NASA Science Mission Directorate
Research Opportunities in Space and Earth Sciences – 2010
NNH10ZDA001N-ESDRERR
A.32 Earth System Data Records Uncertainty Analysis

NASA's new EARTH SYSTEM DATA RECORDS UNCERTAINTY ANALYSIS PROGRAM will extend and enhance the use of Earth System Data Records, including Climate Data Records, through rigorous estimation of error in ESDRs used by NASA communities. The selected projects will increase the science value of Earth System science measurements by identifying and validating systematic errors, and improving error estimations. Projects are expected to provide in-depth analysis of the properties of long-term data sets, with a focus on detecting systematic error, better quantifying error, and properly attributing uncertainty sources. A second focus is to resolve known issues of such data sets. In so doing, projects may orchestrate correct (and appropriate) methodologies, and may utilize advanced algorithms, techniques, and technologies that advance the understanding of uncertainties in Earth system science measurements. Resultant tool development is a third focus of the program.

The scope of problems that Earth System Data Records Uncertainty Analysis projects address include:

- Estimating, validating, and conveying measurement differences between sensors or between sensors, validation measurements and/or models;
- Estimating, validating, and conveying measurement errors in merged data products;
- Estimating, validating, and conveying systematic errors in long-term Earth system science data records; and
- Other contributions to the determination, validation, and conveyance of Earth science measurement quality and quantification of uncertainties.

This program will support 21 projects with funding of approximately $6 M per year for a nominal period of 3 years, although some projects have shorter or longer periods of performance. As these projects complete their investigations, NASA expects to build a Best Practices capability for consultation and use by Earth system science researchers involved in building Earth system data records. NASA will evaluate the need of future ESDR Uncertainty Analysis solicitations as these projects near completion.

Steven Ackerman/University of Wisconsin-Madison
Estimating Uncertainties in MODIS Cloud Data Records

In response to ROSES 2010 solicitation A.32, Earth System Data Records Uncertainty Analysis, support is requested to implement a methodology to analyze systematic error and uncertainties in the MODIS cloud fraction, top pressure, effective radius, and optical thickness. The goal of this proposal is to derive uncertainty estimates for the long-term MODIS cloud climatologies. The products of this error analysis will enable the proper use of MODIS cloud climatologies by climate modelers and climate researchers. It will allow for the meaningful use of the MODIS cloud products in global radiation budget and water cycle analyses. While the MODIS cloud top and optical property algorithm developers have characterized the pixel-level performance of their algorithms using various empirical and/or analytic approaches, a full accounting of the uncertainties due to observation geometry, underlying surface type, and cloudiness characteristics remains.

We propose to attribute aggregation error to specific sources and convey the uncertainty introduced into the long-term mean cloud properties from these errors. We will build on work that has already been done to assess the absolute error in the MODIS cloud fraction and cloud top pressure using comparisons to CALIPSO. We will further use compositing and other statistical approaches to analyze the relative errors reported in the MODIS pixel-level cloud optical and microphysical properties data sets. These analyses will be cognizant of the inherent correlations between error sources in the data record and that the magnitude of the pixel-level error is not constant across spatial domains or atmospheric cycles (e.g., the annual, seasonal, diurnal cycles, and atmospheric oscillations like ENSO). The analysis proposed here will be the extension of multiple piecemeal efforts already undertaken to characterize error in the satellite data records.

Robert Adler/NASA Goddard Space Flight Center
Uncertainties in Global and Regional Precipitation Using the GPCP and TMPA Data Sets

Although satellite-based precipitation retrievals have advanced significantly over the past thirty years and are being used for climate analysis and for hydrological applications, the estimate of associated errors for easy use by users has lagged. This proposal focuses on defining useful estimates of the systematic and random errors of two, very highly used merged data sets of surface precipitation, for which the proposers are responsible. These important data sets are the Global Precipitation Climatology Project (GPCP) monthly and daily analyses and the 3-hr resolution TRMM Multi-satellite Precipitation Analysis (TMPA [often known by its product designation of 3B42]). Formulas for error estimates or uncertainties will be developed for both type products at instantaneous/3-hr to daily to monthly time scales. Both random errors (including sampling) and systematic or bias errors will be included. The error formulas will be developed and included in a tool so users can employ the results over areas and times of their choosing and also will be attached to the products as additional fields. The approach used will be a combination of theoretical error models with parameters adjusted based on empirical analysis against ground data. Important parameters to be included in deriving the error models will include satellite data type, rainfall system structure based on the spatial and temporal variation of the precipitation, surface temperature, land/ocean/coast and other factors. The end result of the research will be tested error models for both random and systematic (bias) error that will be applied to the data sets in question for scientific and application use. The error models will also be provided as tools for use or adjustment by the users and tools will also be supplied for time and space averaging of errors, a typical need of users.

