NATIONAL POLAR-ORBITTING OPERATIONAL
ENVIRONMENTAL SATELLITE SYSTEM (NPOESS)
INTEGRATED PROGRAM OFFICE
8455 Colesville Road
Suite 935
Silver Spring, MD 20910

NPOESS Community Collaborative Calibration/Validation Plan for the NPOESS Preparatory Project
CrIS/ATMS EDRs

DATE: 15 September

No. I30004
VER. 1 REV. B

PREPARED BY: ____
Dr. Christopher Barnet, NOAA/NESDIS/STAR

ELECTRONIC APPROVAL SIGNATURES:

________________________________ ________________________________

________________________________ ________________________________

Prepared by
NPOESS Integrated Program Office

Prepared for
U.S. Department of Commerce
National Oceanic and Atmospheric Administration (NOAA)
National Environmental Satellite, Data, and Information Services (NESDIS)
National Polar-orbiting Operational Environmental Satellite System
(NPOESS) Integrated Program Office (IPO)

Table of Contents

Table of Figures

Table of Tables

1.0 Introduction

This plan describes a coordination strategy for validating the Environmental Data Records (EDRs) generated by the the Cross-track Infrared Sounder (CrIS) and Advanced Technology Microwave Sounder (ATMS). Together known as the Cross-track Infrared Microwave Sounding Suite (CrIMSS), these instruments are being flown on the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP).

The investigators in the Calibration/Validation (C) Team, identified in Table 1, below, will lead activities to demonstrate that the CrIMSS produces EDRs that are within specification and useful to the community.

Table 1 - Members of the CrIS/ATMS EDR Cal/Val Team

Team Member	Funding Source	Activity
Chris Barnet	NOAA/NESDIS/STAR	Cal/val coordination, scientific campaigns of opportunity.
Steven M. Beck	Aerospace	*In situ* LIDAR and radiosonde measurements and evaluation.
William Bell	ECMWF	Global characterization of CrIS/ATMS biases
Gail Bingham	USU/SDL	SDR cal/val lead, SDR proxy datasets.
Bill Blackwell	MIT	ATMS SDR/EDR issues, NAST-M issues, ATMS proxy datasets.
Changyong Cao	NOAA/NESDIS/STAR	Integrated instrument cal/val system for NPP/NPOESS
Stephen English	UKMET	Global characterization of CrIS/ATMS biases
Steve Friedman	NASA/JPL	NASA Sounder PEATE
Mitch Goldberg	NOAA/NESDIS/STAR	Operational utilization of advanced sounders, CEOS/GCOS
Denise Hagan	NGAS	NGAS cal/val activities, POC for CrIMSS SDR
Degui Gu	NGAS	NGAS cal/val activities, POC for CrIMSS EDR
Allan Larar	NASA/LaRC	NAST-I preparations, intensive validation campaigns, LaRC-EDR.
Xu Liu	NASA/LaRC	NGAS code assessments, LaRC-EDR algorithm issues, CrIS proxy datasets.
Anthony Reale	NOAA/NESDIS/STAR	NOAA Product Validation System (NPROVS) for NPOESS
Hank Revercomb	Univ. of Wisconsin	SDR issues, S-HIS preparation/campaigns, Atmospheres PEATE
Lars Riishojgaard	JCSDA	Preparation for ingesting SDR and EDR products
Ben Ruston	NRL	Global characterization of ATMS and CrIS biases
Larrabee Strow	UMBC	SDR issues, radiative transfer issues, pre-flight instrument cal/val
Joel Susskind	GSFC	CrIS/ATMS proxy datasets derived from Aqua

		AIRS/AMSU
Fuzhong Weng	NOAA/NESDIS/STAR	Microwave Integrated Retrieval System (MIRS)
Sid Boukabara	NOAA/NESDIS/STAR	Microwave Integrated Retrieval System (MIRS)

1.1 Data Record Requirements to Ensure Mission Success

The minimum requirement for NPP mission success is to achieve EDR performance within the design specifications of the CrIMSS instruments. Simulation experiments using CrIMSS instrument specifications have shown that the CrIS/ATMS system should be similar to that of other, recently launched, hyperspectral systems, such as the Atmospheric Infrared Sounder (AIRS)/Advanced Microwave Sounding Unit (AMSU) suite on NASA's Earth Observing System (EOS)-era Aqua platform, and the Infrared Atmospheric Sounding Interferometer (IASI)/AMSU/Microwave Humidity Sounder (MHS) suite on the European Meterological Satellite (EUMETSAT).

The satellite data product user community has evolved over the last decade. The traditional user base has changed their needs and new users have emerged. While some weather centers use mature, processed products such as Atmospheric Vertical Moisture Profiles (AVMP) and Atmospheric Vertical Temperature Profiles (AVTP), the current approach of most such centers is to use the apodized radiances directly. Some off-line experiments have demonstrated that cloud-cleared radiances, an EDR intermediate product (IP), have a positive impact on forecast models. It is important to identify exactly the user context for the EDRs—both current and expected in the 2010 time-frame—before we begin assessing their performance. In some sense, the EDRs are simply a transformation from radiance space to geophysical space and EDR validation is an implicit validation of the instrument radiances. Therefore, a requirement for this cal/val plan is to demonstrate that CrIMSS radiances are performing according to EDR specification and to demonstrate that performance for global scenes. The validation of CrIS/ATMS EDR products, as described in this document, can be used to assist in validating the Sensor Data Records (SDRs), themselves necessary for generating clean EDRs.

Beyond simply validating retrieval algorithms, validating EDRs is also an implicit validation of cloud-cleared radiances, forward models, cloud detection capabilities, and the ability to select radiance subsets for numerical weather prediction (NWP) applications. For a selected subset of scenes that are cloud-free (where cloud-cleared radiances are equal to instrument radiances) EDR validation is an implicit characterization and validation of the forward model(s) and SDRs. Therefore, validating EDRs is important for *all* users.

The importance of EDR validation for the global community notwithstanding, in this plan we will address the needs of the traditional user community by characterizing the NPP mission key performance parameters (KPPs) using the Integrated Operational Requirements Document (IORD)-II and Northrop Grumman Aerospace Systems

(NGAS) specifications, hereinafter referred to as the *Sys Spec* [REF] (see Table 1 in Section 1.2, below).

In this plan validation activities are designed to characterize the performance of the EDRs in various ensembles of cases. It is most likely that a “roll-up” of regional assessments will be made to determine whether the EDRs have met their global performance specifications. Specifically, we expect to stratify the specifications using several bins: day/night, latitude bands (*i.e.,* polar, mid-lattude, tropical), land/ocean, and (possibly) altitude and surface characteristics. In this context, it is important to include assessments of current capabilities using heritage sensors and associated algorithms. Most significant will be to use the the High-resolution Infrared Radiation Sounder (HIRS)/AMSU operational products as part of our validation efforts to demonstrate the value of the hyperspectral measurements to the user community. NOAA/NESDIS has worked with the operational HIRS/MSU products for decades and has performed detailed assessments of hyperspectral products from these and other sensors (*e.g.,* AIRS/AMSU, IASI/AMSU/MHS) as compared to current Advanced Tiros Operational Vertical Sounder (ATOVS) products. Heritage hyperspectral sounding systems, such as AIRS/AMSU and IASI/AMSU/MHS, have a number of scientific users that are currently exploiting those data products for weather, air-quality, and climate applications. We expect that at CrIS/ATMS launch user requirements will be more sophisticated than today; therefore, part of the validation activities described herein will be to characterize the CrIS/ATMS EDR products for those applications in the same manner as AIRS and IASI products are currently characterized.

In summary, we will leverage integrated product validation system activities to ensure that we properly assess the EDRs in an operational context. Therefore, this plan employs a comprehensive approach that integrates satellite products, global *in situ* measurements, aircraft observations, and other intensive field campaign data sets.

1.2 Summary of Requirements and Response Approaches

This section will focus on the EDR requirements as levied in the Sys Spec, how the requirements must be tempered with several practical constraints and findings more recent than the publication date of the Sys Spec, and how the cal/val team will address EDRs validation. This section is intended only as an upper-level discussion, however. More detailed analyses of the activities required to validate the EDR requirements and the EDRs themselves will be found in Appendix 1 and Appendix 2 of this document, respectively.

1.2.1 IORD and NGAS Requirements

The IORD requirements and an example of the NGAS requirements (as of [TBD]; these may have changed) are illustrated in Table 2, below. In many cases, they differ considerably. This will be discussed in more detail below (see Appendix 1, Section

A1.2.4.1). Further, the IORD and NGAS specifications are sufficiently vague such that the computation of a statistic is left up to the evaluators to define in detail. We assume that the specification should be met globally, that is, all cases except excluded cases. For CrIMSS, these are only precipitating cases with more than 2 mm of liquid water. The validation dataset would be split into scenes that are "partly cloudy" (less than or equal to 50% cloudiness) and "cloudy" (greater than 50% cloudiness), and the statistic would be computed for all scenes. However, *in situ* datasets are rarely available over a large portion of the geophysical domain, such as polar regions, high altitude regions, open-ocean regions, desert regions, etc. Thus, the *in situ* measurements are "data-limited" and do not necessarily represent a global ensemble. In practice, detailed analysis can be performed with data from dedicated radiosondes launched at satellite overpass times at a few representative locations (*e.g.*, polar, mid-latitude, tropical regions; *see* Tobin 2006). Other datasets, such as those provided by operational radiosonde observations (RAOBs), intensive field campaigns, comparison with other satellite products, and NWP analysis fields, must then be used to fill in the gaps.

The IORD specification is also vague in the sense that some "partly cloudy" scenes can still be difficult for observations in the infrared (*e.g*., uniform broken clouds or low clouds), and precipitating scenes are problematic for microwave sensors in cloudy scenes. Thus, there is a strong desire to remove these from the ensemble before the statistic is computed. For example, the AIRS Science Team has adopted an approach in which any parameter that is "rejected" in the infrared retrieval process can be removed from the statistic. Therefore, the complete atmospheric state is rarely assessed, since only 20% of the lowest part of the troposphere (*i.e.*, 700 mb to the surface) is "accepted", while 90% of the stratosphere is "accepted". This becomes a serious issue when attempting to validate the cloud-cleared radiances, since radiances are a vertical integral of the geophysical state, and a full state is required to compute radiance.

The cal/val team needs to define in detail how the EDR statistic will be computed, and prepare and demonstrate the validity of an agreed upon "recipe" using all cases for this computation prior to launch.

Table 2 - IORD and NGAS Requirements for EDRs

Parameter	IORD-II (Dec. 10, 2001)	NGAS SY15-0007 (Oct. 18, 2007)
AVMP Partly Cloudy, surface to 600 mb (KPP[1])	Greater of 20% or 0.2 g/kg	14.1% ocean, 15.8% land and ice
AVMP Partly Cloudy, 600 to 300 mb	Greater of 35% or 0.1 g/kg	15% ocean, 20% land and ice
AVMP Partly Cloudy, 300 to 100 mb	Greater of 35% or 0.1 g/kg	0.05 g/kg ocean, 0.1 g/kg land and ice
AVMP Cloudy, surface to 600 mb (KPP)	Greater of 20% of 0.2 g/kg	15.8%
AVMP Cloudy, 600 mb to 300 mb	Greater of 40% or 0.1 g/kg	20%
AVMP Cloudy, 300 mb to 100 mb	Greater of 40% or 0.1 g/kg	0.1 g/kg

AVTP Clear, surface to 300 mb (KPP)	1.6 K/1-km layer	0.9 K/1-km ocean, 1.7 K/1-km land & ice
AVTP Clear, 300 to 30 mb	1.5 K/3-km layer	1.0 K/3-km ocean, 1.5 K/3-km land & ice
AVTP Clear, 30 mb to 1 mb	1.5 K/5-km layer	1.5 K/5-km layer
AVTP Clear, 1 mb to 0.5 mb	3.5 K/5-km layer	3.5 K/5-km layer
AVTP Cloudy , surface to 700 mb (KPP)	2.5 K/1-km layer	2.0 K/1-km layer
AVTP Cloudy, 700 mb to 300 mb	1.5 K/1-km layer	1.5 K/1-km layer
AVTP Cloudy, 300 mb to 30 mb	1.5 K/3-km layer	1.5 K/3-km layer
AVTP Cloudy, 30 mb to 1 mb	1.5 K/5-km layer	1.5 K/5-km layer
AVTP Cloudy, 1 mb to 0.05 mb	3.5 K/5-km layer	3.5 K/5-km layer
CH4 (methane) column	1% precision, ±5% accuracy	n/a
CO (carbon monoxide) column	3% precision, ±5% accuracy	n/a

[1] KPP = Key Performance Parameter

1.2.2 Augmenting the Requirements

While we recognize that a global assessment should be made, exactly how the cases are to be weighted needs to be defined explicitly. Since a polar sounder samples a large number of polar cases, and polar cases are traditionally difficult to deal with, the statistic has to be weighted properly to highlight the regions of interest to the user community. Therefore, we need to explicitly define the "roll-up" process and document how this assessment is done. In this sense, a weather user and a climate user could have radically different requirements. A major effort is underway within the AIRS Science Team to define a metric for sounding products that can identify skill in meteorologically significant regimes. This activity will take place concurrently with NPP planning; if appropriate, we will incorporate the lessons learned from AIRS into our analysis.

Hyperspectral instruments are capable of more than just temperature and moisture products, and we now have a user base for cloud-cleared radiances and many trace gas products. Just as EDRs are a transformation of radiances, the trace gas products can be used to illustrate subtle biases in cloud-cleared radiances and are, in and of themselves, a necessary "byproduct" for meeting temperature and moisture product specifications.

In the five-and-a-half years between the launch of AIRS and IASI, atmospheric carbon dioxide has increased about 10 ppmv. Failure to include that in our validation activities and algorithms would induce unacceptable biases of up to 0.3 K in the mid-troposphere. Similarly, carbon dioxide varies by about ± 5 ppmv seasonally and regionally, and induces small time and/or regional biases in our temperature products. Methane, nitric acid, and nitrous oxide have similar impacts on the water vapor retrievals. While the

NGAS algorithm does not include these components, we can utilize off-line CrIS/ATMS EDR processing using the Langley Research Center (LaRC) algorithm and the NOAA Unique CrIS/ATMS Processing System (NUCAPS) to evaluate to what extent they these results are relevant to validating the NGAS system.

Another evolution that has taken place in the last decade is what we mean by “a product”. User requirements for AIRS/AMSU and IASI/AMSU/MHS products have become more sophisticated than previously in that they include averaging kernels and error covariance matrices. The averaging kernels are also important in the validation context because they allow partitioning of instrument and first-guess errors. We have added averaging kernels and error estimates to the AIRS Science Team package (Maddy and Barnet 2008a) and are supplying this information to our users. Averaging kernels are also extremely valuable in the validation context because they can help determine if an algorithm is extracting the full information content of the instrumental radiances. NUCAPS-generated EDRs can also be used to evaluate the information content of the CrIS/ATMS processing system.

Significant pre-launch effort will be allocated to quantify exactly how the statistics will be computed and what ensembles of cases will be considered, including stratification of conditions and degraded and excluded conditions. Part of the activity in the first year will be to explicitly define and demonstrate how these performance specifications can be demonstrated with heritage systems (AIRS/IASI) using available models and *in situ* datasets. Based on our experience with AIRS and IASI, it is safe to say that the specification will be first demonstrated by comparisons with model analysis fields (*e.g.,* those generated by the European Centre for Medium-range Weather Forecasting (ECMWF) and/or the National Centers for Environmental Prediction (NCEP)). The specification would be generated both as a global statistic and as statistics within regimes (*e.g.,* day/night, land/ocean, tropical/mid-latitude/polar, small- and large-scale meteorology, inversions, dust, etc.). The next level of assessment will use operational and dedicated radiosonde launches. It typically takes several months of data collection until a reasonable sample size is achieved; given the small size of the dataset, the ability to reprocess radiosonde data is critical. It is also important to characterize the radiosondes themselves, since a number of launch locations and radiosonde types must be excluded due to poor statistics from the sondes themselves.

1.2.3 Acquiring Data over Problem Areas

Special attention is placed on studying problem areas (*e.g.*, deserts, regions of heterogeneous emissivity, polar ice, high-altitude clouds, dust, and large inversions); however, the *in situ* datasets are very sparse in these regions, and we have found it necessary to supplement *in situ* datasets with intensive field campaigns and scientific campaigns-of-opportunity in these regions. Scientific campaigns-of-opportunity may or may not occur at appropriate times in our schedule. Therefore, this aspect is most important to address with focused intensive field campaigns designed for the NPP mission. Aircraft instruments, along with multiple radiosonde launches, LIDARs, upward-looking FTIR, microwave radiometers, etc., can provide the opportunity to study the characteristics of the CrIS/ATMS SDRs and EDRs in detail. If these intensive field campaigns are centered about an existing ground-site facility and/or can address interests of other agencies, we can employ some level of cost-sharing, develop greater community involvement, and stimulate new user interest.

The importance of aircraft campaigns for CrIMSS is possibly more relevant than it was for the AIRS/AMSU or IASI/AMSU/MHS missions. The ATMS employs a new sampling strategy that oversamples the spatial domain. Aircraft with NPOESS Atmospheric Sounder Testbed–Microwave (NAST-M) support can help verify that the Backus-Gilbert resampling is working properly and that the side-lobe contamination is characterized. The CrIS calibration is more complex than either AIRS or IASI, and recent issues with the internal calibration target (ICT) [REF] may require a more careful analysis of SDRs early in the mission. Aircraft-mounted hyperspectral sensors, such as the NAST–Interferometer (NAST-I) and Scanning High-resolution Interferometer Sounder (S-HIS), provide an independent, NIST-traceable, direct verification of CrIS SDRs.

We plan to exploit the World Meteorological Organization (WMO) Global Climate Observing System (GCOS) international effort to build the GCOS Reference Upper Air Network (GRUAN) of sites. A set of 12 sites has been organized that includes the German Lindenberg site, the Atmospheric Radiation Measurement (ARM) Climate Research Facility (ACRF) sites at Southern Great Plains (SGP), North Slope of Alaska (NSA), Tropical Western Pacific (TWP; Darwin, Manus, Nauru), the Howard University site in Beltsville, MD, and other proposed sites around the world (*see* GCOS-122). Chris Barnet is a member of the GRUAN working group, and has initiated discussions for coordination with the NPP cal/val activities. A GRUAN implementation meeting was held in March 2009; support of the NPP mission was discussed. At this time, the ACRF sites remain our most viable for validation. Futher, as Changyong Cao of NESDIS/STAR, is the chair of the Committee on Earth Observation Satellites (CEOS) working group on calibration and validation, we can leverage his support in these multiple satellite platform intercomparison activities.

2.0 Technical Approach Summary

Given the long-term successful heritage of legacy systems, the basis of our approach is to utilize lessons learned from validating the AIRS/AMSU and IASI/AMSU/MHS sounding systems and to focus on those activities that have had the most impact in their respective validation efforts. This activity will leverage "user community" cal/val efforts and experience as much as possible. This section is intended only as a summary. A more detailed discussion of EDR validation activities is found in Appendix 2 of this document.

