

# JCSDA Quarterly

NOAA | NASA | US NAVY | US AIR FORCE

[CLICK HERE TO READ ONLINE](#)

## IN THIS ISSUE

- 1 NEWS IN THIS QUARTER**  
Assimilation of Geostationary Hyperspectral InfraRed Sounders (GeoHIS): Opportunities and Challenges  
  
Approximating Observation Error Statistics using Two Machine Learning Models
- 17 EDITOR'S NOTE**
- 18 NEWS ANNOUNCEMENTS**
- 19 PEOPLE**
- 22 SCIENCE CALENDAR**
- 22 CAREER OPPORTUNITIES**

## NEWS IN THIS QUARTER

# Assimilation of Geostationary Hyperspectral InfraRed Sounders (GeoHIS): Opportunities and Challenges

Wei Han<sup>1</sup>, Robert Knuteson<sup>2</sup>, Jun Li<sup>2</sup>, Dick Dee<sup>1</sup>, Thomas Auligne<sup>1</sup><sup>1</sup>JCSDA <sup>2</sup>Space Science and Engineering Center

## 1 Background

Hyperspectral infrared (IR) sounders feature thousands of channels, which collectively provide high vertical resolution and the capability to accurately measure atmospheric temperature and humidity vertical structure information (Smith, et. al., 2009). The assimilation of high spectral resolution infrared sounder observations from polar orbit satellites has been widely used in global and regional numerical weather prediction (NWP) models and has large positive impact on NWP. However, Polar hyperspectral IR sounders have inadequate temporal coverage which limits their capability on rapid evolving weather systems analyses and forecasts. High temporal geostationary (Geo) hyperspectral IR sounder (GeoHIS) radiance measurements enable continuous sounding of the atmospheric temperature and moisture, and thus capture the temporal and spatial variability for high impact weather or rapid changing weather events.

In 2019, theWMO 2040 vision for the space-based component of the Integrated Global Observing System (WIGOS) for the operational geostationary satellite constellation included "Geostationary core constellation with a minimum of five satellites providing complete Earth coverage, Multi-spectral VIS/IR imagery with rapid repeat cycles, IR hyperspectral sounders, lightning mappers, and UV/VIS/NIR sounders." The GeoHIS is mainly for atmospheric temperature, humidity, wind (by tracking cloud and water vapour features), identification of rapidly evolving mesoscale features, sea/land surface temperature; cloud amount and top height/ temperature; atmospheric composition (aerosols, ozone, greenhouse gases, trace gases).

---

**JOINT CENTER FOR SATELLITE  
DATA ASSIMILATION**

5830 University Research Court  
College Park, Maryland 20740

3300 Mitchell Lane  
Boulder, Colorado 80301

[www.jcsda.org](http://www.jcsda.org)

---

**EDITORIAL BOARD**

**Editor:**  
James G. Yoe

**Assistant Editor:**  
Kat Shanahan

**Director:**  
Thomas Auligné

**Chief Administrator Officer:**  
James G. Yoe

On 10 December 2016, the successful launch of China's Fengyung FY-4A satellite into geostationary orbit initiated a new era in Earth observation by providing the first time-continuous observations of the upwelling thermal infrared at high spectral resolution with the Geostationary Interferometric Infrared Sounder (GIIRS). GIIRS is a Michelson interferometer that measures the atmospheric infrared radiation in a spectral range 700-2250cm<sup>-1</sup> at a spectral resolution of 0.625cm<sup>-1</sup>. It has 1650 spectral channels covering longwave infrared (LWIR) (700-1130cm<sup>-1</sup>, 689channels) and middle-wave infrared (MWIR) (1650-2250cm<sup>-1</sup>, 961channels) bands (Yang et al., 2017.) A subset of GIIRS longwave temperature sounding channels has been assimilated in China's global NWP system GRAPES (Global/Regional Assimilation and PrEdiction System) since December 2018 and improve the forecast over East Asia (Han et al., 2019b), especially for high impact weather forecasting, such as Typhoons. The European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) is developing an operational advanced GEO hyperspectral IR sounder (IRS) as a part of Meteosat Third Generation (MTG-3) in the mid 2020's. The Japan Meteorological Agency (JMA) also has started discussions of a follow-on program for the operational geostationary satellites Himawari-8 and -9,

which are scheduled to operate until 2029. The impact of the GeoHIS on numerical weather predictions has been demonstrated using an Observing System Simulation Experiment (OSSE) approach, which illustrated the improvement on the forecasts of the representative meteorological field and the typhoon position (Okamoto, et. al.,2020.)

NOAA recently has evaluated a range of space architecture options to select one that will provide the highest priority observations effectively and efficiently for a future Geosynchronous and Extended Orbit (GEO-XO) program. Ideally, an imager, lightning mapper, infrared hyperspectral sounder, and ocean color instrument are best suited on two spacecraft near the current GOES-East and GOES-West positions, while the atmospheric composition instrument is recommended to reside in a central location. A day/night spectral channel is recommended as part of either the imager or the sounder. An infrared sounder will provide real-time atmospheric temperature and water vapor profile data to be assimilated into advanced numerical weather prediction models and improve short-term severe weather forecasting.

Below is a figure illustrating one notional GEO-XO constellation concept, as of January 27, 2021. This constellation is preliminary,

**Figure 1:** A preliminary proposed GEO-XO constellation concept. (NOAA, 2021)



pending program approval. The first GEO-XO launch, planned for around the year 2030, will maintain and advance NOAA's critical geostationary observations through 2055.

Based on the evaluation and assimilation of GIIRS in Joint Effort for Data assimilation Integration (JEDI) framework (Han and Knuteson, 2020), this paper discusses the opportunities and challenges of GeoHIS. The opportunities include targeted observing for high impact weather, and improvement in convective storm forecasts, including tornadoes and hurricanes. The challenges include accurate spectral and radiometric calibration, considering the possible diurnal variation, fast radiative transfer model for large satellite zenith angles, and continuous data assimilation for application to high temporal observations.

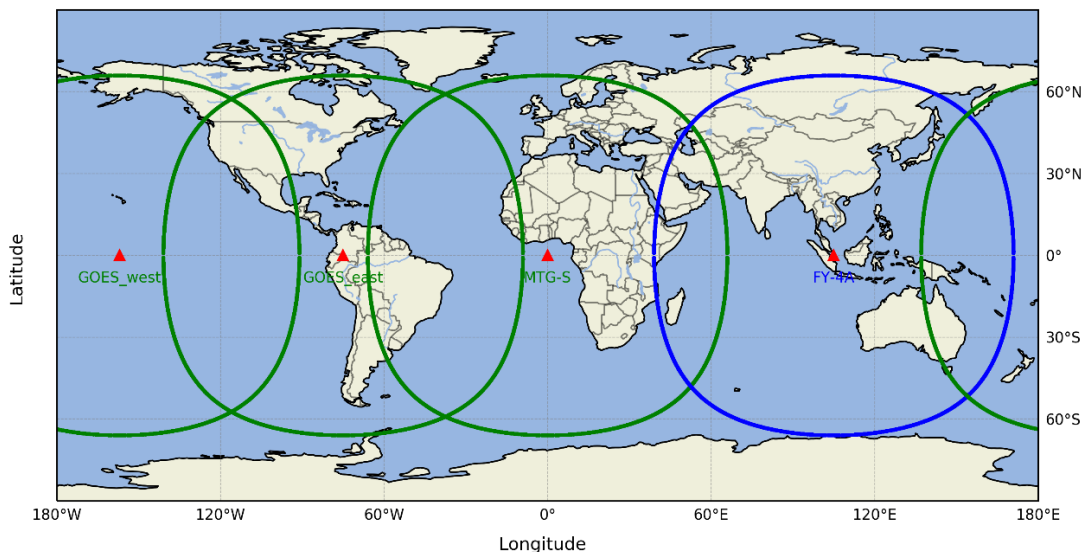
## 2 Opportunities for NWP

FY-4A GIIRS is the only GeoHIS currently in orbit, while EUMETSAT will have hyperspectral IR sounders (IRS) in GEO orbit in 2023 time frame and NOAA is planning to have Geo-XO around 2030. A global GeoHS ring will be formed at that time, covering around 60 degrees north and south latitudes around the world,

providing global high-temporal and high-spectral resolution measurements as shown in Figure 2. This will benefit the regional and global numerical weather prediction, especially for the improvement on high impact weather forecasts. Such a global GeoHIS ring will dramatically improve our ability to monitor storms and hurricanes through their complete life cycles, and could provide high spatial and temporal information of atmospheric temperature and moisture structures. Together with current Hyperspectral Infrared Sounder (HIS) onboard polar orbiting satellites, a more complete global observation picture of weather system life cycles can be monitored leading to a greater understanding of their atmospheric dynamics as well as to more accurate predictions and warnings using NWP or nowcasting.

