Stability and Accuracy Requirements for Satellite Remote Sensing Instrumentation for Global Climate Change Monitoring

Raju Datla
NIST, Gaithersburg, MD 20899
Bill Emery
University of Colorado
George Ohring
NOAA/NESDIS (Consultant)
Roy Spencer
NASA Marshall Space Flight Center
Bruce Wielicki
NASA Langley Research Center

ABSTRACT: In the remote sensing community, stability and accuracy are two critical terms that need to be used in a consistent and clear way for their use to be of maximum benefit. Their definitions in combination with the vocabulary of the ISO Guide on uncertainty analysis have been applied to long time series measurements using satellite instruments for climate monitoring. Based on the definitions of these terms, requirements have been developed for satellite instrumentation at a calibration workshop on November 12 – 14, 2002 organized by NIST, NOAA, NPOESS IPO and NASA. The background analyses that lead to some of these requirements are discussed.

1 INTRODUCTION

To assess the impact of anthropogenic effects on global climate and to quantitatively distinguish them from natural effects is a daunting scientific task. To meet this challenge, satellite instruments, for example, must be capable of observing atmospheric temperature trends as small as 0.1^0 C/decade, ozone changes as small as 1%/decade, and variations in the sun's output as small as 0.1%/decade.

The importance of understanding and predicting climate variation and change has increased significantly in the last decade. In 2001, the White House requested the National Academy of Sciences (NAS) National Research Council (NRC) (NRC, 2001a) to review the uncertainties in climate change science. One of the three key recommendations from the NRC's report is to "ensure the existence of a long-term monitoring system that provides a more definitive observational foundation to evaluate decadal- to century-scale changes, including observations of key state variables and more comprehensive regional measurements". To accelerate Federal research and reduce uncertainties in climate-change science, in June 2001, President George W. Bush created the Climate Change Research Initiative (CCRI).

To develop recommendations for improving the calibration of satellite instruments to meet the challenge of measuring global climate change, the National Institute of Standards and Technology (NIST), National Polar-orbiting Operational Environmental Satellite System-Integrated Program Office (NPOESS-IPO), National Oceanic and Atmospheric Administration (NOAA), and National Aeronautics and Space Administration (NASA) organized a workshop at the University of Maryland Inn and Conference Center, College Park, MD, November 12-14, 2002. Some 75 scientists, including researchers who develop and analyze long-term data sets from satellites, experts in the field of satellite instrument calibration, and physicists working on state of the art calibration sources and standards, participated in the workshop. Workshop activities consisted of keynote papers, invited presentations, breakout groups, and preparation of draft input for a workshop report. The keynote papers and invited presentations provide extensive background information on issues discussed at the workshop and are posted on the NIST web-site: http://physics.nist.gov/Divisions/Div844/global/mgcc.html. (Please Note: To access this site, you have to input user name: mgccoutline, and password: div844mgcc)

The final report of the workshop, to be published in March 2004 by NIST

- Defined the required absolute accuracies and long-term stabilities of global climate data sets;
- Translated the data set accuracies and stabilities to required satellite instrument accuracies and stabilities; and
- Evaluated the ability of current observing systems to meet these requirements.

The focus of the report is on passive satellite sensors that make observations in spectral bands ranging from the ultraviolet to the microwave. The climate change variables of interest include the following:

- Solar irradiance, Earth radiation budget, and clouds (total solar irradiance, spectral solar irradiance, outgoing longwave radiation, net incoming solar radiation, cloudiness);
- Atmospheric variables (temperature, water-vapor, ozone, aerosols, precipitation, and carbon dioxide); and
- Surface variables (vegetation, snow cover, sea ice, sea surface temperature, and ocean color).

This list is not exhaustive. The variables were selected based on the following criteria: 1) importance to decadal scale climate change, 2) availability or potential availability of satellite-based climate data records, and 3) measurability from passive satellite sensors. The workshop breakout groups were aligned with the above three groups of climate variables.