2.1 Pre-launch Phase

There are five important activities in the pre-launch phase. They are to:

(1) Use simulated and proxy data to exercise the NGAS algorithm and mitigate as many issues as possible prior to launch. This requires the ability to modify an off-line version of the NGAS code to ingest SDRs that are derived from model simulations and proxy datasets such as AIRS/AMSU or IASI/AMSU/MHS.
(2) Demonstrate, using IASI and/or AIRS EDR products, that we can use the validation datasets described in this plan to "roll-up" a statistic that demonstrates acceptable (or better) performance of those instruments.
(3) Prepare the infrastructure for acquisition of validation datasets. This includes upgrades to NAST-I, NAST-M, and S-HIS aircraft sensors, coordination of other payload sensors, aircraft platforms, and ground infrastructure, including ground-site coordination for instrumentation, launching of radiosondes at satellite overpass times, and Intensive Periods of Operations (IOPs).
(4) Package simulated or proxy datasets into SDR (or RDR) format and distribute to NWP users in standard formats (*i*.e., channel subsets in Binary Universal Form of Representation (BUFR) format) so they can begin to prepare their systems for CrIS/ATMS.
(5) Perform a "dress rehearsal" of an intensive field campaign to test our ability to communicate rapidly and to iron out problems with file formats, data access, etc., to assure readiness to evaluate the performance of the CrIS/ATMS EDR algorithms.

2.2 Early-orbit Checkout Phase

The Early Orbit Checkout (EOC) phase encompasses roughly the first 90 days post-launch, as sensors are activitated. Our plan assumes that the ATMS sensor will be activated within about 3 weeks of launch, and that the CrIS sensor will be activated within about 8 weeks after launch. EDR processing and validation activities cannot fully commence until first light data are available in SDR format. It is imperative that all early SDRs (including first light) are made available to the cal/val team as quickly as possible, even if there are significant uncertainties and calibration problems.

At first light, the CrIS/ATMS SDRs will be compared to synthetic radiances (*i.e.,* computed using a forward model) derived from ECMWF and NCEP/Global Forecast System (GFS) forecast and analysis fields. Data from the ATMS sensor will also be compared with data from in-orbit AMSU/MHS sensors from NASA's Aqua, NOAA's Polar-orbiting Environmental Satellites (POES), and ESA's METOP satellites. The CrIS sensor will be compared with in-orbit hyperspectral sensors, specifically AIRS (if still operational) and IASI. Simultaneous nadir overpass (SNO) comparisons will be particularly useful in this phase to understand if the radiances are performing as expected. Changyong Cao's proposal for *Integrated Cal/Val for NPP/NPOESS* is primarily an SDR activity, but the EDR cal/val team can use these scenes for detailed analysis across multiple sensors.

The CrIS/ATMS radiances will be processed by the MIT algorithms, LaRC algorithms and the NUCAPS system (*i.e.,* the AIRS/IASI heritage algorithm; see Appendix 2). Geophysical retrieval products from these systems (*i.e.*, the LaRC-EDRs, NUCAPS-EDRs and STAR offline-EDRs) can be compared to model fields, radiosonde data, and other *in situ* measurements to help validate the SDRs from both CrIS and ATMS. If operational EDR products are available from the instrument data processing system (IDPS EDRs), then off-line retrieval products can be directly compared to the operational products.

2.3 Intensive Calibration and Validation Phase

The overall approach to Intensive Calibration and Validation (ICV) phase, assumed to be some 18 months long, is to compare EDR products with a number of available datasets. Validation of KPPs will be done using validated SDRs: A multiple intercomparison of data from the ECMWF, the NCEP/GFS, radiosondes, AIRS/AMSU products (if available), ATOVS operational HIRS products, NOAA IASI/AMSU/MHS products, off-line CrIS/ATMS retrieval products and, of course, the IDPS products. Cloud-cleared radiances will be validated both explicitly and implicitly, as described below.

Explicit validation of cloud-cleared radiances can be done in two ways. First, cloud-cleared radiances from the NUCAPS system can be compared directly with NGAS cloud-cleared radiances, an IP product. In addition, cloud-free CrIS fields-of-view (FOVs) can be compared with cloud-cleared radiances from the same field-of-regard (FOR) to verify that cloud clearing represents the cloud-free part of the scene. These analysis techniques have recently been used at NOAA/NESDIS/STAR to validate AIRS and IASI cloud-cleared radiances. Implicit validation of cloud-cleared radiances can, to a degree, result from the validation of EDR products using *in situ* or model profiles, given that the EDRs are derived from the cloud-cleared radiance IP.

The datasets that can be used are defined below in the order of increasing relevance to demonstration of meeting the IORD specifications.

(1) *Forecast and analysis datasets.* These datasets provide a global view of temperature, moisture, cloud, and surface properties and can be used to assess quickly the EDRs and identify problem areas. They include European Center for Medium Range Forecast (ECMWF) data and data from the National Center for Environmental Prediction Global Forecasting System (NCEP/GFS). NWP centers (*i.e.*, NCEP, ECMWF, UKMET) will also prepare to ingest CrIS/ATMS radiances and initially will compare observed radiances to their own analysis (*i.e.*, observation minus background). These analyses will be useful in identifying regional or spectral biases.

(2) *Operational radiosonde observations.* RAOBs and use of the NPROVS analysis tools developed at NESDIS/OSDPD by Anthony Reale can be used to quantify the skill of the NUCAPS-EDRs. This approach is preferred over use of models but it takes a significant number of months before enough RAOBs are acquired to compute meaningful statistics. Typically, 200-300 operational radiosondes are co-located within ±3 hours and 100 km radius on any given day. For a particular orbit these tend to have a particular regional sampling. For example, AIRS 01:30 orbit tends to capture west coast of USA and IASI 09:30 orbit tends to capture more of the east coast of USA and China as well as the southern hemisphere. Both tend to do well over Europe. Over a 6 month period a large global sample of 10’s of thousands of radiosondes can be acquired. Figure 1 shows operational RAOB sampling collocated with AIRS and IASI accepted infrared retrievals (≈ 50%) from Feb. 21, 2008 through Mar. 15, 2008 as an example of this kind of comparison.

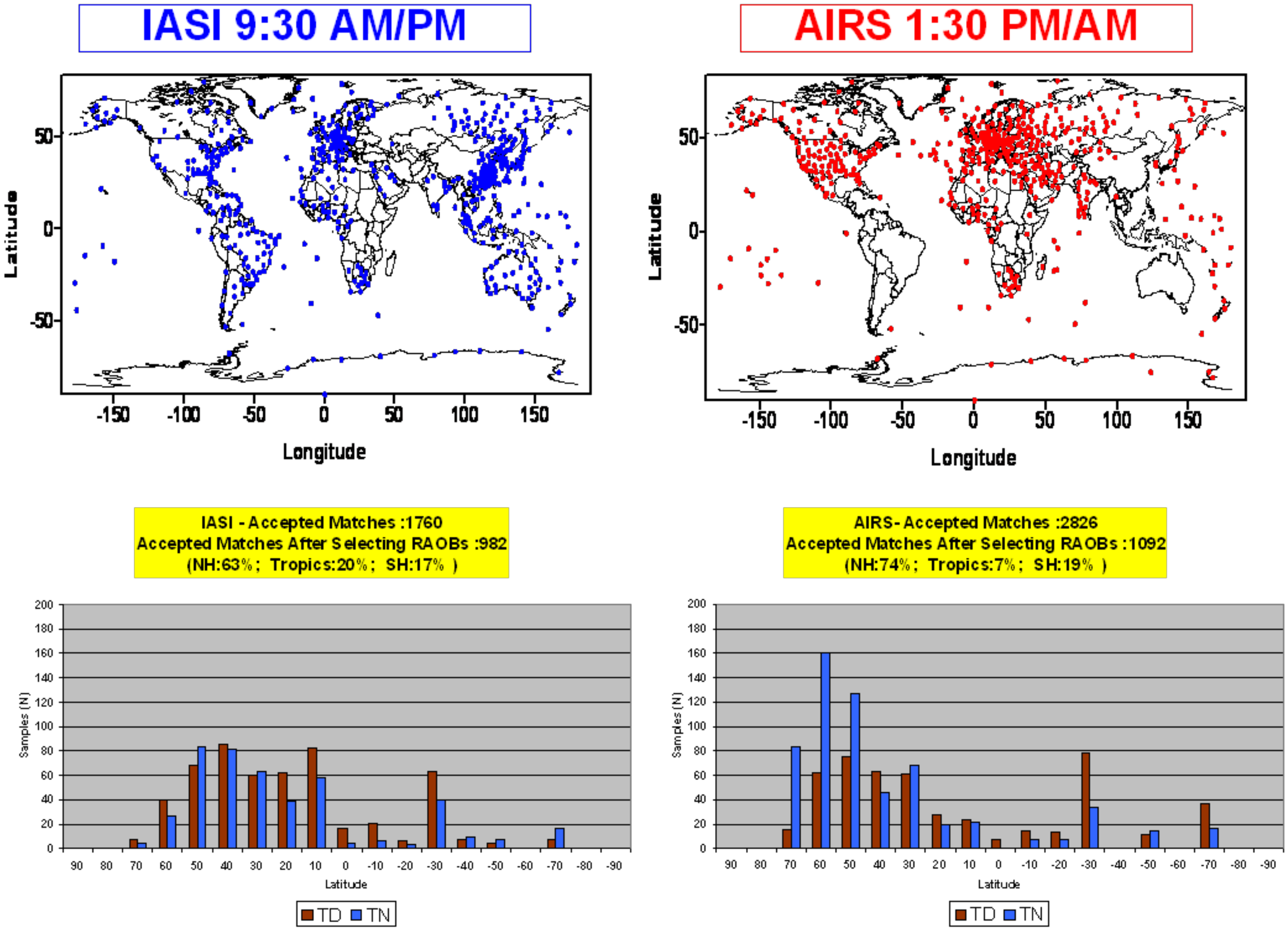


Figure 1 - Operational radiosonde sampling of AIRS and IASI matchups with IR retrievals (accepted cases only) for February 21, 2008 through March 15, 2008.

(3) *Dedicated high-quality radiosonde data.* Radiosonde launches coordinated with satellite overpass times are very important for demonstrating the performance of the CrIMSS AVTP and AVMP EDR algorithms. For example, for the AIRS/AMSU mission two radiosondes were launched per overpass during an intensive 90-day period. Multiple sondes improved the accuracy of the *in situ* data; two sondes were in the air during the 23-millisecond AIRS sampling time), and the two sonde profiles could be compared to eliminate any errant sondes. Additional instrumentation and models were used to product a “best atmospheric state estimate” (*e.g.*, see Tobin 2006a). We can exploit the following to demonstrate global skill in AVTP and AVMP profiles:

- ACRF radiosondes at Southern Great Plains (SGP), Tropical Western Pacific (Manus Island, TWP), and North Slope of Alaska (NSA). These are now part of GRUAN reference network.
- Howard University Beltsville Campus (HUBC) GCOS Reference Upper Air Network (one of the 12 reference GRUAN sites) located in the suburban Washington DC area (within a half-hour of NOAA/STAR and the IPO).

- Water vapor Validation Experiments (WAVES, Dave Whiteman, GSFC). These are intensive field campaigns at HUBC that focus on satellite validation. Additional instrumentation (*e.g.,* ground-based and aircraft-based LIDAR) are typically employed.
- Aerosol and Ocean Science Expeditions (AEROSE) in the tropical Atlantic (AEROSE, Nicholas Nalli, NESDIS/STAR) focus on tropical cyclo-genesis regions. These campaigns are over the open ocean in regions that are also typically impacted by African dust and smoke from biomass burning.

(4) *Satellite products.* Radiance and/or geophysical data from other satellites (*e.g.,* ATOVS, Aqua, METOP), in conjunction with the use of SNOs, can demonstrate radiometric performance, and intercomparison of EDR products can isolate problems in algorithms. While SNOs are only available for polar high latitudes, other regions can be studied using satellite spatial co-registration by applying a "double difference" to partially remove the effects of the diurnal cycle.

(5) *Field Campaigns Data.* Data from NAST-I, NAST-M, Scanning HIS, other aircraft sensors, radiosondes, dropsondes, etc. and data from IPO-funded aircraft validation field campaigns will be of significant aid. Previous validation experience with AIRS (during the European Aqua Thermodynamic Experiment, EQUATE) and IASI (during the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT's) Joint Airborne IASI Validation Experiment, JAIVEx) has shown these datasets to be of high value in the early phases of validation as well as throughout the mission. We anticipate that early field campaigns will focus on simple scenes (*e.g.,* uniform geophysical fields with high probability of cloud-free scenes similar to JAIVEx), and then move toward more challenging scenes in later field campaigns.

(6) Intensive scientific field campaigns of opportunity and ancillary measurements. Examples of such field campaigns include:

- NOAA Earth System Resource Laboratory/Global Monitoring Division (ESRL/GMD) surface and aircraft observations of trace gases and the *CarbonTracker* assimilation model of surface observations.
- Stratospheric Troposphere Analysis of Regional Transport (START, Laura Pan, NCAR) for ozone, which also includes multiple instruments for temperature, moisture and other traces gases. The focus is on stratospheric/tropospheric exchange regions in the mid-latitudes.
- Hiaper Pole-to-Pole Observations (HIPPO; Steve Wofsy, Harvard) for trace gases—specifically CO_2 and CH_4—along Northern to Southern Hemisphere transects.
- NASA and NAS Airborne Field Experiments
- International campaigns conducted by Metop colleagues (*e.g.,* Jonathan Taylor, UK Met. Office)
- Other scientific experiments of opportunity.

Items 1 through 4 are low-cost, high-impact observations and provide the backbone of EDR evaluation. In particular, item 3 is a very high priority and requires additional funding. At a minimum we need best-state estimates from the three ARM sites for intensive periods. The launch of multiple radiosondes at overpass times from NSA, SGP, and TWP will form the backbone for demonstrating AVTP and AVMP performance, and it would be reasonable to consider a long-term commitment to having continuous measurements from these three sites. Alternatively, we could have monthly intensive observation periods every 3-4 months.

We expect that IPO-funded intensive field campaigns (item 5) will be required for SDR validation and that these will be used for IDPS EDR validation activities. There is a benefit to have field campaigns coordinated with other instrument cal/val efforts (*e.g.*, VIIRS, OMPS) since those instruments can help characterize the scene. For example, there is a synergy between CrIS and VIIRS in that CrIMSS can provide atmospheric profiles to assist evaluating the VIIRS atmospheric correction, whereas VIIRS can provide valuable sub-pixel cloud and surface information to assist CrIS in evaluating cloud clearing and surface retrieval issues. The initial priority is to verify CrIMSS SDRs and EDRs in open ocean, cloud-free scenes at night, and then move on to progressively more difficult scenes. The earliest aircraft campaign(s) will be designed to capture a mixture of easy and moderately difficult scenes so that a variety of components for which the algorithm can be tested are completed within one campaign.

The field campaigns also need to be planned in detail to study algorithmically challenging regimes. Some areas of interest to CrIS/ATMS cal/val might be polar regimes with heterogeneous surface emissivity, low-level inversions and associated cloudiness, and nearly isothermal structure; tropical regimes with high cirrus and convective cloudiness, dust (African and Asian), and desert surface emissivity; and mid-latitude stratospheric/troposphere exchange. Given the difficulties of coordinating *in situ* aircraft campaigns, it is desirable to have both S-HIS and NAST-I funded. This ensures that if one instrument has problems on a given day then the other instrument can still obtain measurements. When both instruments are functional, the intercomparison of the two instruments is valuable to ensure radiometric accuracy and precision, and to improve characterization of the sub-pixel footprint with different sampling strategies (*e.g.*, aircraft at different altitudes, different scan geometries, etc.).

Intensive scientific campaigns-of-opportunity are relatively of low cost, but they tend to be focused on specific scientific issues (*e.g.,* air quality, dust/aerosols, stratospheric mixing) and are not necessarily available at optimal periods within the CrIS/ATMS validation schedule. They can, however, be of high impact in that the user community is engaged and the satellite products can be employed for experiment planning purposes in addition to the comparisons against *in situ* observations. This has significant value because of the large number of interdisciplinary scientists viewing and evaluating the data complement.

Whenever possible the cal/val team should support scientific campaigns-of-opportunity with instruments capable of validating CrIS/ATMS. In this way, we can engage the scientific community while acquiring the data necessary for validation purposes.

2.4 Long-Term Monitoring Phase

During the Long-Term Monitoring (LTM) phase, considered to extend from the end of ICV to the end of operational lifetime, characterization of all EDR products and long-term demonstration of performance can be performed with respect to models, operational radiosondes, dedicated radiosondes, intensive field campaigns, and scientific campaigns-of-opportunity. As discussed earlier, there should be a commitment to continue the bes state estimates from the three ARM sites (Tobin 2006) throughout the NPP mission. A significant component of the LTM phase will be to encourage national and international partnerships in validating NPP and NPOESS products. This component requires the ability to reprocess a collection of historical datasets. This is discussed in more detail in Appendix 2, Section A2.3.

As stated earlier, we plan to make use of the GRUAN sites, of which approximately 12 should be operational world-wide before NPP launch, to collect a globally representative set of radiosonde data that can be used to characterize the EDRs as a function of season, and to characterize long-term stability of these datasets. The AIRS ACRF measurements that were taken during the first four years of the AIRS/AMSU mission have tapered off in recent times. The lack of significant statistics (approximately 15,000 radiosondes have been launched in total) makes detailed characterization of the AIRS products difficult. We have relied on the operational radiosonde database for long-term monitoring; however, this approach is plagued with difficulties because of variable sonde data quality and asynchronous launch times and locations. The NPP/NPOESS program will benefit substantially from a program of continuous, dedicated radiosonde launches over the NPP mission lifetime.

3.0 Schedules and Milestones

The schedules and milestones for CrIMSS cal/val activities as described here assume a Spring 2011 launch. Top-level tasks include the following, with identification of the cal/val activity phase provided:

3.1 Year 1 (4/2008 to 3/2009) (Pre-Launch)

- Hold meetings with the IPO and NGAS to define ensembles and methodologies to measure EDR performance. In particular, this will include scheduling technical interchange meetings (TIMs) with NGAS to coordinate validation activities.
- Prepare and deliver a detailed calibration and validation plan, including the detailed "recipe" for computation and roll-up of regional product assessments to a global statistic to demonstrate acceptable characterization of EDR products and to ensure that the products meet or exceed IORD specifications.
- Organize the cal/val team and hold planning meetings (*e.g.,* Sounder Operational Algorithm Team (SOAT) meetings). Prepare a detailed task list, with milestones for each cal/val team member.
- Develop CrIS and ATMS proxy datasets to exercise the NGAS algorithm using real data to prepare the NSGT algorithm for product generation with real instrument measurements. We recommend continuing the AIRS/AMSU-derived proxies and also developing IASI/AMSU/MHS proxy datasets with correct polarization for ATMS. These datasets can be used to correct algorithm deficiencies, for example to make recommendations to fix the non-local thermodynamic equilibrium (NLTE) channel use, emissivity hinge point issue, and empirical bias correction discussed in more detail in Section 6 of this document.
- Provide NWP centers with simulated datasets so they can begin their cal/val preparations. NCEP, ECMWF, and UKMet have all expressed interest in these early datasets.