FY-4A GIIRS provides an opportunity for researchers to use the real GeoHIS observations to develop and evaluate algorithms and methods to exploit this new class of data for NWP applications via data assimilation. The JCSDA and the University of Wisconsin Space Science and Engineering Center (UW-SSEC) began to cooperate on a research project on the GIIRS data assimilation in JEDI framework

**Figure 2:** Global GeoHIS ring anticipated around 2030. Areas within 75 degrees local zenith angle from future global ring of GHS. FY-4A was launched in 2016, each geostationary satellite covers around 60 degrees in both latitude and longitude directions.



since October 2019. GIIRS L1 data has been reprocessed at SSEC and introduced in the Joint Effort for Data assimilation Integration (JEDI) framework. The spectral and radiometric biases and uncertainty are estimated in JEDI-FV3. The differences between observations and simulations using JEDI-FV3(HOFX) have been produced and analyzed since May 2020. Assessments of Geostationary hyper-spectral IR Sounder (Geo-Sounder) are being carried out using GIIRS in the JEDI framework, with the objective of reducing or avoiding redundant work within the community and increasing efficiency of research and of the transition from development teams to operations focusing the Geo-Sounders. Integration of GIIRS observations into the JEDI framework, includes the preprocessing of GIIRS to JEDI Interface for Observation Data Access (IODA), the observation operator using CRTM, and quality control through JEDI filters (Han et al., 2021a.)

### 2.1 Targeted observing to improve high impact weather forecasts

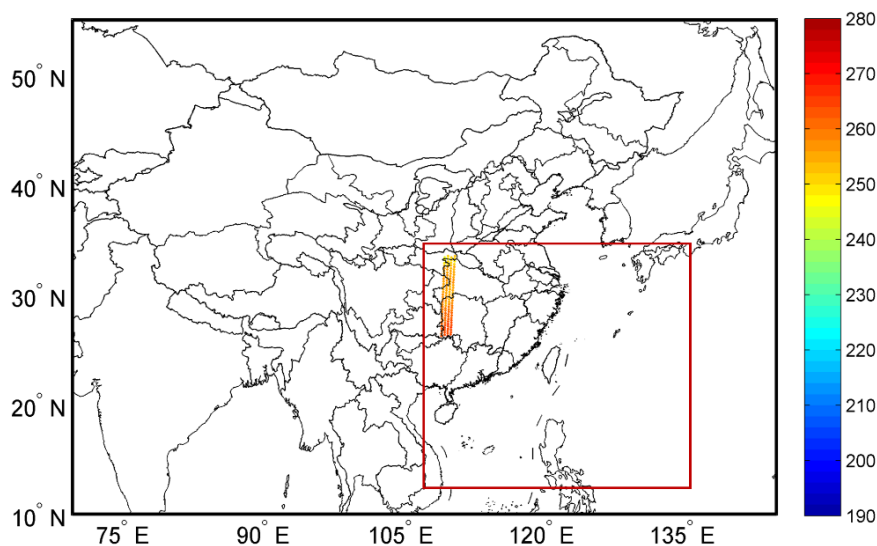
Targeted observing is a technique to reduce rapidly growing forecast errors with the assimilation of observations in the sensitive regions at initial time (Han W. et al, 2019a). The efforts of using targeted observations to

improve numerical forecasts of high-impact weather events has been carried out over the past two decades, particularly during the THORPEX era (2005–14). The target regions were often of synoptic scale, which is usually beyond the range of coverage of a single aircraft mission.

For the first time, GIIRS on board FY-4A has the capability of providing observations every 15 minutes for selected regions where active weather events occur, which demonstrate the unique opportunity for targeted observations to improve high impact weather. In 2018, three targeted observing experiments were carried out for the Typhoon Maria, Ambil, and Mangkhut cases. The target region was determined using GRAPES global singular vector (SV) framework and then GIIRS was controlled to observe the target area every 30 minutes for the Ambil and Mangkhut cases, while the targeted area was chosen by chief forecaster for Maria case with 15 minutes temporal resolution as shown in Figure 3. The assimilation experiments of the high temporal targeted GIIRS measurements in GRAPES 4D-Var show significant positive impact on Typhoon track forecasts (Han et al., 2019a.)

Future global GeoHIS ring will offer a great

**Figure 3:** The FY-4A GIIRS targeted observing region for Typhoon Maria (2018) with 15 minutes temporal resolution in the 36 hours period, 00Z 10 July – 12Z 11 July, 2018.





opportunity for the targeted observing due to the big advantages of HIS on geostationary orbit, as the sounder could be quickly targeted to a specific and relative large area to measure the three dimensional atmospheric temperature and moisture with very high temporal resolution using the “stare and step” observing mode. The GeoHIS sampling period could range from minutes to an hour, depending upon the area coverage selected for the measurement, which makes the GeoHIS an ideal global targeted observing system to improve high impact weather (such as hurricanes and storms) forecast in real time.

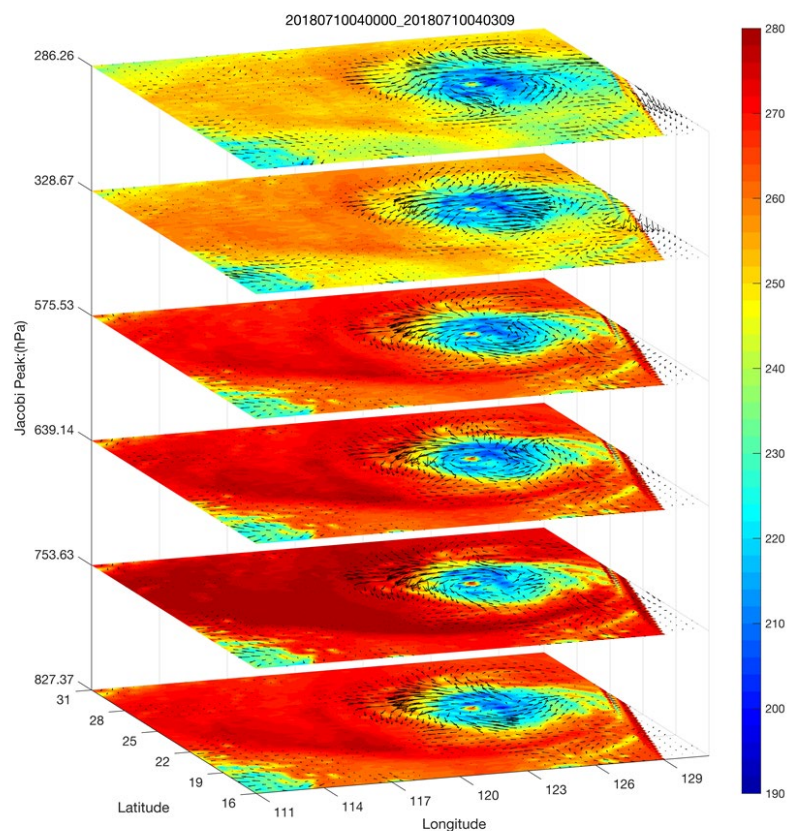
## 2.2 Four-dimensional wind field from geostationary hyperspectral infrared sounder radiances

One of the key advancement that GeoHIS achieves beyond current operational geostationary capabilities is that the water

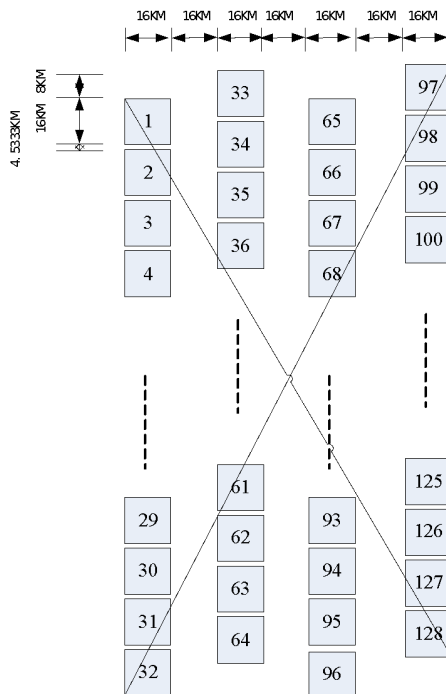
vapor (and other trace gases such as ozone) winds will be altitude-resolved throughout the troposphere to stratosphere. The traditional concept uses a two-step approach: First, get the three-dimensional water vapor retrievals, and then deduce the winds using water vapor sequence which is similar to the atmospheric motion vector (AMV) algorithm (Smith et. al., 2001a, 2001b.)

However, a novel approach based on combined optical flow (OF) and machine learning (ML) techniques has been developed for retrieving wind fields from high temporal resolution GEO hyperspectral IR sounder radiance measurements (Li et al., 2021.) Four-dimensional (4D) wind fields (Figure 4) (3D wind in time sequence), derived from GIIRS radiance measurements with 15-minute temporal resolution during Typhoon Maria (2018) indicate reasonable accuracy and precision when compared

**Figure 4:** 4D-Winds from GIIRS high-temporal water vapor sounding channels at 04Z 10 July 2018.



**Figure 5:** FY-4A GIIRS uses a two dimensional Large Focal Plane detector Array (LFPA) (32X4) covering an area on Earth of 648x112 km.



with independent ERA5 reanalysis and dropsonde wind profiles, also the temporal variation of the wind field from GIIRS with 15-minute interval is consistent with that from the ERA5 with 1 hour time interval. Further experiments show that higher temporal resolution for geostationary IR sounder measurements could provide better dynamic information. Together with the thermodynamic information, the high temporal resolution 4D dynamic information from geostationary hyperspectral IR radiance measurements can be extracted reliably and efficiently for various quantitative applications such as NWP data assimilation, near real-time (NRT) weather monitoring, situation awareness and now-casting.