There have been a number of previous reports that have also discussed accuracy and stability measurement requirements for long-term climate data sets (for example, Hansen et al., 1993; Jacobowitz, 1997; NPOESS, 2001) and calibration issues (Guenther et al., 1997; NRC, 2000; NRC, 2001b). In the next section, we will discuss this terminology of accuracy and stability as an extension of the vocabulary of the Guide to the Expression of uncertainty in Measurement (ISO, 1993) or simply referred to in this article as the ISO Guide. In the section following it, we will describe the requirements of accuracy and stability for climate change variables reported by the Workshop of November 12 – 14, 2002. In the final section before the summary, we will discuss the analysis that led to requirements, giving examples of those variables such as solar irradiance, total column water-vapor, total column ozone and sea surface temperature.

We would like to reiterate and expand upon two important points mentioned above. Firstly, the workshop focused on passive satellite instruments, with very little discussion of active sensors such as radar or lidar. There were several reasons for this decision: 1) We wanted to constrain the workshop to a manageable size and produce a significant output within the 21/2 day time limit of the deliberations, 2) A largely different group of remote sensing and calibration experts is associated with active instruments, and 3) Current long-term satellite data sets are all based upon passive instruments and will continue to be so into the future, probably for at least a decade. However, it is not too early to begin planning for the use of active instruments, with their enhanced capabilities, in climate monitoring, and this could be the subject of a follow-up workshop. Secondly, the number of climate variables dealt with at the workshop was limited and selected according to pre-specified criteria. Once again, this limitation was imposed to constrain the scope of the workshop to a manageable and productive event. As a result, some climate variables thought to be important may not have been discussed, e.g., soil moisture, sea level, and sea surface salinity. All three variables would receive focused attention at any workshop dealing with active instruments.

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

Measuring small changes over extended time periods necessarily involves the concepts of accuracy and stability of time series. Accuracy is defined by the International Organization for Standardization (ISO) as the "closeness of the agreement between the result of the measurement and the true value of the measurand" (ISO, 1993). Specifically the quantity to be measured is called the measurand and truth is the true value of the measurand. So the measurands in our case are the various climate variables. This same definition is also given in the International Vocabulary of Basic and General Terms in Metrology, commonly abbreviated as VIM. (ISO, 1993). In the notes on the definition of the term accuracy, the ISO guide observes it as a qualitative concept. As such, Barry Taylor & Chris Kuyatt (1994) commented "Because "accuracy" is a qualitative concept, one should not use it quantitatively, that is associate numbers with it: numbers should be associated with measures of uncertainty instead." However, the community of remote sensing experts has already established the usage of the term accuracy as defined above, but used in a quantitative fashion (see, for example, Hansen et al. (1993)). It has been widely used in the scientific as well as policy making and contractual documents for various satellite missions of NOAA, NPOESS and NASA.

Rather than attempting to change this now widely accepted and used terminology, we opted to re-emphasize its commonly accepted definition and use it as an extension of the terminology of ISO guide. So, the term accuracy may be thought of as the closeness to the truth and is measured by the bias or systematic error of the data, that is, the difference between the short-term average measured value of a variable and the truth. The short- term average value is the average of a sufficient number of successive measurements of the variable under identical conditions such that the random error is negligible relative to the systematic error. The ISO guide elaborates the meaning of the systematic error as that which would "result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus the value of the measurand." Elaborating further, "Repeatability conditions include the same measurement procedure, the same observer, the same measuring instrument, used under the same conditions, with repetition over a short period of time." The systematic error is called bias in the case of a measuring instrument. The ISO guide takes the position that since the value of the measurand cannot be completely known, the systematic error and its causes cannot be completely known. However, the measurement goal is to arrive at the true value of the measurand. Therefore, we should follow the procedure of the ISO guide to determine the systematic error, for example, "that if a device is tested through a comparison with a known reference standard and the uncertainties associated with the standard and the comparison procedure can be assumed to be negligible relative to the required uncertainty of the test, the comparison may be viewed as determining the error of the device" (ISO,1993). The methods to establish the true value of a variable (the measurand) should be consistent with the internationally adopted methods and standards, thus establishing the System of Units (SI) traceability (BIPM, 1998; NIST, 1995). According to the resolution adopted by The 20th Conference Generale des Poids et Measures (CGPM) – the international standards body in Paris – adopted the following resolution: "that those responsible for studies of Earth resources, the environment and related issues ensure that measurements made within their programs are in terms of well-characterized SI units so that they are reliable in the long term, be comparable worldwide and be linked to other areas of science and technology through the world's measurement system established and maintained under the Convention du Metre" (CGPM, 1995).

