Hyperspectral Infrared Measurements Simulation and Collection for GOES-R Sounder Lossless and Lossy Data Compression Study and Beyond

HUNG-LUNG HUANG, BORMIN HUANG, KEVIN BAGGETT, ALOK AHUJA
Cooperative Institute for Meteorological Satellite Studies, CIMSS
University of Wisconsin-Madison
TIMOTHY J. SCHMIT and ROGER W. HEYMANN
NOAA, National Environmental Satellite, Data, and Information Service, NESDIS

CIMSS, UW-Madison
1225 West Dayton Street, Madison, Wisconsin 53706
UNITED STATES OF AMERICA

Abstract: Geosynchrous hyperspectral infrared remote sensing measurements are to be collected by the future satellite environmental sensors to make real-time ecological monitoring, daily weather forecasting and long-term climate study, just to name a few. Due to its measurement characteristic, a few order-of-magnitude increase of data volume is anticipated.
For example, Geostationary Operational Environmental Satellite (GOES)-R sounding instrument called Hyperspectral Environmental Suite (HES) is under designed to be the NOAA's next-generation operational sounding sensor capable of making infrared measurements over wide spectrum (~3.5 to 15 microns) at spectral resolution better than ($\nu/\Delta\nu \geq \sim 1000$ or $\Delta\nu \leq \sim 1$ cm^{-1}), and also improved signal to noise ratio (NEdT≤~0.25K), field of view spatial sampling and size (<10 km), and most of all, fast (every half hour to hours) and broad area coverage (sub-continental to continental US scale).
Lossless and lossy data compressions are the only ways to effective reduce data volume before satellite downlink and data re-distribution at ground sites. This paper reviews a government-university team's status of current approach and future plan of providing a comprehensive hyperspectral infrared database to support government-lead data compression study. Beyond that, it is intended to elevate its effort to back industry's design and build of HES on time for GOES-R operations as well.

Key-Words: Hyperspectral, geosynchrous, data compression, lossless, lossy.

1 Introduction

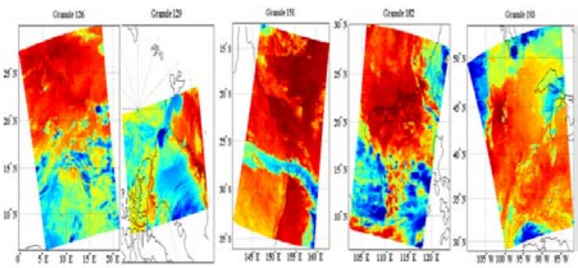
New instruments such as the Atmospheric Infrared Sounder (AIRS), Interferometric Monitor for Greenhouse Gases (IMG), Geosychronous Imaging Fourier Transform Spectrometer (GIFTS), Cross-track Infrared Sounder (CrIS), Infrared Atmospheric Sounding Interferometer (IASI), the Hyperspectral Environmental Suite (HES) on GOES-R, and others provide researchers with the potential to make significant advancements in hyperspectral data distribution, processing and science applications. In this paper, we will overview the current hyperspectral end-to-end system infrastructure development in University of Wisconsin-Madison. Emphasis is on the end-to-end system component of the Hyperspectral Sounder Simulator and Processor (HSSP), under development by

CIMSS UW-Madison [1], which is directly geared to support the government-lead data compression for on-board and on-ground data compression algorithm development as well as operational implementation consideration. Section 2 describes the functional components of HSSP, an end-to-end system that provides interfaces and links between each individual component for specific hyperspectral study. Due to page limitation, although the detail of HSSP won't be discussed in this paper it can be found in the companion paper – *"Global Effective Hyperspectral Cloud Property Atlas for Hyperspectral Sounder Simulator and Processor"* by Li Guan and her co-authors. In Section 3, we review the high level approach of measurement simulation generation to support all weather, i.e. clear and cloudy, measurements. Section 4 is discussing all possible available space and airborne hyperspectral observations that can be utilized for conducting data compression performance evaluation at different level of data, i.e. raw data – level 0, calibrated data – level 1, or other intermediate domains such as interferogram, radiance or brightness temperature. In the last section, section 5, we discuss the overall objective of this effort to support data compression project and for the subsequent uses in the HSSP end-to-end support to the GOES-R HES program.