Mark Bourassa/Florida State University
Uncertainty in the Winds Climate Record

The short-term goal is to identify biases and spurious trends in the satellite winds record. Winds are an essential climate variable. Winds (more accurately stresses, but wind is commonly used as a proxy for stress) are related to air/sea exchanges of momentum, energy (proportional to latent and sensible heat fluxes, and a modifier of surface emissivity), moisture (evaporation), and gas fluxes (e.g., CO_2). Consequently, changes in winds contribute to changes in the budgets for energy, momentum, moisture, and CO_2. Throughout the QuikSCAT period the annual average global ocean and tropical Pacific Ocean winds have been remarkably consistent. There are now several products (e.g., HOAPS, IFREMER, and CCMP) that combine winds from multiple satellites, and show multi-decadal trends. In one of these products, the trends are due to incorrect handling of the seriously rain contaminated data; this problem is due to an error in the development of the merged product. In the cases of the CCMP product, the satellite winds have been very carefully intercalibrated (except for very high wind speeds, which will be improved in the next release of these data sets). This suggests that the trend for this product is either correct or due to either sampling or characteristics of the gridding process.
We propose to investigate the consequences of sampling in the CCMP wind product and the COAPS gridded wind product (currently nearly completion of the development phase; covering a similar period as the CCMP product). Analysis of wave data strongly suggests that changes in satellite winds contribute to changes in the resulting surface water waves. We will determine if these changes are also evident in the CCMP and COAPS wind products. For the COAPS product, a very careful error analysis will be done on a pixel-by-pixel basis; errors related to observation error, sampling, and representation (interpolation in space and time) will be considered. The PO.DAAC will collaborate in distribution and visualization of the COAPS product including the error analyses.

Amy Braverman/Jet Propulsion Laboratory
Likelihood-based Quantification of Agreement between Climate Model Output and NASA Data Records

This proposal responds to that portion of the ROSES 2010 A.32 opportunity calling for estimating and conveying measurement differences between sensors and models. We will apply statistical methods based on likelihoods for quantifying how well observational Earth system and climate data records (ESDR's and CDR's, hereafter collectively called data records) agree with corresponding climate model output. This type of evaluation and diagnosis has traditionally been based on simple summary statistics that do not account for important distributional characteristics of observational or model generated data. For instance, many comparisons are based on means and standard deviations of time series, without accounting for extreme values or spatial and temporal dependence. The work we propose uses formal statistical likelihoods to capture and assess the consistency of an observed data record with model predictions of it on the basis of more complete distributional information.

The likelihood function is the conditional probability distribution of an unknown quantity as a function of the conditioning quantity. For instance, if P(A|B) (read ``the probability of A for a given value of B) is normal with mean B, then the likelihood function for the mean is a function of different candidate values of B: L(b)=P(A|B=b). In other words, it shows how the probability of A changes when we assume different values of B are true. We will let A be statistics based on data records of water vapor, primarily from the AIRS instrument, and B be a climate model identifier. We will use model-generated data to estimate these conditional probability distributions of the statistics given different models. Then we ``score'' the agreement between observations and models by the relative values of the model-derived probabilities of the observed statistics. The model for which this relative probability is highest is the most likely model in the sense that its output is most consistent with the observations.

We will apply this methodology using two different observational water vapor data records, and model output for the CMIP5 near-term, decadal prediction core experiments which cover the decade of the 2000's. The first water vapor record consists of observations from the AIRS Level 2 Standard Data Product. It provides multiple observations per day at fine spatial resolution, and allows us to grade consistency with model output on the basis of high-frequency, process-level behavior. The second data record is from the current, funded MEaSUREs project, ``A Multi-Sensor Water Vapor Climate Data Record Using Cloud Classification'' (PI Eric Fetzer). This record provides water vapor measurements at coarse spatial and temporal resolution (i.e. five or ten degrees and monthly), and will allow us to assess consistency of observations with models on the basis of longer term, climatological phenomena. In both cases, we will iterate our analysis on problematic segments of the time series in order to score the models in a gross sense, and also to diagnose possible reasons for discrepancies relative to the data records.

Jeff Dozier/University of California
Error Analysis of MODIS Fractional Snow-Covered Area and Snow Albedo in Mountain Regions

Several MODIS-derived Earth System Data Records have been developed to assess snow resources in mountain areas. However, they have not been rigorously validated, and their uncertainties and systematic errors are not well characterized. We propose to validate: (i) daily MODIS fractional snow cover, calculated by both normalized difference snow index (NASA standard products MOD10A1 and MYD10A1), and spectral unmixing; (ii) daily snow albedo; and (iii) spatiotemporally smoothed fractional snow cover and albedo, as a monthly product.

For several wet and dry years, we will assemble a set of products and their validation datasets. The validation data will include a comprehensive (~1000) set of Landsat scenes, augmented at some locations and times by higher resolution data (digitized aerial photographs, Quickbird imager, and declassified images from national reconnaissance satellites) from which accurate binary maps of snow cover can be created. We will validate snow in the forests with existing LiDAR snow survey data, to characterize snow depth variability at sub-meter scales, including sub-canopy areas. We will use also use existing ultrasonic snow depth sensors to characterize snow accumulation, melt, and duration in open, canopy-edge, and sub-canopy conditions.

MODIS looks ±55° from nadir, yielding (because of Earth curvature) a ±65° ground zenith angle. During persistent cloud cover, the only view of the ground for a week or two might well be at such a highly oblique angle. To evaluate fractional snow cover in those situations, we will use Landsat TM or ETM+ scenes that coincide with the edges of snow-covered MODIS images.

We will compare time-space smoothing and interpolation to a sequence of Landsat TM or ETM+ scenes over the Sierra Nevada from 2000 to present, focusing especially on locations and times when the smoothed and interpolated product differs from the raw MODIS fractional snow cover. We will validate both the time-space continuity product's predictive mode (estimate snow-cover on a given day using only data prior to that day) and its retrospective mode (reconstruct the history of snow properties for a prior period.)

We have access to snow albedo data and coincident AVIRIS imagery from field sites at the Senator Beck Basin in the San Juan Mountains, Colorado and at Mammoth Mountain in the Sierra Nevada, California. We will use these data to validate AVIRIS-derived snow albedo, and then coincident AVIRIS/MODIS data to validate MODIS-derived albedo.