3.2 Year 2 (4/2009 to 3/2010) (Pre-Launch)

- Hold meetings with the IPO and NGAS to define stratification of statistics and to define how these will be "rolled-up" into a statistic that can be used to demonstrate that the IORD specifications have been met or exceeded.
- Hold meetings with the IPO and NGAS to define procedures for start-up of NGAS code. Specifically, determine which datasets will be used for look-up tables and empirical bias corrections, and which datasets will be used for validation.
- Begin focused SOAT meetings on cal/val topics. Expect to hold two to three SOAT meetings (one might be via teleconferencing) to

coordinate activities and to discuss progress on specific milestones.
- Acquire GRUAN (including ARCF sites) “best-estimate states” at Aqua and/or METOP overpass times.
- Continue using CrIS and ATMS proxy datasets to exercise and improve the NGAS IDPS EDR algorithm.
- Develop and test software to compute statistics to demonstrate ability to meet IORD specifications. This software will be tested on ATOVS, AIRS/AMSU, and IASI/AMSU/MHS systems to demonstrate performance on heritage instruments/algorithms within the IPO IORD specifications.
- Participate in one or two externally funded scientific campaigns-of-opportunity (*e.g.,* AEROSE in July 2009).
- Create an initial METOP proxy dataset (based upon one or more orbits), and deliver to the cal/val team.

3.3 Year 3 (4/2010 to 3/2011) (Pre-Launch)

- Hold meetings with the IPO and NGAS in anticipation of launch. Specifically, we need to finalize the communication pathway for algorithm upgrade recommendations.
- Continue SOAT meetings to discuss pre-launch activities. We anticipate two or three face-to-face SOAT meetings and monthly teleconferences between SOAT meetings.
- Complete CrIS/ATMS SDR and EDR evaluation methodology using ATOVS, AIRS/AMSU, and IASI/AMSU/MHS product data sets.
- Generate CrIS/ATMS product validation readiness reports based on proxy datasets. A preliminary evaluation of NUCAPS EDRs using early NPP SDR data can be made at first light; a more-detailed evaluation of IDPS EDRs can be made toward the end of this period.
- Perform a “dress rehearsal” of the NPP cal/val intensive campaign using IASI/AMSU/MHS data. Generate a set of statistics (using proxy radiances derived from IASI and/or aircraft measurements in the NGAS algorithm and NUCAPS algorithm) to demonstrate that the CrIMSS algorithm can meet IORD and NGAS specifications.
- Prepare for acquisition of GRUAN “best-estimate” measurements. It is highly desired that radiosonde launches begin pre-launch to demonstrate ability to compute statistics for AIRS/AMSU and/or IASI/AMSU/MHS overpasses. Once SDRs are stable, the acquisition of CrIS/ATMS overpass times will begin.
- Create METOP (and Aqua) marine proxy datasets (*e.g.,* based upon AEROSE campaigns) and stage for the cal/val team.

3.4 Year 4 (4/2011 to 3/2012) (EOC to ICV)

- Hold meetings with the IPO and NGAS to reassess priorities and

deliverables on an annual or more-frequent basis, if necessary (*see* Section 5 of this document for Reporting information).
- Hold SOAT meetings to discuss preliminary validation results. Expect two or three SOAT meetings with monthly teleconferences in-between.
- Provide CrIS/ATMS product validation report based on early NPP data from IDPS, if available. Work with Atmospheric and Environmental Research (AER)/NGAS to resolve algorithm issues as they arise.
- Acquire and process GRUAN (including ACRF) "best-estimate" states at NPP overpass times.
- Participate in two intensive field campaigns-of-opportunity where aircraft observations, ground truth (GRUAN site preferred), AIRS/AMSU (if available), IASI/AMSU/MHS, and CrIS/ATMS products will be intercompared. These campaigns require airborne support by NAST-M, NAST-I, and S-HIS.
- Participate in one or two externally funded scientific campaigns-of-opportunity.

3.5 Year 5 (4/2012 to 3/2013) (ICV)

- Hold meetings with the IPO and NGAS to reassess priorities and deliverables on an annual or more-frequent basis, if necessary.
- Hold three to four SOAT meetings to discuss details of validation results and to coordinate writing the validation report.
- Continue assessments and troubleshooting.
- Generage CrIS/ATMS product validation report, including global assessment.
- Acquire and process GRUAN "best estimate" states at NPP overpass times.
- Participate in two intensive field campaigns-of-opportunity where aircraft observations, ground truth (GRUAN site preferred), AIRS/AMSU (if available), IASI/AMSU/MHS, and CrIS/ATMS products will be inter-compared.
- Participate in one or two externally funded scientific campaigns-of-opportunity.

4.0 Resource Requirements

Resources of several types are required to undertake the activities described in this plan. These include personnel, funding, coordination with cognizant entities, and access to data, each of which will be discussed here.

4.1 Personnel

The EDR cal/val effort is a team activity that combines expertise from members of the government, research institutions, academia, and the NPOESS contractor, NGAS. Many of the activities described in this plan are already underway through externally funded projects. There is a strong synergy between the activities underway and planned activities; however, there will be no duplication of effort. The assumption in this plan is that funding for the external activities will remain at requested levels, and that milestones within those activities are met.

The support envisioned in this plan is to cover a five-year period to allow significant contributions in the first two years after NPP launch (*i.e.,* during the Pre-launch, EOC and ICV performance assessment phases). The budget will also be stretched over the same period, leveraging current projects that are already performing overlapping tasks. The cost of this key NOAA support is kept to a minimum by leveraging existing capabilities (*e.g.,* International Global Navigation Satellite System (GNSS) Service (IGS)-funded efforts, AIRS and IASI product development, NPOESS Data Exploitation (NDE) CrIS product development, and GOES-R Algorithm Working Group (AWG) activities) .

Our integrated efforts will result in a low-risk, cost-saving plan for the IPO to demonstrate CrIS and ATMS SDR and EDR performance. We will draw on community expertise using *in situ* validation, tightly coupled with SDR validation. One of the principal risks to the EDR cal/val effort will be the ability to run, understand, and/or modify the NGAS IDPS EDR algorithm in off-line mode (offline EDRs), hosted on local machines at cal/val member institutions. For this reason, it is highly desirable to have some of the original authors of the CrIS/ATMS algorithm actively participate in the cal/val effort.

4.1.1 Team Members within NOAA

Members of the NOAA community involved with the cal/val program described in this document are identified in Table 3 (below), along with their funding sources and responsibilities in this cal/val effort.

Since the early 1990s, the NOAA/NESDIS center for Satellite Applications and Research (STAR) has made significant contributions to the calibration, validation, and distribution of near-real-time products from hyperspectral infrared (IR) sounder data.

Chris Barnet and Mitch Goldberg are active members of the science team for NASA's AIRS, EUMESAT's IASI, and the WMO/CEOS. Because of our unique collaboration with the Joint Center for Satellite Data Assimilation (JCSDA), the NOAA Climate Program Office, and NESDIS operational centers, we will be able to contribute significantly to the success of CrIS/ATMS for NWP and climate applications. NOAA has provided initial performance validation at "first light" for both AIRS/AMSU and IASI/AMSU/MHS SDRs and EDRs, and has continued to provide critical algorithm components (*e.g.*, radiance tuning, regression coefficients, algorithm upgrades) and validation of all products throughout these missions. These products include cloud-cleared radiances, temperature, moisture, ozone, carbon monoxide, methane, carbon dioxide, nitric acid, nitrous oxide, volcanic sulfur dioxide, cloud products, and surface products. The investigators have migrated the AIRS/AMSU near-real-time algorithms, and are currently producing and validating EDRs from the IASI/AMSU/MHS that are now operational NOAA products, available from both NESDIS/Environmental Satellite Processing Center (ESPC) (near real-time products) and NOAA/National Climate Data Center (NCDC) (archived products). We are also implementing the AIRS/IASI algorithm and validation approach with the CrIS/ATMS on NPP to produce "NOAA-unique" cloud-cleared radiance and trace gas products for the NWP and science communities. As the algorithm evolves or as new algorithms are employed, we have the capability to reprocess all of the validation and globally archived datasets. The NUCAPS system will be an emulation of the AIRS/IASI system that is in-place at the time of the CrIS/ATMS launch; it will also incorporate all of the lessons learned from both the AIRS/AMSU and IASI/AMSU/MHS missions.

Table 3 -

Table 3 - NOAA Cal/Val Team Members, Funding Sources, and Cal/Val Responsibilities

Team Member	Funding Source	Responsibility
Chris Barnet	EDR Cal-Val	Cal/Val coordination; scientific campaigns-of-opportunity.
Mitch Goldberg	EDR Cal-Val	Operational use of advanced sounders and use of proxy datasets to demonstrate performance using NGAS operational code.
Mitch Goldberg	NESDIS/PSDI	NOAA-unique CrIS/ATMS Product System (NUCAPS).
Mitch Goldberg	NESDIS/PSDI	IASI Product System
Anthony Reale	IPO	NOAA Product Validation System (NPROVS) for NPOESS.
Changyong Cao	IPO	Development of an integrated instrument cal/val system for NPP/NPOESS.
Lars Riishojgaard JCSDA/NCEP	EDR Cal-Val	Preparation for ingesting SDR and NUCAPS EDR products.; global characterization of ATMS and CrIS biases with respect to GFS analysis and forecast fields.

Fuzhong Weng	NESDIS/PSDI	Microwave Integrated Retrieval System (MIRS) activities related to ATMS.
Sid Boukabara	NESDIS/PSDI	MIRS activities related to ATMS.

The motivation for using NUCAPs, including the underlying spectroscopy, for the AIRS, IASI and CrIMSS systems is that we can study and mitigate the impact of instrument differences on our products, thus making a significant advancement towards climate-quality EDRs. The advantage of exploiting this system in the cal/val context is that the system has a community of users and has been already validated for temperature and moisture (Hagan *et al.* 2004, Tobin *et al.* 2006, Divakarla *et al.* 2006, Susskind *et al*. 2006), ozone (Divakarla *et al.* 2008, Pan *et al*. 2007, Monahan *et al*. 2007), carbon monoxide (McMillan *et al.* 2008, McMillan *et al.* 2005, Warner *et al.* 2007, Warner *et al.*2008) , carbon dioxide (Maddy *et al*. 2008b) and methane (Xiong *et al.* 2008). This reduces the risk in demonstrating that CrIS/ATMS is capable of performing within specification, given that the CrIS/ATMS radiances will be the only components that are new. Similarly, MIT, LaRC and UW-Madison have developed EDR algorithms that exploit common spectroscopy; these can be used to compare aircraft sensor sub-pixel retrievals with off-line retrievals from CrIS/ATMS. These have been used successfully in validating AIRS data (*e.g*., Taylor *et al*. 2008, Zhou *et al.* 2007) and IASI (*e.g.*, Larar *et al*. 2008; Zhou *et al*. 2008).

Chris Barnet, Mitch Goldberg, Fuzhong Weng, Changyong Cao, Sid Boukabara and Anthony Reale are cal/val team members who are also NOAA/NESDIS/STAR scientists. Chris Barnet will provide coordination and reporting for all cal/val member activities, use of other satellite datasets (AIRS and IASI), and use of the NUCAPS system to provide low-cost, low-risk validation using operational radiosondes in coordination with Tony Reale's NPROVS effort, best-estimate states from dedicated radiosondes, and global gridded subset datasets. Chris Barnet will also coordinate the participation in scientific campaigns-of-opportunity—such as START and AEROSE—discussed elsewhere in this plan.

In situ data is extremely valuable and must be archived and used to evaluate upgrades to the algorithm over time. A unique component of the NOAA activity is to provide the ability to reprocess these datasets. Mitch Goldberg's IPO-provided funding will be used to integrate the NGAS operational system into all of these datasets to allow reprocessing and evaluation of performance and upgrades. Mitch Goldberg's development activities are externally funded; however, it is assumed that NESDIS/STAR will continue to fully fund these activities. Loss of development efforts (*e.g.*, Mitch Goldberg's IPO and Product System Development and Implementation (PSDI) funding), instrument cal/val and SNO analysis (Changyong Cao's IPO funding), and EDR product evaluation and validation (Anthony Reale's IPO funding) would be a risk and would require reworking this plan.

MIRS product development (Fuzhong Weng's PSDI funding) gives us an independent assessment of the ATMS product with an operational system, and will be useful—at no cost to the IPO. The team lead will develop and coordinate participation in the intensive validation and scientific campaigns. As mentioned in Section 3, in Year 2 we plan to participate in these intensive validation campaigns prior to NPP launch to construct proxy datasets as a risk-reduction effort—in effect, a “dress-rehearsal” for ICV activities—while simultaneously encouraging community participation in these processes. The operational RAOB database and match-up system is already funded by other projects to the extent of about 1.5 full time equivalents of cost-sharing, so we have not included levels of effort for this activity. We would expect one to two scientific experiments-of-opportunity per year during the ICV phase, and are prepared to support those analyses during both the Pre-Launch and ICV phases for NPP risk reduction and system development and NPP characterization and analysis, respectively.

4.1.2 Team Members External to NOAA

A number of individuals outside of the NOAA community bring unique experience to the cal/val program; they are identified in Table 4, along with their funding sources and responsibilities in this cal/val effort. The EDR validation team includes a number of SDR validation team members (*e.g.*, Gail Bingham, Dave Tobin, Larrabee Strow, Allen Larar for CrIS, and Bill Blackwell for ATMS).

Specifically, Larrabee Strow is critical for EDR validation in that he can provide analysis of the radiometric statistics (*i.e.*, biases, covariances, and spectroscopic analyses) for the CrIS. He is also our interface to the pre-launch calibration analyses. Dave Tobin is also critical in post-launch characterization of the CrIS instrument and in prepararing “best-estimate states” from the ACRF sites. Bill Blackwell is an expert on microwave issues in general; he also has knowledge of ATMS pre-flight testing and other instrument-specific issues.

Table 4 - Cal/Val Team Members External to NOAA

Team Member	Organization	Funding Source	Responsibilities
Denise Hagan	NGAS	NGAS	Coordination with NGAS cal/val activities
Degui Gu	NGAS	NGAS	NGAS IDPS-EDR cal/val activities
Gail Bingham	USU/SDL	SDR cal/val	SDR cal/val Lead
Allen Larar	NASA/LaRC	SDR & EDR cal/val	NAST-I preparations, aircraft intensive validation campaigns.
Xu Liu	NASA/LaRC	EDR cal/val	NGAS code assessments, LaRC-EDR algorithm issues
Larrabee Strow	UMBC	SDR cal/val	SDR issues, radiative transfer issues, pre-flight instrument cal/val issues.
Hank Revercomb	UW-Madison	SDR & EDR cal/val	SDR issues, S-HIS preparation and aircraft intensive validation campaigns.
Bill Blackwell	MIT	SDR & EDR cal/val	ATMS SDR/EDR issues, NAST-M issues, ATMS proxy datasets.
Joel Susskind	GSFC	EDR cal/val	CrIS/ATMS proxy datasets derived from Aqua AIRS/AMSU

Aircraft intensive campaigns designed for CrIS/ATMS SDR validation are also desirable for validation of EDR retrieval products derived from those campaigns. These are valuable only if we can coordinate their availability in a reasonable time frame over a large portion of the intensive campaign. Members of those team(s) should be part of this validation activity (*e.g.*, Allen Larar, Hank Revercomb, Bill Smith, etc.) so that we can leverage and participate in SDR validation. We also have members with expertise in the theoretical understanding and implementation of the AER algorithm within NGAS. Xu Liu (LaRC), one of the original authors of the CrIMSS code, is a critical member with the ability to modify quickly the NGAS code for proxy datasets and to make changes as required to the post-launch algorithm to enable independent assessments of performance and to suggest improvements.In this regard, part of the first-year activity is to define how algorithm recommendations can be tested and passed back to NGAS.

NGAS cal/val activities will be conducted in parallel with the government cal/val activities discussed in this plan. Denise Hagan is the lead for the NGAS SDR efforts. Degui Gi is the lead for the NGAS EDR efforts, which will be primarily focused on the cal/val of the CrIMSS EDRs against requirements levied in the Sys Spec for the NPP and NPOESS programs. Some details of this activity are discussed in Appendix 1, while details of the NGAS cal/val plans specifically are available in a separate document (NGAS D44198_A). Note that the overall EDR cal/val tasks envisioned by the broader team cover a wider range of activities than those needed solely to assess EDR performance against the NPOESS Sys Spec. As a result, NGAS plans to leverage activities benefiting the larger community, identified in Appendix 1.

The NWP community, the NPP PEATE community, and the Department of Defense (DoD) community are also active members of the NPP cal/val team. Externally funded members of the community who have expressed interest in working with the SDR and EDR cal/val teams are listed in Table 5.

Table 5 - NWP, DoD, and Sounder PEATE Cal/Val Team Members

Participants	Organization	Planned Activities
Stephen English	UKMET	Global characterization of ATMS and CrIS biases with respect to UKMET analysis.
William Bell	ECMWF	Global characterization of ATMS and CrIS biases with respect to UKMET analysis.
Steve Friedman	NASA/JPL	Global characterization of ATMS and CrIS biases with respect to AIRS and IASI.
Ben Ruston	NRL	Global characterization of ATMS and CrIS biases.
Steven M. Beck	Aerospace	*In situ* LIDAR and radiosonde measurements and evaluation.

Since the first draft of this plan (July 15, 2008) other dedicated validation assets have also become available. Steven M. Beck has offered to provide radiosonde and LIDAR validation support for no cost. If we use sites within reasonable distance of their facility (*e.g*., Edwards Air Force Base, California and in Hawaii) they can provide water vapor and temperature validation via LIDAR, balloonsondes, and an upward-looking radiometer. Their LIDAR system can also provide aerosol optical thickness data. Some of these pre- and post-launch measurements can potentially piggyback on Defense Meteorological Satellite Program (DMSP) F18 Special Sensor Microwave Imager/Sounder (SSMIS) measurements. See Wessel *(*2000) for a description of their activities with DMSP.

4.2 Funding

The cal/val activities described here are leveraging a significant amount of funding for external activities. As discussed earlier in Section 4.1.1, we expect NOAA/PSDI activities to remain funded for a number of the team members. Use of NWP analysis fields, operational radiosondes, the WMO GRUAN sites, and scientific experiments-of-opportunity represent low-cost, high-visibility efforts within the community. These opportunities notwithstanding, cal/val funding is necessary for overall coordination and for specific NPP calibration campaigns. We expect that the PIs listed in Table 6 will be funded by the IPO CrIS EDR cal/val effort at the approximate full-time equivalent (FTE) for their staff.

Table 6 - PI-led Efforts (Full-time Equivalent for Staff) for EDR Evaluation
(Superscripts on PI names refer to notes below the table.)

PI	FY08	FY09	FY10	FY11	FY12

Barnet[3]	1.1	2.0	2.0	2.0	2.0
Blackwell[1]	1.0 (5.0)	1.0 (5.0)	1.0 (5.0)	1.0 (5.0)	1.0 (5.0)
Goldberg[3]	2.3	2.3	2.3	2.3	2.3
Larar[1,2,3]	2.5 (5.0)	2.5 (5.0)	2.5 (5.0)	2.5 (5.0)	2.5 (5.0)
Liu[3]	2.0	2.0	2.0	2.0	2.0
Revercomb[1,2,3] & Tobin[1]	1.1 (3.3)	1.1 (3.3)	1.1 (3.3)	1.1 (3.3)	1.1 (3.3)
Riishojgaard/JCSDA[1,3]	0	0.0 (3.2)	0.0 (3.2)	0.0 (3.2)	0.0 (3.2)
Strow[1]	1.5 (3.0)	1.5 (3.0)	1.5 (3.0)	1.5 (3.0)	1.5 (3.0)
Susskind[3]	0	0.5	0.5	0.5	0.5
Cao[1,3]	0.0 (2.0)	0.0 (2.0)	0.0 (2.0)	0.0 (2.0)	0.0 (2.0)
Reale[3]	2.0	2.0	2.0	2.5	2.5

[1] For PIs that also have SDR efforts (Larar, Revercomb, Blackwell, Strow, Cao) the total FTE is shown in parenthesis.