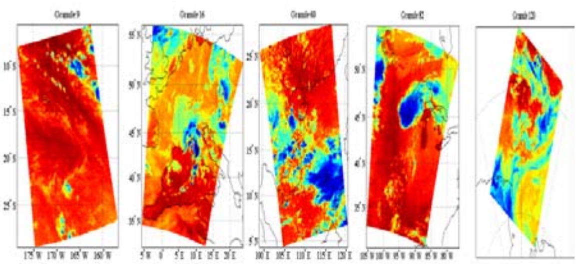
2 Brief HSSP Overview – Data Compression Component

University of Wisconsin's Hyperspectral Sounder Simulator and Processor (HSSP) [1] includes NWP modeling, radiative transfer modeling [2], instrument effect simulation, compression data set generation and impact analysis, sensor design trade study, calibration implementation, atmospheric retrieval algorithm development, water vapor tracking for wind derivation, and conduct simulation verification and validation.

Data compression component in the HSSP is a special element of its own right. Where data downlink, rebroadcast and archive for both short and long term that are related to data compression algorithms and implementation approaches all need to be studied. They are no optimal compression algorithms for infrared hyperspectral sounder data and no rigorous studies conducted to demonstrate compression impacts on retrieval products. However, several optimize algorithms are developed for compress hyperspectral infrared sounder data their compression impacts on data themselves and derived products have just began [3,4]. So far only granting type of data in calibrated radiance domain is tested [5,6]. The raw level 0 senor digital counts compression is also been tested [5]. The needs of simulating geosynchrous class data that closely emulating HES measurements are approaching. The uses of many other types of hyperspectral sounding measurements such as observations from airborne sensors are of great desire too. Preliminary estimates of compression ratio for practical data downlink and rebroadcast under GOES-R bandwidth allocation requires the uses of both lossless (~2x compression ratio) and lossy (~6x) in order to meet the operational demands. Compression studies using real or simulated measurements in different domains are necessary. Their lossy compression impact on global observations (i.e. including extreme measurements and outliers) themselves and their related impacted on the derived products will also need to be rigorously evaluated. Fig. 1 is example hyperspectral infrared standard dataset used to evaluate compression performance of different compression approaches and algorithms.



Five of the sample AIRS granules (IR window)



Five of the sample AIRS granules (IR window)

Fig. 1. Example standard dataset used by government-lead GOES-R compression team.

3 Hyperspectral Measurement Simulations

The following simulated measurements classified according to simulation modeling are outlined. All these components will need to be modeled as accurate as possible. As part of HSSP efforts, some of these modeling components are developed except improved and realistic handling of land surface property, multiple layer of clouds, cloud micro-property and high temporal and high spatial cloud resolving atmosphere need to be developed as well.

- **Top of Atmosphere Radiance (TAR)**
  - Clear Sky
    - Gas Component
    - Surface Component
      - Land types
      - Ocean/Water body
  - Cloudy Sky
    - Clouds
      - Ice
      - Water
      - Single vs. Multiple Layer

To simulate accurate sensor data the following components of instrument effects need to included at minimum:

- **Sensor Data**
  - Sensor Signal
    - Off-axis
    - Self-apodization
    - Non-linearity
  - Sensor Noise
  - Sensor Calibration
    - Spectral
    - Radiometric

The NWP model data that emulates the high temporal and high spatial resolution of geosynchrous sounder measurements can provide the needed input for the simulation of top of atmosphere radiances (TARs). Current meso-scale community model MM5 and in the future, Weather Research and Forecasting (WRF) models that provides the following atmospheric state are also listed below:

- **NWP Model Data (MM5/WRF)**
  - Model Atmosphere State
    - Profiles: Temperature; Water Vapor; Wind
    - Surface parameters: Air and Skin Temperature; Emissivity
    - Clouds: Types; Altitude; Phase; Optical Property

In short, to simulate GOES-R hyperspectral sounder measurements realistically at proper temporal, spatial, spectral resolution it needs to start with the generation of thermal dynamic consistent profiles of temperature, water vapor, clouds, winds and surface property. With these multiple variables as inputs to the comprehend radiative transfer model to simulate clear gases absorption, surface emission/reflection, clouds absorptions and scattering at filed of view of a sensor to be simulated. After top of atmosphere radiances at each sensor filed of view are created, it needs to be further transformed into sensor measurement space to include all instrument effects to emulate raw sensor output counts for variety of design trade-off, data processing, and sensor data compression studies. The UW HSSP is designed to possess most of the capability mentioned above to support the ongoing HES broad scope of activities including the most recent government funded industry formulation phase study, ground processing system design and prototyping, data and products processing algorithms research and development, and highlighted in this paper, the simulation for data compression study. Fig. 1 to 4 are series of figures demonstrated current hyperspectral simulations capability where complicated gases spectroscopy, cloud micro-physical property and fast parameterized model are developed to realistically emulate current and future space-borne infrared measurements. Fig. 1 are simulated spectra of clear and cloudy sky (with different cloud optical thicknesses, size distributions and habits) for nominal tropical, polar and mid-latitude atmospheric conditions.

Fig.2

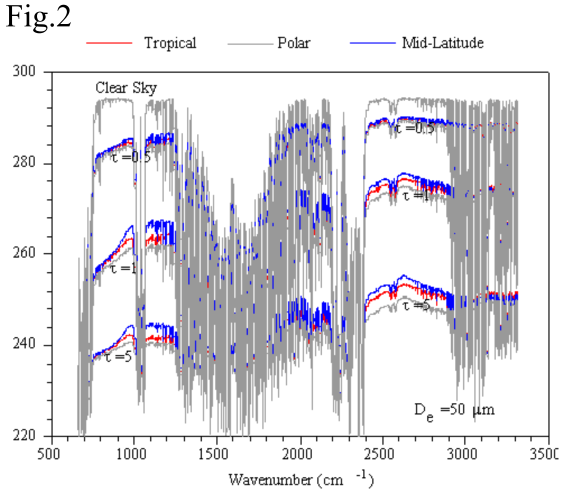


Fig. 2 Calculated clear and cloudy sky spectra representing tropical, polar and mid-latitude profile conditions.

Fig. 3 compares three spectral computed for clear (in red), mid-level cloud (in cyan), and very high cold ice cloud (in blue) conditions. Thousand of spectral images like this are simulated every 30 minutes for very large special domain (few thousands km by few thousands km) for geostationary or geosynchrous imaging sounders' measurement to be provided by the next generation of environmental satellites [1].

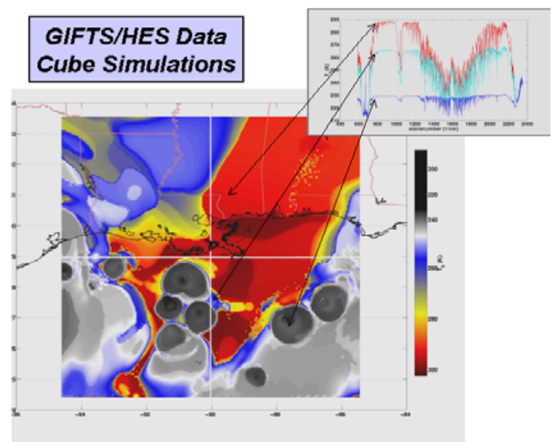


Fig. 3 simulated clear and cloudy sky spectra representing near by clear, mid-level and high-level cloudy conditions.

Fig. 4 outlines large database is designed to carried out the cloud absorption and scattering modeling to account for the proper top of atmosphere radiance measurements due to the cloud effects.