Another approach to retrieve the 4D wind field is through 4D-Var using the tracer effect. Peubey and McNally (2009) demonstrated that the impact of high temporal geostationary clear-sky radiances (CSRs) on 4D-Var wind analyses through “tracer advection effect.” The wind field is adjusted in order to fit observed humidity features via the minimization of the 4D-Var

cost function. Comparing to the assimilation of CSRs from the two water-vapour channels on Meteosat-9 in that study (Peubey and McNally, 2009), there are hundreds of humidity sounding channels on GeoHIS, which lead to altitude-resolved wind field through 4D-Var.

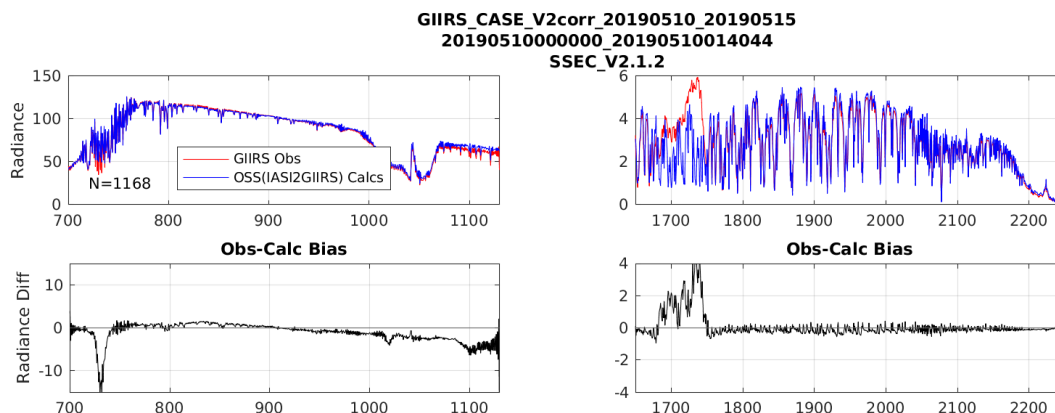
### 3 Challenges of GeoHIS Assimilation

#### 3.1 Spectral and radiometric calibration

In order to measure near instantaneous large geographical area of the earth with high horizontal resolution, the GeoHIS usually acquires a number of spectral soundings using a two-dimensional Large Focal Plane detector Array (LFPA). For instance, FY-4A GIIRS uses a two-dimensional LFPA (32X4) covering an area on Earth of 648x112 km (Yin et al., 2020) as shown in Figure 5, while MTG IRS will use a 160X160 LFPA covering an area on Earth of 640x640 km (Coppens et al., 2017.)

To use GeoHIS data in NWP, it is important that its spectral and radiometric accuracy be understood and well-characterized. As both NWP and instruments improve, the field of view (FOV)-to-FOV differences become of greater concern in data assimilation for both bias and observation estimation. For the GeoHIS with thousands of FOVs in LFPA, the spectral and radiometric uniformisation is of vital importance for the NWP users community. The evaluation of GIIRS also shows that there are very complex biases and uncertainties in spectral and radiometric for each detector, from diurnal to seasonal time scales (Yin et al., 2020; Han et al., 2021.) Satellite radiance assimilation requires uniform spectral calibration among FOVs, so the FOV dependent spectral biases must be corrected. Thus, investigation on the calibration needs to be enhanced to make better use of the GeoHIS data in data assimilation. The close collaboration

**Figure 6:** Example LBLRTM simulated GIIRS radiance ( $mW/(m^2 sr cm^{-1})$ ) spectrum (blue) and observed GIIRS (red) (upper panels). Observation minus calculation radiometric bias for clear sky cases on 10 May 2019 between 0-2 UTC (lower panels).



between calibration and data assimilation will lead to better understanding of the data and better use of the data.

The National Satellite Meteorological Center (NSMC) of the China Meteorological Administration (CMA) publicly released data from the FY4A GIIRS sensor starting on 24 January 2018. GIIRS radiance spectra from the FY4A satellite at about 105 E have been available in near real-time from CMA since that date. The Space Science and Engineering Center (SSEC) at the University of Wisconsin-Madison has been acquiring and processing GIIRS radiance data and evaluating it against simulated radiance observations and reference satellite observations.

The GIIRS is a Fourier Transform Spectrometer (FTS) which uses an on-board metrology laser to provide a spectral reference for the radiance spectrum. The design of the FTS spectrometer is that all the channels within a spectral band have a wavenumber grid defined by a constant wavenumber interval. The FTS provides superior channel-to-channel spectral integrity compared to alternate designs such as a grating spectrometer (Smith et al., 2002.) In order to achieve the level of accuracy in the wavenumber scale desired for NWP data assimilation a refinement of the on-orbit spectral calibration using known atmospheric absorption lines

simulated using an atmospheric radiative transfer model is required.

Figure 6 shows an example GIIRS spectrum with simultaneous measurements in two broad spectral bands. The longwave (LW) band covers the 700-770  $cm^{-1}$  temperature sounding band, the 700-1000  $cm^{-1}$  longwave window containing information about surface emission and low-level water vapor, and the 1030-1070  $cm^{-1}$  ozone emission band. The midwave-shortwave (MSW) band contains spectral channels sensitive to upper and mid-tropospheric water vapor (1650-1900  $cm^{-1}$ ) and near surface temperature as well as carbon monoxide (1900-2300  $cm^{-1}$ ).

Assessment of the NSMC calibration of the FY4A GIIRS sensor by SSEC indicates radiometric biases in selected wavenumber regions in each of the two GIIRS spectral bands. Detailed evaluation suggests these radiometric biases are caused by an error in the NCSM implementation in the ground calibration.

The calibration uncertainty of GIIRS has also been studied. In order to assess the spectral calibration accuracy of GIIRS, an efficient and accurate spectral shift estimation and correction algorithm was developed called Iterative Spectral Shift Estimation and Correction (ISSEC) (Han et al., 2021b.) ISSEC is used to estimate the GIIRS spectral shift of each detector based on CrIS and

GIIRS simultaneous nadir overpass (SNO) and JEDI NWP simulations HOFX. The spectral accuracy of resampled GIIRS L1 radiance is greatly improved with a spectral uncertainty less than 10 ppm for most of the detectors. The long term trend and diurnal variation of spectral shift are analyzed based on SNOs and NWP simulations respectively. Based on lessons learned from FY-4A GIIRS, recommendations are proposed for future geostationary hyperspectral InfraRed sounders in terms of homogeneity of large detector array, diurnal variation in geostationary orbit and calibration challenges and solutions.

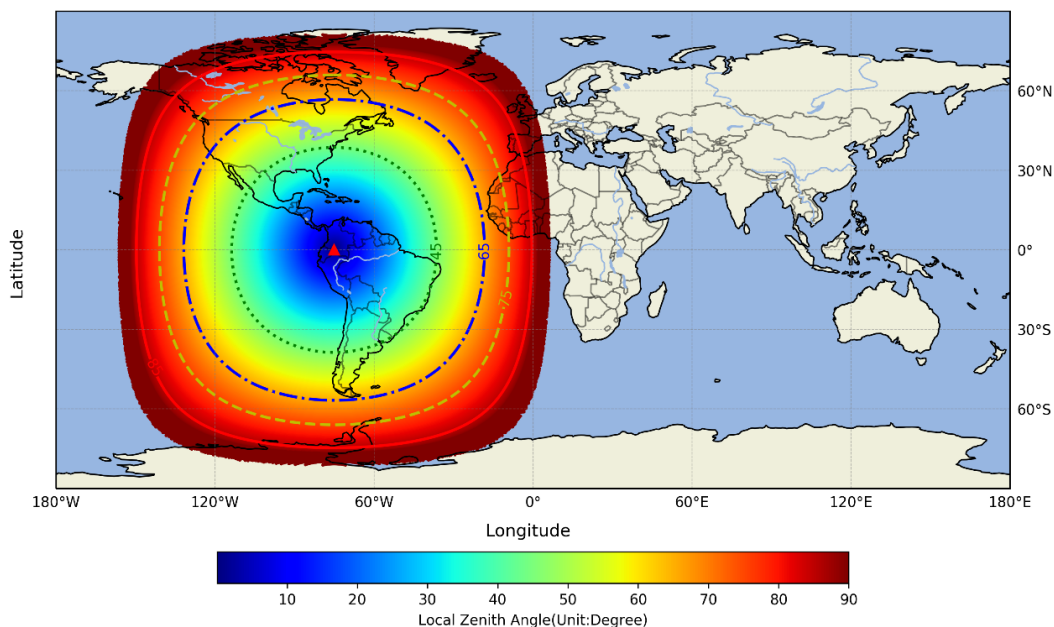
### 3.2 Observation operator improvement for GeoHIS with large local zenith angle

In the most of the current operational data assimilation system, the satellite radiance assimilation is using a one-dimensional vertical model background (short-term forecast) profile, located at the ground footprint of the observation location, to compute the model equivalent of the observation using fast radiative transfer models (such as RTTOV and CRTM.) Since in general, the lines of satellite sensor view

are not vertical, the observation operator can be improved by interpolating the trial field to the slant path. Bormann (2017) shows that taking the viewing geometry better into account leads to significant improvements in the simulation of brightness temperatures from model fields compared to the traditional approach, which is particularly noticeable for larger viewing zenith angle (VZA), for channels that peak in the upper troposphere or higher and have relatively low noise, and for high and mid-latitudes.