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

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

It should be pointed out that achieving the instrument measurement requirements does not guarantee determining the desired long- term trends. Superimposed on these trends is climatic noise – short-term climate variations – that may mask the signal we are trying to detect or reduce our confidence in the derived trend.

Well-validated stability and accuracy statements are critical to the success of a satellite sensor. Achieving the required stability and accuracy on orbit has been plagued by a myriad of problems, from degradation of plaques to unforeseen stray-light effects. An accuracy claim supported by an uncertainty statement with traceability to the SI units as maintained by the National Institute of Standards and Technology and validated by comparisons with other instruments ideally provides the necessary information to establish the absolute value of a climate variable for absolutely calibrating climate-change models. It also facilitates the search for the unforeseen or for new physics (for instance, the discrepancy between model and measured atmospheric transmission), and provides traceability to measurements made by past and future generations and by completely different instruments. A reliable accuracy statement also provides a route to make comparisons between two different sensors when satellite overlap is not possible due to premature failure or policy decision.

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

3 OUTCOME OF THE WORKSHOP: ACCURACY AND STABILITY REQUIREMENTS

3.1 *Climate Variables*

The required accuracies and stabilities of the climate variable data sets were established with consideration of changes in important climate signals based on current understanding and models of long-term climate change. Such signals include the following:

- Climate changes or expected trends predicted by models
- Significant changes in climate forcing or feedback variables (e.g., radiative effects comparable to that of increasing greenhouse gases)
- Trends similar to those observed in past decades.

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

The factor 1/5, or 20 %, is somewhat arbitrary. It should be periodically reevaluated. If the climate change signal is one unit per decade, a 20 % stability would imply an uncertainty range of 0.8 to 1.2, or the ratio of minimum to maximum excursion is a factor of 1.5, in our estimate of the signal. One basis for choosing such a factor is related to the uncertainty in climate model predictions of climate change. Thirty-five climate model simulations yield a total range of 1.4 K to 5.8 K, or factor of about 4, in the change in global temperature by 2100 (IPCC, 2001). Thus, a stability of 20 % should lead to a considerable narrowing of the possible climate model simulations of change. Achieving the stability requirement does not guarantee determining these long-term trends. Superimposed on these trends is climatic noise – short-term climate variations – that may mask the signal we are trying to detect or reduce our confidence in the derived trend.

Although excellent absolute accuracy is not critical for trend detection, it is crucial for understanding climate processes and changes. Continuous efforts should be undertaken to constantly improve the accuracy of satellite instruments. It is also essential when satellite overlap is not possible, for comparing measurement across generations, for providing an absolute scale, for models, etc.

Table 1 summarizes the required accuracies and stabilities of the data sets for the solar irradiance, Earth radiation budget, and cloud variables; atmospheric variables; and surface variables. The table also indicates which one of the above climate signals – climate changes, climate forcings, climate feedbacks, or trends similar to recent trends – forms the basis for the requirement.

Table 1. Required accuracies and stabilities for climate variable data sets. Column labeled signal indicates the type of climate signal used to determine the measurement requirements.