The MODIS cloud product contains errors of both omission and commission, and all MODIS snow products sometimes confuse clouds and snow. The time-space continuity product attempts to correct this by forcing snow that is too fine-grained and covers the whole pixel to be classified as cloud, while clouds that are too coarse-grained are classified as snow. We will validate this approach using Landsat TM and ETM+, which can identify clouds both by texture and by snow/cloud contrast in band 5.

Snowmelt runoff forecasting in mountainous areas has historically used empirical models forced by sparse, in situ measurements. These models already have large errors in some years, and since they rely on a data record that assumes stationarity, they are theoretically ill-suited for water management in a changing climate. Coupling fractional snow cover and albedo to snowmelt models and reservoir operations would be significantly advanced by our proposed investigation, which would validate the products, analyze the structure of errors, and advise users of caveats and likely accuracy. Moreover, the validation datasets will be useful after the investigation is finished, as comprehensive test cases for future researchers.

Eric Fetzer/Jet Propulsion Laboratory
Uncertainty Estimates in the A-Train Water Vapor Climate Data Record

We are currently creating a Climate Data Record from several sensors in the A-Train satellite constellation under the NASA-funded Making Earth Science Data Records for Use in Research Environments (MEaSUREs) project "A Multi-Sensor Water Vapor Climate Data Record Using Cloud Classification." In that effort we are developing a water vapor climatology conditioned on cloud state, with an emphasis on mean and variability, based on standard water vapor products from AIRS, AMSR-E and MLS, and cloud products from CloudSat and MODIS. In the proposed work we will incorporate comprehensive error estimates into our MEaSUREs Climate Data Record. We will assess the usefulness of the standard data product error estimates through intercomparison of the satellite data sets, and also through direct validation with in situ radiosondes observations. We will also provide estimates of uncertainties in a hierarchy of quantities produced from the A-Train data products, varying from individual retrievals (where possible) to uncertainties on trends at regional and global scales. We will provide a coarse resolution Climate Data Record, with rigorously developed uncertainty estimates, for studies of decadal-scale changes in the Earth water vapor. This data record will be designed to test for water vapor increases required for a strong and positive water vapor feedback response to anthropogenic warming.

Simon Hook/NASA/Jet Propulsion Laboratory
Estimating, Validating and Conveying Measurement Differences in the Land Surface Temperature and Emissivity Products from ASTER, MODIS and AIRS

Land surface temperature and emissivity (LST&E) has been identified by the NASA Earth Science Division as a key Earth System Data Record (ESDR). An ESDR is defined by NASA as a unified and coherent set of observations of a given parameter of the Earth system. Currently there are significant discrepancies among the LST&E products provided by NASA that arise due to the different methods used to derive them, and the different spatial and spectral scales associated with each sensor. The goal of this study is to reconcile and quantify these differences to enable the development of a LST&E ESDR with a clearly defined set of uncertainties.

Sensors such as ASTER, MODIS and AIRS on NASA's Terra and/or Aqua satellites produce LST&E products at a variety of spatial, spectral and temporal scales. While validation studies have been undertaken for a particular product from a given sensor, there have been relatively few studies that have examined and conveyed the sources of uncertainty between the products with a view to producing a single LST&E ESDR. The lack of such studies limits the usefulness of satellite-derived LST&E data as an ESDR, and as inputs to ecosystem and climate models. For example, an error of 0.1 in the LSE will result in current climate models having errors of up to 7 Wm-2 in their upward longwave radiation estimates - which is a much larger term than the surface radiative forcing due to an increase in greenhouse gases (~2-3 Wm-2), making accurate knowledge of the LST&E a critical component for climate change studies. Such errors can occur under certain circumstances with existing NASA LST&E products. An in-depth study on error and uncertainty analysis between current LST&E products is essential to provide an uncertainty metric by which LST&E climate datasets and trends from future sensors can be measured.

The main goals of this study will be to 1) identify and resolve the largest sources of uncertainty for each of the LST&E products from ASTER, MODIS and AIRS, 2) produce a comprehensive, land-cover dependent error budget for the standard LST&E products from ASTER, MODIS and AIRS and 3) evaluate the implications of these uncertainties when the products are used in ecosystem and climate models.

This study will address the Earth System Data Records Uncertainty Analysis Program by 'Estimating, validating and conveying measurement differences between sensors'. We will investigate sources of uncertainty between the LST&E products from ASTER, MODIS and AIRS and their impact on existing climate and ecosystem models. We will provide a clearly defined uncertainty budget for LST&E which will enable the creation of a LST&E ESDR for use by the modeling community and in data assimilation systems at multiple spatial, spectral and temporal scales.

Yongxiang Hu/NASA Langley Research Center
Using CALIPSO/CloudSat Observations for Assessing Uncertainties in Atmospheric Radiative Fluxes and Their Vertical Distributions

Active remote sensing of aerosols and clouds by CALIPSO and CloudSat can help reduce the uncertainties in atmospheric radiative flux calculations. Together, the lidar aboard CALIPSO and the radar aboard CloudSat are uniquely capable of providing unambiguous information about (a) the vertical distributions and profiles of clouds and aerosols; (b) night-time cloud properties; (c) the spatial distributions and optical properties of aerosols and thin ice clouds located above bright water clouds and/or snow/ice/land surfaces; (d) multi-layer cloud occurrence and the determination of cloud thermodynamic phase; and (e) retrievals of aerosol and cloud optical depths without assuming single scattering properties. Using the backscatter from well-characterized targets (e.g., molecular backscatter, ocean surfaces, and water clouds) below aerosols and transparent ice cloud layers, layer optical depths can be derived accurately without making assumptions about aerosol and ice cloud microphysics. Atmospheric radiative fluxes and their vertical distributions can then be accurately computed by combining the aerosol/cloud observations from CALIPSO and CloudSat with data from the passive remote sensing instruments aboard other A-Train satellites. These newly derived and more accurate fluxes will be used for assessing the uncertainties of the atmospheric radiative fluxes in the existing climate record.
The proposed study includes,

1. Developing an improved aerosol and cloud product using optical depths derived from combined active and passive measurements made by CALIPSO/CloudSat and other collocated A-train satellite observations.