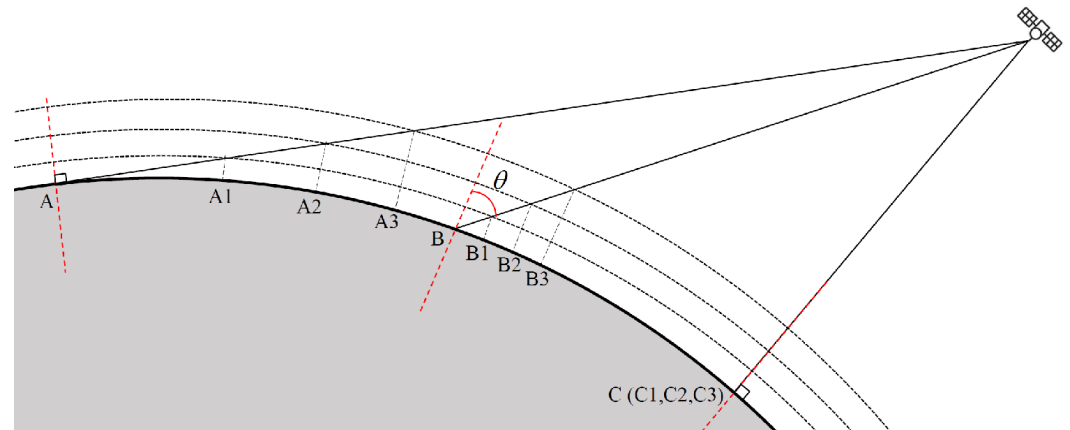
The maximum VZA of hyperspectral IR sounders aboard polar orbiting satellites is usually less than  $60^\circ$ , while that of the GeoHIS could be as large as  $90^\circ$  as shown in Figure 7. The geo-location difference is around 50 km for observations peaking in the mid-stratosphere and viewed with a zenith angle of  $60^\circ$  (as shown in Figure 8). Given the advent of sounders with considerably lower noise levels and the steady increase of NWP model resolution and accuracy, the temperature gradients in the slant path could not be neglected for GeoHIS with larger VZA, especially for high peaking channels. In current

**Figure 7:** The local view zenith angle (VZA) of GOES<sub>west</sub>. The VZA of  $45^\circ$  (dot),  $65^\circ$  (dash dot),  $75^\circ$  (dash),  $85^\circ$  (solid) degree are shown in the contour lines.





**Figure 8:** Schematic illustration of the satellite viewing geometry, viewing a location on Earth with a local zenith angle of  $\theta$  along a slanted path through the atmosphere (A,B) and nadir (C). The vertical profile currently used for radiative transfer calculations is indicated in red (dashed).



operational NWP, the observations with VZA larger than a threshold (such as  $65^\circ$ ) are eliminated, partly due to the limitation of the fast radiative transfer model (RTM) where the zenith angles only up to about 65 degrees in the coefficient training (Di et al., 2018). So in order to use the GeoHIS with large VZA to cover more area as shown in Figure 7, there are two important aspects that need to be improved: (1) the fast RTM need to be trained with larger VZA accordingly; (2) the slant path interpolation need to be considered in the radiance simulation.

### 3.3 How to best use the high temporal GeoHIS in NWP

The biggest advantage of GeoHIS is the high temporal resolution. How to make best use of the GeoHIS high-temporal and high-spectral measurements in data assimilation presents an opportunity and also a challenge. There has been lots progress in data assimilation methodologies in the last 20 years, including some recent machine learning based techniques, and the 4D-Var is still one of the most advanced data assimilation methodology for the use of temporal near-continuous observations theoretically and practically (Lean et al. 2020.) For regional NWP systems which are intended to forecast sub-synoptic and mesoscale features that have correspondingly short predictability time scales, there is a clear benefit to generate

analyses and forecasts frequently using the high temporal observations. Typically, these use hourly cycling data assimilation systems with a 1 hour assimilation window (Benjamin et al., 2016; Smith et al. 2020) shows that geostationary satellite hyperspectral soundings [i.e., FY-4A GIIRS] improve hazardous precipitation forecasts when used, in addition to the combined polar hyperspectral and geostationary multispectral satellite profile data, to initialize the numerical forecast model using hourly assimilation cycle. This improvement results from the provision of hyperspectral resolution vertical soundings, rather than solely multispectral resolution soundings, within the space and time gaps of the polar hyperspectral data. Improvements in global hazardous weather predictions can be expected when high-spatial-resolution and high-temporal-resolution hyperspectral sounding instruments are carried on the future international system of geostationary satellites.

### References

Benjamin, S.G., Weygandt, S.S., Brown, J.M., Hu, M., Alexander, C.R., Smirnova, T.G., Olson, J.B., James, E.P., Dowell, D.C., Grell, G.A., Lin, H., Peckham, S.E., Smith, T.L., Moninger, W.R., Kenyon, J.S. and Manikin, G.S. (2016) A North American hourly assimilation and model forecast cycle: The

- Rapid Refresh. *Monthly Weather Review*, 144, 1669–1694. <https://doi.org/10.1175/MWR-D-15-0242.1>
- Bormann, N., 2017. Slant path radiative transfer for the assimilation of sounder radiances. *Tellus A: Dynamic Meteorology and Oceanography* 69, 1272779. <https://doi.org/10.1080/16000870.2016.1272779>
- Coppens, D., Theodore, B., Klaes, D., 2017. MTG-IRS: from raw measurements to calibrated radiances, in: *Earth Observing Systems XXII. Presented at the Earth Observing Systems XXII, International Society for Optics and Photonics*, p. 1040203. <https://doi.org/10.1117/12.2272553>
- Di, D., J. Li, W. Han, W. Bai, C. Wu, and W. P. Menzel, 2018: Enhancing the fast radiative transfer model for FengYun-4 GIIRS by using local training profiles, *Journal of Geophysical Research - Atmospheres*, 123, [doi:10.1029/2018JD029089](https://doi.org/10.1029/2018JD029089).
- Han W. et al, 2019a: Target observing experiments using high temporal geostationary sounder : Typhoon Ambil case, the 99th American Meteorological Society Annual Meeting, Phoenix, Arizona, 6-10 January 2019.
- Han W. et al, 2019b: Assimilation of high temporal resolution GIIRS in 4D-Var, 2019 Joint Satellite Conference, Boston, Sept.28-Oct.4,2019. <https://ams.confex.com/ams/JOINTSATMET/videogateway.cgi/id/505317?recordingid=505317>
- Han W. and R. Knuteson, 2020, Assessment of Geostationary Hyper-spectral Sounders in JEDI, The 7th WMO Workshop on the Impact of Various Observing Systems on NWP, 30 November - 3 December 2020. <https://meetings.wmo.int/impact-workshop-7/English/1-12-P-Wei.pdf> (accessed 3.22.21)
- Han W., S. Hui, R. Knuteson, D. Dee, 2021b: Assessment and Integration of the Geostationary Hyperspectral Infrared Sounder in JEDI, Ninth AMS Symposium on the Joint Center for Satellite Data Assimilation (JCSDA), 2021.
- Han W. et. al., 2021b, ISSEC: An fast and accurate algorithm for hyperspectral infrared sounder spectral shift estimation and correction with application on FY-4A GIIRS, to be submitted.
- Lean, P., Hólm, E.V., Bonavita, M., Bormann, N., McNally, A.P., Järvinen, H., 2021. Continuous data assimilation for global numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* 147, 273–288. <https://doi.org/10.1002/qj.3917>
- Li J., Z. Ma, W. Han, D. Di, Z. Li, R. Yin, W. P. Menzel, and T. J. Schmit, 2021: Four-dimensional wind field from geostationary hyperspectral infrared sounder radiances with high temporal resolution – Typhoon Maria case, EUMETSAT meteorological satellite conference. (submitted).
- Majumdar, S.J., 2016. A Review of Targeted Observations. *Bull. Amer. Meteor. Soc.* 97, 2287–2303. <https://doi.org/10.1175/BAMS-D-14-00259.1>
- NOAA, GEO-XO | NOAA National Environmental Satellite, Data, and Information Service (NESDIS) [WWW Document], URL <https://www.nesdis.noaa.gov/GEO-XO> (accessed 3.22.21).
- Peubey, C., McNally, A.P., 2009. Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Quarterly Journal of the Royal Meteorological Society* 135, 1863–1876. <https://doi.org/10.1002/qj.500>