	Signal	Accuracy	Stability(per decade)
SOLAR IRRADIANCE, EARTH RADIATION BUDGET, AND CLOUD VARIABLES			
Solar irradiance	Forcing	1.5 W/m^2	0.3 W/m^2
Surface albedo	Forcing	0.01	0.002
Downward longwave flux: Surface	Feedback	1 W/m^2	0.2 W/m^2
Downward short-wave radiation: Surface	Feedback	1 W/m^2	0.3 W/m^2
Net solar radiation: Top of atmosphere	Feedback	1 W/m^2	0.3 W/m^2
Outgoing longwave radiation: Top of atmosphere	Feedback	1 W/m^2	0.2 W/m^2
Cloud base height	Feedback	0.5 km	0.1 km
Cloud cover (Fraction of sky covered)	Feedback	0.01	0.003
Cloud particle size distribution	Feedback	TBD*	TBD*
Cloud effective particle size	Forcing: Water Feedback: Ice	Water: 10% Ice: 20 %	Water: 2% Ice: 4%
Cloud ice water path	Feedback	25%	5%
Cloud liquid water path	Feedback	0.025 mm	0.005 mm
Cloud optical thickness	Feedback	10%	2%
Cloud top height	Feedback	150 m	30 m
Cloud top pressure	Feedback	15 hPa	3 hPa
Cloud top temperature	Feedback	1 K/cloud emissivity	0.2 K/cloud emissivity
Spectrally resolved thermal radiance	Forcing/ climate change	0.1 K	0.04 K
ATMOSPHERIC VARIABLES			
Temperature			
Troposphere	Climate change	0.5 K	0.04 K
Stratosphere	Climate change	0.5 K	0.08 K
Water-vapor	Climate change	5%	0.26%
Ozone			
Total column	Expected trend	3%	0.2%
Stratosphere	Expected trend	5%	0.6%
Troposphere	Expected trend	10%	1.0%
Aerosols			
Optical depth (troposphere/ stratosphere)	Forcing	0.01/0.01	0.005/ 0.005
Single scatter albedo (troposphere)	Forcing	0.03	0.015
Effective radius (troposphere /stratosphere)	Forcing	greater of 0.1 or 10%/0.1	greater of 0.05 or 5%/0.05
Precipitation		0.125 mm/hr	0.003 mm/hr
Carbon dioxide	Forcing/ Sources-sinks	10 ppmv/10 ppmv	2.8 ppmv/1 ppmv
SURFACE VARIABLES			
Ocean color		5 %	1 %
Sea surface temperature	Climate change	0.1 K	0.04 K
Sea ice area	Forcing	5 %	4 %
Snow cover	Forcing	5 %	4 %
Vegetation	Past trend	3 %	1 %

* To be determined

3.2 *Satellite Instruments*

The requirements for the data sets must be translated into required accuracies and stabilities of the satellite measurements. In some cases, for example, solar irradiance and top of the atmosphere Earth radiation budget, there is a one to one correspondence. For other climate variables, this translation is more complex. And for a few of the variables, additional studies are needed to determine the mapping of data set accuracies and stabilities into satellite accuracies and stabilities.

Because of the difficulties in achieving necessary accuracies (exo-atmospheric total solar irradiance is one example, (Quinn and Frohlich, 1999)), a key attribute for the satellite instruments is long-term stability. This may be achieved by either having an extremely stable instrument or by monitoring the instrument's stability, by various methods, while it is in orbit. An ideal external calibration source is one that is nearly constant in time and able to be viewed from different orbit configurations. In addition to other attributes, if there is scientific evidence regarding the degree of stability of such a source, and it is believed to be at an acceptable level for long-term climate studies, then the stability of the satellite sensor can be assessed independent of other reference standards. With such monitoring, instrument readings can be corrected for drift. However, this brings up a measurement challenge for establishing the degree of stability of the external reference source. Obviously the methods and instruments testing the stability of those sources must have stability requirements far more stringent than given in this report. One method that has been successfully implemented for the reflected solar spectral interval is lunar observations, from orbit, with the sensor on orbit. One example is the ocean color satellite Sea-viewing Wide Field-of-view Sensor (SeaWiFS), which used lunar observations to correct for degradation in the near-infrared channels (Kieffer et al., 2003). The required lunar data are being supplied by a dedicated ground-based facility (Anderson et. al., 1999).