[2] Aircraft campaign activities (Larar, Revercomb) have shared funding with NASA

[3] Government employee (Barnet, Goldberg, Larar, Liu, Susskind, Reale, Cao) time is not included in these figures and represents a cost-sharing with the IPO.

Funding allocation will be based on needs identified by the Pis, coordinated by the IPO and cal/val leads. Because funding realities have driven the need to dovetail with other projects, funding for the NPP cal/val effort has been kept at less than that of previous programs (*e.g.*, AIRS), but this only assumes externally funded community activities. This situation is partially mitigated by use of community products (*e.g.*, operational radiosondes, GRUAN, scientific campaigns of opportunity); however, a large amount of funding is needed to enable the necessary dedicated intensive field campaigns using NAST-I, NAST-M, and S-HIS. Some of these campaigns are necessary to ensure we have accurate *in situ* data products in the appropriate time frame to meet a tight validation schedule.

One funding issue that needs to be resolved is acquisition of the funding necessary for intensive radiosonde launches from ACRFs in the early post-launch phase, as well as long-term continuous launches. We assume that beginning in 2011 part of Hank Revercomb's funding will be used to fund Dave Tobin's efforts to provide the analysis of the ACRF best-estimate measurements and to coordinate the subcontracts to ACRF sites. A rough order-of-magnitude estimate is that we will need approximately $300K/yr to support the acquisition and analysis for three months of launches from the three ACRF facilities. This cost is addressed in Table 5.

4.3 Coordination

4.3.1 Coordination with Other Disciplines/Sensors

The CrIS/ATMS EDR plan needs to be closely coordinated with the plans of other disciplines and cal/val teams. This is of particular importance in the implementation of intensive campaigns so that the benefits and cost-sharing of a full-platform validation can be realized. The CrIMSS cal/val team needs products from the clouds discipline area—specifically, the cloud flags from VIIRS. Access to VIIRS SDR and EDR products (*i.e.*, radiances, clouds, sea surface temperature (SST)), CERES outoing longwave radiation (OLR), and OMPS ozone is also desirable for focus days and intensive field campaigns.

4.3.2 Coordination with Other Programs

The CrIS/ATMS cal/val effort needs to be coordinated with all user communities. This includes the NWP community (*e.g.,* NCEP/JCSDA, ECMWF, UKMet, and NRL), DoD, and the NASA NPP atmosphere and sounder PEATEs). The EDR cal/val lead will ensure that this is done. In this regard, JCSDA has implemented a Hyperspectral Infrared Working Group, co-chaired by Chris Barnet (STAR) and John Derber (NCEP), and a Microwave Working Group, co-chaired by Sid Boukabara (STAR) and Nancy Baker (NRL), with users from JCSDA, NCEP, GMAO and NRL, to address coordination and post-launch evaluation of CrIS and ATMS, respectively.

4.3.3 Coordination with Other NPOESS Cal/Val Team Members

The individual CrIMSS EDR and SDR cal/val team groups need to closely coordinate their activities. It is recommended that, if possible, intensive aircraft campaigns be coordinated such that all cal/val teams focus on common datasets. It is likely that CrIS radiance observations and data from aircraft *in situ* observations can be used to assist in calibrating VIIRS, in a manner similar to that described in Tobin *et al.* (2006b). This type of coordination would be mutually beneficial since VIIRS high-spatial-resolution data are needed to characterie the CrIS sub-pixel space.

It is of vital importance that all datasets are shared with NGAS in a timely manner, with an appropriate level of documentation. Denise Hagan is the lead and coordinator for NGAS cal/val activities, which are dependent on the government cal/val activities discussed in this plan.

4.4 Validation Data

4.4.1 Existing Data Sets

Proxy datasets derived from AIRS/AMSU and/or IASI/AMSU/MHS are needed for NWP preparation and risk mitigation for the NGAS IDPS operational algorithm *(e.g.,* ACRF, GRUAN, operational RAOBs, JAIVEx, AEROSE, EQUATE). We anticipate generating a set of CrIS/ATMS focus orbits from proxy datasets in FY09. Previous proxy datasets were derived from AIRS and AMSU on the Aqua platform; however, these have a disadvantage in that AIRS has spectral gaps. Further, the 183 GHz channels from the Humidity Sounder for Brazil HSB, also flown on Aqua, were available only for a short time. IASI, on the other hand, has higher spectral resolution, with continuous coverage over the CrIS sampling domain, and we already have more than a year of data on the ground. Although IASI only has four FOVs per AMSU FOR, METOP has an advantage in that both AMSU and MHS are functional.

4.4.2 Additional Data Sets

Data must be acquired and processed for all NPP measurements collocated with *in situ* measurements. This processing should include all associated IPs such that the complete atmospheric state is available for radiance computation—including 100-level $T(p)$ and $q(p)$ at line-of-sight, ozone, surface emissivity, etc.—and cloud-cleared radiances so that observed minus calculated radiances can be computed. We anticipate that these will be maintained by the individual PIs. We anticipate that the following datasets will be needed:

- SNO matchups of CrIS/ATMS with AIRS /AMSU (if available) and IASI/AMSU/MHS.
- Spatially co-located matchups of CrIS with AIRS and IASI using the “double difference” method.
- Individual CrIS/ATMS FOR (3 × 3 CrIS) matchups with operational RAOBs. These datasets are sized at roughly 100 GB/yr.
- Dedicated launches of radiosondes from ACRF and GRUAN . These datasets are sized at roughly 10s of GB/yr).
- Global datasets for focus days. These are sized at approximately 200 GB/focus day for radiances, products, analysis, etc., and about four focus days per year.

4.4.4 Required Field Campaigns

Intensive field campaigns enable detailed characterization of instrument performance, and they can be designed to test the instrument/algorithms in meteorologically important regimes. Coordinated intensive field campaigns with airborne sensor support provide the most accurate, NIST-traceable “truth” for validating SDRs and EDRs. These measurements can be used to characterize the sub-pixel characteristics of the ATMS and CrIS FOVs. They also enable assessment of system performance for challenging cases of meteorological importance. Ensuring that aircraft instruments are functional and ready post-launch, we desire a “dress rehearsal” intensive field campaign prior to launch that focuses on an existing instrument (*e.g.*, IASI). We can also use this dress rehearsal to develop communication pathways between the field campaign participants, cal/val members, and NGAS, and to demonstrate EDR performance using proxy data derived from field campaign instruments.

The first ICV-phase CrIS/ATMS intensive validation campaign would occur approximately six months after launch, when SDRs are reasonably mature. The campaitn would use aircraft support ideally centered near a GRUAN site (*e.g.*, ACRF sites). We then expect to have one or two intensive campaigns a year, focusing on regions of interest to the EDR validation (*e.g.*, urban, mountainous, polar regimes).

Aircraft retrievals will be unique in their ability to characterize the sub-pixel effects and test concepts relating to the sub-pixel instrument line-shape (ILS) issue. As mentioned earlier, aircraft sensors (*i.e.*, NAST-I, S-HIS) provide NIST-traceable measurements. Redundancy of sensors in the field is desired for near-real-time confirmation of sensor performance. Also, specific instrument issues can be addressed via these aircraft campaigns. For example, the ATMS algorithms employs a Backus-Gilbert resampling of the raw ATMS FOVs to AMSU-like FOVs; this can be tested/verified using NAST-M data. Similarly, the CrIS sub-pixel instrument ILS correction for nonuniform scenes can be tested with NAST-I and S-HIS data.

4.5 Sensor Data

A detailed instrument model is necessary for the CrIS/ATMS cal/val team. This includes specification of the ILS and instrument noise model, including both random and systematic noise effects for both ATMS and CrIS. Spatial sampling functions (footprint) are also necessary, along with other measured characteristics needed for accurate simulations (*e.g.*, polarization sensitivity).

4.6 Computing Resources/Technology

The cal/val team members will provide specific validation datasets to the Government Resource for Algorithm Verification, Independent Testing, and Evaluation (GRAVITE) [REF] as a common dataset for use by all teams. Otherwise, the current plan is to leverage in-house resources and diagnostic codes. These will be covered by individual proposals to the IPO.

5.0 Reporting

The results of cal/val activities will be published in the peer-reviewed literature. We will encourage a number of early papers on the capabilities of CrIS/ATMS EDRs by engaging community involvement through intensive and scientific campaigns.

5.1 IPO-defined reports

The cal/val team lead will provide templates for monthly reporting to the cal/val team members. We will utilize the format we developed successfully for the NOAA GOES-R Algorithm Working Group (AWG). The cal/val lead will be responsible for generating monthly synthesis reports to the IPO, derived from reports solicited from the team members**.**

5.2 Reporting responsibility

5.2.1 Spending Plans/Expenditures

Each PI will prepare Earned Value Management (EVM) style reporting (similar to GOES-R AWG). These will be merged into a single monthly report for EDR cal/val activity costs and expenditures. The PIs will report their monthly billing and their estimate of the percentage of tasks completed based on performance metrics and well-defined milestones.

5.2.2 Instrument Performance Reports

Each PI will prepare short reports that will be merged into a synthesis report that is reviewed by the entire cal/val team before submission to the IPO.

5.2.3 Issue Peports for Each Sensor SDR/EDR

Specific issues will be addressed by an internal report from the cal/val lead to IPO cal/val management.

5.2.4 Milestone Documents

Similar to reports described in Section 5.2.1, the PI will provide monthly status of milestone progress that will be summarized into a monthly report to the IPO.

5.2.5 Regular Progress Reports

Per the GOES-R AWG model, the EDR cal/val team will prepare a report to the IPO on activities of all its members, based on the monthly template.

5.2.6 Outstanding Issue Reports

These reports are similar to those described in Section 5.2.3. Outstanding issues will be reported in a separate document submitted to to the IPO; if appropriate, these can be tracked within our monthly reports.

5.2.7 Risk Identification and Mitigation Strategy Reports

As risks or issues are identified and reported (see Section 5.2.3 or 5.2.6), a recommended strategy for mitigation will be presented. In some cases, this might evolve over the course of several monthly reports as the problem is characterized and understood.

6.0 Areas of Concern

There are few areas of concern at this time. These are described below, with suggestions for mitigating the problem areas.

6.1 Long-term launches of radiosondes and analysis of best-estimate states require funding

The ATMS and CrIS IDPS-EDR products need intensive *in situ* measurements post-launch with enough statistics to demonstrate that the IORD specification has been met. While other low-cost, low-risk datasets will be employed (*e.g.*, operational radiosondes, other spaceborne assets, forecast models, etc.), the global assessment of CrIMSS products will require a large number of overpass-coordinated radiosondes to be launched and analyzed.

At present this is not identified in any NGAS, IPO, or NASA budgets and is a significant risk to the program. We estimate that a reasonable number of radiosondes can be launched and a “best atmospheric estimate” can be prepared for approximately $300K/yr.

6.2 CrIS Internal calibration target) and detector nonlinearity issues require radiance validation

The CrIS internal calibration target (ICT) has a known problem that results in an emissivity that is lower than specification shortward of 1700 cm^{-1}. If this problem is not fixed prior to launch it imposes a high risk in terms of using the SW side of the water band and any shortwave infrared (SWIR) channels. This issue would be mitigated by modeling of the instrument thermal characteristics and the SDR algorithm would be modified accordingly.

Characterization of the instrument model and SDR algorithm would require additional resources and planning for radiance validation. Similarly, potential residual errors in nonlinearity that affect CrIS calibration as a function of scene would need more scrutiny during cal/val campaigns.

6.3 CrIS SDR sub-pixel instrument line shapes require correction

The CrIS FOVs sample a spatial integral of the Earth scene; as a result, they have an ILS that is a function of the scene. This effect should be small and can be mitigated by knowing the distribution of brightness within the scene using VIIRS pixels. The impact of this effect can also be measured and quantified by co-located satellite (*e.g.*, SNOs) and aircraft scenes (*e.g.*, using 2-km NAST-I FOVs) and comparing NAST-I retrievals with the spatially integrated retrievals from the CrIS/ATMS system. This will be a part of the SDR/EDR validation.

6.4 CrIS EDR algorithm issues affecting AVTP and AVMP products require incorporating lessons-learned

It has been determined from algorithm tests using AIRS and IASI proxy data obtained by Xu Liu and Joel Susskind that there are issues already identified in the IDPS EDR algorithm. Use of channels sensitive to non-LTE effects in the shortwave can seriously degrade retrievals of AVTP in the upper atmosphere. Also, surface emissivity is a necessary product to retrieve lower tropospheric temperature and moisture. The current algorithm employs a methodology that relies too heavily on simulated data and needs to incorporate lessons learned from AIRS and IASI to be robust at launch (*e.g.,* emissivity over desert or polar regions).

6.5 CrIS IDPS-EDR algorithm's inability to handle empirical radiance bias corrections requires code changes

The operational NGAS IDPS EDR code does not have the ability at this time to derive and apply empirical bias corrections. The cal/val team can develop a methodology and make recommendation for code changes prior to launch using IASI proxy data. Post-launch, the cal/val team can derive empirical bias corrections (see Appendix 2, Section A2.1) with respect to the Optimal Spectral Sampling (OSS) radiative transfer model (RTM), and make recommendations to NGAS.

6.6 CrIS carbon monoxide product of lesser quality than heritage retrievals requires sampling changes.

The truncation of the SWIR interferogram from 0.8 cm to 0.2 cm causes the CO resonance at 2142 cm^{-1} to be unresolved. Simulations show that this product will be significantly degraded relative to AIRS and IASI CO products. To be responsive to IORD 4.1.6.8.3, recommendations for NPP and NPOESS to increase shortwave infrared (SWIR) interferogram sampling have been made numerous times. Apparently this is a command option in the CrIS instrument, but the decision needs to be made prior to completion of ground testing. There are also advantages of having high spectral resolution in the SWIR for frequency calibration. This change does not impact NGAS IDPS EDR capabilities, since temperature sounding lines are not resolved at either spectral resolution, but the lack of a CO product does impact the hyperspectral science community. Operationally, a high-spectral-resolution SDR could be preprocessed into a low-spectral-resolution SDR before presenting it to the IDPS EDR algorithm, thereby eliminating any costs associated with retrofitting the EDR code.

7.0 References

Barnet, C. D., M. Goldberg, T. King, N. Nalli, W. Wolf, L. Zhou and J. Wei 2005. “Alternative cloud clearing methodologies for the Atmospheric Infrared Sounder (AIRS).” *SPIE* **5890** 12 pgs. doi:10.1117/12.615238

Chahine, M. T., T.S. Pagano, H. H. Aumann, R. Atlas, C. Barnet, J. Blaisdell, L. Chen, M. Divakarla, E. Fetzer, M. Goldberg, C. Gautier, S. Granger, S. Hannon, F. W. Irion, R. Kakar, E. Kalnay, B. H. Lambrigtsen, S.Y. Lee, J. Le Marshall, W. W. McMillan, L. McMillin, E. T. Olsen, H. Revercomb, P. Rosenkranz, W.L. Smith, D. Staelin, L .L. Strow, J. Susskind, D. Tobin, W. Wolf and L. Zhou 2006. AIRS: improving weather forecasting and providing new data on greenhouse gases. *Bull. Amer. Meteor. Soc*. **87** p.911-926. doi:10.1175/BAMS-87-7-911

Divakarla, M. G, C. D. Barnet, M. D. Goldberg, L. M. McMillin, E. Maddy, W. Wolf, L. Zhou and X. Liu. 2006 “Validation of AIRS temperature and water vapor retrievals with matched radiosonde measurements and forecasts. “ *J. Geophys. Res*. **111** D09S15 doi:10.1029/2005JD006116.

Divakarla, M, C. Barnet, M. Goldberg, E. Maddy, B. Irion, M. Newchurch, X. Liu, W. Wolf, L. Flynn, G. Labow, X. Xiong, J. Wei, and L. Zhou 2008. “Evaluation of Atmospheric Infrared Sounder (AIRS) Ozone Profiles and Total Ozone Retrievals with Matched Ozonesonde Measurements, ECMWF Ozone Data, and Ozone Monitoring Instrument (OMI) Retrievals.”, *J. Geophys Res*. **113** D15308 doi:10.1029/2007JD009317.

Dunion, J. P. and C. S. Velden 2004. The impact of the Saharan air layer on Atlantic tropical cycle activity. *Bull. Amer. Meteor. Soc*. **85** p.353-365.

GCOS-122 2008. Report of the GCOS Reference Upper Air Network (GRUAN) implementation meeting, Lindenberg, Germany, 26-28 February 2008 available at http://www.wmo.ch/pages/prog/gcos/index.php?name=publications.

Hagan, D. E., *et al.* 2004. Validating AIRS upper atmosphere water vapor retrievals using aircraft and balloon in situ measurements. *Geophys. Res. Lett.*, **31**, L21103, doi:10.1029/2004GL020302.

Larar A. M., D. K. Zhou, X. Liu, and W. L. Smith, “Select Methodology for Validating Advanced Satellite Measurement Systems,” *Proceedings of the International Radiation Symposium (IRS2008)*, Foz do Iguaçu, Brazil, 3-8 August, 2008.

Maddy, E. S. and C. D. Barnet 2008. Vertical resolution estimates in Version 5 of AIRS operational retrievals. *IEEE Trans. Geosci. Remote Sens*. **46** p.2375-2384

Maddy, E. S, C. D. Barnet, and M. Goldberg, C. Sweeney, X. Liu 2008. CO2 retrievals from the Atmospheric Infrared Sounder: Methodology and Validation. *J. Geophys. Res*. **113** D11301 doi:10.1029/2007JD009402.

McMillan, W. W., J. X. Warner, M. L. McCourt-Comer, E. S. Maddy, A. Chu, L. Sparling, E. Eloranta, R. Hoff, G. Sachse, C. D. Barnet, I. Razenkov and W. Wolf 2008. AIRS views of transport from 12-22 July 2004 Alaskan/Canadian fires: correlation of AIRS CO & MODIS AOD with forward trajectories and comparison of AIRS CO retreivals with DC-8 *in situ* measurements during INTEX-A/ICARTT. *J. Geophys. Res*. **113** D20301 doi:10.1029/2007JD009711.

McMillan, W. W., C. D. Barnet, L. Strow, M. T. Chahine, M.L. McCourt, J. X. Warner, P. C. Novelli, S. Korontzi, E. S. Maddy and S. Datta 2005. Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.* **32** p.11801-11801. doi:10.1029/2004GL021821

Morris, V., P. Clemente-Colon, N. R. Nalli, E. Joseph, R. A. Armstrong, Y. Detres, M. D. Goldberg, P. J. Minnett, and R. Lumpkin 2006. Measureing trans-Atlantic aerosol transport from Africa. *Eos. Trans, AGU*, **87**, p.565-571.

Nalli, N. R., P. Clemente-Colón, V. Morris, E. Joseph, M. Szczodrak, P. J. Minnett, J. Shannahoff, M. D. Goldberg, C. Barnet, W. W. Wolf, W. F. Feltz and R. O. Knuteson, 2005: Profile observations of the Saharan air layer during AEROSE 2004, *Geophys. Res. Lett.*, **32**, L05815, doi:10.1029/2004GL022028.

Nalli, N. R., P. Clemente-Colon, P. J. Minnett, M. Szczodrak, V. Morris, E. Joseph, M. D. Goldberg, C. D. Barnet, W. W. Wolf, A. Jessup, R. Branch, R. O. Knuteson and W. F. Feltz 2006. Ship-based measurements for infrared sensor validation during Aerosol and Ocean Science Expedition 2004. *J. Geophys. Res.* **111** D09S04 doi:10.1029/2005JD006385, 13 pgs.

