sensing	teillet@ccrs.emr.ca
Theon, John USA	NASA/HQ	None
Townshend, John USA	University of Maryland College Park	jt59@umail.umd.edu
Walton, Charles USA	NOAA Science Center	None
Wang, Jinxue USA	NCAR/ACD	jwang@eos.acd.ucar.edu
Weinreb, Michael USA	NOAA/NESDIS	mweinreb@nesdis.noaa.gov
Wingham, Duncan UNITED KINGDOM	Mullard Space Science Laboratory	djw@mssl.ucl.ac.uk

Name and Country	Affiliation	Email Address
Winokur, Robert USA	NOAA/NESDIS	rwinokur@nesdis.noaa.gov
Withrow, John FRANCE	Intergovernmental Oceanographic	jwithrow@unesco.org

8.3 Agenda

Wednesday, May 10 (Day 1)

7:30 - 8:30	Steering Committee Working Breakfast	
7:30 - 8:30	Registration	
8:30 - 8:45	Welcoming Remarks and CENR Perspective	Charles Kennel, Robert Winokur, and Robert Schiffer
8:45 - 8:55	Goals and Objectives Workshop	Bruce Guenther
8:55 - 9:00	Plenary Session Opening Remarks	Ichtiaque Rasool
9:00 - 9:40	Ozone Trends Case Study	Robert Hudson
9:40 - 10:20	ISCCP Case Study	Chris Brest
10:20 - 10:35	Coffee Break	
10:35 - 11:15	Blended Sea Surface Temperature Case Study	Richard Reynolds
11:15 - 11:45	Data Requirements for Global Land Surface Modeling	Forrest Hall
11:45 - 12:00	Discussion and Breakout Into Discipline Groups	Ichtiaque Rasool/Susan Till
	Oceanic Global Change Observations Calibration and Validation • Ocean Surface Measurements and Air-Sea Interactions	Richard Reynolds
	Atmospheric Global Change Observations Calibration and Validation • Tropospheric Clouds, Water Vapor, and Precipitation • Structure, Dynamics, and Trace Gases	John Barnett/ Dennis Chesters
	Terrestrial Global Change Observations Calibration and Validation • Ecosystem Dynamics, Land-Cover Classification, and Snow Cover	Philippe Teillet/ Forrest Hall
12:00 - 1:30	Lunch	
1:30 - 5:30	Discipline Session Case Study Presentations and Round-Table Discussions	
5:30	Adjourn	
6:30	Joint Banquet With the GCOS Space Observations Panel Banquet Speaker: Professor Pierre Morel, Visiting Senior Scientist, NASA/HQ, Office of Mission to Planet Earth "Detecting Global Climate Change or Climate Transients?"	

Thursday, May 11, 1995 (Day 2)

7:30 - 8:30	Steering Committee Working Breakfast
7:30 - 8:30	Assembly/Informal Interaction
8:30 - 12:00	Discipline Session Round-Table Discussions (Continued)
12:00 - 1:30	Lunch
1:30 - 5:30	Discipline Session Writing Assignments
5:30	Adjourn
6:30	Steering Committee Working Dinner
7:30	Splinter Session: Satellite Sensor Radiometric Calibration and Intercalibration (Changes in Technology, Continuity Breaks, etc.)

Friday, May 12, 1995 (Day 3)

7:30 - 8:30	Assembly	
8:30 - 10:30	Discipline Session Review and Integration of Writing Assignments	
10:30 - 10:45	Coffee Break	
10:45 - 12:00	Calibration Splinter Presentation to Plenary	TBD
10:45 - 12:00	Atmospheric Discipline Presentation to Plenary	John Barnett
12:00 - 1:30	Lunch	
1:30 - 2:45	Oceanic Discipline Presentation to Plenary	TBD
2:45 - 4:00	Terrestrial Discipline Presentation to Plenary	Philippe Teillet
4:00 - 4:30	Schedule, Action Item Assignments, and Adjourn	Bruce Guenther
4:30	Closing Remarks and Adjourn	NASA

9. REFERENCES

Abel, P., B. Guenther, R. N. Galimore, and J. W. Cooper, 1993: Calibration Results for NOAA-11 AVHRR Channels 1 and 2 Congruent Path Aircraft Observations, *J. Atmos. Oceanic Technol.*, **10**, 493-508.