- Okamoto, K., Owada, H., Fujita, T., Kazumori, M., Otsuka, M., Seko, H., Ota, Y., Uekiyo, N., Ishimoto, H., Hayashi, M., Ishida, H., Ando, A., Takahashi, M., Bessho, K., Yokota, H., 2020. Assessment of the Potential Impact of a Hyperspectral Infrared Sounder on the Himawari Follow-On Geostationary Satellite. *Sola* 16, 162–168. <https://doi.org/10.2151/sola.2020-028>
- Smith, William L.; Harrison, F. Wallace; Revercomb, Henry E. and Bingham, Gail E. Geostationary Fourier Transform Spectrometer (GIFTS) - The New Millennium Earth Observing-3 Mission [Geostationary Imaging Fourier Transform Spectrometer (GIFTS) - The New Millennium Earth Observing-3 Mission]. IRS 2000: Current problems in atmospheric radiation. Proceedings of the International Radiation Symposium, St. Petersburg, Russia, 24-29 July 2000. A. Deepak Publishing, Hampton, VA, 2001, pp.81-84.
- Smith, William; Harrison, Wallace; Hinton, Dwayne; Parsons, Vickie; Larar, Allen; Revercomb, Henry; Huang, Allen; Velden, Christopher; Menzel, Paul; Peterson, Ralph; Bingham, Gail and Huppi, Ronald. GIFTS - A system for wind profiling from geostationary satellites. International Winds Workshop, 5th, Lorne, Australia, 28 February-3 March 2000. Proceedings. European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany, 2001, pp.253-258. Reprint # 3033.
- Smith Sr., W.L., Revercomb, H., Bingham, G., Larar, A., Huang, H., Zhou, D., Li, J., Liu, X., Kireev, S., 2009. Technical Note: Evolution, current capabilities, and future advance in satellite nadir viewing ultra-spectral IR sounding of the lower atmosphere. *Atmos. Chem. Phys.* 9, 5563–5574. <https://doi.org/10.5194/acp-9-5563-2009>
- Smith, W.L., Zhang, Q., Shao, M., Weisz, E., 2020. Improved Severe Weather Forecasts Using LEO and GEO Satellite Soundings. *Journal of Atmospheric and Oceanic Technology* 37, 1203–1218. <https://doi.org/10.1175/JTECH-D-19-0158.1>
- Smith Sr., W. L., H. Revercomb, E. Weisz, D. Tobin, R. Knuteson, J. Taylor, and W. P. Menzel, 2021: Hyperspectral Satellite Radiance Atmospheric Profile Information Content and Its Dependence on Spectrometer Technology, Accepted for publication in the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- Okamoto, K., Owada, H., Fujita, T., Kazumori, M., Otsuka, M., Seko, H., Ota, Y., Uekiyo, N., Ishimoto, H., Hayashi, M., Ishida, H., Ando, A., Takahashi, M., Bessho, K., Yokota, H., 2020. Assessment of the Potential Impact of a Hyperspectral Infrared Sounder on the Himawari Follow-On Geostationary Satellite. *Sola* 16, 162–168. <https://doi.org/10.2151/sola.2020-028>
- WMO, Vision for the WMO Integrated Global Observing System in 2040 [WWW Document], 2020. 2019 edition. WMO- No. 1243. World Meteorological Organization. <https://public.wmo.int/en/resources/library/vision-wmo-integrated-global-observing-system-2040> (accessed 3.22.21).
- Yin, R.,W. Han, Z.Gao and D.Di. 2020, The evaluation of FY4A's Geostationary Interferometric Infrared Sounder (GIIRS) longwave temperature sounding channels using the GRAPES global 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 146, 1459–1476. <https://doi.org/10.1002/qj.3746>.

# Approximating Observation Error Statistics using Two Machine Learning Models

Amanda Burke<sup>1</sup>, Elizabeth Satterfield<sup>2</sup>, and Amy McGovern<sup>1,3</sup>

<sup>1</sup>School of Meteorology, University of Oklahoma, Norman, Oklahoma <sup>2</sup>Naval Research Laboratory, Monterey, California <sup>3</sup>School of Computer Science, University of Oklahoma, Norman, Oklahoma

## Motivation

Data Assimilation (DA) schemes provide the best estimate of the current atmosphere state by optimally blending observational data and short-term background forecasts based on assumed error statistics. Current operational methods of estimating an observation error covariance matrix,  $\mathbf{R}$ , are typically based on innovation (observation-minus-forecast) statistics (e.g., Desroziers et al., 2005). However, innovation based methods are subject to error when assumptions are violated. The ability of such methods to estimate observation error covariance can be impacted by the accuracy of the assumed observation and forecast (background) error statistics, bias, unaccounted for error correlation, model errors, and quality control. In addition, such innovation based methods are limited by data requirements and regionally varying or flow dependent error estimates are not always attainable.

Machine learning models are able to map input data to labeled observations and output skillful predictions of the labeled data that do not require large computational resources. In this work, we compare two different types of machine learning models, a random forest (RF, Breiman, 2001) and 1-Dimensional convolutional neural network (CNN, Kiranyaz et al., 2021), also known as a deep learning (DL) model, with the Desroziers et al. (2005) method to determine the feasibility of using machine learning models to estimate observation error covariance under relaxed assumptions.

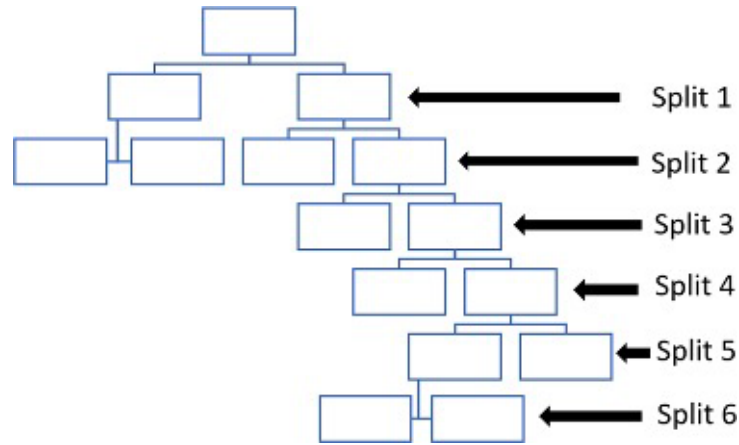
## Experimental Design

The inputs to the random forest and deep learning models are chosen based on the output of the DA schemes, consistent with the inputs to the Desroziers et al. (2005) method. This choice also enables direct comparisons between the methods under different simplified model scenarios. The inputs to the machine learning models are observations ( $\mathbf{y}^o$ ), the background forecast ( $\mathbf{x}^f$ ), analysis ( $\mathbf{x}^a$ ), innovation ( $\mathbf{d}_b^o = \mathbf{y}^o - \mathbf{H}(\mathbf{x}^f)$ ), analysis residual ( $\mathbf{d}_a^o = \mathbf{y}^o - \mathbf{H}(\mathbf{x}^a)$ ), and analysis increment ( $\mathbf{d}_b^a = \mathbf{H}(\mathbf{x}^a) - \mathbf{H}(\mathbf{x}^f)$ ), where  $\mathbf{H}$  denotes the observation operator which maps model space to observation space.

Although a machine learning model may be able to learn the difference terms based on the input observation, forecast, and analysis data, we include these variables to give the models more information about the DA scheme. Other variables, such as the square of the difference



**Figure 1:** Illustrative example of a single decision tree, with a depth of 6, from a random forest.



terms may also offer benefit, however we chose these variables for this feasibility study because of their significance within the DA community.

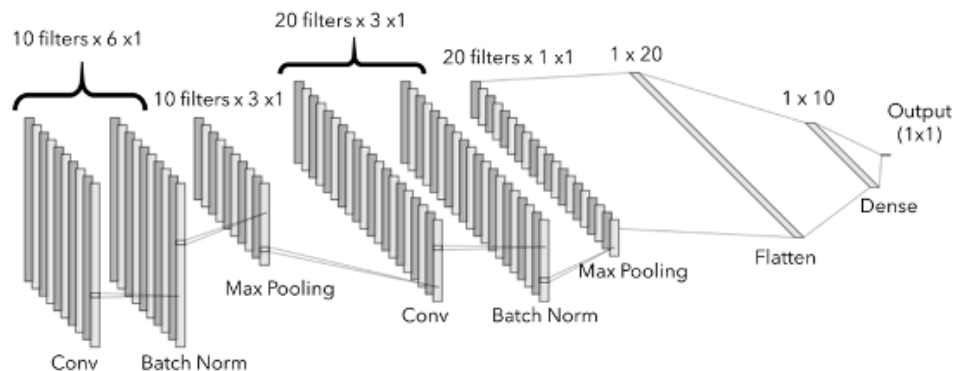
To train both machine learning models, we use model 3 of Lorenz '96 with 96 grid points and standard parameter values (Lorenz, 2005). A nature run, which is used to define true states, is obtained by cycling for 1000 time steps. Observations are created by perturbing the true state with a random draw from a normal distribution with zero mean and a prescribed observation error variance. These observations are assimilated, at each grid point, using an Ensemble Kalman transform Filter (ETKF, Bishop et al., 2001), with an ensemble size of  $K=500$ , to create inputs for the machine learning algorithms. For this initial feasibility study, the machine learning algorithms are trained

with a known true state. The impact of sub-optimal training data will be addressed in a follow-on study.