2. Quantifying the uncertainty of the A-train multi-sensor aerosol and ice cloud products using (a) comparisons with existing aircraft validation data, and (b) comparisons of ice cloud optical depths derived from CALIPSO measurements of transparent layers constrained by backscatter magnitudes from clear air, ocean surfaces, and/or water clouds below the ice cloud layer.
3. Assessing errors in aerosol and ice cloud properties due to passive remote sensing limitations by comparing the A-train multi-sensor aerosol and cloud products with collocated MODIS aerosol and cloud statistics, such as aerosol and cloud detection, layer number and vertical location, cloud amount, aerosol and cloud optical depths, cloud liquid and ice water paths.
4. Evaluating the uncertainties in atmospheric radiative fluxes, their vertical distributions, seasonal and inter-annual variations by comparing estimates derived from the A-train multi-sensor aerosol and cloud properties to those currently computed from collocated MODIS aerosol/cloud products.
5. Estimating the link between uncertainties of the water cycle in perturbed physics ensemble models and the uncertainties of the atmospheric heating/cooling rates.

The objective of this proposed study is to assess and quantify the improvement in radiative flux estimates that result from the inclusion of long term active remote sensing on existing and future satellite missions such as CALIPSO, EarthCare, and ACE.

Charles Ichoku/NASA Goddard Space Flight Center
Coherent Uncertainty Analysis of Aerosol Data Products From Multiple Satellites

Among the known atmospheric constituents, aerosols represent the greatest uncertainty in climate research. Although several NASA and other satellite-borne sensors routinely measure aerosols, there is often disagreement between similar aerosol parameters retrieved from different sensors, leaving users confused as to which sensors to trust for answering important science questions about the distribution, properties, and impacts of aerosols. As long as there is no consensus and the inconsistencies are not well characterized and understood, there will be no way of developing reliable climate data records from satellite aerosol measurements. Although some of the discrepancies could be due to imperfections in the retrieval algorithms or other external factors, part of it could also be due to the fact that the satellite measurements reflect the unique (spectral, spatial, angular, or polarization) characteristics of the individual sensors, which could each represent an advantage, depending on the type of aerosol parameter being retrieved. Fortunately, the most globally representative well-calibrated ground-based aerosol measurements corresponding to the satellite-retrieved products are available from the Aerosol Robotic Network (AERONET). This proposal team was funded under the ACCESS program element of the ROSES 2007 solicitation to collocate the AERONET products with corresponding satellite-retrieved products from multiple sensors, to support the integrated validation and uncertainty analysis of the latter. That project known as the expanded MAPSS (Multi-sensor Aerosol Products Sampling System) is nearing completion, and plans are underway to release the first version by the middle of June 2010. When released, the MAPSS data system will facilitate the joint analysis of Level 2 aerosol products sampled from multiple sensors in a uniform fashion. Therefore, it has become necessary and feasible to provide the scientific community with solid information on the uncertainty characteristics of the aerosol products from the different satellite sensors in a quantitative and comprehensive manner, to enable them to make balanced judgment on the best combination of aerosol products to use for a given application, region, or time period. That is what this proposal seeks to accomplish.

B Johnson/NIST
Uncertainty Analysis of the 13 year Time Series for the in situ Vicarious Calibration of Ocean Color Satellite Sensors

Substantial effort by NASA has been invested in order to create and maintain a long-term, consistent, and calibrated time series of ocean color radiometry over multiple missions and satellite sensors. This is a very difficult measurement problem because the water-leaving radiance, or amount of sunlight scattered out of the oceans, is a small fraction of the total radiance measured by the satellite sensor. As a result, the SI traceability of ocean color radiometric values relies completely on a vicarious calibration approach utilizing reference oceanic sites. The ocean color bands in the SeaWiFS sensor, the MODIS Terra and Aqua sensors, and international sensors such as MERIS utilize the Marine Optical BuoY in situ sensor located off the coast of Lani, Hawaii, for vicarious calibration. The vicarious calibration is performed on the oceanic-atmospheric system as observed during satellite overpasses of the MOBY site. To improve the Type A uncertainty, multiple matchups acquired over an extended time interval are necessary. A robust and rigorous uncertainty analysis of this data set is outstanding. Broadly speaking, there are three aspects to the uncertainty budget for the long-term time series of the global ocean color radiometric data set: the in situ radiometric time series, the in situ to satellite matchup time series for determination of the vicarious calibration gain coefficients, and the global, satellite derived values for water-leaving radiances (or remote sensing reflectances). The uncertainty budget has elements attributed to sensor characterization functions (which change in time), natural in situ variability from environment effects, and the veracity and efficacy of the measurement equations (including models and algorithms) that describe the complete methodology. We propose to carry out a rigorous analysis of uncertainty of the results from MOBY. We will develop a complete uncertainty model, investigate and quantify systematic sources of error, and develop and implement appropriate correction algorithms for these sources of bias. The analysis methodology will comply with international practices, e.g., ISO "Guide to the Expression of Uncertainty in Measurement." We will perform a sensitivity study of the impact on the global ocean color time series, including validation against global in situ measurements, agreement between satellite sensor time-series, and changes in global and regional trends. The results are expected to establish the MOBY data set as a SI traceable climate data records. The improvements will contribute to establishing and reducing uncertainty in the ocean color satellite data products, which will be realized in the next reprocessing of the SeaWiFS and MODIS ocean color products that will likely occur toward the end of FY 2013.