Ahmed, F., 1994 : Etudes du Traitement Secondaire de ScaRaB dans l'identification du type de scène à l'aide de données a priori, Mémoire DEA Univ. Paris-7, 35 pp.

Allen, M. R., J. R. Christy, M. J. Murray, and C. T. Mutlow, 1994: Potential Impact of Satellite-Derived Data Sets on Global Change Detection, *Proc. 19th Climate Diagnostics Workshop*, Washington.

Allen, M. R., C. T. Mutlow, G. M. C. Blumberg, J. R. Christy, R. T. McNider, and D. T. Llewellyn-Jones, 1994: Global Change Detection, *Nature*, **370**, 23-24.

Becker, F., H.-J. Bolle, and P. R. Rowntree, 1988: The International Satellite Land Surface Climatology Project, ISLSCP Report No.10, United Nations Environmental Programme.

Bhartia, P. K., S. Taylor, R. D. McPeters, and C. Wellemeyer, 1995: Application of the Langley Plot Method to the Calibration of SBUV Instrument on Nimbus-7 Satellite, *J. Geophys. Res.*, in press.

Birrer, I. J., E. M. Bracalente, G. J. Dome, J. Sweet, and G. Berthold, 1982: σ° Signature of the Amazon Rain Forest Obtained from the Seasat Scatterometer, *IEEE Trans. Geos. and Rem. Sens.*, GE-20 11-17.

Bojkov, R. D., L. Bishop, W. J. Hill, G. C. Reinsel, and G. C. Tiao, 1990: A Statistical Trend Analysis of Revised Dobson Total Ozone Data Over the Northern Hemisphere, *J. Geophys. Res.*, **95**, 9785-9807.

Brest, C. L. and W. B. Rossow, 1992: Radiometric Calibration of AVHRR Data for ISCCP, *Intl. J. Rem. Sens.*, **13**, 235-273.

Brooks, D. R., E. H. Harrison, P. Minnis, J. T. Suttles, and R. Kandel, 1986: Development of Algorithms for Understanding the Temporal and Spatial Variability of the Earth's Radiation Balance, *Rev. Geophys.*, **24**, 422-438.

Cebula, R. P., E. Hilsenrath, B. Guenther, 1989: Calibration of the Shuttle Borne Solar Backscatter Ultraviolet Spectrometer, *SPIE Proc.*, **1109**, 205-218.

Cebula, R. P., H. Park, and D. F. Heath, 1988: Characterization of the Nimbus-7 SBUV Radiometer for the Long-Term Monitoring of Stratospheric Ozone, *J. Atmos. Oceanic Technol.*, **5**, 215-227.

Chapman, S. and R. S. Lindzen, 1970: *Atmospheric Tides, Thermal and Gravitational*, Gordon and Breach Science Publishers, New York, NY 10011.

Che, N. and J. C. Price, 1992: Survey of Radiometric Calibration Results and Methods for Visible and Near Infrared Channels of NOAA-7, -9 and -11 AVHRRs, *Rem. Sens. Environ.*, **41**, 19-27.

Christy, J. R., R. W. Spencer, and R. T. McNider, 1995: Reducing Noise in the MSU Daily Lower Tropospheric Global Temperature Data Set, *J. Climate*, **8**, 888-896.

Cihlar, J. and P. M. Teillet, 1995: Forward Piecewise Linear Calibration Model for Quasi-Real Time Processing of AVHRR Data, *Canadian J. Rem. Sens.*, **21**, 22-27.

Drummond, J. R., *Measurements of Pollution in the Troposphere (MOPITT)*, in The Use of EOS for Studies of Atmospheric Physics, edited by J. C. Gille and G. Visconti, pp. 77-101, North Holland, Amsterdam, 1992.

Edwards, T. E., R. Browning, J. Delderfield, D. J. Lee, K. A. Lidiard, R. S. Milborrow, P. H. McPherson, S. C. Peskett, G. M. Toplis, H. S. Taylor, I. Mason, G. Mason, A. Smith, and S. Stringer, 1990: The Along Track Scanning Radiometer Measurement of Sea-Surface Temperature from ERS-1, *J. British Interplanetary Soc.*, **43**, 160-180.