The random forest model, based on an ensemble of decision trees that are constructed from randomly selected subsets of the total input data, is made up of 50 trees with a maximum depth of 6 meaning that the trees may split up to 6 times (Fig. 1). A split within each decision tree is learned based on thresholds of the input variable values that best divide the data associated with the different labeled observations.

The deep learning model is a shallow one-dimensional CNN with two layers of convolution, batch normalization, and max pooling (Fig. 2). The number of filters for each layer varies based on the complexity of the experimental design. Each CNN is trained with a batch size of 256 examples

**Figure 2:** One-Dimensional CNN architecture with convolutional (Conv), batch normalization (Batch Norm), and maximum pooling (Max Pooling) layers. Examples of 10 and 20 filters are demonstrated however different numbers of filters are applied for each experiment depending on data complexity.



with varying number of epochs again depending on experiment complexity. The filters, or kernels, are randomized number matrices with a size that's user defined. The filters slide, or convolve, over each individual input predictor and perform a dot product of the input image values and the matrix. The convolution is performed over each input predictor and the resulting dot products are summed together to produce a single feature map. Depending on the user constraints, such as the size of the kernel or if the user wants to keep the dimensionality the same, this resulting feature map is the same size of the input image or slightly smaller.

After the convolution step, a batch normalization layer normalizes the summed dot products to speed up learning by rescaling the newly convolved data. A non-linearity is applied to the convolved data. For this study, the values greater than 1 are not affected, but values less than 0 are multiplied by 0.1. To learn information at different scales, a pooling layer decreases the dimensionality of the convolved layer, with a maximum value filter convolving over the image and only keeping the highest values. Usually the filter is a 2 by 2 matrix, thus decreasing the size of the next input images for convolution by a factor of 2. After multiple convolution, normalization, and pooling layers the data are flattened to a vector and new kernel weights and a non-linearity are applied to the data. The vector is reduced in size and complexity until a single output is learned. The D1 prediction is compared to the known truth. Based on the differences between the prediction and truth, the randomized kernels/filters are slightly changed to reduce the loss between the predictions. This process is repeated a number of times (epochs) until the user is satisfied with the loss.

Using the machine learning models, we

predict the true observations,  $\mathbf{y}^t$ , associated with the input forecasts and observations with errors. We then can calculate the observation error and associated observation error covariance matrix,  $\mathbf{R}$ , based on outputs of the DA system. The Desroziers et al. (2005) method calculates  $\mathbf{R}$  as the expected value of the outer product of analysis residuals and innovations,

$$\mathbf{R} = E[(\mathbf{d}_a^o)(\mathbf{d}_b^o)^T]$$

The machine learning models predict the true state  $y^t$  to compare to the observations with prescribed errors by the following method,

$$\mathbf{R} = E[(\mathbf{y}^o - \mathbf{y}^t)(\mathbf{y}^o - \mathbf{y}^t)^T].$$

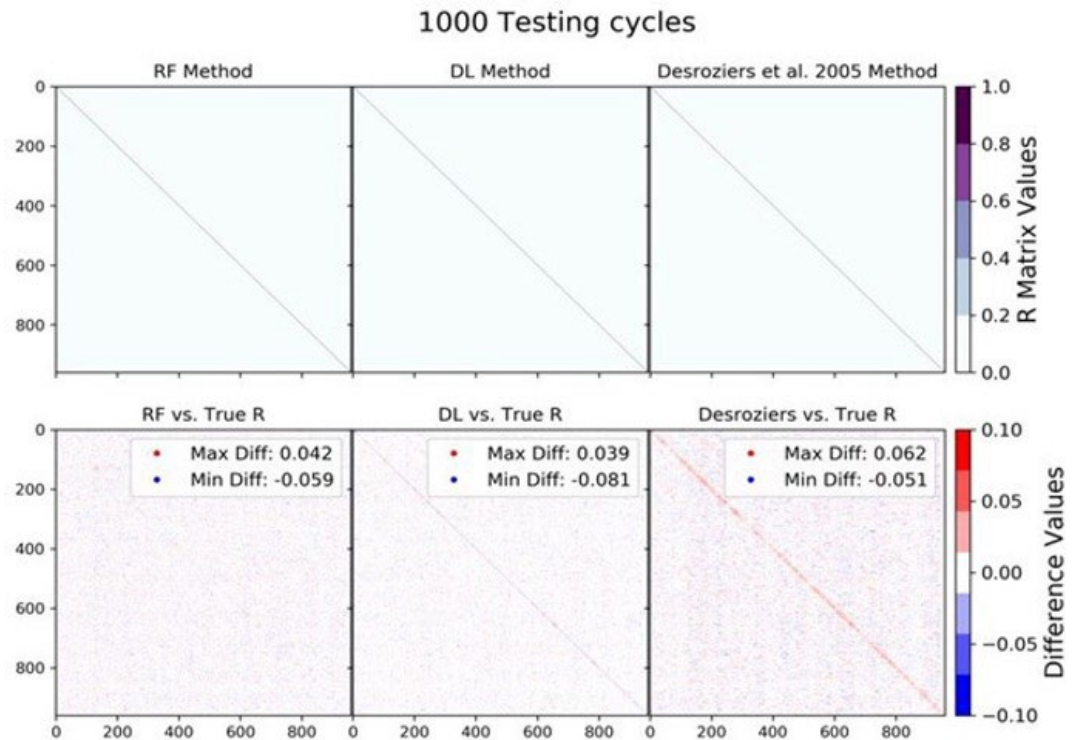
This method allows the machine learning models to better estimate the observation error covariance matrix under scenarios in which the Desroziers et al. (2005) method would be subject to error.

## Results

The first experiment compared the two machine learning models and Desroziers et al. (2005) method, where the assumed and actual observation error covariance was equal to the identity matrix. In this first scenario, observation errors are assumed static in time and accurately prescribed in the DA. This is the initial feasibility case to determine if the machine learning models are capable of emulating Desroziers et al. (2005) under the ideal conditions. The deep learning model is trained over 20 epochs with 16 and 32 filters for each convolutional layer, respectively. Figure 3 shows the predicted  $\mathbf{R}$  matrix from the three different methods as well as the differences between the true  $\mathbf{R}$  matrix (identity in this case) and the predicted error covariances.

All three methods capture the identity matrix, with the random forest method predicting the matrix with the least differences from the true  $\mathbf{R}$  matrix. The deep learning method

**Figure 3:** First experiment results predicting the identity  $\mathbf{R}$  matrix. Top row shows the predicted  $\mathbf{R}$  matrix of a random forest (top left), deep learning model (top center) and Desroziers et al. (2005) method (top right). The bottom row displays differences in the predicted  $\mathbf{R}$  matrix and the true matrix, with maximum and minimum differences shown in the legend.



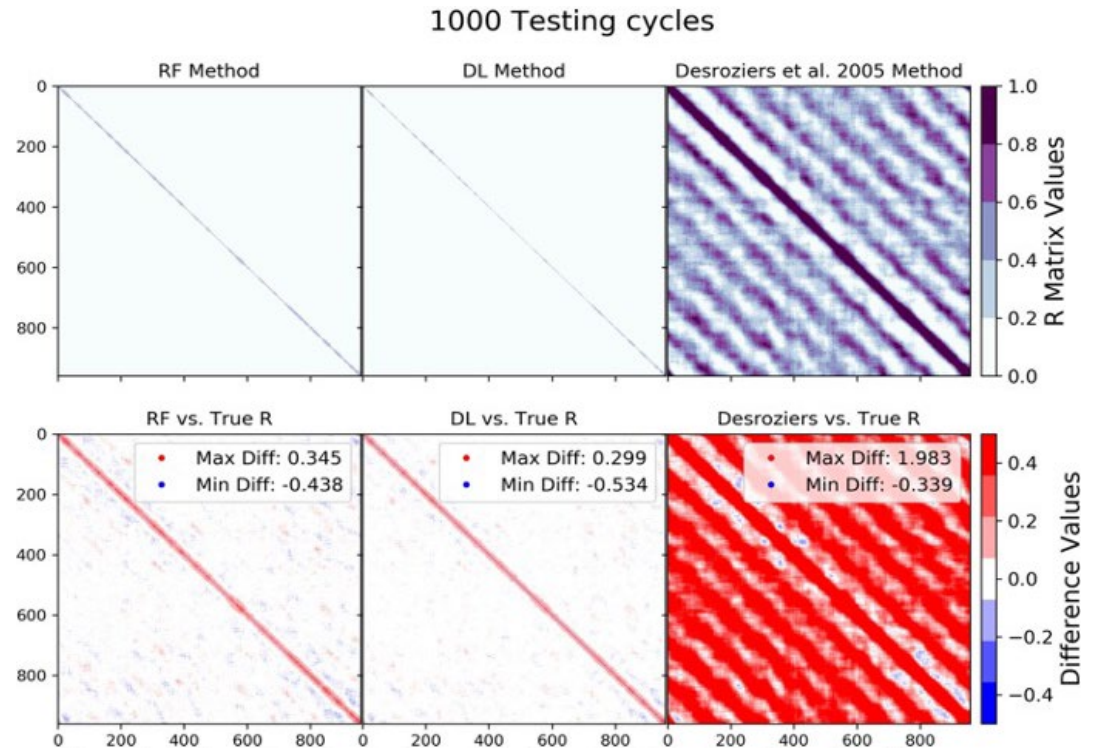
has slightly lower values on the diagonal than the identity matrix. The Desroziers et al. (2005) method produces the largest differences from the identity matrix, with higher predicted values on the diagonal. The larger errors in the Desroziers et al. (2005) method are likely due to a coarsely tuned multiplicative ensemble inflation, an overestimate in observation error variance is consistent with an under-dispersive ensemble (Waller et al., 2016). In general, all three methods correctly capture the identity matrix with varying small differences in predicted value along the diagonal.