John Kimball/The University of Montana
Accuracy Assessment and Attribution of Uncertainty for the Global Freeze-Thaw Earth System Data Record

The freeze-thaw (FT) state parameter derived from satellite microwave remote sensing quantifies the predominant frozen or non-frozen state of the landscape and is closely linked to surface energy budget and hydrological activity, seasonal dynamics of vegetation growth, terrestrial carbon budgets and land-atmosphere trace gas exchange. We are developing a global Earth System Data Record of daily landscape freeze-thaw state dynamics (FT-ESDR) through an existing NASA MEaSUREs funded activity. Approximately three reprocessing and FT-ESDR releases are planned under this effort, involving variable duration satellite records and multiple, overlapping sensors and data series. An initial FT-ESDR release is available online through the NASA NSIDC DAAC. The initial FT-ESDR has been verified using in situ temperature measurements from global weather stations. Other data quality (QC) metrics provide spatially contiguous information on FT-ESDR accuracy. The QC data include potential negative impacts of temporal gaps in sensor data time series, open water, terrain and land cover heterogeneity. The current QC methods account for less than 25% of the variance in classification accuracy. With the remaining accuracy variance being unexplained, there is noteable uncertainty in the FT-ESDR. Available weather station networks are relatively sparse for many areas and may not adequately represent the global range of environmental conditions and FT variability. Potential utility of overlapping, multi-frequency active/passive microwave remote sensing data are also largely undefined due to lack of information regarding relationships between the various sensor retrievals and associated error sources. These methods also don't represent explicit systematic and random error components of the FT retrieval, including sensor footprint, frequency and FT sensitivity, temporal fidelity of the retrievals, signal degradation from atmospheric effects, in situ measurement error and satellite overpass timing.
We propose to conduct detailed sensitivity and error analyses of the FT-ESDR for improved quantification and attribution of FT retrieval uncertainties, involving multiple sensor data records, with variable frequencies, polarizations, overpasses, spatial resolution and temporal sampling. The satellite active/passive data records to be investigated include the Scanning Multi-channel Microwave Radiometer (SMMR), Special Sensor Microwave Imager (SSM/I), Advanced Microwave Scanning Radiometer on EOS (AMSR-E), SeaWinds-on-QuikSCAT, ERS-1/2 scatterometers, the Advanced Scatterometer (ASCAT), and the Soil Moisture and Ocean Salinity (SMOS) mission. Collectively these data enable a potential FT-ESDR record spanning more than 30-years. We will apply a suite of forward radar backscatter and microwave emissions models to determine microwave sensitivity to individual and cumulative error sources, and track resulting error propagation through the FT retrieval algorithms and resulting classification results. We will conduct spatial and temporal scale degradation exercises using existing multi-scale satellite microwave remote sensing records to quantify spatial resolution and temporal fidelity components of FT classification uncertainty. These activities will enable new FT algorithms and classifications using ensemble measurements from multiple active/passive microwave sensors, frequencies and polarizations. We will also exploit new data sources from upcoming FT field campaigns

under the SMOS and NASA Soil Moisture Active Passive (SMAP) Decadal Survey missions; these data will be applied with existing measurement networks for forward model calibration and FT validation. Sensitivity studies will be conducted using both model simulations and satellite and in situ network measurements.
The results of this work will include an improved global FT-ESDR with well quantified accuracy, including a detailed error budget for the FT measurement.

Susan Kulawik/Jet Propulsion Lab
Estimation of biases and errors of CO_2 satellite observations from AIRS, GOSAT, SCIAMACHY, TES, and OCO-2

We propose a rigorous error estimation methodology to unify atmospheric CO_2 data products generated from NASA's AIRS, TES, GOSAT-ACOS and upcoming OCO-2 satellite sensor projects. Errors will be characterized using vertical profiles and total column CO_2 measurements traceable to international reference standards for atmospheric CO_2. Our methodology accounts for the unique measurement attributes of each satellite sensor via its averaging kernel (Rodgers and Connor, 2003) to yield consistent precision and bias estimates for all space-based CO_2 measurements, and assess the impact of clouds, spatial and temporal averaging, and characterize biases as a function of time and location. This work is essential for producing consistent, calibrated, long-term CO_2 data products across multiple satellite sensors, and lays the groundwork for creating the first atmospheric CO_2 Earth Science Data Record (ESDR).

Zhong Liu/GES DISC DAAC and George Mason University, NASA Goddard Space Flight Center
Integrate IPWG Validation Algorithms into TRMM Online Visualization and Analysis System (TOVAS)

Problem Statement: Significant progress has been made over the past decade in satellite precipitation product development. However biases and uncertainties are common among precipitation products and an obstacle exists in quickly gaining knowledge of product quality, biases and behavior at a local or regional scale, namely user defined areas or points of interest. Current online operational inter-comparison and validation services have not addressed this issue adequately. To expedite the use of satellite precipitation products, the obstacle needs to be removed.