Since satellites and their instruments are short-term – NPOESS satellites and instruments have design lives of about 7 years – satellite programs launch replacement satellites to continue the observations. Thus, the long-term data record for any climate variable will consist of contributions from a series of satellite instruments, some using different techniques. To assess the reproducibility of the measurement results, to assist in understanding the differences that arise even with instruments of similar design, and to create a seamless data record, it is essential that the satellites be launched on a schedule that includes an overlap interval of the previous and the new instrument. Acquiring multiple independent space-based measurements of key climate variables would also help insure maintenance of stability in the event of a single instrument failure.

Table 2 summarizes the required accuracies and stabilities of the satellite instruments for solar irradiance, Earth radiation budget, and cloud variables; the atmospheric variables; and the surface variables. The table also indicates the types of satellite instruments used for the measurements.

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

	Instrument	**Accuracy**	**Stability (per decade)**
SOLAR IRRADIANCE, EARTH RADIATION BUDGET, AND CLOUD VARIABLES			
Solar irradiance	Radiometer	1.5 W/m^2	0.3 W/m^2
Surface albedo	Vis radiometer	5%	1%
Downward longwave flux: Surface	IR spectrometer and Vis/IR radiometer	See tropospheric temperature, water-vapor, cloud base height, and cloud cover	See tropospheric temperature, water-vapor, cloud base height, and cloud cover
Downward shortwave radiation: Surface	Broad band solar and Vis/IR radiometer	See net solar radiation: TOA, cloud particle effective size, cloud optical depth, cloud top height, and water-vapor	See net solar radiation: TOA, cloud particle effective size, cloud optical depth, cloud top height, and water-vapor
Net solar radiation: Top of atmosphere	Broad band solar	1 W/m^2	0.3 W/m^2
Outgoing longwave radiation: Top of atmosphere	Broad band IR	1 W/m^2	0.2 W/m^2
Cloud base height	Vis/IR radiometer	1 K	0.2 K
Cloud cover (Fraction of sky covered)	Vis/IR radiometer	See cloud optical thickness and cloud to temperature	See cloud optical thickness and cloud to temperature
Cloud particle size distribution	Vis/IR radiometer	TBD*	TBD*
Cloud effective particle size	Vis/IR radiometer	3.7 μm: Water, 5%; Ice, 10% 1.6μm: Water, 2.5%; Ice, 5%	3.7 μm: Water, 1%; Ice, 2% 1.6μm: Water, 0.5%; Ice, 1%
Cloud ice water path	Vis/IR radiometer	TBD*	TBD*

	Instrument	Accuracy	Stability (per decade)
Cloud liquid water path	Microwave and Vis/IR radiometer	Microwave: 0.3 K Vis/IR: see cloud optical thickness and cloud top height	Microwave: 0.1 K Vis/IR: see cloud optical thickness and cloud top height
Cloud optical thickness	Vis radiometer	5%	1%
Cloud top height	IR radiometer	1 K	0.2 K
Cloud top pressure	IR radiometer	1 K	0.2 K
Cloud top temperature	IR radiometer	1 K	0.2 K
Spectrally resolved thermal radiance	IR spectroradiometer	0.1 K	0.04 K
ATMOSPHERIC VARIABLES			
Temperature			
Troposphere	MW or IR radiometer	0.5 K	0.04 K
Stratosphere	MW or IR radiometer	1 K	0.08 K
Water-vapor	MW radiometer IR radiometer	1.0 K 1.0 K	0.08 K 0.03 K
Ozone			
Total column	UV/VIS spectrometer	2 % (λ independent), 1 % (λ dependent)	0.2%
Stratosphere	UV/VIS spectrometer	3%	0.6%
Troposphere	UV/VIS spectrometer	3%	0.1%
Aerosols	VIS polarimeter	Radiometric: 3% Polarimetric: 0.5 %	Radiometric: 1.5% Polarimetric: 0.25 %
Precipitation	MW radiometer	1.25 K	0.03 K
Carbon dioxide	IR radiometer	3%	Forcing: 1%; Sources/ sinks: 0.25%

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

* To be determined

4 ANALYSIS OF REQUIREMENTS

4.1 *Solar Irradiance, Earth Radiation Budget and Clouds*

How were the requirements set?