As the machine learning methods were capable of reproducing the identity  $\mathbf{R}$  matrix under ideal conditions, we compared the three methods under scenarios where assumptions required by the Desroziers et al. (2005) method are violated. The deep learning model is trained over 30 epochs with 32 and 64 for each layer. The number of filters and epochs are greater than the first experiment to capture the increased data complexity. All three models are tasked with predicting the identity  $\mathbf{R}$  matrix,

with observations at every grid point, no correlation, and the errors which are static in time. To mimic model error, the DA system is cycled with a standard forcing value of  $F=15$ , however the nature run that computes the testing observations has a forcing value of  $F=10$ . The machine learning models are trained with the standard forcing value to see if the methods are able to adapt to the change.

Figure 4 shows the large differences between the machine learning methods and Desroziers et al. (2005). Both machine learning models generally capture the correct identity matrix, with higher values on the diagonal than the truth. The deep learning model shows smaller differences both for the diagonal and the off-diagonal elements compared to the random forest model. This result indicates that if appropriate training data is available, the machine learning models are able to adapt to discrepancies between the actual and assumed error statistics, in this case unaccounted for model error. The Desroziers et al. (2005) method fails to capture the identity matrix.

**Figure 4:** Similar to Figure 3 but predicting the identity  $\mathbf{R}$  matrix under changing model forcing conditions.



## Summary

Through this feasibility study we showed that machine learning and deep learning models can successfully approximate the observation error covariance matrix,  $\mathbf{R}$ , under relaxed assumptions. The caveat is that suitable training data must be available for the machine learning models. In other words, the Desroziers et al. (2005) method will suffer more from errors in the prescribed error covariance, while the machine learning models will suffer from inaccurate "true states" used to train the models. In future work, we will investigate this trade space by using independent analyses or merged data products as a proxy for the true states.

Finally, we only investigated spatially and temporally homogeneous cases, but found the largest errors with all three methods occur when introducing model errors. In future work we plan to investigate the ability of the machine learning models to adapt to the inhomogeneous case, as well as to perform quality control.

## Acknowledgements

The work reported here was performed under the ONR Naval Research Enterprise Internship Program (NREIP).

## References

- Bishop, C. H., Etherton, B. J., & Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review*, 129 (3), 420 - 436. Retrieved from [https://journals.ametsoc.org/view/journals/mwre/129/3/1520-0493\\_2001\\_129\\_0420\\_aswtet\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/129/3/1520-0493_2001_129_0420_aswtet_2.0.co_2.xml) doi: 10.1175/1520-0493(2001)129(0420:ASWTET)2.0.CO;2
- Breiman, L. (2001). Random forests. *Machine Learning*, 45 (1), 5-32. doi: 10.1023/A:1010933404324
- Desroziers, G., Berre, L., Chapnik, B., & Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of*



*the Royal Meteorological Society*, 131 (613), 3385-3396. doi: <https://doi.org/10.1256/qj.05.108>

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0888327020307846> doi: <https://doi.org/10.1016/j.ymssp.2020.107398>

Lorenz, E. N. (2005). Designing chaotic models. *Journal of the Atmospheric Sciences*, 62 (5), 1574 - 1587. Retrieved from <https://journals.ametsoc.org/view/journals/atsc/62/5/jas3430.1.xml> doi: 10.1175/JAS3430.1

Waller, J. A., Dance, S. L., & Nichols, N. K. (2016). Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Quarterly Journal of the Royal Meteorological Society*, 142 (694), 418-431. doi: <https://doi.org/10.1002/qj.2661>

---

#### EDITOR'S NOTE



**Welcome to the Spring 2021 issue of the JCSDA quarterly.** This is an exciting time of year in the Joint Center, as the Annual Operating Plan is completed and begins to be executed on April 1. The Director's staff, project leaders, executive team, and management oversight board have been busy aligning JCSDA activities for the coming year with strategic priorities and resources.

Of course, the scientific and technical work continued unabated even as the planning was being completed. The calendar of upcoming events presented in this issue indicates that the pace shows no signs of slackening as we move into the summer. An event of particular note is the annual science and technical workshop will be conducted in early June, using a virtual format. This will be an excellent opportunity to get a comprehensive view of the recent and on-going work in the JCSDA, made all the more compelling since our 2020 workshop was cancelled as a consequence of the COVID-19 pandemic.

Another recent opportunity for sharing work with JCSDA colleagues and the broader science community was the 9th JCSDA Symposium at the Annual American Meteorology Society (AMS) meeting in January, 2021. The authors of two of the contributions to the Symposium accepted the invitation to provide science articles for this edition of the newsletter. The first of these, by Wei Han and co-authors, describes challenges and opportunities associated with the assimilation of observations from hyperspectral infrared sounders in geosynchronous orbits. The second by Amanda Burke and co-authors, summarizes their feasibility study of using machine learning models to approximate observation error statistics to support data assimilation. These methods and these observations are expected to be of importance in the near future, and I am confident that you will find these articles of great interest and value.

Finally, a number of new colleagues and visitors have joined us since the last newsletter. Mini-biographies for Rachel Schweiker, Kriti Bhargava, Fabio Diniz, and Claude Gilbert are included in this issue to let you know a bit about their experience and the roles they are playing in the Joint Center, and a little about themselves, too. Please take time to read these, and to make them welcome, virtually, and I hope, in person before too long.

Jim Yoe, Editor

## NEWS ANNOUNCEMENTS

**2021 AMS Short Course: Introduction to the Joint Effort for Data assimilation Integration (JEDI)**

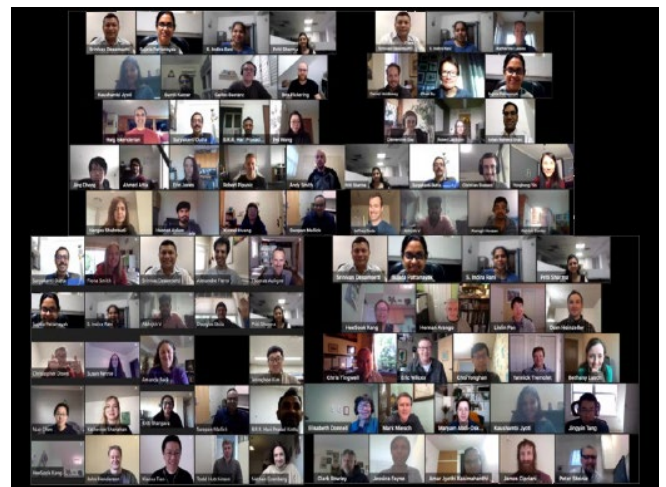
*By Kat Shanahan, JCSDA*

The Joint Effort for Data assimilation Integration (JEDI) project thrives through a partnership of core and community contributions. A strong community focus allows software engineers and scientists alike to download and run the public code, access support forums, and contribute to new developments. Sharing resources enables a faster transfer of research to operations so that JCSDA stakeholders and the data assimilation community can utilize JEDI in transitions to next-generation prediction systems. Ultimately, the hope is to leverage collaborative community development to craft a truly state-of-the-art unified data assimilation system.

National and international interest in effectively utilizing this cutting-edge system has increased over the past several years in anticipation of the first public release in October 2020. After the public release, desire for JEDI training opportunities has surged. Users and developers from the scientific to academic communities are eager to learn about how to apply JEDI to their own projects. To keep up with demand, the JCSDA has expanded support and training through the launch community-monitored forums, lecture videos on our [YouTube channel](#), more JEDI Academies, and a new JEDI Short Course offering at the AMS 2021 Annual Meeting.

The “Introduction to the Joint Effort for Data assimilation Integration (JEDI)” AMS short course took place last week on March 10, 2021. Spanning only four hours, the short course was designed to be a condensed version of what one might learn on the first day of a typical JEDI Academy. By the end of the short course, participants were able to understand how to get JEDI up and running, how to simulate observations with the Unified Forward Operator (UFO), and practice data assimilation with JEDI. The course consisted of four brief lectures and two practical sessions where attendees worked through tutorials and met with instructors one-on-one if needed.