Cloud Database - albedo & tansmissivity function

For ice clouds:

✓ **Optical thickness: 0.04-50**

✓ **Effective size: 10-157µm**

✓ **Effective Shape:** Aggregates, solid hexagonal columns, Spheres, Bullet-rosettes, Droxtals, Hollow columns, Plates, and Spheroids (8)

✓ **Zenith angle: (0-80°)**

✓ **Wavenumber:(500-2500cm^{-1})**

For water clouds:

✓ **Optical thickness: 0.06-150**

✓ **Effective size: 2-20 µm**

✓ **Zenith angle: (0-80°)**

✓ **Wavenumber:(500-2500cm^{-1})**

Fig. 4. Database required to model cloud micro-physical property for proper cloudy radiances calculation.

Fig. 5 shows the cloud habit (i.e. size and, shape of cloud particle) effects on the simulated hyperspectral infrared measurements.

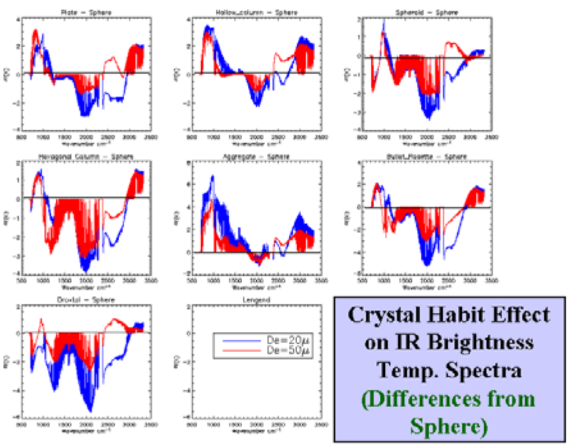


Fig.5 is the cloud habits effects on hyperspectral radiance measurements.

4 Space and Air -borne Measurements Collections and Utilization

In this section, we are creating an atlas of the real measurements required to supplement to data compression needs, where due to the limitation of simulated data, both airborne and space borne hyperspectral measurements are needed to perform all necessary data compression and products impact analysis. The following is our current understanding of available datasets that are useful for the use of data compression study:

- **AIRS data**
  - Level 1-B (calibrated and navigated SDR) Radiances
  - Cloud Cleared Level 2 (Cloud Cleared) Radiances
  - Level 2 (Sounding Retrieval) Profiles
    - Temperature, Water Vapor
    - Cloud Height; Amount; Emissivity
    - Surface Parameters: Air and Skin Temperature; Emissivity
- **NAST-I and S-HIS Data**
  - Level 1-B (calibrated and navigated SDR) Radiances
  - Level 2 (Sounding Retrieval) Profiles
    - Temperature, Water Vapor
    - Cloud Height; Amount; Emissivity
  - Surface Parameters: Air and Skin Temperature; Emissivity
- **In-situ Data**
  - Co-located clear Sounding
    - Temperature, Water Vapor Profile
  - Co-located NWP Model Analysis
    - NCEP or ECMWF Model Analysis
  - Field Experiment Special measurements
    - Soundings
    - Cloud Parameters
    - Surface Parameters

So far limited AIRS data has been selected at raw (sensor count) and calibrated radiance domain to demonstrate data compression performance and processing requirement. NAST-I and S-HIS data that have different but unique hyperspectral characteristic can also supplement AIRS data for compression interferometric class of measurements. Other data types such as time, space, and scale co-registered NWP model and in-situ data can again augment compression impact verification and validate processing concept and performance in representative setting where not only data themselves but also their derived products need to be assessed. Fig. 6 displays example spectra of global space-borne hyperspectral infrared measurements collected by AIRS.