Overall, the variables in this section are linked in their role in the energetics of the climate system. Accurate measurements of solar irradiance are key to defining climate radiative forcing, and its accuracy requirements are specified in that context. Changes in surface albedo can represent both changes in climate forcings due to human caused land-cover change, and climate feedbacks due to changes in ecosystems and in snow and ice cover resulting from climate changes. Cloud feedback remains the largest single factor in the current large uncertainty in climate sensitivity (IPCC, 2001). Cloud properties are critical to understanding and defining the role of clouds as feedback mechanisms in the climate system. The earth radiation budget is critical to the climate system, and is a key diagnostic for a wide range of climate forcings (aerosol), feedbacks (clouds, ice/snow), and climate responses (heat transport). Accuracies for clouds and radiation budget are defined at levels sufficient to be at or above estimates of unforced natural climate variability in current climate models; these accuracies must also be sufficient to directly observe decadal changes in clouds and radiation budget that would constrain potential cloud feedback mechanisms in climate models.

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

(ERBE, Scanner for Radiation Budget (ScaRaB), CERES) from satellites have indicated significant interannual to decadal variability in the tropics from latitude 20S to latitude 20N. This variability is not shown in current climate model simulations and is representative of changes that are critical to assess accurately from observations, and to be able to predict from climate models. A climate observing system that cannot rigorously observe such changes with high confidence is very unlikely to be able to constrain and verify cloud feedbacks within climate prediction models.

Accuracy requirements can also be determined by considering the amount of climate change likely over the next few decades. For example, many climate change models use a 1 %/year increase in carbon dioxide to simulate a nominal doubling of CO_2 in 70 years. This doubling is a radiative forcing of the climate system of about 4 Wm^{-2}, or about 0.6 Wm^{-2} per decade. A change in global average cloud fraction sufficient to offset this radiative forcing would be about 0.015 if all other cloud properties remained fixed. This would be a cloud feedback so strong that climate change due to greenhouse forcing would become negligible. We suggest that a minimum signal-to-noise ratio of at least 5 is needed to detect such change, suggesting a requirement for stability per decade in global cloud cover of 0.003. This would be sufficient to detect a cloud feedback. This approach essentially follows that used by Hansen et al. (1993) in a workshop report that summarized accuracies required for long-term monitoring of global climate forcings and feedbacks. The accuracy requirements in this section are in general similar to those in Hansen et al. where the same climate variable was evaluated. As in that report, this workshop concluded that the appropriate scaling for climate requirements is the radiative flux changes that can potentially alter climate: either forcing or feedback.

The NPOESS project convened a workshop to assess climate measurement requirements for the Integrated Operational Requirements Document (IORD) variables (Jacobowitz, 1997). The report influenced the IORD to add or change stability requirements, but had little effect on other IORD requirements, which were focused on instantaneous observations and, often, high spatial resolution. Climate spatial scales run from 50 km through global, and climate time scales run from a few weeks to centuries for current global change concerns. The requirements in this report and in Hansen et al. (1993) for clouds and radiation budget are often more stringent than in the NPOESS climate workshop. The NPOESS workshop does not appear to have used a consistent definition of the radiative forcings and feedbacks. Many of its threshold stability values would not be able to detect the decadal changes expected for forcings and feedbacks. Following Hansen et al. (1993), the current report tries to address the requirements in a consistent radiative forcing or feedback metric. It also assumes that the forcing or feedback must be detected accurately enough to assess decadal change at the level of 20% of the anticipated greenhouse gas forcings per decade. If four forcing and/or feedback mechanisms are found to be significant at this level and the data verifies that a future climate model predicts them to this accuracy, then in the simplest sense the uncertainty in future predictions by the climate model is composed of four likely independent errors, each of which is 20% of the base greenhouse forcing. We might anticipate in this scenario that the uncertainty in future predictions would be: 20% x square root (4) = 40%. This would be a dramatic improvement over the current factor of 4 or larger uncertainty. But it also suggests that the stricter stability requirements in the current document and in Hansen et al. (1993) are to be thought of as thresholds or minimum values, not as desired objectives. The objectives should be set even tighter by a factor of 2 to 4 (10% to 5% of the greenhouse forcing).