We have developed a TRL-7 prototype (http://disc2.nascom.nasa.gov/Giovanni/tovas/rain.ipwg.shtml) in the TRMM Online Visualization and Analysis System (TOVAS) to address this issue (Liu et al. 2009). Despite its limited functionality and datasets, users can use this tool to generate customized plots within the United States for 2005. In addition, users can download customized data for further analysis, e.g. comparing their gauge data.
Analysis Method: To meet increasing demands, we propose to integrate community recognized IPWG (International Precipitation Working Group) Validation Algorithms/statistics into the TRMM Online Visualization and Analysis System

(TOVAS), allowing users to generate customized plots and data. In addition, we will expand the current coverage from the United States to global. TRMM monthly, other precipitation products and their derived climatology products will be added as well.

Whenever an algorithm team changes their product version number, users would like to know the differences between the new and the old by inter-comparing both versions of products in their areas of interest. Making this service available to users will help them to better understand associated changes, such as, biases, which is particularly important to application users. We plan to implement this inter-comparison in all TRMM Level-3 monthly standard products.

Significance: Precipitation is hard to measure and difficult to predict. Each year droughts and floods cause severe property damages and human casualties around the world. Accurate measurement and forecast are important for mitigation and preparedness efforts. The proposed additions will complement and accelerate the existing and ongoing validation activities in the community and enhance existing data services for TRMM and the future Global Precipitation Mission (GPM).

Relevance to the Program Element: The proposed additions response to the 3rd focus of the program for tool development. This proposal directly addresses NASA's Earth Science Division research themes by providing estimating, validating and conveying measurement differences between sensors, in merged data products (e.g. 3B42, 3B43), and in long-term Earth system science data records. The proposed IPWG algorithms have been widely used and recognized by the precipitation community.

Anthony Mannucci/Jet Propulsion Laboratory
Radio Occultation Climate Records

The objective of the proposed effort is to create a long-term, consistent, calibrated GPS radio occultation (RO) data record spanning 18 years, from 1995 to 2013. This effort is urgently needed by the Earth science community and by projects within the Earth Science Data Systems (ESDS) program. The Decadal Survey for Earth Science has identified radio occultation as a critical measurement for observing decadal-scale climate change and has recommended RO for the CLARREO mission. The Climate Virtual Observatory Project within ESDS is using RO to recalibrate AIRS, AMSU and MODIS. Both of these projects require well-calibrated RO records: CVO requires 0.1 K temperature accuracy, and CLARREO requires 0.03% fractional refractivity accuracy. This level of calibration accuracy has not yet been achieved for RO measurements. We present a detailed plan to push the state-of-the-art in RO uncertainty analysis, and develop algorithms that correct the retrievals themselves so that they are accurate to 0.1 K and 0.03%. We will apply these algorithms to the RO data spanning 18 years, from GPS/MET through to the present COSMIC era. A calibrated RO data set will be distributed to the Earth science community, including CVO researchers, for their free use.

Kyle McDonald/Jet Propulsion Laboratory
Assessment of Accuracy and Uncertainty of the Inundated Wetlands Earth System Data Record

Wetlands cover less than 5% of Earth's ice-free land surface but exert major impacts on global biogeochemistry, hydrology, and biological diversity. The extent and seasonal, interannual, and decadal variation of inundated wetland area play key roles in ecosystem dynamics. Despite the importance of these environments in the global cycling of carbon and water, there is a scarcity of suitable regional-to-global remote-sensing data for characterizing their distribution and dynamics. Through a current NASA MEaSUREs project, we are assembling a global-scale Earth System Data Record of natural Inundated Wetlands (ESDR-IW) to facilitate investigations on their role in climate, biogeochemistry, hydrology, and biodiversity (McDonald, PI; Chapman, Hess, Kimball, Moghaddam, Co-Is). The ESDR comprises (1) Fine-resolution (100 meter) maps, delineating wetland extent, vegetation type, and seasonal inundation dynamics for regional to continental-scale areas covering crucial wetland regions, and (2) global multi-temporal 25 km mappings of inundated area fraction (Fw) across multiple years. The fine-scale ESDR component is constructed from synthetic aperture radar (SAR) data from JAXA's JERS satellite and from their Phased Array L-Band SAR (PALSAR) on-board the Advanced Land Observing Satellite (ALOS). The global maps of inundated area fraction are derived at 25 km scale from remote sensing observations from active/passive microwave instruments. We have identified key issues which contribute to uncertainty in the ESDR data sets. Error sources include radiometric inconsistency of the remote sensing data sources, paucity of ground validation datasets available for implementation of classification algorithms, temporal undersampling relative to hydrologic variability, and ambiguities associated with implementation of coarse-resolution mixture models. We propose to conduct systematic analysis of error sources related to all aspects of ESDR-IW assembly, including uncertainties associated with remote sensing and ground training and validation data sets employed, algorithms applied, and cross-product harmonization. To accomplish this, we will (1)Estimate error sources associated with derivation of SAR-based wetlands maps, analyzing uncertainty associated with JERS and PALSAR radiometric attributes, data set compositing times, training data used in implementing retrieval algorithms, accuracy of retrieval algorithms, and application of those algorithms across broad landscape regions where ground-based measurements are not available for training and validation.(2)Assess error sources associated with generation of the 25 km global inundation data sets, analyzing uncertainties associated with the multi-platform remote sensing time series that support generation of the long-term global data record, with the temporal compositing employed in generating the time series data record, with calculation of Fw through mixture model and radiometric retrievals, and with atmospheric effects and radio frequency interference.(3)Employ the single normal equation simulation (SNESIM), a state-of-the-art geostatistical method for sub-pixel mapping, to quantify in a spatially explicit manner the uncertainty in our global coarse-resolution inundation area fraction retrievals. (4)Evaluate consistency of our products with similar global products over globally representative locations and periods. This work will create an enhanced ESDR of inundated wetlands with statistically robust uncertainty estimates. The ESDR documentation will include a detailed breakdown of error sources and associated uncertainties within the data record. This effort will ensure that the ESDR-IW inundation products will be the best available data sets for global-scale modeling that involves a surface water component. The ESDR data sets will also benefit preparation for NASA Decadal Survey missions including Soil Moisture Active-Passive (SMAP) and Surface Water Ocean Topography (SWOT).