Nalli, N. R., H. Xie, C. D. Barnet, and M. D. Goldberg 2008. Using AEROSE-domain SEVERI data for legacy sounding products from the GOES-R Advanced Baseline Imager. In *Proc. EUMETSAT Meteorological Satellite Conference*, Darmstadt, Germany EUMETSAT, *EUM P.52*, ISBN 978-92-9110-082-8, ISSN 1011-3932.

Monahan, K. P., L. L. Pan, A. J. McDonald, G. E. Bodeker, J. Wei, S. E. George, C. D. Barnet and E. Maddy 2007. “Validation of AIRS v4 ozone profiles in the UTLS using ozonesondes from Lauder, NZ and Boulder, USA.” *J. Geophys. Res*. **112** D17304 doi:10.1029/2006JD008181, 11 pgs.

Pan, L., K. P. Bowman, M. Shapiro, W. J. Randel, R. Gao, T. Campos, A. Cooper, C. Davis, S. Schauffler, B. A. Ridley, J. C. Wei and C. Barnet 2007. “Chemical behavior of the tropopause observed during the Stratosphere-Troposphere Analyses of Regional Transport experiment.” *J. Geophys. Res*. **112** D18110 doi:10.1029/2007JD008645, 13 pgs.

Susskind, J., C. D. Barnet, J. M. Blaisdell, L. Iredell, F. Keita, L. Kouvaris, G. Molnar and M. T. Chahine 2006. "Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover." *J. Geophys. Res*. **111** doi:10.1029/2005JD006272, 19 pgs.

Taylor, J. P., W. L. Smith, V. Cuomo, A. M. Larar, D. K. Zhou, C. Serio, T. Maestri, R. Rizzi, S. Newman, P. Antonelli, S. Mango, P. Di Girolamo, F. Esposito, G. Greico, D. Summa, R. Restieri, G. Masiello, F. Romano, G. Pappalardo, G. Pavese, L. Mona, A. Amodeo and G. Pisani 2008. "EAQUATE An international experiment for hyperspectral atmospheric sounding validation. " *Bull. Amer. Meteor. Soc*. **89** p.203-218. doi:10.1175/BAMS-89-2-203

Tobin, D. C., H. E. Revercomb, R. O. Knuteson, B. L. Lesht, L. L. Strow, S.E. Hannon, W. F. Feltz, L. A. Moy, E. J. Fetzer and T. S. Cress 2006a. "Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water retrieval validation." *J. Geophys. Res.* **111** D09S14 doi:10.1029/2005JD006103, 18 pgs.

Tobin, D. C., H. E. Revercomb, C. C. Moeller and T. S. Pagano 2006b. "Use of Atmospheric Infrared Sounder high spectral resolution spectra to assess the calibration of Moderate resolution Imaging Spectroradiometer on EOS Aqua." *J. Geophys. Res*. **111** D09S05 doi:10.1029/2005JD006095, 15 pgs.

Warner, J., C. Barnet, M. Luo and G. Sachse 2008. Improved Tropospheric Carbon Monoxide Profiles Using Collocated AIRS and TES Satellite Measurements. submitted to Geophys. Res. Lett.

Warner, J., M. McCourt-Comer, C. D. Barnet, W.W. McMillan, W. Wolf, E. Maddy and G. Sachse 2007. "A comparison of satellite tropospheric carbon monoxide measurements from AIRS and MOPITT during INTEX-A." *J. Geophys. Res*. **112** D12S17 doi:10.1029/2006JD007925, 12 pgs.

Wessel, J., S. M. Beck, Y.C. Chan, R. W. Farley and J. A. Gelbwachs 2000. "Raman Lidar calibration for DMSP/SSM/T-2 microwave water vapor sensor." *IEEE Trans. Geosci. Remote Sens*. **38** p.141-154.

Xiong, X. C. D. Barnet, E. S. Maddy, C. Sweeney, X. Liu, L. Zhou, M. Goldberg 2008. "Characterization and Validation of Methane Products from the Atmospheric Infrared Sounder (AIRS)." *J. Geophys. Res.* **113** G00A01 doi:10.1029/2007JG000500.

Zhou, D. K., W. L. Smith, V. Cuomo, J. P. Taylor, C. D. Barnet, P. Di Girolamo, G. Pappalardo, A. M. Larar, X. Liu, S. M. Newman, C. Lee and S. A. Mango 2007. "Retrieval validation during the European Aqua Thermodynamic Experiment." *Quart. J. Roy. Meteor. Soc*. **133** p.203-217.

Zhou, D. K., W. L. Smith, A. M. Larar, X. Liu, J. P. Taylor, P. Schlüssel, L. L. Strow, and S. A. Mango (2008), All weather IASI single field-of-view retrievals: Case study - Validation with JAIVEx data, *Atmos. Chem. Phys. Discuss*., **8**, 21001–21035.

Relevant calibration/validation documents for CrIS/ATMS:

DeAmici, G. 2003. NPOESS calibration and validation plan volume 3: Cross-track Infrared Microwave Sensor Suite (CrIMSS). NGAS Document No. D31409-03, 272 pgs.

Mango, S.A., R.E. Murphy, J.D. Cunningham and P. Sablehouse 2001. NPOESS NPP Calibration and Product Validation Plan. IPO Technical Document Version 11, Dec. 30, 2001, 216 pgs.

Gu, Degui (POC) Feb. 8, 2007. CrIS Volume II, EDR ATBD (ref P1196-TR-I-4-0-ATBD-01-04). NGAS Document No. D43772, Rev. A, 305 pgs.

York, T. (POC) Oct. 18, 2007. NPOESS System Specification. NGAS Document No. SY15-0007, Rev. M, 221 pgs.

TRW Space & Electronics Oct. 28, 2004. NPOESS calibration and validation plan volume 12: Joint NPP plan. Draft Version 4.0, 475 pgs.

EDRPR: NPP EDR Production Report, Document Number: D37005 Revision D, 11 Feb 2009

https://collab2.st.northropgrumman.com/eRoom/npoess/ConfigurationManagement/0_2a46ef

APPENDIX 1: Validation of the NPOESS System Specification for NPP CrIMSS EDRs

The objective of this section is to identify activities and procedures needed to conduct calibration/validation (cal/val) of the CrIMSS EDRs against requirements levied in the NPOESS System Specification (Sys Spec) for the NPP and NPOESS programs.

A1.1 Scope

The EDR cal/val effort is a team activity that involves expertise from members of the government, research institutions, academia, and the NPOESS contractor (NGAS). Since the EDR cal/val tasks envisaged by the team cover a broader range of activities than those needed solely to assess EDR performance against the NPOESS Sys Spec, this section should focus only on those tasks necessary to validate the NPOESS Sys Spec. As a result, this plan everages on the activities benefiting the larger community and identified in other sections of this plan, and identifies any additional tasks needed to accomplish this objective.

A1.2 General Requirements and Assumptions

Requirements for this version of the cal/val Plan of the CrIMSS EDRs are listed in Revision N of the NPOESS Sys Spec (SY15-0007). Appendix D of the Sys Spec identifies impacts of weather and measurement conditions on EDR requirements' lists climate product long-term stability requirements; specifies accuracy, precision, and uncertainty (APU) measurement standards for EDR performance; discusses ground-truth/correlative data (GTCD) for EDR APU assessments; and defines individual attribute values—to be interpreted as upper bounds anywhere in the area where measurements are obtained—for each sub-requirement of the EDR.

The NPOESS Sys Spec provides guidance that is applicable to cal/val of all EDR products. Of particular emphasis are the following points:

- Except for cloud EDRs, which must be generated regardless of cloud cover, EDR requirements should apply to clear conditions only, unless otherwise specified.
- Under Degradation Conditions (DC), a product is generated that has utility, but is degraded in performance. For each DC, a performance value is provided as an estimate, but there is no requirement for validation of the estimate.
- Under Exclusion Conditions (EC), a data product may be so degraded as to have reduced utility, measurement may not be possible, or processing may not converge for some EDRs. Unless otherwise indicated for an EDR, the data product will be produced—when possible—to satisfy select users or EDR processing support, but no estimate is made of performance, nor is performance characterization required during validation.

- EDR Performance may be specified for different background conditions. These Stratification Conditions (SC) involve nominal environmental and measurement conditions, and are listed for each product. Performance should be “validated” within a stratification condition, and should not be averaged across SC boundaries. The term “validated” in this context implies that this plan needs to characterize performance and should address mitigation plans to bring performance into compliance.

  - The Sys Spec also indicates that SC and DC apply to the aggregate of retrievals at the horizontal cell size (HCS) level, not on individual retrievals in the aggregation.
  - Unless otherwise specified in specific EDR sections, an HCS performance category is assigned based on conditions associated with at least 50% (TBR) of the individual retrievals used in the aggregation. In some cases, it may be appropriate to exclude entire grid cells when mixed performance categories apply to the individual Ips, or assign a worst-case performance category, based on the worst-case IP category.

- Climate Long Term Stability (LTS) attributes, designated as “Long Term Stability (C)”, are not intended to drive stressing design changes on spacecraft sensors or require increased IDPS processing capabilities. These attributes require characterization of the precision and uncertainty bias of the sensor’s measurements and the impact on the resulting data products. The intent is to be able to characterize Long-term Stability (C) and the associated bias in the measurements, not meet the absolute threshold value of the EDR attribute.
  - The Sys Spec flags some EDR LTS requirements as Climate [“(C)”] related, but no specific guidance is given for the non-climate flagged attributes.
  - Therefore, all LTS attributes will be treated the same with the general approach to be to characterize the stability, as defined for Climate LTS, rather than validating a specific level of performance.

- The APU definitions, and Probability of Correct Typing (PCT) are applied to assess EDR performance under the assumption that every value of ground truth for the EDR data sample is well defined and generally accepted in the user community. When GTCD are used, estimates of their known errors will be factored into the comparisons between the GTCD and the EDRs to produce estimates of the measurement APU, PCT, and LTS of the EDRs. The confirmation of meeting the measurement APU, PCT, and LTS requirements for an EDR may be accomplished by analysis, laboratory measurements, simulations, and comparisons to direct or indirect observations, including observations taken from aircraft or spacecraft platforms.
  - Unless otherwise indicated in the specific EDR sections, the approaches discussed in this plan will be based on comparisons to “direct or indirect observations,” which will be identified as correlative data in this document.
  - Note that the correlative data is not exactly the same as “truth”, in that it has some error characteristics that must be accounted for when computing the APU. It is assumed that the error characteristics will be provided by either the data provider or the matchup provider. In some cases a single organization is collecting matchups that can be used by the entire team, including NGAS.

  - The NPOESS Science Investigator-led Processing System (NSIPS) will be the repository for correlative data collected by all members participating in the cal/val program, and these data will be available for use by cal/val teams in near real-time. Procedures used to collect, process, and distribute correlative data will be provided not documented not later than 18-months (TBD) prior to NPP launch, along with data formats to ensure that all members can exploit these datasets.

- Note that the correlative data is not the same as "truth", in that is has some error characteristics that must be accounted for when computing the APU. The general assumption is that the error characteristics will be provided by either the data provider or the matchup provider. In some cases a single organization is collecting matchups that can be used by the entire team including NGAS

- An EDR or an EDR attribute may be listed as a performance parameter. Key EDR requirements and Key EDR attributes are in **bold type** and marked with an asterisk (*) in the NPOESS Sys Spec.

The requirements for each EDR are typically listed in tables with some guidance provided in the narrative associated with each EDR. Unless otherwise specified in the individual EDR sections that follow, this document will only refer to the validation of those attributes identified as measurement range, measurement APU, PCT, and LTS.

We assume that other attributes listed the EDR requirements tables, such as HCS, reporting interval, refresh rates, mapping uncertainty, etc., are handled by inspection/ analysis of the SDR performance from which these EDR characteristics are derived. EDR latency requirements are validated by inspection of the IDPS processing timelines, and are outside the scope of this document.

The system specification requirements for the CrIMSS EDRs (AVMP, ATVP, Pressure Profile) are found in the Sys Spec, and will not be provided here. Another product, the Ozone Profile IP, derives from the CrIMSS retrieval algorithm. It is defined as the ozone volume mixing ratio at the internal OSS RTM pressure grid along CrIS line-of-sight, with units of parts per million by volume (ppmv). Since it is a byproduct of the CrIMSS EDR retrieval algorithm, it will have the same HCS and reporting interval as the CrIMSS temperature and moisture profile EDRs. There is no specific performance requirement on the IP, but the performance should be assessed and adequately characterized during post-launch cal/val activities.

With these caveats in mind, the following discussion tailors the cal/val plan for the CrIMSS product group to the requirements listed in the NPOESS Sys Spec.

A1.3 CrIMSS EDR Production

The CrIMSS EDR retrieval algorithm (including post-retrieval processing) was designed to produce CrIMSS EDRs that are in compliance with the definitions and requirements specified in the NPOESS System Spec. The NPP EDR Production Report (EDRPR 2009) and CrIMSS EDR algorithm theoretical basis document (ATBD; Gu 2007) contain detailed descriptions of the EDR data produced by the IDPS. Only summaries of component parameters and related processing requirements are provided here.

A1.3.1 Vertical Grid

The vertical reporting grids consist of a set number of vertical cells, which are defined based on the vertical reporting interval and vertical cell size specified in the System Spec.

Specifically, each vertical cell is uniquely defined by its thickness and the pressure at its center. The grid for AVTP consists of 41 vertical cells, with center pressures and thickness listed in Table 7. For AVMP, the grid has 21 vertical cells (Table 8). The retrieved temperature and moisture profiles, which are defined at the internal OSS RTM pressure grid, are interpolated and averaged to compute the EDR values for each of the vertical cells.

Table 7 - Vertical Reporting Grid for AVTP

Layer Index	Pressure (mb)	VCS* (km)	Layer Index	Pressure (mb)	VCS (km)	Layer Index	Pressure (mb)	VCS (km)
1	0.5	5	16	150	3	31	750	1
2	0.7	5	17	175	3	32	800	1
3	0.9	5	18	200	3	33	850	1
4	1	5	19	225	3	34	870	1
5	3	5	20	250	3	35	890	1
6	5	5	21	275	3	36	900	1
7	7	5	22	300	3	37	920	1
8	9	5	23	350	3	38	940	1
9	10	5	24	400	3	39	960	1
10	30	3	25	450	3	40	980	1
11	50	3	26	500	1	41	1000	1
12	70	3	27	550	1	42	1020	1
13	90	3	28	600	1			
14	100	3	29	650	1			
15	125	3	30	700	1			

* Vertical Cell Size

Table 8 - Vertical Reporting Grid for AVMP

Layer Index	Pressure (mb)	VCS (km)	Layer Index	Pressure (mb)	VCS (km)	Layer Index	Pressure (mb)	VCS (km)
1	100	2	10	550	2	19	910	2
2	150	2	11	600	2	20	930	2

3	200	2	12	650	2	21	950	2
4	250	2	13	700	2	22	970	2
5	300	2	14	750	2			
6	350	2	15	800	2			
7	400	2	16	850	2			
8	450	2	17	870	2			
9	500	2	18	890	2			

The pressure profile EDRs are reported on a vertical grid of 0-30km from a reference level at 1100mb. There is no vertical averaging for this EDR.

If a vertical cell is partially or completely below the local terrain, no EDRs are produced and a fill is reported.

A1.3.2 Horizontal Grid

The CrIMSS EDRs are not reported on a typical horizontal grid. Instead, the EDRs are reported for each and every retrieval attempted by the algorithm. The corresponding cell size and geolocation information are part of the product.

By design, the HCS for the CrIMSS EDRs can be single FOV (14km at nadir), 2x2 FOVs (28km at nadir), and 3x3 FOVs (40km), for clear, partly cloudy and cloudy conditions respectively. For the NPP mission, all EDRs will have the cell size of 3x3 FOVs, regardless of cloudiness.

A1.3.3 QC and Merged EDRs

There are a number of quality flags implemented in the CrIMSS EDR algorithm to summarize data product type and quality, and provide additional information to users for quality control and to stratify the data for their specific applications. Table 9 lists all quality flags currently implemented in the algorithm for data and intermediate products.

Table 9 - CrIMSS EDR/IP Quality Flags

Flag Name *	Description	Values	Type	Bits
Product Yield	Percent of retrievals within granule with high quality of retrieval.	0% - 100%	Granule	8 (granule)
CrIS Input Data Quality	Percent of retrieved pixels with high-quality input values for CrIS SDR.	0% - 100%	Granule	8 (granule)
ATMS Input Data Quality	Percent of retrieved pixels with hig-quality input values for ATMS SDR.	0% - 100%	Granule	8 (granule)
CrIS SDR Detector Failure	Flags if (and which) CrIS detectors have failed.	0: good 1: failed	Granule	27 (granule)

Apodization Flag	Indicates the apodization applied to the CrIS SDR.	0: no apodization 1: Hamming 2: Blackman	Granule	2 (granule)
Day/Night Flag	Flags whether a granule is daytime, nighttime, or terminator.	0: Daytime 1: Nighttime 2: Terminator	Granule	2 (granule)
AVTP Overall Quality	Overall quality of the reported AVTP EDR: Converged IR+MW, converged IR only if ATMS not available, converged MW only if IR or IR+MW didn't converge, and non-converged. Applies to AVTP only.	0: (IR + MW) 1: (IR only) 2: (MW only) 3: (non-converged)	Retrieval	2
AVMP Overall Quality	Overall quality of the reported AVMP EDR: Converged IR+MW, converged IR only if ATMS not available, converged MW only if IR or IR+MW didn't converge, and non-converged. Applies to AVMP only.	0: (IR + MW) 1: (IR only) 2: (MW only) 3: (non-converged)	Retrieval	2
PP Overall Quality	Overall quality of the reported pressure profile EDR: Converged IR+MW, converged IR only if ATMS not available, converged MW only if IR or IR+MW didn't converge, and non-converged. Applies to pressure profile EDR only.	0: (IR + MW) 1: (IR only) 2: (MW only) 3: (non-converged)	Retrieval	2
Ozone Overall Quality	Overall quality of the reported Ozone Profile IP: Converged IR+MW, converged IR only if ATMS not available, or non-converged. Applies to ozone profile IP only.	0: (IR + MW) 1: (IR only) 2: (non-converged)	Retrieval	2
ATMS SDR Quality	Quality of the re-mapped ATMS SDR - pass through from remapped ATMS SDR.	0: good 1: invalid	Retrieval	22 (22x1)
CrIS SDR Quality	Quality of the CrIS SDR; pass-through from CrIS SDRs.	0: good 1: degraded 2: invalid	Retrieval	54 (9x3x2)
ATMS Availability	Flags if ATMS data are available.	0: Available 1: Not available	Retrieval	1
IR-MW Convergence	Flags if the IR + MW retrieval has converged.	0: Converged 1: Did not Converge	Retrieval	1
MW Only Convergence	Flags whether the ATMS MW-only retrieval has converged.	0: Converged 1: Did not Converge	Retrieval	1

Temperature Difference	Flags if the statistical difference between the MW+IR temperature profile and the MW-only temperature profile exceeds a threshold.	0: Within threshold or over land 1: Not within threshold	Retrieval	1
Cloudiness	Indicates cloud conditions within a CrIS FOR as clear, partly cloudy, or cloudy per definition of the NPOESS Sys Spec, Appendix D.	0: Clear 1: Partly Cloudy 2: Cloudy	Retrieval	2
Rain Flag	Indicates precipitation within the CrIS FOR exceeding 2 mm/hr.	0: No rain 1: Rain	Retrieval	1
Sun Glint Flag	Flags sun glint within the CrIS FOR. There is only one flag per FOR, and the sensor angles for the center FOV (FOV 5) are used for these calculations.	0: No sun glint 1: Sun glint	Retrieval	1
Coast Flag	Fraction of land in the CrIS FOR.	0-100%	Retrieval	16
Cell Size	Indicates the number of FOVs in a cluster (cell) on which the retrieval is performed. Can be 1, 4, or 9, with corresponding HCS of approximately 15, 30, or 45 km, respectively	0: 1 FOV 1: 4 FOVs 2: 9 FOVs 3: No retrieval	Retrieval	2
Temperature Out of Range Flag	This flag indicates that the atmospheric temperature at one or more of the pressure levels, or the surface skin temperature, is out of the expected range.	0: all in range 1: one or more out of range	Retrieval	1
Non-LTE Flag	Indicates a NLTE condition, which occurs when the local solar zenith angle is equal to or less than 90° (*i.e.,* daytime). When the non-LTE is set to 1, all channels from 2250-2380cm^{-1} should be excluded from the retrieval.	0: no Non-LTE 1: Non-LTE	Retrieval	1
Ice Mass Flag	This flag indicates ice from a water surface. It is set based on the surface temperature retrieved in the first stage MW-only retrieval: Flagtged as ice if less than 273.15K	0: No Ice 1: Ice on water surface	Retrieval	1
IR-MW Chi Square	Statistical difference between cloud-cleared (measured for MW channels) radiances and computed radiances - Stage 2	Real	Retrieval	32

MW Chi Square 1	Statistical difference between observed and computed microwave radiances - Stage 1	Real	Retrieval	32
MW Chi Square 2	Statistical difference between observed and computed microwave radiances - Stage 2	Real	Retrieval	32
IR Noise Amplification Factor	IR noise amplification factor due to cloud clearing	Real	Retrieval	32
Ozone Spectral Signature Flag (OSSF)	Difference between the ozone absorption radiance and the background radiance. Defined as: OSSF = (R(529)+R(673) – R(625) – R(646))/2. The center frequencies of these bins are 980cm^{-1}, 1070 cm^{-1}, 1040 cm^{-1}, and 1053.125 cm^{-1}, respectively.	Real	Retrieval	32

The algorithm performs two sequential retrievals in processing each set of CrIS/ATMS SDR measurements. In the first stage, the MW-only retrieval, only ATMS data is used. In the second stage, both ATMS and CrIS data are processed to solve for the state parameters. The algorithm computes a number of quality control measures to determine whether to report the first stage or the second stage retrieval results. In general, the second-stage retrieval result will be reported if it can match both the MW and cloud-cleared IR radiances to within the predetermined thresholds; if not, the MW-only retrieval results will be reported. In the rare case when both retrievals fail to pass the quality control, the initial guess (based on climatology) will be reported.