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

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

4.1.1 *Solar Irradiance as an example*

The IORD-II (NPOESS, 2001) requirements were reviewed and were endorsed for both total irradiance and spectral irradiance accuracy and stability. The threshold for absolute accuracy of total irradiance is 1.5 Wm^{-2} (0.1%), and for stability 0.02%/decade over many decades. As for many instruments, the stability of the active cavity radiometers greatly exceeds the absolute accuracy. At least a one-year overlap of observations is needed to remove instrument differences in absolute calibration. A 0.02%/decade stability requirement is sufficient to detect a 0.3 Wm^{-2} change in solar irradiance over a decade. This stability will constrain solar radiative forcing of the Earth's climate to within $(0.3)(0.25)(0.7) = 0.05$ Wm^{-2} per decade. The factor of 0.25 converts solar constant to the global average insolation over the Earth's surface, while the factor of (0.7) is the approximate fraction of energy absorbed by the Earth. This stability requirement will also allow rigorous tests of decade to century time scale variability in solar output as the length of the data record grows. The system would be capable of detecting 0.5 Wm^{-2} per century change in solar forcing. Even this subtle change would be a significant fraction of anticipated greenhouse gas forcing over the next century.

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

4.2 *Atmospheric Variables*

How Were the Requirements Set?
The expected decadal changes in a variety of atmospheric variables were used to determine accuracy and stability requirements. This usually involved using the expected response to global warming estimated from general-circulation model experiments. As in the previous section, we assume that a signal-to-noise ratioof at least 5 is required to reliably detect these changes from an instrument stability standpoint. The instrument accuracy, as has been discussed above, is less of an issue. As long as overlapping satellite records can be constructed to determine the calibration offsets between instruments, we can relax the absolute accuracy requirements to what is expected (and indeed already achievable) from a variety of sensor technologies in the coming decade. This is not to minimize the importance of understanding the sources of absolute accuracy errors, since some of these sources could conceivably affect the stability we require for climate monitoring. For many of the passive microwave or infrared technologies, instrument absolute accuracies of 0.5^{o} C can meet our requirements, as long as these accuracy numbers are dominated by a systematic bias that can be accounted for during satellite overlap periods.

4.2.1 *Total Column Water-vapor as an Example*

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

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

4.2.1.1 *Instrumentation requirements*

Weak water-vapor absorption lines in the infrared (on the wings of the 6.7 μm band or in the water-vapor continuum from 11μm to12 μm) or the microwave (around the 22 GHz water-vapor line) are used to observed emission from the lower atmosphere. Discrimination of water-vapor in the lower troposphere is dependent on the relative contrast between the surface emission and the atmospheric emission. In the infrared, both ocean and land surfaces have emissivities near 1.0, creating a low sensitivity to lower tropospheric water-vapor. In the microwave near 22 GHz, the ocean emissivity ranges from 0.5-0.6 but the land emissivity is near 1.0. Because of this, there is good contrast in the microwave and a greater sensitivity to changes in lower tropospheric water-vapor over the ocean versus infrared techniques. In the microwave, the water-vapor weighting function (i.e., change in transmittance with change in logarithm of pressure) is stable and the radiance is linearly related to brightness temperature. In contrast, in the infrared, the weighting function is more highly variable (and is a function of the water-vapor profile) and the radiance is a non-linear function of temperature (about T^8 near 6.7 μm).