Flittner, D. E. and P. N. Slater, 1991: Stability of Narrow-Band Filter Radiometers in the Solar-Reflective Range, *Photogram. Eng. Rem. Sens.*, **57**, 165-171.

Frederick, J. E., X. Niu, E. Hilsenrath, 1990: An Approach to the Detection of Long-Term Trends in Upper Atmospheric Ozone From Space, *J. Atmos. Oceanic Technol.*, **7**, 734.

Gellman, D. I., S. F. Biggar, M. C. Dinguirard, P. J. Henry, M. S. Moran, K. J. Thome, and P. N. Slater, 1993: Review of SPOT-1 and -2 Calibrations at White Sands from Launch to the Present, *Proceedings of SPIE Conference 1938*, Orlando, Florida, pp. 118-125.

Gille, J. C., and J. J. Barnett, "The High Resolution Dynamics Limb Sounder (HIRDLS): An Instrument for the Study of Global Change," in The Use of EOS for Studies of Atmospheric Physics, edited by J. C. Gille and G. Visconti, pp. 433-450, North Holland, Amsterdam, 1992.

Gleason, J.F., P. K. Bhartia, J. R. Herman, R. McPeters, P. Newman, R. S. Stolarski, L. Flynn, G. Labow, D. Larko, C. Seftor, C. Wellemeyer, W. D. Komhyr, A. J. Miller, and W. Planet, 1993: Record Low Global Ozone in 1992, *Science*, **260**, 523-526.

Guyot, G. and X. F. Gu, 1994: Effects of Radiometric Corrections on NDVI Determined from SPOT-HRV and Landsat-TM Data, *Rem. Sens. Environ.*, **49**, 169-180.

Hall, F. G., J. R. Townshend, R., and E. T. Engman, 1995: Status of Remote Sensing Algorithms for Estimation of Land Surface State Parameters, *Rem. Sens. Environ.*, **51**, 138-156.

Hall, F. G., D. E. Strebel, J. E. Nickeson, and S. J. Goetz, 1991: Radiometric Rectification: Toward a Common Radiometric Response Among Multidate, Multisensor Images, *Rem. Sens. Environ.*, **35**, 11-27.

Hartmann, D. L., V. Ramanathan, A. Berroir, and G. E. Hunt, 1986: Earth Radiation Budget Data and Climate Research, *Rev. Geophys.*, **24**, 439-468.

Heath, D. F., Z. Wei, W. K. Fowler, N. V. Nelson, 1993: Comparison of Spectral Radiance Calibration of SBUV-2 Satellite Ozone Monitoring Instruments Using Integrating Sphere and Flat Plate Diffuser Techniques, *Metrologia*, **30**, 259-264.

Herman, J. R., R. Hudson, R. McPeters, R. Stolarski, Z. Ahmad, X. Y. Gu, S. Taylor, and C. Wellemeyer, 1991: A New Self-Calibration Method Applied to TOMS and SBUV Backscattered Ultraviolet Data to Determine Long-Term Global Ozone Change, *J. Geophys. Res.*, **96**, 7531-7545.

Hilsenrath, E., R.P. Cebula, M.T. DeLand, K. Laamann, S. Taylor, C. Wellemeyer, P.K. Bhartia, 1995: Calibration of the NOAA-11 SBUV/2 Ozone Data Set From 1989 to 1993 Using In-Flight Calibration Data and SSBUV, *J. Geophys. Res.*, in press.

Hilsenrath, E., D. E. Williams, R. T. Caffrey, R. P. Cebula, and S. J. Hynes, 1993: Calibration and Radiometric Stability of the Shuttle Solar Backscatter Ultraviolet (SSBUV) Experiment, *Metrologia*, **30**, 243-248.

Hood, L. L, R. D. McPeters, J. P. McCormack, L. E. Flynn, S. M. Hollandsworth, and J. F. Gleason, 1993: Altitude Dependence of Stratospheric Ozone Trends Based on Nimbus-7 SBUV Data, *Geophys. Res. Lett.*, **21**, 2667-2670.