A1.4 Requirement Validation

This plan will focus on the requirements of measurement APU only.

A1.4.1 Requirement Interpretation

The CrIMSS EDR requirements are significantly different between the NPOESS Sys Spec and the IORD, as first demonstrated in Table 2 in Section 1.2.1 of the main document.

According to the IORD, CrIMSS EDRs will be stratified based on the amount of clouds present in the scene. If the cloud coverage is less than 50%, it is classified as “clear”; if the cloud coverage is over 50%, it is called “cloudy”. The requirements are usually tighter under clear *vs.* cloudy conditions. This in general agrees with the understanding that the capability of an infrared sounder is limited by clouds. However, it fails to take into consideration what the retrieval algorithms will do with the clouds.

Similar to the AIRS algorithm, the CrIMSS EDR retrieval algorithm compares sensor-measured radiances in nearby FOVs to detect and remove cloud contamination to get the so-called "cloud-cleared" radiance, the radiance that would have been observed if the clouds were not present. How well this cloud-clearing process works depends on not only the presence and extent of clouds in the scene, but also the cloud contrast between adjacent FOVs: The larger the contrast the more accurate the estimated cloud radiance is expected to be.

With this understanding, three cloud stratifications have been defined in the NPOESS Sys Spec: C*lear, Partly Cloudy,* and C*loudy*. As explicitly described in the Sys Spec, whether a cell is clear, partly cloudy, or cloudy will be determined by analyzing radiances from CrIS spectral bands, and will include consideration of ancillary data, algorithm convergence criteria, and quality control in performance tuning. For the CrIMSS EDRs, clear conditions should include "cloud-free" (full IR sensing capability) as a minimum, and "predominantly cloud-free," to the extent permitted by uncertainty requirements, described below, tuned to increase yield above the minimum cloud-free case. Cloudy specifications are applicable only to conditions when the FOR, which contains 3 x 3 CrIS FOVs, is "predominantly overcast", resulting in loss of IR sensing capability. Partly cloudy conditions exist if neither clear nor cloudy conditions are met. We anticipate that most of the measurements will be classified as partly cloudy when CrIS is still a major contributor to generating the product, but with a significant amount of capability lost due to cloud contamination.

To characterize CrIMSS EDR product quality and to score performance against the Sys Spec, we need to quantitatively define the boundaries between clear and partly cloudy conditions and between cloudy and partly cloudy conditions. Currently, the CrIMSS EDR performance has been measured and reported with the following definitions:

- *Clear* - the CrIMSS EDR retrieval algorithm detected no cloud within a FOR;
- *Cloudy* - the CrIMSS EDR algorithm detected overcast cloud or more than three layers of clouds within a FOR;
- *Partly Cloudy* - the CrIMSS algorithm detected one to three layers of clouds.

There is one exception, i.e., when an IR and MW retrieval does not converge or pass the internal quality control, the EDR product will be produced from the MW-only retrieval. In that case, the EDR products should be re-lassified as cloudy regardless of the initial cloud detection results.

There are differences in CrIMSS component usage, depending on the situation: With this definition, CrIS will have full capability and produce the best quality EDRs under clear conditions. Under cloudy condition, it is just the opposite: CrIS loses most of its capability, so the EDRs will be produced from ATMS data alone. Most of the measurements are expected to fall between the two extremes and have the quality somewhat improved over the ATMS alone, but less than the cloud-free CrIS results. It should be pointed out that the cloud classification is flexible, and the drawing lines can be adjusted post-launch based on the actual performance and/or users' needs.

A1.4.2 Degraded and Excluded Conditions

EDRs produced under degraded and exclusion conditions should be identified and excluded when assessing the CrIMSS EDRs performance against the Sys Spec. These EDRs should be characterized for performance separately. EDR performance degradation and exclusion can be resulted from three categories discussed below.

A1.4.2.1 Environmental Conditions

Degraded or excluded environmental conditions are clearly specified in the NPOESS Sys Spec. There is one excluded condition for CrIMSS temperature and moisture EDRs: When precipitation rate is over 2 mm/hr, these two EDRs are produced, but have no performance requirements.

There is no degraded condition specified for the temperature and moisture EDRs.

There is no exclusive condition specified for the pressure profile EDR, but two degraded conditions are specified (1) when precipitation is within 13 km, and (2) when surface pressure uncertainty >2.5mb.

Note: The pressure EDRs derive from the temperature and moisture EDRs. There could be some inconsistency in the specifications. This potential inconsistency needs to be reviewed and resolved.

A1.4.2.2 Sensor Hardware Conditions

We expect that under certain unfavorable operating conditions, the CrIMSS sensors may produce degraded or even invalid SDR data. When such sensor performance anomalies occur, we expect that the SDR algorithms will catch the condition and flag the SDRs accordingly. The CrIMSS EDR algorithm has implemented the logic to check CrIS and ATMS SDR quality flags to ensure proper handling of degraded and invalid SDR data. The SDR overall quality flags are retained in the EDR products (table).

A1.4.2.3 Algorithm and Ancillary Conditions

Even with normal SDRs, the algorithm may not always produce valid EDRs. This could occur, for example, when the algorithm fails to converge. Also, ancillary data required to run the EDR algorithm may be missing, degraded, or unavailable, causing further EDR quality degradation or exclusion. Similar to the degradations and exclusions related to sensor operating conditions or environmental conditions, these invalid or degraded EDRs should also be excluded from the performance validation.

A1.4.3 Algorithm Tuning

Although the CrIMSS EDR algorithm will continue to be tested with simulated and proxy test data to demonstrate performance compliance before launch, some algorithm modules will not become fully functional until they are tuned with real data. Without the required algorithm tuning, the CrIMSS EDRs are not expected to be fully compliant with the NPOESS Sys Spec.

A1.4.3.1 ATMS Precipitation Detection Algorithm

In the CrIMSS algorithm, scenes with precipitating clouds are detected using the NOAA/NESDIS Day 1 algorithm. Over land, the algorithm applies a scattering index test using the 23- and 89-GHz channels. Over ocean, the algorithm relies on both a scattering index test and an emission-based test. The emission test uses cloud liquid water information obtained from the 23-, 31-, and 50-GHz channels. For both land and ocean, thresholds are applied to determine if precipitation is present or not. Further analysis of the scene is performed to differentiate precipitation from snow cover, sea ice, and deserts.

Since the NESDIS algorithm currently implemented only applies to AMSU channels, it must be trained after launch with real data to make it work with ATMS. This training requires global precipitation data to be collected and collocated with the ATMS footprint.

A1.4.3.2 Cloud Detection Algorithm

Cloud detection is based on the CrIS radiance contrast within a CrIS FOR. Radiances in the cloud-clearing channels are analyzed using a principal component analysis approach to determine the number of distinguishable cloud formations (including background radiance) within the FOR. Two tests are performed: One is to check the eigenvalues; the other to check the residual error in the reconstructed radiances using the “significant” eigenvectors. In the process, two empirical coefficients are introduced to improve cloud detection accuracy. Since they are initially tuned and set with the simulated CrIS data and they show some degree of sensitivity to the radiance error characteristics, these two empirical parameters need to be tuned with real data after launch to optimize performance.

Some cloud truth data is desirable to verify and validate the tuned parameters. However, since these parameters are optimized for the CrIMSS EDR performance, it is sufficient to have the tuning done and verified with the EDR performance validation truth dataset.

A1.4.3.3 Local Angle Adjustment Algorithm

The CrIS cloud-clearing algorithm uses observations collected in a 3 × 3 FOV array to correct the observed radiances for the presence of clouds. This method assumes that the observations collected at the 9 FOVs are equivalent except for the cloud amount. In fact, the observations are viewed along slightly differing paths through the atmosphere., which modifies the weighting functions for each of CrIS channels. The resulting radiance differences are a source of error in the retrieval if they are incorrectly attributed to cloud. This problem can be addressed by adjusting the observations to a common, central zenith angle. The radiance adjustment for a given observation (*i.e.*, at each FOV) is computed for each CrIS channel as a linear function of the observed radiances for all channels. The coefficients used in the correction are determined through a regression analysis, based on a representative set of radiances generated for the appropriate geometry. The correction process is referred to as the Local Angle Adjustment (LAA), and is performed in the preprocessing module prior to the retrievals. The CrIS LAA approach is similar to that outlined in the AIRS ATBD [REF].

The LAA algorithm coefficients should be tested and adjusted with real data to ensure performance compliance. Ineffective or erroneous LAA module may lead to incorrect cloud detection and cause degradation in the retrieval profiles.

A1.4.3.4 Updates to the *a priori* and Empirical Orthogonal Functions

To ensure the retrieved geophysical parameters are physically plausible, the CrIMSS EDR algorithm uses climatology as constraints in the inversion of radiances to estimate state parameters. In addition, some of the state parameters are represented with Empirical Orthogonal Functions (EOF) to improve the numerical stability and computation speed.

The climatology and EOFs are derived from a diversified global dataset intended to reduce the algorithm’s dependence on this training process. These EOFs—and the corresponding climatological means and covariance—need to be updated with the best available representative global dataset. The update is especially needed for the surface emissivity (both IR and MW) due to lack of more comprehensive data at the time they were initially derived. AIRS or IASI emissivity product—if validated—could be very useful for this purpose.

Training (*i.e.*, EOFs, means, and covariance) is needed for the following parameters: temperature profile, surface skin temperature, moisture profile, ozone profile, surface MW emissivity, and surface IR emissivity.

A1.4.3.5 OSS RTM tTning and Error Characterization

The OSS RTM has recently been updated based on AER's improved and more accurate line-by-line RTM (LBLRTM) [REF]. A limited validation is done with AIRS and IASI data to characterize the RTM error. We anticipate that additional validation and updates should be conducted to improve accuracy post-launch. As a result of this validation effort, the RTM uncertainty will be updated and used by the retrieval algorithm. In addition, there is a good possibility that a bias term may need to be quantified and corrected in the CrIMSS EDR algorithm.

To characterize the OSS RTM error and possibly make improvements, a complete set of state parameters of the best quality—and collocated with CrIS/ATMS radiances—is needed. Dedicated radiosondes, and possibly numerical weather predication data, should be collected for this purpose. The truth-state data may need to be extended to all OSS RTM pressure levels and augmented with other state parameters (*e.g.,* ozone profile) that are needed to carry out the radiative transfer calculation.

A1.4.3.6 Sensor Error Characterization

The error characteristics of the CrIS and ATMS SDRs are specified based on simulated sensor effects. They need to be updated once real data become available; this can be combined with the SDR cal/val effort. For ATMS SDRs, the actual noise reduction factor achieved with real data for each of the 22 channels must be characterized

These tasks are critical if the CrIMSS EDR algorithm is to achieve the desired performance. cal/val activities should be planned to complete these tasks with the highest possible priority before and/or after launch to optimize the algorithm. Without these expected algorithm updates, the CrIMSS EDR products may not meet all requirements in the Sys Spec. Once the CrIMSS EDR algorithms are successfully tuned with real data, we expect that CrIMSS EDR performance should be in full compliance with the requirements. We recognize several performance risk areas that may have large negative impact on the CrIMSS EDR performance. These potential risk areas and the corresponding mitigation approach will be discussed in section 1.5.6.

A1.5 Validation Truth Data Required

A1.5.1 [Title TBD]

The truth datasets discussed in Section 2.3 in general meet the needs for verifying the NPOESS system specifications. In the order of relevance they are relisted below:

- ECMWF (or NCEP) datasets,
- Operational radiosondes,
- Dedicated, high-quality radiosondes and other high-quality *in situ* instrument data (*e.g.,* NAST-I, NAST-M, S-HIS, etc.)

- Satellite products from other missions.

It needs to be clarified that how these datasets should be used either individually or in a combined fashion to characterize the CrIMSS EDR performance and to demonstrate NPOESS Sys Spec compliance. In addition, these datasets must be collected/generated following agreed requirements on contents, quality, formats, etc. At a minimum, these data should meet the following requirements:

- Each dataset should contain a sufficiently large number of samples for each stratification to render the computed performance statistics significant and meaningful;
- Each dataset should be collocated with the CrIMSS EDRs;
- Each dataset should cover the full range of the required measurement conditions;
- Each dataset should be converted to EDRs or be convertible to EDRs, per definition described in Section 1.4.1.
- Each dataset should be fully quantified as to uncertainty. If the data in a dataset have different quality, they should be quality-flagged with quantitative measure of uncertainty for each subgroup;
- Each dataset show be fully characterized as to spatial and temporal mismatches between the truth and the EDRs.

The truth data should also be in the agreed format, fully documented, and deposited to the data server to which NGAS has access. Detailed format requirements must be defined as soon as possible to facilitate the ongoing development of cal/val analytical tools.

A1.5.2 Performance Characterization and Validation

The matchup datasets should be stratified and filtered based on the System Spec to form sub-datasets corresponding to different measurement conditions. The EDR products and the matchup truth for degraded and excluded conditions should form separate sub-datasets. Nonconverged and fill EDRs should be excluded in performance assessments.

EDR performance statistics should be computed by comparing the EDRs and the matchup truth data for each EDR stratification group, including degradation and exclusion groups.

A1.5.3 Performance Statistics

The uncertainty of the temperature and pressure profile EDRs should be computed using the following equation:

$$U_T = \sqrt{\sum_{i=1}^{N} \frac{(T_i - T_i^0)^2}{N}}$$

where: U_T is the uncertainty at a given temperature, , T_i is the retrieved temperature profile EDR (*i.e.,* mean temperature in a cell), and T_i^0 is the error-free truth value of the measured EDR.

In practice, there is always error in the validation truth data, and one can only compute the RMS “difference” between the retrieved EDRs and the “truth”

$$\hat{U}_T = \sqrt{\sum_{i=1}^{N} \frac{(T_i - \hat{T}_i^0)^2}{N}}$$

where: T is the EDR value computed from the validation truth data. Obviously, such calculated RMS error is not the same as EDR uncertainty because it also contains the error inherent in the validation truth data. An error model needs to be developed to infer from the RMS difference the true EDR uncertainty. Additional discussion of this error model follows.

For the moisture profile EDR, the uncertainty should be calculated using the following equation

$$\hat{U}_Q = \frac{\sqrt{\sum_{i=1}^{N} (Q_i - \hat{Q}_i^0)^2}}{\sqrt{\sum_{i=1}^{N} (\hat{Q}_i^0)^2}} \times 100$$

where: Q_i, Q and U are the retreived moisture EDR value, the truth value, and the error-relative EDR error (percentage), respectively. Again, such computed uncertainty is the total RMS difference between the retrieved moisture profile EDR and the validation truth EDR.

The pressure profile EDR is specified as to precision and accuracy. They should be calculated using the following equations, respectively:

$$\hat{P}_P = \frac{\sqrt{\sum_{i=1}^{N} (P_i - \hat{P}_i^0)^2 / (N-1)}}{\sum_{i=1}^{N} \hat{P}_i^0 / N} \times 100$$

$$\hat{A}_P = \frac{1}{N} \sum_{i=1}^{N} (P_i - \hat{P}_i^0)$$

There is no APU requirement specified for the Ozone Profile IP, but these values should be computed in a manner similar to that for the pressure profile EDR.

A1.5.4 Performance Error Budget Model

The calculated APU contain errors in the truth data and the EDR products, as well as errors due to spatiotemporal mismatch. An error model should be developed to analyze the resulted performance statistics. At a minimum, the model should contain three distinct components: Uncertainty in the truth data, uncertainty in the EDRs, and the expected difference between truth and EDR due to spatiotemporal mismatches. The uncertainty in the EDRs should compared with the Sys Spec requirement to demonstrate performance compliance.

A1.6 EDR Performance Risks

In the best scenario, the CrIMSS EDRs should meet the NPOESS Sys Spec after the required algorithm tuning, as discussed in Section A1.4.3. In case of a performance shortfall, there are areas in the CrIMSS algorithm that could be further improved to have a potentially positive impact on EDR performance quality. These potential algorithm tasks can be added to the post-launch algorithm risk-watch list.

A1.6.1 Cloud Clearing

The cloud-clearing approach that is based on radiance contrast in adjacent FOVs has been widely adopted by the sounding community to remove cloud contamination in the infrared sounding data [REF]. It has proven effective by application to hyperspectral sounders (*e.g.,* AIRS and IASI) of spectral and radiometric characteristics that are similar to CrIS. The cloud-clearing method implemented in the CrIMSS algorithm is the same as the AIRS algorithm, so we expect it to work well.

Nevertheless, as a risk reduction effort NGAS has examined at alternative approaches to improve cloud-clearing performance. One inserts clear-sky VIIRS IR radiances into the cloud-clearing process. A preliminary study, based on simulated data, has demonstrated the positive impact [REF]. How much improvement can be realized with real data is not yet clear, but may worth the effort to plan for in advance.