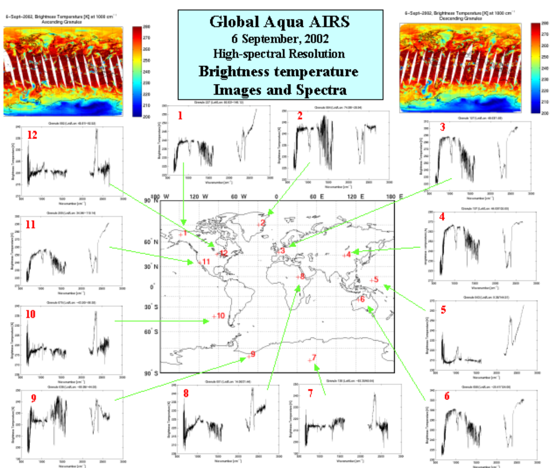


Fig. 6 global AIRS spectra depicting complicated infrared measurements representing diversified atmospheric and surface phenomena.

Fig. 7 is an example of co-located airborne – NAST-I and space-borne hyperspectral infrared spectra making a high quality and consistent absolute measurements. Although at different measurement spectral resolution they agree with each other with high accuracy.

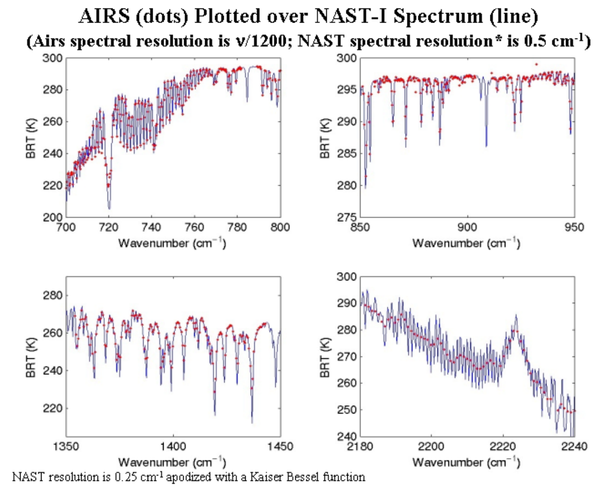


Fig 7. Example spectra collected from aircraft and satellite separately but showing excellent measurement agreement and accuracy.

5 Data Analysis and Visualization Tools

Data analysis and visualization is an essential part of the integrated system of HSSP. The flowing tools will also be developed to aid the compression study:

- I/O Reader and Writer
- IR Fast Radiative Transfer Model (RTM)
    - Clear and/or Cloudy
    - GIFTS/HES
    - AIRS
    - NAST-I
    - S-HIS
- Data/Products Display
- Noise Estimation and Filtering
- SDR (Level 1) to EDR (Level 2) Converter
- Documents
- Users' Support POC

As a truly end-to-end system emphasis not only modeling and processing capabilities, the so-called "last mile" can be bridged by the availability of data analysis and visualization tools. For example IR fast RTM will be available for HSSP users to develop their own unique hyperspectral applications. Noise estimate and filtering software can help engineers and scientists to enter the insight of single to noise ratio of a sensor system and optimize their unequalled engineering sensor design or scientific utilities.

6 Discussions and Summary

Up to date, a government lead GOES-R hyperspectral data compression team that partners with industry and universities is making some significant progress and can be briefly summarized as:

- Several optimized algorithms for hyperspectral IR sounder data are developed.
- Begins of studies demonstrating compression impacts on products.
- Generated and have shared limited standard data compression datasets, among team members, for comparisons of various methods.
- Gain experience and continue improving principle component compression work.
- Preliminary estimates for 2X lossless compression ratio on hyperspectral data using algorithms developed so far, is possible.
- Preliminary estimates for 6X lossy compression ratio on hyperspectral data using algorithms developed so far, is achievable.
- Modified wavelet-based algorithms to deal with sensor data of "imperfect" number (non-2^n, i.e. non- 256; 512; 1024; 2048, etc) of field of views per scan line.
- Characterizing compression ratio as a function of data bit depth.
- Has developed a new pre-processing scheme to improve the compression performance using existing or newly developed compression methods.
- Begin to design and conduct study to characterize the impact of lossy compression on data and products.
- Near real-time AIRS PC compressed data being made available to global NWP centers.
- Investigated several state-of-the art algorithms on hyperspectral data (ISO image compression standards – JPEG2000, JPEGLS)
- Adopted and tested several lifting schemes that include: 45 in 1-D; 9 of which were also converted to 3-D compression.
- Compression team expands to include member from newly established NOAA cooperative institute – CREST.