Janz, S., E. Hilsenrath, J. Butler, and R. P. Cebula, 1994: Uncertainties in Radiance Calibrations of Backscatter Ultraviolet (BUV) Instruments as Determined From Comparisons of BRDF Measurements and Integrating Sphere Calibrations, submitted to *Metrologia*.

Jaross, G., A. Krueger, R. P. Cebula, C. Seftor, D. Hartmann, R. Haring, and D. Burchfield, 1995: Calibration and Postlaunch Performance of the Meteor 3/TOMS Instrument, *J. Geophys Res.*, in press.

Justice, C. O. and J. R. G. Townshend, 1994: Data Sets for Global Remote Sensing: Lessons Learned, *Int. J. Rem. Sens.*, **15(7)**, 3621-3639.

Kandel, R. S., J.-L. Monge, M. Viollier, L. A. Pakhomov, V. I. Adasko, R. G. Reitenbach, E. Raschke and R. Stuhlmann, 1994: The ScaRaB Project: Earth Radiation Budget Observations From the Meteor Satellites, World Space Congress (Washington, 1992) - COSPAR Sympos. A.2-S, *Adv. Space Research*, **14**, 147-154.

Kandel, R. S., 1990: Satellite Observations of the Earth Radiation Budget and Clouds, *Space Sci. Rev.*, **52**, 1- 32.

Khalil, M. A. K., and R. A. Rasmussen, 1988: Carbon Monoxide in the Earth's Atmosphere: Indications of a Global Increase, *Nature*, **332**, 242-245.

Lee III, R. B. and B. R. Barkstrom, 1991: Characterization of the Earth Radiation Budget Experiment Radiometers, *Metrologia*, **28**, 183-187.

Markham, B. L. and J. L. Barker, 1987: Radiometric Properties of U.S. Processed Landsat MSS Data, *Rem. Sens. Environ.*, **22**, 34-71.

McClain, E. P., W. G. Pichel, and C. C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperatures, *J. Geophys. Res.*, **90**, 11587-11601.

McPhaden, J. W., 1993: TOGA-TAO and the 1991-93 El Niño-Southern Oscillation Event, *Oceanography*, **6**, 36-44.

Monge, J. L., R. S. Kandel, L. A. Pakhomov, and V. I. Adasko, 1991: ScaRaB Earth Radiation Budget Scanning Radiometer, *Metrologia*, **28**, 261-284.

MOPITT Algorithm Theoretical Basis Document, Version 2, 1995. Available from NCAR MOPITT Team, NCAR/ACD, Boulder, CO 80307.

Morasse, C. R. and M. A. DIorio, 1992: Développement dune base de données AVHRR pour le continent nord-américain dans le cadre des études sur les changements planétaires, *Canadian J. Rem. Sens.*, **18**, 180-181.

Mutlow, C. T., A. M. Zavody, I. J. Barton, and D. T. Llewellyn-Jones, 1994: Sea-Surface Temperature Measurements by the Along-Track Scanning Radiometer on the ERS-1 Satellite: Early Results, *J. Geophys. Res.*, **99**, 22,575-22,588.

National Plan for Stratospheric Monitoring and Early Detection of Change, 1988-1997, U.S. Dept. of Commerce/NOAA, FCM-P17-1989, July 1989.

Ohring, G. and J. Dodge, 1992: The NOAA/NASA Pathfinder program, in *IRS'92: Current Problems in Atmospheric Radiation*, (Eds: S. Keevalik and O. Karner), A. Deepak Publishing, Hampton, U.S.S., 404-408.

Planet, W. G., J. H. Lienesch, A. J. Miller, R. Nagatani, R.D. McPeters, E. Hilsenrath, R.P. Cebula, M.T. DeLand, C.G. Wellemeyer, K. Horvath, 1994: Northern Hemisphere Total Ozone Values From 1989-1993 Determined With the NOAA-11 Solar Backscatter Ultraviolet (SBUV/2) Instrument, *Geophys. Res. Lett.*, **21**, 205-208.