A1.6.2 Surface Emissivity

CrIMSS EDR performance may be impacted if surface emissivity is improperly handled.

Currently, the algorithm simultaneously retrieves surface emissivity at selected frequency hinge points, and then linearly interpolates to all other frequencies between the hinge points. We expect that this approach will work fine over ocean or surfaces that are largely covered by vegetation. Over semiarid or desert regions, however, surface emissivity can exhibit large spectral variability; this may cause problems. Although the hinge points can be increased and adjusted to better represent the emissivity in these challenging conditions, it remains to be verified how effective the adjustment can be. There are alternative approaches to deal with surface emissivity. For example, it is possible to use EOFs to represent emissivity spectral variability. It may also be worth some effort to modify the algorithm to rely more on the initial-guess emissivity and the *a priori* under these challenging conditions.

A1.6.3 Thin Cirrus Clouds

There is a concern that thin cirrus depression may pass through cloud clearing and be present in the cloud-cleared radiances. Because of the prevalence of thin cirrus clouds, it may worth some effort to investigate their impact on EDR performance if this not properly handled, and to devise a method to minimize the impact if deemed necessary

A1.6.4 Non-local Thermodynamic Equilibrium (NLTE)

The OSS RTM in the CrIMSS EDR algorithm does not model radiative effects due to NLTE. Therefore, when processing daytime data the algorithm turns off the channels that are mostly affected by NLTE (*i.e.*, 2250-2380 cm^{-1}). This may not be desirable and performance may be somewhat impacted. Since a technique to incorporate NLTE effects into the RTM exists [REF], its use in updating the OSS RTM should be considered either before or after launch.

A1.6.5 Trace Gas Retrieval

The baseline CrIMSS retrieval algorithm assumes a fixed amount of trace gases (*e.g.*, CH_4, N_2O, CO, etc.) Deviation of the actual profile shape and amount from assumed values will cause some retrieval errors. There are two straightforward algorithm modifications that could mitigate this performance risk: One is to turn off the channels that are the most affected by these trace gases; the other is to simultaneously retrieve total column measurements of some of the trace gases (*e.g.*,CH_4 and N_2O). During the cal/val period, studies should be performed to assess actual performance improvement requirements, and implement the changes if needed.

A1.7 Procedures for Calibration and Validation of CrIMSS EDRs

It is of paramount importance that procedures are developed to acquire, segment, and archive data needed to conduct the cal/val process for EDRs generated by the NPOESS Operational System. Data generated by the NOAA IDPS at the NOAA Satellite Operations Facility (NSOF) will be distributed to the cal/val user community through NSIPS. All EDRs and deliverable IPs will be archived permanently in the Comprehensive Large Array-data Stewardship System (CLASS). Deliverable IPs include the VIIRS Cloud Mask IP, the Quarterly Surface Type IP, the Ozone Nadir Profiler Deliverable IP, and the IR Ozone Product Deliverable IP. Retained IPs , *e.g.* pixel-level cloud top temperatures, are provided solely for the purposes of cal/val, and will be available on NSIPS for only seven days (TBD), although special procedures can be implemented if required. A list of the retained IPs is given in the NPP EDRPR. Further, NSIPS serves as a repository for correlative data, which allows users subscribe to match-up datasets through the Product Generation Executive (PGE).

Figure 2 provides an overview of the process of conducting cal/val on EDR products generated by the NOAA IDPS, using correlative data provided by a variety of user groups and distributed via NSIPS. In general, EDR data products have embedded data fields that already contain many of the necessary QC flags and performance conditions necessary to identify stratified, degraded, and excluded conditions. There may be instances where additional flags, found in the delivered IPs, retained IPs or other IDPS data products, may be useful to generate additional internal QC flags that one would want to store with the matchup data between EDRs and correlative data. A process is needed to map this additional data into the EDR grid and decide how the aggregation of the individual quality control flags is performed (*e.g.* min, max, average, count).

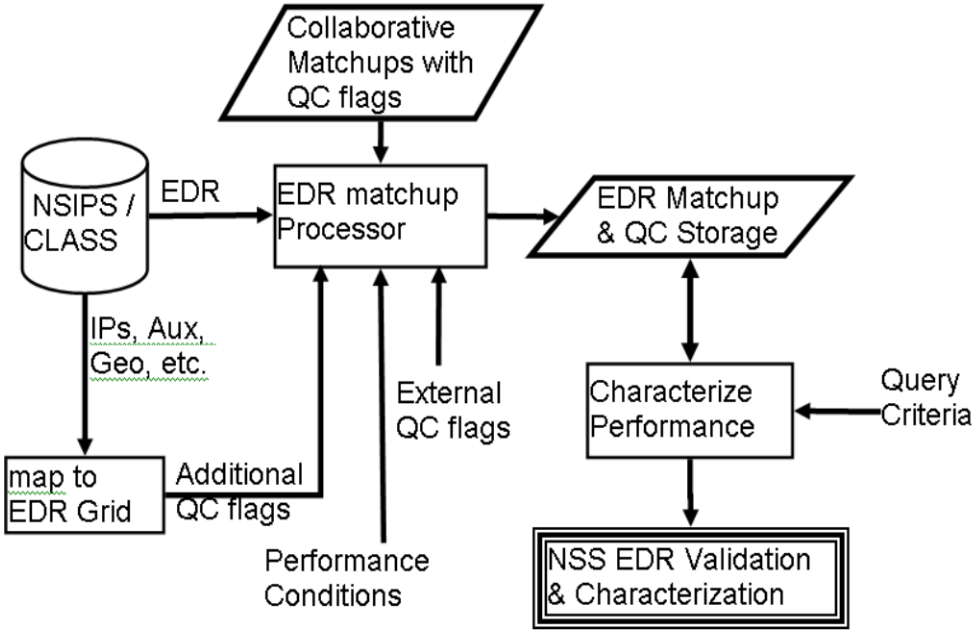


Figure 2 - Process to establish performance of the IDPS on the CrIMSS EDRs

QC flags can be used to indicate an issue with any source of EDR errors, which are usually attributed to three basic sources: the sensor, algorithm/phenomenology, or ancillary/auxiliary data. External QC flags can include many forms of measure that may be useful in isolating performance impacts. The SDR cal/val analysis may indicate certain anomalous conditions that could impact EDR performance that should be flagged (*e.g.* unplanned stray light at certain positions). External QC flags can also be used to track software or reprocessing versions, look-up tables (LUTs), or table uploads related to EDR processing. These flags can also indicate issues related to ancillary data quality.

The EDR matchup processor will ingest correlative matchups along with any QC flags associated with the matchups. The QC flags may also include time and distance offsets in the matchup, and spatiotemporal measures of matchup stability. The matchup processor will need to deal with issues of time, distance, and FOV mismatches. We assume that matchups will be in a form that converts the raw or indirect form into EDR data units appropriate for IP comparisons, since a number of users are interested in characterizing IP performance. This would allow matchup files to be used to support both IP and EDR characterization. The QC matchup processor will be responsible for handling issues related to the EDR grid size and how to convert point or line measurements into area measurements. We assume that matchup files would identify the corresponding IP pixel in a CrIMSS granule, which can then be used to identify the corresponding EDR grid cell. The QC matchup will copy over the entire EDR data structure (including all flags) as well as generating additional flags (*e.g.* the standard deviation of surrounding pixels).

All QC flags, whether input to or generated by the QC EDR matchup processor, will be stored in EDR and QC matchup storage. This will include all available matchups, irrespective of whether they are a subject to stratified, degraded, or excluded conditions. A separate query process can then be used to characterize performance under arbitrary conditions, *e.g.,* a specific stratification called for in the NSS. Queries can also be used to identify finer levels of stratification or to set thresholds for various QC flags to understand performance issues. Queries can also be used to assess performance based on individual correlative data sources to identify inconsistencies with various data sources.

When comparisons are made with correlative data, the error characteristics of the correlative data used to provide an estimate of truth must be considered. The estimate of correlative data error can be stored as a QC flag, and may be stored on a per-source or per-measurement basis, depending on what information is available. In many cases, comparisons with truth estimates will produce outliers that will be out of a family of typical differences. These outliers will be flagged so that a post-analysis process can determine if the outliers are due to the EDR performance or a particular truth measurement.

Approaches such as an Allan variance method [REF] can be used to find optimal QC screening. In an Allan variance approach bias and variance of the cross-comparisons may be examined as a function of some matchup criteria, *e.g.*, time offset. As the allowable time offset is increased, more matchups are created. This reduces noise, but this increase in time offset decreases correlations between the VIIRS measurement and the correlative measure. The resulting bias and variance can be examined in the matchups as a function of time offset to select the optimal balance. In some cases, the effects of time offsets can be very dependent on the temporal variability of the matchups, which in turn can depend local time due to diurnal effects. All of these effects can be tracked through the use of appropriate QC metrics.

Performance estimates will be produced by applying various queries to the matchup database. These queries will allow data subsetting according to any of the QC flags or data sources. The query will also allow creation of measurement bins so that results can be characterized across the required measurement range. An agreed-to set of QC conditions (*e.g.*, temporal mismatch) can be decided post-launch to provide

A1.8 Cal/Val Risks and Mitigation Activities

A1.8.1 Correlative Truth Data Collection

It is vitally important to have required correlative truth data collected, processed, and distributed to cal/val team member in a timely manner to allow them to carry out their planned tasks. Equally important is that the quality and quantity of correlative truth data be sufficient to allow tuning the CrIMSS EDR algorithm to optimize its performance and for eventual validation of EDR performance.

As a means of mitigating this risk, the cal/val team should define clear and verifiable requirements for correlative truth datasets as early as possible. We recommend that a giver-receiver type of schedule be developed to ensure that related tasks are being carried out as planned.

A1.8.2 Cal/Val Infrastructure Development and Data Management

Both correlative truth data and SDR and EDR products. Data storage, access, match up, quality control. Interface and communication between different systems and subsystems (IDPS, CLASS, GRAVITY, NSIPS). Educate cal/val team members to effectively use the system.

A1.8.3 Calval Tools Development and Validation

Tools need to be verified and validated before they are applied.

A1.8.4 Task Coordination

Identify lead for each task; IMS; G-R

A1.8.5 Resources

Secure adequate funding for team members to carry out the planned task

Appendix 2: Detailed Plan for CrIS and ATMS EDR Validation

As noted earlier, this plan describes a coordination strategy for validating the EDRs generated by the the Cross-track Infrared Sounder (CrIS) and the Advanced Technology Microwave Sounder (ATMS), together known as the Cross-track Infrared Microwave Sounding Suite (CrIMSS). While EDR validation for these instruments are most commonly addressed as the suite, there are circumstances that make it appropriate for the instruments to be considered separately. These circumstances will be called out explicitly in the discussion that follows.

A2.1 Validation of "First Light" Spectra

SDR and EDR validations are highly synergistic and cannot be done in isolation. While EDR validation tends to operate under the assumption that the SDRs are validated, this process is anything but sequential.

For example, with both the AIRS and IASI systems the "first light" was an early release of just a few granules (June 13, 2002 for AIRS; Jan. 15, 2007 for IASI). The earliest possible release of data to the cal/val teams must be accommodated at startup by NGAS and the IPO. In the cases of AIRS/AMSU and IASI/AMSU/MHS, the data streams were sporadic for several months after first light and, of course, instrument calibrations changed significantly during this time. However, significant progress was made during this period with the small handful of granules that were available. There was also constant communication between instrument engineers, the SDR team, and the EDR team. The key to success is thus to have open dialogue between discipline teams, cal/val teams and members, and the user community. In the NPP context, we plan to facilitate such communication by TIMs, SOAT meetings, and cal/val telecoms. Such formal approaches notwithstanding, we would like to note that early and informal pathways for communication were identified at the Aqua "End of Mission" review (December 2008) as one of the most critical reasons for that mission's success.

Several components of first-light spectra validation are discussed below.

A2.1.1 Compute Empirical Orthogonal Functions from SDRs

One of the first activities in CrIS and ATMS EDR validation is to compute EOFs from the SDRs and to begin to assess the stability and information content of the SDRs. This enables monitoring of the SDRs and building a statistical database (which is covered under Mitch Goldberg's IGS funding). We will also begin capturing SNO matchups with other satellites (which is covered under Changyong Cao's IGS funding), and begin to test forward models, cloud detection algorithms, NWP channel sub-setting, etc. This activity helps radiance users within the assimilation community and is funded as part of NOAA's PSDI CrIS/ATMS system planning.

At first light, the "easiest" scenes are typically analyzed first. The term "easiest" usually means less complicated scenes that have been identified as cloud-free, nighttime, over oceans, and within ± 50 degrees latitude; however, there are many exceptions. In a cloud-clearing algorithm scenes with high contrast among the infrared FOVs within a microwave footprint (nine FOVs for AIRS/AMSU and CrIS/ATMS; four for IASI/AMSU) might also be selected. Other important scenes might be scenes that match up with other observations, such as other satellite observations, operational radiosondes, and intensive scientific campaigns-of-opportunity. The initial priority at first light, then, is to acquire as many scenes as possible that to ensure that the SDRs are performing as expected and to use the EDR algorithms (*i.e.*, IDPS, NUCAPS, LaRC, STAR-offline) as early as possible to begin to diagnose the myriad issues that might arise.

A2.1.2 Reprocess Critical Datasets

At NOAA/NESDIS/STAR we have invested in the ability to employ reprocessing of these critical datasets with an "off-line" science version of our operational NUCAPS algorithm. We also plan to develop this capability with the NGAS code (*i.e.*, Xu Liu's, Mitch Goldberg's tasks) to generate CrIS/ATMS STAR-offline EDRs. The fact is that there are *always* post-launch issues with new instruments that will require reprocessing. Some of these issues might be as mundane as missing fields or misaligned entries in data files. Usually, we find algorithm implementation issues involving extreme circumstances (*e.g.*, desert emissivity, extremely cold or warm scenes, high altitude terrain, etc). We expect the NOAA AIRS/IASI-heritage approach to be "seamless" because in the NUCAPS system the only new elements are the CrIS and ATMS radiances. The radiative transfer model and algorithm are mature and fully equipped for ingesting CrIS/ATMS observations. Our experience with AIRS and IASI is that this system can be up and running within a few hours of first light (albeit with limited performance).

Once the NUCAPS system is running on SDRs it opens a large variety of diagnostic tools to help study and isolate issues and rapidly improve the results. The operational version of the NUCAPS EDR algorithm is literally a filtered version of a single-science code (STAR-offline EDR) that is capable of running all instrument systems. This is an important concept: At NOAA, we literally have one "plug-and-play" code for AIRS/AMSU/HSB, IASI/AMSU/MHS, and CrIS/ATMS. The "filter" is an automated process that strips out diagnostic code and code that is not relevant to CrIS/ATMS. Therefore, rapid redeliveries can be made to our operational system in the early time frames to address any issues that might arise with CrIS/ATMS. Since *in situ* or selected datasets are sparse and valuable, it does not make sense only to use them once. If a problem is detected or if algorithm improvements are made, the ability to rerun the algorithm on exactly the same ensemble is critical to evaluate performance and validate algorithm changes. Our system employs complete reprocessing capability of *all* datasets. For both AIRS/AMSU and IASI/AMSU/MHS, selected datasets were reprocessed hundreds of times before the NUCAPS system became operational.

It is not clear that the IDPS and the NGAS cal/val group will have a reprocessing capability. Therefore, our strategy is to use AIRS/AMSU and IASI/AMSU/MHS offline-EDR products to help isolate problems with the NGAS IDPS-EDR products. In addition, once we have the NUCAPS system up and running we can intercompare NUCAPS-EDRs and IDPS-EDRs directly, thereby isolating many SDR and EDR issues. Offline, we will have the capability to reprocess cal/val datasets with NUCAPS and—hopefully—the NGAS code. Some of the PIs at other centers (*e.g*., Xu Liu and Daniel Zhou at LaRC) have developed off-line EDR capabilities from both aircraft and satellite sensors. These aircraft sensor algorithms can be employed to help identify problems with the operational EDR code as described above.

A2.1.3 Check SDR and EDR Reasonableness

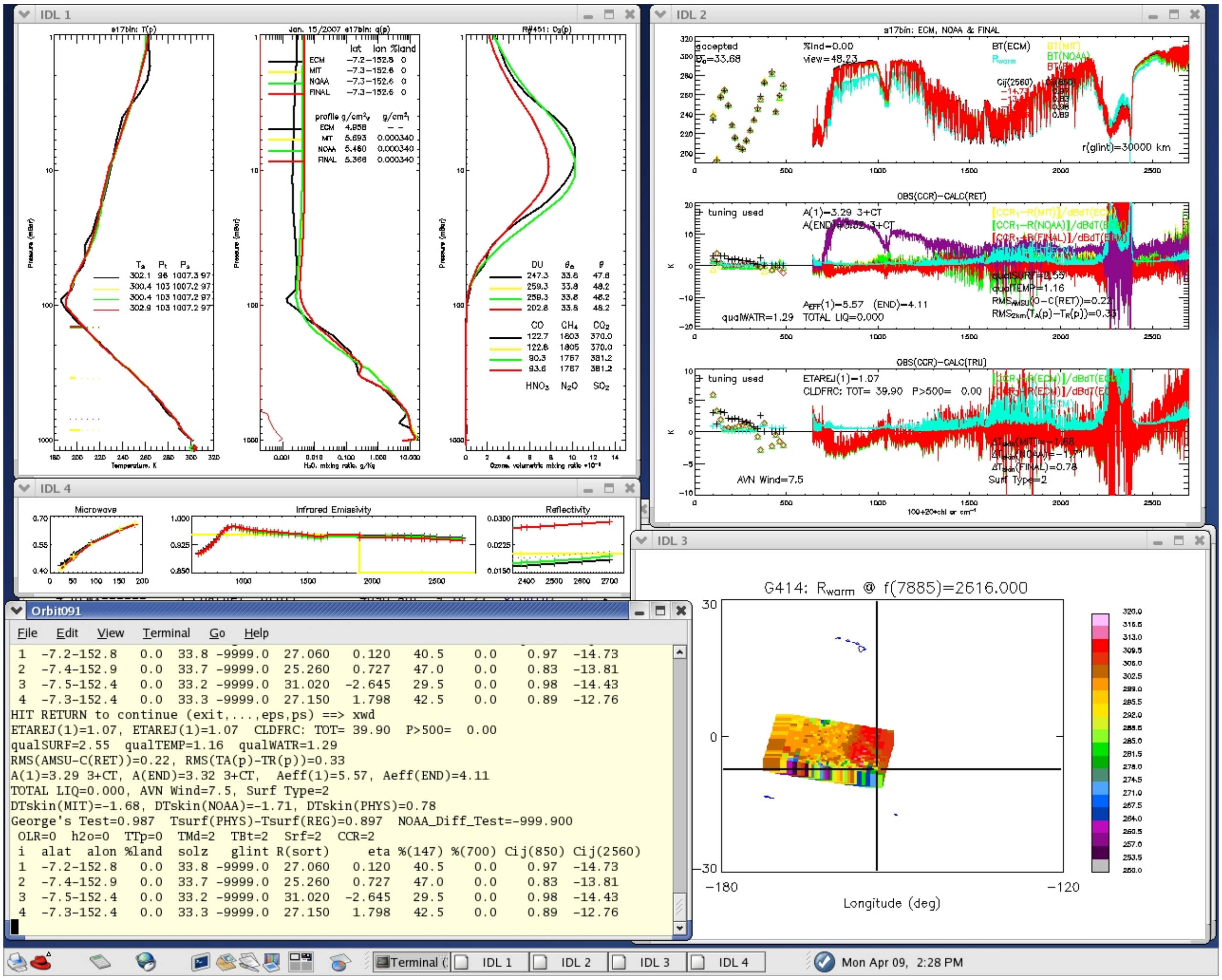
Beginning at first light, forecast models can be used to check for reasonableness in the SDR and EDR products. Radiances can be computed from the United Kingdom Meteorology Ofiice (UKMet), ECMWF, and the NCEP GFS analysis or forecast state. This will be done by a number of cal/val team members. Larrabee Strow will compare CrIS radiances to ECMWF cloud-free night scenes with an eye on instrument performance. NWP centers will be monitoring background-removed observed intensities (obs-background) for non-precipitating ATMS and cloud-free CrIS channels, and to provide analysis of global biases. NOAA/NESDIS/STAR will compare cloud-free and cloud-cleared radiances with respect to forward computations from ECMWF, NCEP GFS, and radiosonde atmospheric states to quantify EDR performance.