Over 100 community members participated in the AMS short course, making it one of the largest training sessions the JCSDA has offered to date. Despite the large numbers and online format, attendees stated that the short course was valuable to their understanding of JEDI. The JCSDA is looking forward to hosting more “mini” JEDI Academies, including an academy for management, in the near future. To stay on top of future training opportunities, please sign up to receive news via [www.jcsda.org](http://www.jcsda.org).



*Several short course attendees pose for a virtual group photo.*

## PEOPLE



## Claude Gilbert

Claude Gilbert joined the JCSDA in August 2020 as a Senior Software Engineer working in the JEDI core team. “Senior” might refer to his age and not as much to his level of maturity. He will be focusing on providing end-users with tools to manipulate and archive data by using Research Repository for Data and Diagnostics (R2D2), to create, execute and reproduce research experiments with different workflow managers Experiments and Workflows Orchestration Kit (EWOK). His work with the JEDI team will also enable near real-time observation ingestion pipelines.

Following university where he studied mathematics and physics, Claude joined Meteo France where he was trained in digital electronics. After designing intel Z80/8085 motherboards, he maintained the software for a precipitation radar written in assembler without an operating system. He then introduced Object-Oriented Programming at the technical services in Meteo France using Turbo Pascal. At the same time, he taught C and C++ programming at evening courses for an Open University (CNAM). He then moved into the private sector, still in meteorology, when he travelled mostly Africa when he met his wife in Zimbabwe. The fact that neither of them could speak the other’s language did explain the lack of arguing and the quick marriage. Claude still claims that he hadn’t understood what the question meant, so he said “yes”. That was 30 years and two grown up children ago.

At age 28, Claude went back to University to do an MSc. in Human-Computer Interaction at the DeMontfort University in Leicester, UK, where he focused on Artificial Intelligence and published a paper in the American Journal of Artificial Intelligence (AAAI) on using the vertebrate immune system as a model for machine learning, particularly to build contents- addressable systems in imagery.

Later, joining the European Centre for Medium-Range Weather Forecasts (ECMWF), Claude automated the huge production of web site plots and developed a forecast verification package. He introduced Python at ECMWF in 2004 despite a lot of resistance, and built a Meteorological Python package called MetPy still used in their forecast verification system. In 2010 he joined GMAO at NASA to develop an ad-hoc verification and observation monitoring system. While working there, he adapted the ECMWF verification system for the Bureau of Meteorology in Melbourne.

For the next few years, he worked for different commercial companies and corporations as mostly Lead Enterprise Architect, Data Architect or Vice-President of Data, but he disliked not being close enough to the technical teams. Although company Christmas parties and senior managers leaving dos in London were really pretty excellent, it was time to go back to “doing.”

Claude claims to have a French national title in 4x100m relay in 1979. Since all this occurred before the invention of Radio, TV and Internet, we couldn’t verify this piece of information. Claude likes to play music, he is a bass player and has been a member of different rock/reggae bands. He really enjoys gigging.



## Fabio Diniz

Fabio Diniz joined the JCSDA in August 2020 as a Project Scientist working with the observations team. His role is to contribute to the development of Forecast Sensitivity-based Observation Impact (FSOI) capabilities in the Joint Effort for Data Assimilation Integration (JEDI) system.

Fabio is originally from the South of Brazil, where he earned a B.Sc. in Meteorology from the Brazilian Federal University of Pelotas (UFPEl) and was first introduced to data assimilation. Later he moved to the Southeast region of Brazil to pursue an M.Sc. in Meteorology at the Brazilian National Institute for Space Research (INPE), where he started to contribute to various data assimilation activities. At that time, INPE was developing a local ensemble transform Kalman filter (LETKF) which provided motivation for a M.Sc. degree dissertation on the development of an ensemble-based FSOI tool. This effort led him to visit NASA's Global Modeling and Assimilation Office (GMAO) and to earn a Ph.D. in Meteorology for investigating the impact of over 40 years of observations assimilated in the MERRA-2 reanalysis using an FSOI methodology.

Apart from science, Fabio enjoys spending time with his family and friends. He also enjoys outdoor activities and playing sports.



## Kriti Bhargava

Dr. Kriti Bhargava joined UCAR/JCSDA in February 2021 and will be working with the Sea ice Ocean Coupled Data Assimilation (SOCA) project that aims at developing marine data assimilation (DA) systems within the Joint Effort for Data Assimilation Integration (JEDI). She will be mainly working on the development of high-resolution regional ocean data assimilation in the context of hurricane forecasting with special focus on the implementation of the MOM6 interface to the regional North Atlantic domain of NOAA's Hurricane Analysis Forecast System (HAFS). She will be based in the NOAA-NCWCP building in College Park, MD that is, post-COVID.

Dr. Bhargava earned her PhD and MS in Atmospheric and Oceanic Sciences at the University of Maryland (UMD), College Park. For her PhD, she implemented an adaptive online scheme to estimate and correct systematic errors that could be used with the operational scale global models. She tested this technique using the NCEP's operational Global Forecast System (GFS). During and after her PhD, her research interests included strongly coupled data assimilation. After that she worked as a Postdoctoral Associate at the UMD, extending her methods to the ocean-atmosphere coupled system using the weakly coupled Climate Forecasting System. At UMD, she also collaborated with the Indian Monsoon Mission and investigated the impact of assimilating surface observation in the CFS system.

Kriti is from India where she earned her Bachelor's degree at the Indian Institute of Technology, Bombay (IIT-B) in Civil Engineering, with a minor in Energy Sciences. She is actively involved in alumni student mentorship programs, mentoring graduate students at UMD and undergraduate students at IIT-B. In her leisure time, Kriti enjoys hiking, camping, traveling, painting (especially acrylics), and baking/cooking.





## Rachel Schweiker

Rachel Schweiker joined the JCSDA in December, 2020 as a scientific visitor on the observations team. Her primary responsibility is to develop and maintain diagnostic data visualization tools to access the quality and impact of observations in JEDI applications.

Ms. Schweiker earned her Bachelor's degree in Marine Science with minors in Biology and Geography from Boston University in 2013. She spent the majority of her time in ecology and genomics labs including a tropical ecology semester in Ecuador, and took as many remote sensing courses as possible. After she graduated, she moved to Berlin, Germany to work in a diabetes research lab at the Max Delbrück Center for Molecular Medicine.

After realizing that her favorite part of research was data analysis and communication, she made a plan to pursue data visualization as a career back in Boston. She worked at several startups including Understory Weather, Zoba, and VocaliD, where she honed her R programming and graphic design skills.

Ms. Schweiker received her Master's in Statistics from Harvard University in 2017 after two semesters of study (she does not recommend this challenging fast track!) She also conducted several graphic design and data analysis projects during her graduate studies, and upon completing her degree she formed a data visualization company called Datascribe, LLC ([datascribeconsulting.com](https://datascribeconsulting.com)). She has worked with Airbus, L3Harris, GNS Healthcare, and Brigham and Women's Hospital to create data analysis tools and visualizations. Rachel is excited to join the JCSDA to work as part of a team again, and communicate the phenomenal work done here through data visualization.

Outside of work, Rachel likes to play with her cat, Daphne, and climb, trail run, and travel with her husband, Joel. They moved to Boulder just last year to be closer to the mountains. When obstacle course races resume, that's where you'll find them on the weekends!

## SCIENCE CALENDAR

MEETINGS OF INTEREST			
DATE	LOCATION	WEBSITE	TITLE
May 25-27, 2021	Virtual	<a href="https://www.events-force.net/osos2021">https://www.events-force.net/osos2021</a>	Second International Operational Satellite Oceanography Symposium (OSOS-2)
June 24-30, 2021	Virtual	<a href="https://cimss.ssec.wisc.edu/itwg/itsc/">https://cimss.ssec.wisc.edu/itwg/itsc/</a>	International TOVS Study Conference - ITSC-XXIII
September 13-18, 2021	Bonn, Germany	<a href="https://symp-bonn2021.sciencesconf.org/">https://symp-bonn2021.sciencesconf.org/</a>	Joint WCRP-WWRP Symposium on Data Assimilation and Reanalysis, hosted by DWD, in conjunction with the ECMWF Annual Seminar
September 20-24, 2021	Romania	<a href="https://www.eumetsat.int/eumetsat-meteorological-satellite-conference-2021">https://www.eumetsat.int/eumetsat-meteorological-satellite-conference-2021</a>	EUMETSAT Meteorological Satellite Conference 2021
October 18-20, 2021	Boulder, CO & Virtual	<a href="https://usclivar.org/meetings/reanalysis-2021">https://usclivar.org/meetings/reanalysis-2021</a>	Workshop on Future US Earth System Reanalysis



MEETINGS AND EVENTS SPONSORED BY JCSDA			
DATE	LOCATION	WEBSITE	TITLE
June 7-11, 2021	virtual	<a href="https://www.jcsda.org/18th-jcsda-technical-review-meeting-and-science-workshop-home">https://www.jcsda.org/18th-jcsda-technical-review-meeting-and-science-workshop-home</a>	18th JCSDA Technical Review Meeting and Science Workshop

## CAREER OPPORTUNITIES

Opportunities in support of JCSDA may be found at <https://www.jcsda.org/opportunities> as they become available.