Qi, J., Y. Kerr, and A. Chehbouni, 1994: External Factor Consideration in Vegetation Index Development, *Proceedings of the Sixth International Colloquium on Physical Measurements and Signatures in Remote Sensing*, Val dIsère, France, pp. 723-730.

Ramanathan, V., 1987: The Role of Earth Radiation Budget Studies in Climate and General Circulation Research, *J. Geophys. Res.*, **92**, 4075-4095.

Randel, W. J. and Fei Wu, 1995: Climatology of Stratospheric Ozone Based on SBUV and SBUV/2 Data: 1978-1994, NCAR Technical Note TN-412+STR.

Rao, C. R. N., and J. Chen, 1995: Inter-satellite calibration linkages for the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on NOAA-7, -9, and -11 spacecraft, to appear in *Int. J. Rem. Sens.*

Rao, C. R. N. and J. Chen, 1994: Post-Launch Calibration of the Visible and Near-Infrared Channels of the Advanced Very High Resolution Radiometer on NOAA-7, -9, and -11 Spacecraft, NOAA Technical Report NESDIS 78, NOAA/NESDIS, E/RA22, 5200 Auth Road, Washington, D.C., 20233, U.S.A., 22 pages.

Rao, C. R. N., J. Chen, F. W. Staylor, P. Abel, Y. J. Kaufman, E. Vermote, W. R. Rossow, and C. Brest, 1993: Degradation of the Visible and Near-infrared Channels of the Advanced Very High Resolution Radiometer on the NOAA-9 Spacecraft: Assessment and Recommendations for Corrections, NOAA Technical Report NESDIS 70, U.S. Department of Commerce, Washington, D.C. 20233

Rao, C. R. N., J. T. Sullivan, C. C. Walton, J. W. Brown, and R. H. Evans, 1993: Nonlinearity Corrections for the Thermal Infrared Channels of the Advanced Very High Resolution Radiometer: Assessment and Recommendations, NOAA Technical Report NESDIS 69, U.S. Department of Commerce, Washington, D.C. 20233

Reichle, Jr., H. G., V. S. Connors, J. A. Holland, R. T. Sherrill, H. A. Wallio, J. C. Casas, E. P. Condon, B. B. Gormsen, and W. Seiler, 1990: The Distribution of Middle Tropospheric Carbon Monoxide During Early October 1984, *J. Geophys. Res.*, **95**, 9845-9856.

Reichle, Jr., H. G., V. S. Connors, J. A. Holland, W. D. Hypes, H. A. Wallio, J. C. Casas, B. B. Gormsen, M. S. Saylor, and W. D. Hesketh, 1986: Middle and Upper Tropospheric Carbon Monoxide Mixing Ratios as Measured by a Satellite-Borne Remote Sensor During November 1981, *J. Geophys. Res.*, **91**, 10865-10887.

Reynolds, R. W. and T. M. Smith, 1994: Improved Global Sea Surface Temperature Analysis Using Optimum Interpolation, *J. Climate*, **7**, 929-948.

Reynolds, R. W., 1993: Impact of Mount Pinatubo Aerosols on Satellite-Derived Sea Surface Temperatures, *J. Climate*, **6**, 768-774.

Reynolds, R. W. and D. C. Marsico, 1993: An improved real-time global sea surface temperature analysis, *J. Climate*, **6**, 114-119.

Reynolds, R. W., C. K. Folland, and D. E. Parker, 1989: Biases in satellite derived sea-surface-temperatures, *Nature*, **341**, 728-731.

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis, *J. Climate*, **1**, 75-86.

Rossow, W. B., E. Kinsella, A. Wolf, and L. C. Garder, 1987: International Satellite Cloud Climatology Project (ISCCP) Description of Reduced Resolution Radiance Data, World Climate Research Programme, *World Meteorological Organization Technical Document No.58*, 132 pp.

Running, S. W., C. Justice, V. Salomonson, D. Hall, J. Barker, Y. Kaufman, A. Strahler, A. Huete, J. -P. Muller, V. Vanderbilt, A. M. Wan, P. Teillet, and D. Carneggie, 1994: Terrestrial Remote Sensing Science and Algorithms Planned for EOS/MODIS, *Int. J. Rem. Sens.*, **15**, 3587-3620.