Carl Mears/Remote Sensing Systems
Complete Error Characterization of the DISCOVER Earth System Data Records

We propose a detailed error analysis of the ocean products provided by the DISCOVER project, which is funded through NASA MEaSUREs program. The ocean products all come from an array of operational and research satellite microwave (MW) radiometers. The prod-ucts are surface wind speed, columnar atmospheric water vapor, columnar cloud liquid water, rain rates and sea surface temperature (SST). In addition to products derived from individual satellites, DISCOVER also produces merged, multi-instrument wind, vapor, cloud, rain and SST products as well as derived hydrological products such as evaporation.
We will perform a comprehensive error analysis that considers measurement noise, algorithm sensitivity, geophysical model accuracy, and the influence of contamination parameters such as rain, land, sea ice, and RFI. The project will have two main deliverables.
1. A pixel-by-pixel error estimate for each and every geophysical retrieval. These will be reported to users in the form of additional layers in our standard daily, 3-day average, and monthly products. These estimates will primarily describe short-term errors.
2. Estimates of long-term errors (i.e., decadal trend error bars) of each ocean product. These long-term error estimates will be communicated to the Users via a web-based Annual Validation Report
The addition of error information to the DISCOVER products will greatly increase their usefulness to all users. For example, the availability of error estimates will enable the accurate assimilation of the measurements into derived products and the assessment of the statistical significant of conclusions reached when using the products in research applications.

Timothy Moore/University of New Hampshire
Optical Water Type-Based Uncertainties of Satellite Ocean Color Products

Earth science data records derived from ocean color satellite products form an important source of information for studying and monitoring global ocean biogeochemical processes. These products, such as the chlorophyll-a concentration, are generated from the primary ocean color product - spectral water-leaving radiance - through bio-optical algorithms. The validity or accuracy of these derived products has historically been evaluated by comparing them with in situ measurements of the same property at discrete locations. Factors which collectively contribute to the overall uncertainty budget of the products include algorithm errors, radiance errors from atmospheric correction and sensor calibration, and match-up errors resulting from time and space mismatches. These error sources are not constant over space and time, and vary among optical environments in the oceans. We have developed a method based on the concept of optical water types for mapping ocean color product uncertainties, and we have demonstrated its application to estimate the MODIS chlorophyll product uncertainty. We are proposing to adapt this method to other ocean color satellite products with the end result of producing uncertainty maps for a suite of products. The advantage gained with our proposed method is that it can be applied to any ocean color satellite data to map uncertainty at the same scale as the radiance measurements.

Alexander Ruzmaikin/Jet Propulsion Laboratory
The Data We Use and Should Use for Climate Research

The goal of our proposal is to assist the development of long-term satellite data products suitable for climate research. The need for high quality climate data originates from the assessment that the Earth's global temperature trend due to increase in CO_2 is about 0.1K/decade. However larger changes can be seen on decadal scales in satellite data. These changes may be caused by low-frequency natural variability, such as ocean or solar variability. A single point measurement from a satellite infrared instrument typically has an accuracy of about 1K. Thus, large samples and statistical methods must be brought to bear.

Currently, most climate data productions are reduced to means and standard deviations within predefined grid cells. However climate change is driven by many physical processes and the resulting climate variables have complex probability distributions, in contrast to the Gaussian distributed standard errors of measurements, which are fully characterized by the means and standard deviations. Given this, to what extent is the use of means and standard deviations useful or correct? What information is lost?

On the user end, concern arises from the way the data are used to compare models output to observations. Commonly used methods of identification of climate variability based on linear and parametric approaches such as regression, correlation analyses and band-pass filtering do not properly take into account non-linear and non-stationary nature of climate variability and impose strong, unnecessary demands on accuracy and precision of data.

Our proposal addresses the following problem formulated in the AO: "Estimating, validating, and conveying systematic errors in long-term Earth system science data records". We investigate the uncertainties and errors in the construction and use of climate data records.

First, we will examine the validity of means and standard deviations as a basis for climate data products. We will explore the conditions under which these two simple statistics are inadequate summaries of the underlying empirical probability distributions by contrasting them with a nonparametric, method called Entropy-Constrained Vector Quantization technique designed to preserve the statistical properties of the underlying data and to provide a better way to summarize large volumes of remote sensing data for climate studies.

Second, we will carry out in-depth analyses of the properties of long-term data sets using data-adaptive methods, with a focus on detecting systematic errors caused by the use of linear and parametric assumptions. We will investigate applications of 1) Empirical Mode Decomposition, which accounts for non-linearity and non-stationarity without parametric or functional assumptions, and 2) Recurrent Synchronization to investigate how averaging and filtering influence phase relationships between two data sets.

We will use the Level 1B data and investigate the procedure of production of the Level 3 out of Level 2 (retrieved) data based on the measurements by the Atmospheric Infrared Sounder and Atmospheric Microwave Sounding Unit (AIRS/AMSU-A), a set of two instruments operating in the infrared and microwave regions on Aqua satellite. The instruments have been in operation about eight years and are projected for 14 years of operation thus providing a sufficiently long time series of data.