In support of continuing GOES-R hyperspectral data compression effort and based on progress so far we presume that the following tasks need to be carried out in time for making intelligent recommendations and proposes proven approaches for GOES-R data management:

- Continue to study the impact of lossy compressions on data (level 1) and core and potential new products (level 2 to level 3).
- Validate compression results (both data and products) on near global observations (except poles).
- Conduct compression process in data domain spaces such as senor "raw count-space", calibrated radiance and/or brightness temperature space.
- Process on interferometric (or interferogram), as well as grating-type data.
- Reach out to other space agencies and industries with similar hyperspectral data compression interests/needs.
- Conduct trade-off studies between hardware requirement (CPU, memory, speed etc) and compression performance to recommend which types of compression algorithms are suitable

candidates for GOES-R downlink, rebroadcast, and data archive utilities.

- Continue work on new and improved preprocessing and compression methods.
- Continue Principal Component Analysis (PCA) work using independent (historical) and dependent (data themselves) approaches for both lossless and lossy compression.
- Begin to analyze, model and demonstrate error propagation and containment.

***Acknowledgements*. This research was supported by NOAA grant NA07EC0676.**

References:

[1] Huang, H.-L. C. Velden, J. Li, E. Weisz, K. Baggett, J. E. Davies, J. Mecikalski, B. Huang, R. Dengel, S. A. Ackerman, E. R. Olson, R. O. Knuteson, D. Tobin, L. Moy, J. A. Otkin, D. J. Posselt, H. E. Revercomb, and W. L. Smith, 2004: Infrared hyperspectral sounding modeling and processing: An overview. Preprints, Eighth Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IAOS-AOLS), Amer. Meteor. Soc., Boston, MA, Seattle, WA.

[2] Huang, H.-L., D. Tobin, J. Li, et, al., 2003: Hyperspectral Radiance Simulator – Cloudy Radiance Modeling and Beyond. SPIE Optical Remote Sensing of the Atmosphere and Clouds III, Vol, 4891, 180-189. 25-27 Oct., 2002, Hangzhou, China.

[3] Huang, B., H.-L. Huang, H. Chen, A. Ahuja, T. J. Schmit, and R. W. Heymann, 2003: Lossless Data Compression Studies for NOAA Hyperspectral Environmental Suite using 3D Integer Wavelet Transforms with 3D Embedded Zerotree Coding. SPIE Third International Symposium on Multispectral Image Processing and Pattern Recognition, 20-22 Oct. 2003, Beijing, China, Proc. SPIE, 5286, 305-315, 2003.

[4] Huang, B., H.-L. Huang, T. J. Schmit, H. Chen, K. Soundarapandian, K. Baggett, and R. W. Heymann, 2003: Data Compression Studies for NOAA Hyperspectral Environmental Suite (HES) using 3D Integer Wavelet Transforms with 3D Set Partitioning in Hierarchical Trees. SPIE Remote Sensing, 8-12 Sep. 2003, Barcelona, Spain, Proc. SPIE, 5238, 255-265, 2003.

[5] Huang, B., A. Ahuja, H.-L. Huang, T. J. Schmit, and R. W. Heymann, 2004: Improvements to Predictor-based Methods in Lossless Compression of 3D Hyperspectral Sounding Data via Higher Moment Statistics. WSEAS Transactions on Electronics, Issue 2, Vol. 1, 299-305.

[6] Huang, B., A. Ahuja, H.-L. Huang, T. J. Schmit, and R. W. Heymann, 2003: Lossless compression of 3D hyperspectral sounder data using context-based adaptive lossless image codec with bias-adjusted reordering. Accepted for publication in Optical Engineering.