Santer, R., X. F. Gu, G. Guyot, J. L. Deuze, C. Devaux, E. Vermote, and M. Verbrugghe, 1992: SPOT Calibration at the La Crau Test Site (France), *Rem. Sens. Environ.*, **41**, 227-237.

Schiffer, R. A. and W. B. Rossow, 1985: ISCCP Global Radiance Data Set: A New Resource for Climate Research, *Bull. Amer. Meteorol. Soc.*, **66**, 1498-1505.

Schiffer, R. A. and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The First Project of the World Climate Research Programme, *Bull. Amer. Meteorol. Soc.*, **64**, 779-784.

Schroeder, L. C., et al., 1982: The Relationship Between Wind Vector and Normalized Radar Cross-Section Used to Derive SEASAT-A Satellite Scatterometer Winds, *J. Geophys. Res.*, **87**, 3318-3336.

Sellers, P. J., B. W. Meeson, F. G. Hall, G. Asrar, R. E. Murphy, R. A. Schiffer, F. P. Bretherton, R. E. Dickinson, R. G. Ellingson, C. B. Field, K. F. Huemmrich, C. O. Justice, J. M. Melack, N. T. Roulet, D. S. Schimel, and P. D. Try, 1995: Remote Sensing of the Land Surface for Studies of Global Change: Models—Algorithms—Experiments, *Rem. Sens. Environ.*, **51**, 3-26.

Slater, P. N., 1988: Review of the Calibration of Radiometric Measurements from Satellite to Ground Level, *Proceedings of the ISPRS*, Congress, Kyoto, Japan.

Slater, P. N., S. F. Biggar, R. G. Holm, R. D. Jackson, Y. Mao, M. S. Moran, J. M. Palmer, and B. Yuan, 1987: Reflectance- and Radiance-Based Methods for the In-Flight Absolute Calibration of Multispectral Sensors, *Rem. Sens. Environ.*, **22**, 11-37.

Slutz, R. J., S. J. Lubker, J. D. Hiscox, S. D. Woodruff, R. L. Jenne, D. H. Joseph, P. M. Steurer, and J. D. Elms, 1985: COADS: Comprehensive Ocean-Atmosphere Data Set, Release 1, 262 pp. [Available from Climate Research Program, Environmental Research Laboratories, 325 Broadway, Boulder, CO 80303.]

Smith, T. M., R. W. Reynolds, R. E. Livezey, and D. C. Stokes, 1995: Reconstruction of Historical Sea Surface Temperatures Using Empirical Orthogonal Functions, *J. Climate*, (In press).

Smith, G. L., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, and B. A. Wielicki, 1994: Radiation Budget Measurements for the Eighties and Nineties, World Space Congress (Washington, 1992) - COSPAR Sympos. A.2-S, *Adv. Space Res.*, **14**, 131-146.

Smith, G. L., R. N. Green, E. Raschke, L. M. Avis, J. T. Suttles, B. A. Wielicki, and R. Davies, 1986: Inversion Methods for Satellite Studies of the Earth's Radiation Budget: Development of Algorithms for the ERBE Mission, *Rev. Geophys.*, **24**, 407-421.

Spencer, R. W. and J. R. Christy, 1993: Precision Lower Stratospheric Temperature Monitoring with the MSU: Validation and Results 1979-91, *J. Climate*, **6**, 1194-1204.

Spencer, R. W. and J. R. Christy, 1992a: Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies, Part I: MSU Channel 2, *J. Climate*, **5**, 847-857.

Spencer, R. W. and J. R. Christy, 1992b: Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies, Part II: A Tropospheric Retrieval and Trends 1979-90, *J. Climate*, **5**, 858-866.

Spencer, R. W., J. R. Christy, and N. C. Grody, 1990: Global Atmospheric Temperature Monitoring with Satellite Microwave Measurements: Method and Results, 1979-84, *J. Climate*, **3**, 1111-1128.

Staylor, W. F., 1990: Degradation Rates of the AVHRR Visible Channel for the NOAA-6, -7 and -9 Spacecraft, *J. Atmos. Oceanic Technol.*, **3**, 411-423.