The proposed research and resulting technical recomendations will pave the way to improve data products that preserve the statistical information obtained in the measurements. It will significantly contribute to answering the critical questions on accuracy and time span of data that are sufficient to reliably evaluate climate variability using spacecraft data. This will be beneficial for the NASA climate change research overall, and specifically for formulation and evaluation of science goals for NASA missions, such as the CLARREO mission recommended by the Decadal Survey.

Yudong Tian/GEST/UMBC

Measurement Uncertainty and Error Propagation of Satellite-based Precipitation Sensors

The objective of this proposal is to quantify the uncertainties in NASA's satellite-based precipitation sensor ESDRs, and to identify the propagation of their systematic errors into merged multi-sensor precipitation measurements. We will analyze both the random errors and the systematic errors in 9 long-term microwave precipitation sensor measurements for complete uncertainty quantification. The approach for this proposal is to employ, improve and expand the scientifically rigorous, peer-recognized methods from our previous studies, and adopt well-established methods from other fields, with the support of a high-throughput data processing infrastructure and an extensive precipitation ESDR archive. Specifically, we will:

1. Determine the systematic errors in 9 major precipitation sensor ESDRs over a time period of 8-10 years, to establish the error characteristics of these sensors over the longest time-span possible. Such long-term error quantification will lay the foundation for uncertainty analysis on both the weather and climate time scales;
2. Quantify the random errors in the 9 precipitation sensor ESDRs with well-established methodologies adopted from existing studies and from other fields which do not require the use of reference data and their uncertainties (e.g., Stoffelen 1998; Caires and Sterl 2003; Tian and Peters-Lidard 2010); and

3. Understand and quantify the propagation of the sensor errors to downstream merged precipitation products, to connect and attribute the error characteristics documented in existing studies (e.g. Tian et al. 2009, 2010a) to the upstream sensor measurement errors determined in 1 and 2.

The proposed work will be critical in quantifying the uncertainties in NASA's precipitation ESDRs by determining both the systematic and random errors, and in tracking down the error sources and their relative contributions. These results on uncertainty analysis will be as significant as the data records themselves, and will be indispensible for a wide range of applications such as Earth system model data assimilation, sensor calibration and validation, multi-sensor algorithm development, climate trend analysis and decision making.

Juying Warner/University of Maryland, Baltimore County
Uncertainty Analysis of Tropospheric Carbon Monoxide Data Records Using AIRS and IASI from a Uniform Algorithm

The research objectives of this proposal are: 1) to quantify the errors and uncertainties in the tropospheric Carbon Monoxide (CO) products from recent and upcoming satellite missions, and 2) to provide a uniform algorithm that generates tropospheric CO retrievals from the Atmospheric Infrared Sounder (AIRS)/EOS/Aqua (Aumann, et al., 2003) and the Infrared Atmospheric Sounding Interferometer (IASI)/MetOp (Coheur et al., 2009) to address known issues. This product with known quality will provide twice daily global coverage of tropospheric CO for a period from 2002 through the lifetime of AIRS and continue with IASI instruments for the current and planned missions of 15 years started in late 2006 to obtain consistent CO climate records.

Under previous funding by ROSES, we developed an alternative algorithm for AIRS CO products using the Optimal Estimation (OE) method, described by Rodgers (2000), which is different from AIRS operational method (Warner et al., 2010). We use AIRS operational L2 meteorological and ozone profiles and the cloud-cleared radiances, which are provided by NASA/GES DISC (http://disc.gsfc.nasa.gov/AIRS/index.html) as input. The output from the new retrieval system not only includes global CO profiles as does by AIRS operational products, but also provides the Averaging Kernels (AKs), the error covariance matrices, and the degrees of freedom for signals (DOFS) that are computed using a similar formulation as in the other CO sensors such as MOPITT and TES. The new CO products have undergone validations against in situ measurements and have been inter-compared with MOPITT and TES CO. IASI on MetOp has similar temporal and spatial coverage as AIRS, and they are both hyperspectral thermal sensors. Therefore, the CO data products from both sensors can be retrieved very similarly. These products, with outputs that the user communities are more accustomed to, will improve community understanding of the data records and increase data usability.

Dong Wu/Jet Propulsion Lab
Evaluation of Cloud Ice Measurement Uncertainties in MLS, AMSU-B, MHS, CloudSat, and CALIPSO Data

Cloud ice content, a key Earth system data record (ESDR) in determining cloud radiative and dynamical impacts on weather and climate, remains associated with large uncertainties, not only among different observations but also among different climate models. The modeled cloud ice can differ by a factor 10-20, which now depends on quality global observations to improve and constrain some of the underlying physics. Large dynamic ranges of ice clouds require multiple sensors to measure their physical variables, and the recent observations from NASA's A-Train and NOAA sensors provide an unprecedented opportunity to evaluate, validate, and convey uncertainties of cloud ice measurements. This project focuses primarily on cloud ice measured by the advanced microwave sensors (MLS, AMSU-B, MHS and CloudSat) that have good capability of penetrating clouds for ice mass from different viewing geometry. We will investigate, characterize, and quantify the leading causes of the measurement uncertainty through statistical analysis, radiative transfer modeling, and intercomparisons of these sensors. Based on the uncertainty analysis from this study, corrections will be suggested for merging the long-term MLS, AMSU-B and MHS ice water path (IWP) data since 1991.