A2.1.4 Compute Empirical Bias Correction

One of the first extremely critical activities at first light will be to compute the empirical bias correction for the ATMS and CrIS instrument radiances—a collaborative effort between NOAA/STAR, MIT, and NGAS. For ATMS, this will most likely be a function of view angle due to side-lobe contamination. The microwave bias correction can be initially computed using ECMWF or NCEP information since it takes a number of years before enough statistics can be acquired at each view angle using *in situ* datasets. For CrIS, we expect the bias correction to be very small, but we have found that knowledge of ozone and carbon dioxide are critical components of the 15-μm band radiance computation, methane is a critical component of the 6.6-μm band computation, and nitrous oxide, carbon monoxide, and carbon dioxide are critical for the 4-μm region. ECMWF ozone has not been of high enough quality for these early assessments, and methane and carbon dioxide are not currently variables within NWP models. Initially, this bias correction will be computed using ECMWF or NCEP analyses along with other measurements (*e.g.,* AIRS, IASI); however, the final bias correction requires the ability to reprocess scenes using the retrieval algorithm in an iterative procedure in which a self-consistent answer is achieved. It is for this reason that a large number of high-quality radiosonde matchups from multiple sites is desired early in the mission, and a collaborative relationship with the GRUAN sites would certainly help in acquiring these radiosondes in a timely fashion.

A2.1.5 Implement Performance Diagnostics

Finally, a diagnostic system must be employed that allows algorithm developers to visualize how the algorithm is performing. At NOAA, this is a system that simultaneously views SDRs, EDRs, IPs such as cloud-cleared radiances and trace gas products or geophysical state assumptions (such as surface pressure, viewing angles, etc.). An example of our diagnostic visualization system for AIRS/IASI is shown in Fig. **[TBD]**, below. The desire is to show everything at once, so that a cal/val investigator can diagnose what is working and what is not. In the case shown, the ECMWF, microwave-only, regression, and physical algorithm retrieved states are shown together in the upper-left windows (“IDL1” and “IDL4”: temperature, moisture, and ozone profiles, trace gas column amounts, skin temperature, infrared and microwave spectral emissivity, cloud amount and fractions). In the upper right window, the observed (warmest FOV and cloud-cleared) and computed radiances from ECMWF and retrieval states are shown in the top plot. Observed (cloud-cleared) radiances minus retrieval states are shown for a number of states in the middle plot. This confirms that the algorithm minimized the differences. Finally, in the bottom plot, observed minus “truth” is shown. Here the ECMWF analysis is taken as a proxy for the “truth”; it isolates spectral regions that were problematic. This system can also use other *in situ* data, such as from radiosondes, as truth. In all the radiance plots the cyan color represents the instrument noise at this scene temperature, so “red” curves within the “cyan” envelope are performing reasonably well. In the lower-left window a number of quality parameters are tabulated; in the lower-right window the location of the current profile is shown within the IASI granule. Individual PIs have developed many tools such as this one, which will be used to diagnose problems that may arise in the early months of validation. Case studies will be shared within the CrIS/ATMS cal/val team so that we can identify instrument and algorithm issues.



```
 1  -7.2-152.8   0.0  33.8 -9999.0  27.060   0.120   40.5    0.0     0.97  -14.73
 2  -7.4-152.9   0.0  33.7 -9999.0  25.260   0.727   47.0    0.0     0.83  -13.81
 3  -7.5-152.4   0.0  33.2 -9999.0  31.020  -2.645   29.5    0.0     0.98  -14.43
 4  -7.3-152.4   0.0  33.3 -9999.0  27.150   1.798   42.5    0.0     0.89  -12.76
HIT RETURN to continue (exit,...,eps,ps) ==> xwd
ETAREJ(1)=1.07, ETAREJ(1)=1.07  CLDFRC: TOT= 39.90  P>500=  0.00
qualSURF=2.55  qualTEMP=1.16  qualWATR=1.29
RMS(AMSU-C(RET))=0.22, RMS(TA(p)-TR(p))=0.33
A(1)=3.29 3+CT, A(END)=3.32 3+CT,  Aeff(1)=5.57, Aeff(END)=4.11
TOTAL LIQ=0.000, AVN Wind=7.5, Surf Type=2
DTskin(MIT)=-1.68, DTskin(NOAA)=-1.71, DTskin(PHYS)=0.78
George's Test=0.987  Tsurf(PHYS)-Tsurf(REG)=0.897  NOAA_Diff_Test=-999.900
 OLR=0  h2o=0  TTp=0  TMd=2  TBt=2  Srf=2  CCR=2
 i  alat  alon %land  solz   glint R(sort)     eta %(147) %(700) Cij(850) Cij(2560)
 1  -7.2-152.8   0.0  33.8 -9999.0  27.060   0.120   40.5    0.0     0.97  -14.73
 2  -7.4-152.9   0.0  33.7 -9999.0  25.260   0.727   47.0    0.0     0.83  -13.81
 3  -7.5-152.4   0.0  33.2 -9999.0  31.020  -2.645   29.5    0.0     0.98  -14.43
 4  -7.3-152.4   0.0  33.3 -9999.0  27.150   1.798   42.5    0.0     0.89  -12.76
```

Figure 3 - An example of a diagnostic tool used to view first light spectra from AIRS and IASI (see text).

A2.2 Validation of Key Performance Parameters using Operational RAOBs

For AIRS/AMSU, a multi-year validation of AVTP and AVMP EDRs based upon global RAOB collocations was published in the special *JGR* issue on AIRS product validation (Divakarla *et al.* 2006). In that paper we compared data from over 80,000 radiosondes (mostly land-based) within ±3 hr and ±100 km from the Aqua-AIRS retrievals, forecast data from the NCEP GFS and ECMWF, and operational retrievals from the NOAA-16 Advanced TIROS Operational Vertical Sounder (ATOVS) to verify that AIRS EDRs were performing well. Using radiosondes as the reference, bias and RMS differences were computed for a variety of conditions (*e.g*, land/sea/coast, day/night, polar/mid-latitude/tropical, all seasons, and cloud-cleared products *vs.* cloud-free products) for AIRS and other collocated data sets. The matchup system was expanded to include the

World Ozone and Ultraviolet Radiation Data Centre (WOUDC) ozonesondes and Brewer-Dobson ground station measurements to validate the AIRS ozone profile and total ozone column products (Divakarla *et al.* 2008). We have duplicated and applied this methodology for IASI and analyzed data from tens of thousands of radiosondes prior to declaring the product operational in July 2008. The matchup system developed for AIRS and IASI and the associated infrastructure to produce other correlative data sets (*i.e.*, ECMWF, NCEP GFS, and ATOVS retrievals) can be adapted via NUCAPS and the NGAS off-line code to validate CrIS/ATMS AVTP and AVMP EDRs, along with the ozone IP using global radiosondes, WOUDC ozonesondes, and Brewer-Dobson total ozone measurements.

The AIRS results with global operational RAOBs are similar to those obtained from high-quality radiosondes launched during Aqua overpass times at the ACRF sites (Tobin *et al.* 2006a); however, the radiosonde network is a global network that enables validation of the AIRS/AMSU products in many geophysical conditions and over long times. For example, we found that the temperature bias is highly influenced by a larger bias contribution from “land” samples, and shows a month-to-month and annual variation that correlates with CO_2 variations. Therefore, operational radiosondes data complement dedicated high-quality radiosonde data acquired at ACRF sites.

As before, reprocessing is a key element of our validation approach. For AIRS and IASI, we continue to reprocess the earliest focus day(s), and all of the dedicated and operational radiosondes are periodically reprocessed to analyze long-term biases. Of particular note is that we can reprocess in experimental or operational configurations (*e.g.*, Barnet *et al.* 2005). The ability to easily switch components of the algorithm on/off or the ability to run without certain modules present is critical to bringing up a system rapidly. Post-launch start-up is very difficult if all components of the system must be present to run. For example, at first light the regression coefficients are based on prelaunch simulations and may need to be ignored, as was the case at the AIRS first light. The science code was up and running in a matter of hours, while the operational system took over six months before it was running. In addition, data files that were necessary for operation (*i.e.*, radiance tuning, error covariance, regression, etc.) were generated by the first light science code.

For NPP CrIS/ATMS we will utilize this validation system and intercompare retrieval products from the IDPS with all available datasets that were listed in Section 2.3, as well as a heavy reliance on the NUCAPS system to demonstrate that the IDPS algorithm is functioning reasonably.

A2.3 Characterization of All EDR Products and Long-term Validation

Some of this activity will utilize scientific campaigns-of-opportunity that are, for the most part, of small cost to IPO; however, these will typically be driven by external schedules and will not be specific to the needs of EDR validation. It would be desirable to have a periodic set of IPO-coordinated intensive campaigns with aircraft support to independently characterize the long-term behavior of the CrIS/ATMS IDPS-EDRs. Aircraft campaigns with instruments such as NAST-M, NAST-I, and S-HIS are the only NIST-traceable sensors that can directly characterize the CrIS/ATMS SDRs in detail. While this may be construed as an SDR activity, the combination of aircraft radiance observations and retrievals in the CrIS/ATMS subpixel space, along with other correlative measurements, such as dedicated radiosonde launches, moisture LIDARs, surface characterization, etc., are the only traceable way to measure the long-term stability of the CrIS/ATMS instrument and EDR algorithms. Therefore, we strongly recommend that dedicated radiosondes and aircraft campaigns be supported in the long term. This might be accomplished by adding NAST-like sensors to scientific campaigns –of-opportunity.

Long-term monitoring and evaluation of AIRS/AMSU EDRs has also been accomplished using the operational RAOBs described in section A2.2. At present, we have over 375,000 collocated operational RAOBs in the six-year AIRS dataset; we continually reprocess these datasets looking for ways to characterize and improve the algorithm. For AIRS, we have noticed a small trend in the temperature product that can be analyzed within this reprocessing environment. Here we can test the impact of using different *a priori* states, sensitivity to trace gases (*e.g.,* carbon dioxide impact on AVTP, and methane, nitric acid impact on AVMP), etc.

Involvement in intensive campaigns allows us to work with the user community to characterize and demonstrate our product performance within a focused scientific context. Among other things, we contribute our near-real-time NUCAPS-EDR product fields, and in return gain complete access to the *in situ* measurements. Everyone benefits. For example, our participation in the Stratospheric and Troposphere Analysis of Regional Transport (START) enabled us to study the AIRS products in stratospheric/tropospheric exchange regions (STE; Pan *et al*. 2007). While the scientific focus of this campaign was ozone, participation in this experiment has led to a better understanding on the impact of ozone biases on our cloud-cleared radiances and, therefore, on our temperature and moisture products. Many of the upgrades to the AIRS temperature and, of course, ozone products were developed by participating in this campaign. It was also interesting to validate temperature profiles in situations with double tropopause and atmospheric folds. Correlations of trace gas products can also be used to select interesting dynamical regimes that can then be studied in detail to ensure that temperature and moisture profiles are behaving as expected.

Another campaign that has proven to be useful is the Water Vapor Validation Experiment (WAVES) at the Howard University Reference Upper Air Network site. We have acquired many dedicated radiosondes and ozonesondes as well as LIDAR measurements of moisture in this urban-polluted setting. Again, it is the scientific interaction—in this case, an air quality focus—which helped us to quantify our skill in a specific situation with a large volume of ground truth. Feedback from science users to algorithm developers enables us to characterize and improve our products.

NOAA/STAR has also made an investment in acquiring radiosondes in a region that is sparsely covered. The AEROSE activity—to-date comprised of five, four-week trans-Atlantic cruise legs held in 2004, 2006, 2007 and 2008)—also in collaboration with Howard University's NOAA Center for Atmospheric Sciences (HU/NCAS), have provided a set of *in situ* measurements to characterize the impacts and microphysical evolution of continental African aerosol outflows, including both Saharan dust and sub-Saharan smoke from biomass burning, across the Atlantic Ocean (Nalli *et al.* 2006; Morris *et al.* 2006). This activity is a "piggyback" on other NOAA missions with the NOAA Ship *Ronald H. Brown*, and has contributed unique validation datasets for both AIRS/AMSU and IASI/AMSU/MHS (Nalli *et al.* 2006, 2008) at an extremely low cost—that is, the cost of radiosondes, ozonesondes, and approximately six person-months per cruise for sonde launch support onboard and post-experiment analysis. We have another opportunity in summer of 2009 with IASI and we plan, if funding allows, to continue growing this dataset with future cruises. These campaigns also provide the ground truth radiosondes for assessing the ability of hyperspectral sounders for observing the Saharan air layer (SAL), a dry ,warm air layer hypothesized to suppress tropical storm formation in the Atlantic (see, *e.g.,* Dunion and Velden 2004; Nalli *et al.* 2005)

Participating in these intensive campaigns takes time to nurture. We propose to continue participating in these activities prior to the launch of NPP (*i.e.*, validation of IASI and AIRS EDR products) to construct CrIS/ATMS proxy datasets as a risk reduction measure for the CrIS/ATMS, similar to what has already been achieved for GOES-R. In this way, we can develop methodology, promote coordination and collaboration, develop web page interfaces for the *in situ* and satellite products, etc. In this way, once NPP is launched, the switch-over from IASI validation to NPP validation would be completely transparent.

Appendix 3: Acronyms and Abbreviations

ACRF	ARM Climate Research Facility
AER	Atmospheric and Environmental Research
AEROSE	Aerosol and Ocean Science Expeditions
AIRS	Atmospheric Infrared Sounder
AMSU	Advanced Microwave Sounding Unit
APU	Accuracy, Precision, and Uncertainty
ARM	Atmospheric Radiation Measurement
ATBD	Algorithm Theoretical Basis Document
ATMP	Atmospheric Vertical Temperature Profile
ATMS	Advanced Technology Microwave Scanner
ATOVS	Advanced Tiros Operational Vertical Sounder
AVMP	Atmopsheric Vertical Moisture Profile
AWG	Algorithm Working Group
Cal/Val	Calibration/Validation
CEOS	Committee on Earth Observation Satellites
CERES	Clouds and the Earth's Radiant Energy System
CrIMSS	Cross-track Infrared Microwave Scanning Suite
CrIS	Cross-track Infrared Scanner
DC	Degradation Condition
DMSP	Defense Meteorological Satellite Program
DoD	Department of Defense
EC	Exclusion Condition
ECMWF	European Centre for Medium-range Weather Forecasting
EDR	Environmental Data Record
EDRPR	EDR Performance Requirement
EOC	Early-orbit Checkout
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
EQUATE	
ESA	European Space Agency
ESPC	Environmental Satellite Processing Center
ESRL/GMD	Earth System Resource Laboratory/Global Monitoring Division
EUMETSAT	European Meteorological Satellite
EVM	Earned-value Management
FOR	Field-of-Regard
FOV	Field-of-View
FTE	Full-time Equivalent
FTIR	Fourier-transform Infrared spectroscopy
GB	Gigabytes
GCOS	Global Climate Observing System
GFC	Global Forecast System
GNSS	Global Navigation Satellite System
GOES	Geostationary Operational Environmental Satellite
GRAVITE	Government Resource for Algorithm Verification, Independent Testing, and Evaluation

GRUAN	GCOS Reference Upper Air Network
GSFC	Goddard Space Flight Center
GTCD	Ground Truth/Correlative Data
HCS	Horizontal Cell Size
HIAPER	High-performance Instrumented Airborne Platform
HIPPO	Hiaper Pole-to-Pole Observations
HIRS	High-resolution Infrared Radiation Sounder
HUBC	Howard University – Beltsville Campus
IASI	Infrared Atmospheric Sounding Interferometer
ICT	Internal Calibration Target
ICT	Internal Calibration Target
ICV	Intensive Calibration/Validation
IDPS	Interface Data Processing Segment
IOP	Intensive Period of Operations
IORD	Integrated Operational Requirements Document
IP	Intermediate Product
IPO	Integrated Program Office
IR	Infrared
JAIVEx	Joint Airborne IASI Validation Experiment
JCSDA	Joint Center for Satellite Data Assimilation
JPL	Jet Propulsion Laboratory
KPP	Key Performance Parameter
LAA	Local Angle Adjustment
LaRC	Langley Research Center
LBLRTM	Line-by-Line Radiative Transfer Model
LIDAR	Light Detection and Ranging
LTM	Long-term Monitoring
LTS	Long-term Stability
MHS	Microwave Humidity Sounder
MW	Microwave
NASA	National Aeronautics and Space Administration
NAST-I	NPOESS Atmospheric Sounder Testbed -Interferometer
NAST-M	NPOESS Atmospheric Sounder Testbed – Microwave
NCAR	National Center for Atmospheric Research
NCDC	National Climate Data Center
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite, Data, and Information Service
NGAS	Northrop Grumman Aerospace Systems
NIST	National Institute of Science and Technology
NLTE	Nonlocal Thermodynamic Equilibrium
NOAA	National Oceanic and Atmospheric Administration
NPOESS	National Polar-orbiting Operational Environmental Satellite System
NPP	NPOESS Preparatory Project
NRL	Naval Research Laboratory
NSA	North Slope of Alaska
NUCAPS	NOAA Unique CrIS/ATMS Processing System
NWP	Numerical Weather Prediction
OLR	Outgoing Longwave Radiation
OMPS	Ozone Mapping and Profiler Suite

OSS	Optimal Spectral Sampling
OSSF	Ozone Spectral Signature Flag
PCT	Probability of Correct Typing
PEATE	Product and Evaluation Test Element
POES	Polar-orbiting Environmental Satellites
PP	Pressure Profile
ppmv	Parts Per Million by Volume
PSDI	Product System Development and Implementation
QC	Quality Control
RAOB	Radiosonde Observations
RTM	Radiative Transfer Model
SC	Stratification Condition
SDL	Space Development Laboratory
SDR	Sensor Data Record
SGP	Southern Great Plains
S-HIS	Scanning High-resolution Interferometer Sounder
SNO	Simultaneous Nadir Overpass
SOAT	Sounder Operational Algorithm Team
SSMIS	Special Sounder Microwave Imager/Sounder
STAR	Center for Satellite Applications and Research
START	Stratospheric Troposphere Analysis of Regional Transport
STE	Stratospheric/Tropospheric Exchange
SWIR	Shortwave IR
TBD	To Be Determined
TBR	To Be Reviewed
TIM	Technical Interchange Meeting
TWP	Tropical Western Pacific
UKMET	United Kingdom Meteorological Office
USU	Utah State University
VCS	Vertical Cell Size
VIIRS	Visible/Infrared Imaging Radiometer Suite
WAVES	Water Vapor Validation Experiments
WMO	World Meteorological Organization
WOUDC	World Ozone and Ultraviolet Radiation Data Centre