Stolarski, R., R. Bojkov, L. Bishop, C. Zerofos, J. Stachelin, and J. Zawodny, 1992: Measured Trends in Stratospheric Ozone, *Science*, **256**, 342-349.

Stowe, L. L. (ed.), 1988: Report of the Earth Radiation Budget Requirements Review - 1987, NOAA Technical Report. NESDIS-41, NOAA/NESDIS/SRL, Washington, 103 pp.

Strow, L., 1994: CO Algorithm Status - CAMEX Measurements, AIRS Science Team Meeting, November 2-4, Lexington, MA.

Stubenrauch, C. J., J.-Ph. Duvel, and R. S. Kandel, 1993: Determination of Longwave Anisotropic Emission Factors From Combined Broad- and Narrow-band Radiance Measurements: *J. Appl. Meteor.*, **32**, 848-856.

Suits, G.H., W. A. Malila, and T. M. Weller, 1988: The Prospects for Detecting Spectral Shifts Due to Satellite Sensor Aging, *Rem. Sens. Environ.*, **26**, 17-29.

Suttles, J. T., R. N. Green, G. L. Smith, B. A. Wielicki, I. J. Walker, V. R. Taylor, and L. L. Stowe, 1988: Angular Radiation Models for the Earth-Atmosphere System, NASA RP-1184, Vol. 1 (SW Radiation), 147 pp., vol. 2 (LW radiation), 87 pp.

Teillet, P. M. and G. Fedsejevs, 1995: On the Dark Target Approach to Atmospheric Correction of Remotely Sensed Data, *Canadian J. Rem. Sens.*, in press.

Teillet, P. M., G. Fedosejevs, D. Gauthier, M. A. DIorio, B. Rivard, P. Budkewitsch, and B. Brisco, 1995: An Initial Examination of Radar Imagery of Optical Radiometric Calibration Sites, *Proceedings of the Europto-95 Symposium*, Paris, France, in press.

Teillet, P. M., K. Staenz, and D. J. Williams, 1995: Effects of Spectral, Spatial and Radiometric Characteristics on Remote Sensing Vegetation Indices, in preparation.

Teillet, P. M., K. Staenz, and D. J. Williams, 1994: Effects of Spectral and Spatial Resolutions on NDVI, *Proceedings of the Second International Symposium on Spectral Sensing Research (ISSSR '94)*, San Diego, California, pp. 365-374.

Teillet, P. M. and B. N. Holben, 1994: Towards Operational Radiometric Calibration of NOAA AVHRR Imagery in the Visible and Near-Infrared Channels, *Canadian J. Rem. Sens.*, **20**, 1-10.

Teillet, P. M. and K. Staenz, 1992: Atmospheric Effects Due to Topography on MODIS Vegetation Index Data Simulated From AVIRIS Imagery Over Mountainous Terrain, *Canadian J. Rem. Sens.*, **18**, 283-291.

Teillet, P. M., 1992: An Algorithm for the Radiometric and Atmospheric Correction of AVHRR Data in the Solar Reflective Channels, *Rem. Sens. Environ.*, **41**, 185-195.

Teillet, P. M., 1990: Effects of Spectral Shifts on Sensor Response, *Proceedings of the ISPRS Commission VII Symposium*, Victoria, B.C., Canada, pp. 59-65.

Teillet, P. M., P. N. Slater, Y. Ding, R. P. Santer, R. D. Jackson, and M. S. Moran, 1990: Three Methods for the Absolute Calibration of the NOAA AVHRR Sensors In-Flight, *Rem. Sens. Environ.*, **31**, 105-120.

Thomas, D., J. Ph. Duvel, and R. Kandel, 1995: Diurnal Bias in Calibration of Broad-band Radiance Measurements from Space, *IEEE Trans. Geosci. and Rem. Sens.*, **33**, in press.

Thome, K. J., D. I. Gellman, R. J. Parada, S. F. Biggar, P. N. Slater, and M. S. Moran, 1993: In-Flight Radiometric Calibration of Landsat-5 Thematic Mapper From 1984 to Present, *Proceedings of SPIE Conference 1938*, Orlando, Florida, pp. 126-131.