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

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

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

$$a + bT_b = \ln (UTH\ P_{ref}) \qquad (3)$$

where T_b is brightness temperature, UTH P_{ref} is upper tropospheric humidity at P_{ref}, a reference pressure level, and b = -0.115. Although UTH depends on both water-vapor mixing ratio and atmospheric temperature, observations indicate that the main variations are due to the water-vapor. This equation indicates that to detect a 0.3 % change in water-vapor requires a stability of 0.03 K in brightness temperature.

4.3 *Surface Variables*

Surface variables include land vegetation, snow cover, sea ice, ocean color, and temperature. One problem with defining the requirements for the satellite measurements of the Earth’s surface is the wide range of surface types covered. A fundamental concern is the need for both accurate pre-launch calibration and post-launch validation of all satellite instruments. By their very nature satellite measurements do not directly sense the parameter of interest and it is only through these calibration and validation efforts that we can develop methods to estimate the desired parameters from the satellite data. These concerns apply both to the present and future satellite measurements.

4.3.1 *Sea Surface Temperature as an Example*

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

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

4.3.1.1 *Instrumentation Requirements*

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

$$SST = T_1 + 2.5(T_1 - T_2) \quad (4)$$

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

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

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

These values ignore the influence of sensor pointing angle on SST accuracy for the passive microwave sensors. Both the passive microwave and the thermal infrared sensors will require an in-situ calibration/validation program to insure that these requirements are met. This in-situ program must include both skin and bulk measurements of SST and should be continuous.

5 SUMMARY

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

7 REFERENCES

Anderson, J. M., K.J. Becker, H. H. Kieffer, and D.N. Dodd, 1999: Real-time control of the Robotic Lunar Observatory Telescope, *Astronomical Society of the Pacific*, 111, 737-749.

BIPM, 1998: The International System of Units (SI), 7th edition. Bureau International des Poids et Measures, Paris.

CGPM, 1995: Resolutions adopted by the 20th Conference Generale des Poids et Measures. Bureau International des Poids et Measures (BIPM), Paris.

Guenther, B., J. Butler, and P. Ardanuy, 1997: Workshop on Strategies for Calibration and Validation of Global Change Measurements. *NASA Reference Publication 1397*, 125 pp.

Hansen, J., W. Rossow, and I. Fung (Eds.) 1993: Long-Term Monitoring of Global Climate Forcings and Feedbacks. *NASA CP-3234*. National Aeronautics and Space Administration. Washington, D.C.

IPCC, 2001: Intergovernmental Panel on Climate Change: Third Assessment Report – Climate Change 2001. World Meteorological Organization, United Nations Environment Program.

ISO, 1993: International Vocabulary of Basic and General Terms in Metrology, 2nd edition. International Organization for Standardization.

Jacobowitz, H., 1997, ed.:Climate Measurement Requirements for the National Polar-orbiting Operational Environmental Satellite System (NPOESS): Workshop Report. UCAR, 71 pp.

Kieffer, H.H., T.C. Stone, R.E. Eplee Jr. and R.A. Barnes, 2003: On-orbit response history of SeaWiFS based on lunar calibration. In preparation.

National Research Council (NRC), 2000: Improving Atmospheric Temperature Monitoring Capabilities, Letter Report from the Panel on Reconciling Temperature Observations, Climate Research Committee, 17 pp

National Research Council (NRC), 2001a: Climate Change Science: An Analysis of Some Key Questions, National Academy Press, 29 pp.

National Research Council (NRC), 2001b: Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Part II. Implementation. National Academy Press, 98 pp.

NIST SP 811, 1995: Guide for the use of the International System of Units (SI). *NIST Special Publication 811, 1995 edition.* U.S. Government Printing Office, Washington, DC.

NPOESS, 2001: Integrated Operational Requirements Document II (IORD II) – http://npoesslib.ipo.noaa.gov/

Quinn, T. and C. Fohlich, 1999: Accurate radiometers should measure the output of the sun. *Nature*, **401**, 841.

Taylor, B. and C. Kuyatt, 1994: Guidelines for evaluating and expressing the uncertainty of NIST measurement results. *NIST Technical Note 1297, 1994 edition.* U.S. Government Printing Office, Washington, DC.