Townshend, J. R. G., 1994: Global Data Sets for Land Applications from the Advanced Very High Resolution Radiometer: An Introduction, *Int. J. Rem. Sens.*, **15**, 33196332.

Townshend, J. R. G., C. O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski, and S. Ruttenberg, 1994: The 1 km Resolution Global Data Set: Needs of the International Geosphere-Biosphere Programme, *Int. J. Rem. Sens.*, **15**, 3417-3441.

Townshend, J. R. G., C. Justice, C. Gurney, and J. McManus, 1992: The Impact of Misregistration on Change Detection, *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 1054-1060.

Viollier, M., R. Kandel, and P. Raberanto, 1995: Inversion and Space-time Averaging Algorithms for ScaRaB (Scanner for the Earth Radiation Budget), Comparison with ERBE, Annales Geophysicae, in press.

Walton, C. C., 1988: Nonlinear Multichannel Algorithms for Estimating Sea Surface Temperature with AVHRR Satellite Data, *J. Appl. Meteor.*, **27**, 115-124.

Weinreb, M. P., G. Hamilton, S. Brown, and R. J. Koczor, 1990: Nonlinearity Corrections in Calibration of Advanced Very High Resolution Radiometer Infrared Channels, *J. Geophys. Res.*, **95**, 7381-7388.

Whitlock, C. H., W. F. Staylor, G. Smith, R. Levin, R. Frouin, C. Gautier, P. M. Teillet, P. N. Slater, Y. J. Kaufman, B. N. Holben, W. B. Rossow, C. L. Brest, and S. R. LeCroy, 1990: AVHRR and VISSR Satellite Instrument Calibration Results for Both Cirrus and Marine Stratus IFO Periods, FIRE Science Report 1988, *NASA Conference Proceedings 3083*.

Wielicki, B. A. and R. N. Green, 1989: Cloud Identification for ERBE Radiative Flux Retrieval, *J. Appl. Meteor.*, **28**, 1133-1146.

World Meteorological Organization, Scientific Assessment of Ozone Depletion, 1991: Global Ozone Research and Monitoring Project, Report 25.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 1997

3. REPORT TYPE AND DATES COVERED
Reference Publication - May 10-12, 1995

4. TITLE AND SUBTITLE
Workshop on Strategies for Calibration and Validation of Global Change Measurements: May 10-12, 1995

5. FUNDING NUMBERS
C-NAS5-32350
Code 520

6. AUTHOR(S)
Bruce Guenther, James Butler, and Philip Ardanuy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS (ES)
Goddard Space Flight Center
Greenbelt, Maryland

8. PEFORMING ORGANIZATION REPORT NUMBER
97B00024

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS (ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA RP-1397

11. SUPPLEMENTARY NOTES
Guenther and Butler: Goddard Space Flight Center, Greenbelt, Maryland;
Ardanuy: Hughes STX, Greenbelt, Maryland

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 47

Availability: NASA CASI (301) 621-0390.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Committee on Environment and Natural Resources (CENR) Task Force on Observations and Data Management hosted a Global Change Calibration/Validation Workshop on May 10-12, 1995, in Arlington, Virginia. This Workshop was convened by Robert Schiffer of NASA Headquarters in Washington, D.C., for the CENR Secretariat with a view toward assessing and documenting lessons learned in the calibration and validation of large-scale, long-term data sets in land, ocean, and atmospheric research programs. The National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC) hosted the meeting on behalf of the Committee on Earth Observation Satellites (CEOS)/Working Group on Calibration/Validation, the Global Change Observing System (GCOS), and the U. S. CENR.

A meeting of experts from the international scientific community was brought together to develop recommendations for calibration and validation of global change data sets taken from instrument series and across generations of instruments and technologies. Forty-nine scientists from nine countries participated. The U. S., Canada, United Kingdom, France, Germany, Japan, Switzerland, Russia, and Kenya were represented.

14. SUBJECT TERMS
Calibration; validation; workshop; measurements; global change; trends; ground truth; error budget; AVHRR; TOMS; GOME; CERES; EOS; MODIS; ATSR; SeaWiFS; SBUV/2; lunar observations; direct broadcast data.

15. NUMBER OF PAGES
152

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL