Exploring the Use of Artificial Intelligence (AI) to Optimize the Exploitation of Big Satellite Data in NWP and Nowcasting

S. Boukabara, E. Maddy, N. Shahroudi, R. N. Hoffman, T. Connor, S. Upton, A. Karpovich, C. Sprague, and K. Kumar

NOAA/NESDIS Center for Satellite Applications and Research (STAR) , College park, MD, USA

Riverside Technology Inc. (RTI) @ NOAA/STAR, College park, MD, USA

University of Maryland (UMD), College Park, MD, USA

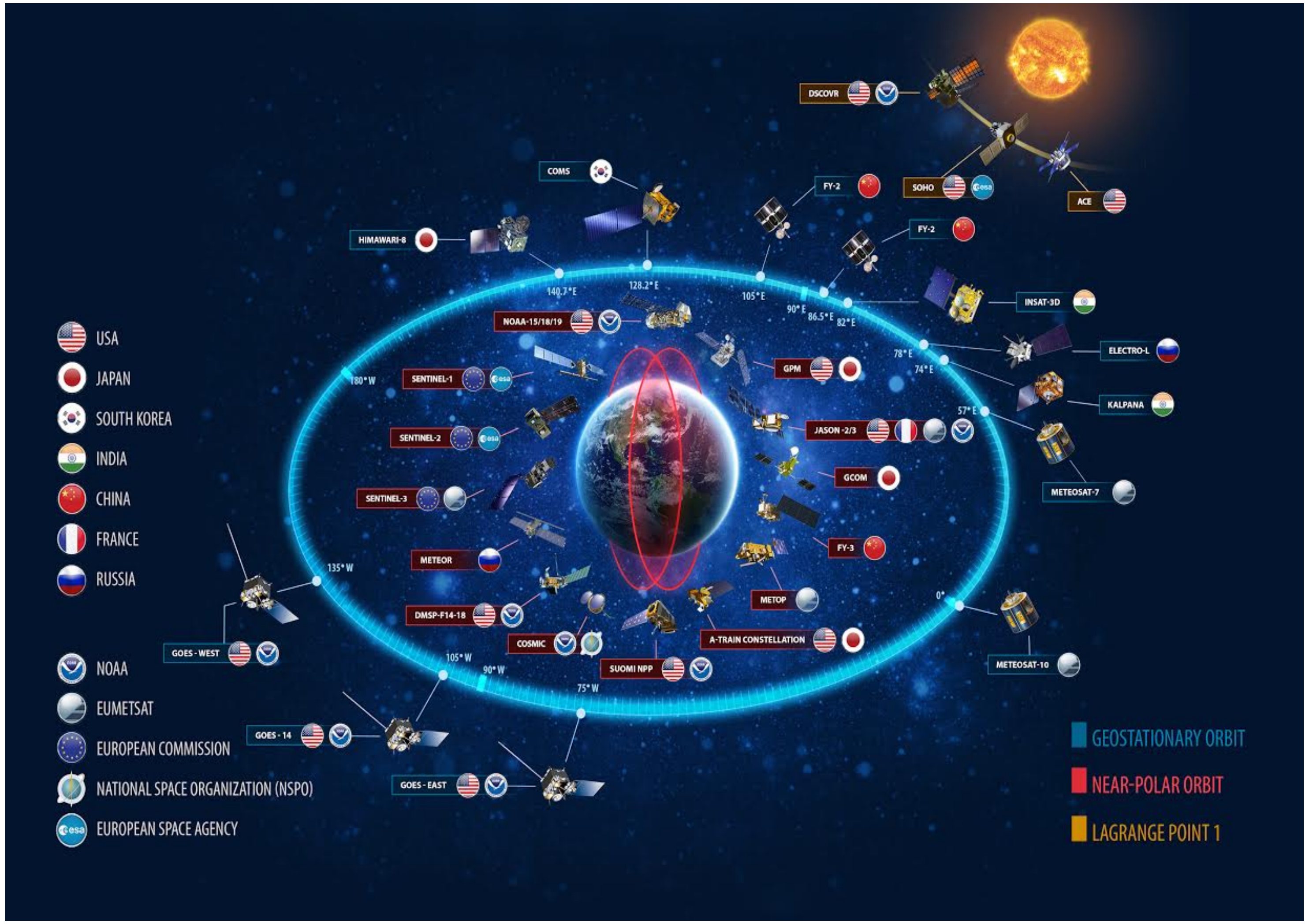
Agenda

1. **Why Artificial Intelligence (AI) ? Background and Motivations**
2. **Methodology & Description**
3. **AI for Remote Sensing and Data Assimilation/Fusion/NowCasting**
4. **Conclusions**

Trends in Global Earth Observation Systems

- **GOS Trends:**
  - New Players in GOS (international, commercial, etc)
  - New Sensors (higher resolutions, etc)
  - New technologies (small sats, etc)
  - Emergence of New GOS (IoT, etc)
  - **Significant Increase in volume and diversity of data**

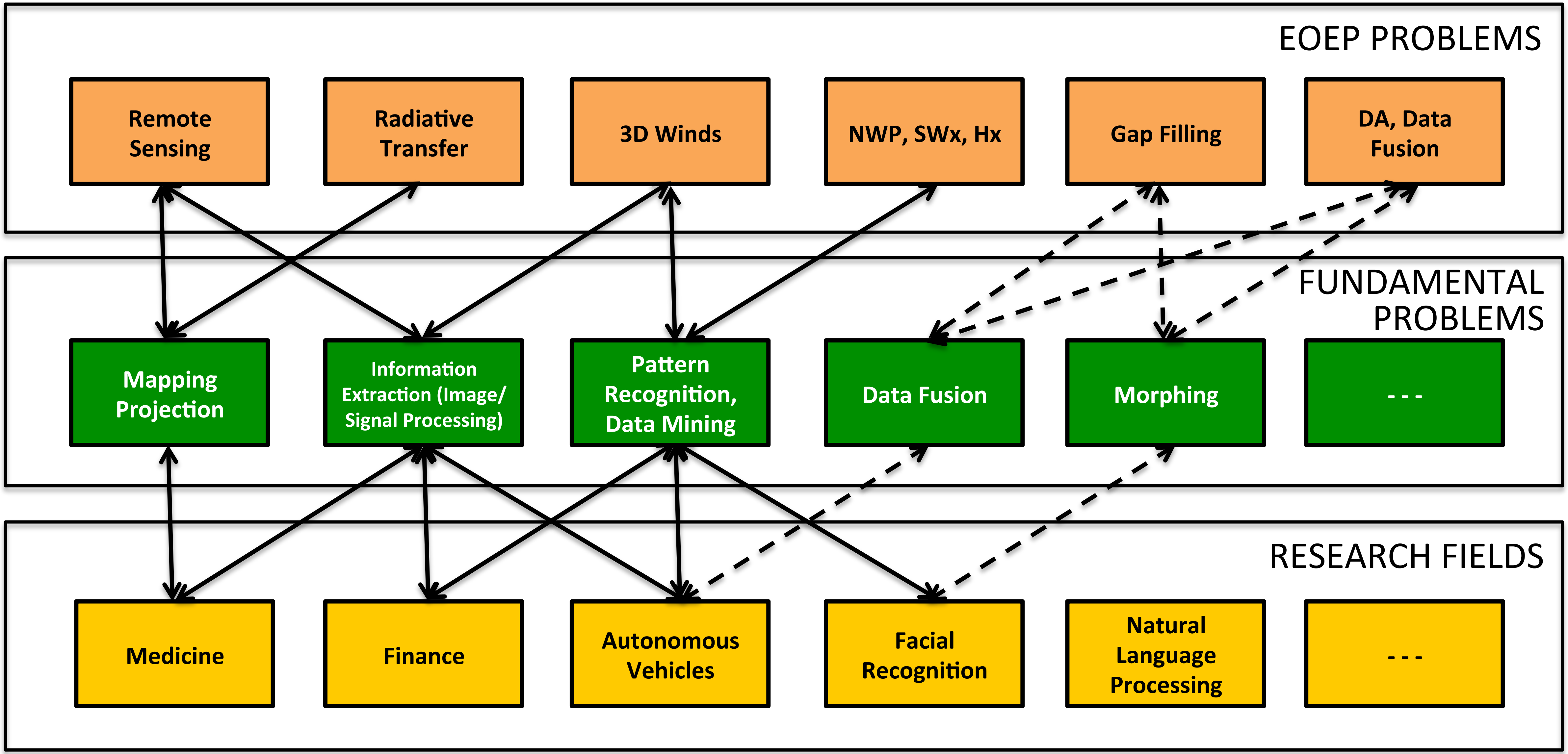
- **Parallel Trends**
  - Budget, HPC Constraints
  - Higher societal impact and expectations
  - Higher users expectations
  - Demand for Increase in quantity of data assimilated (5% currently assimilated)



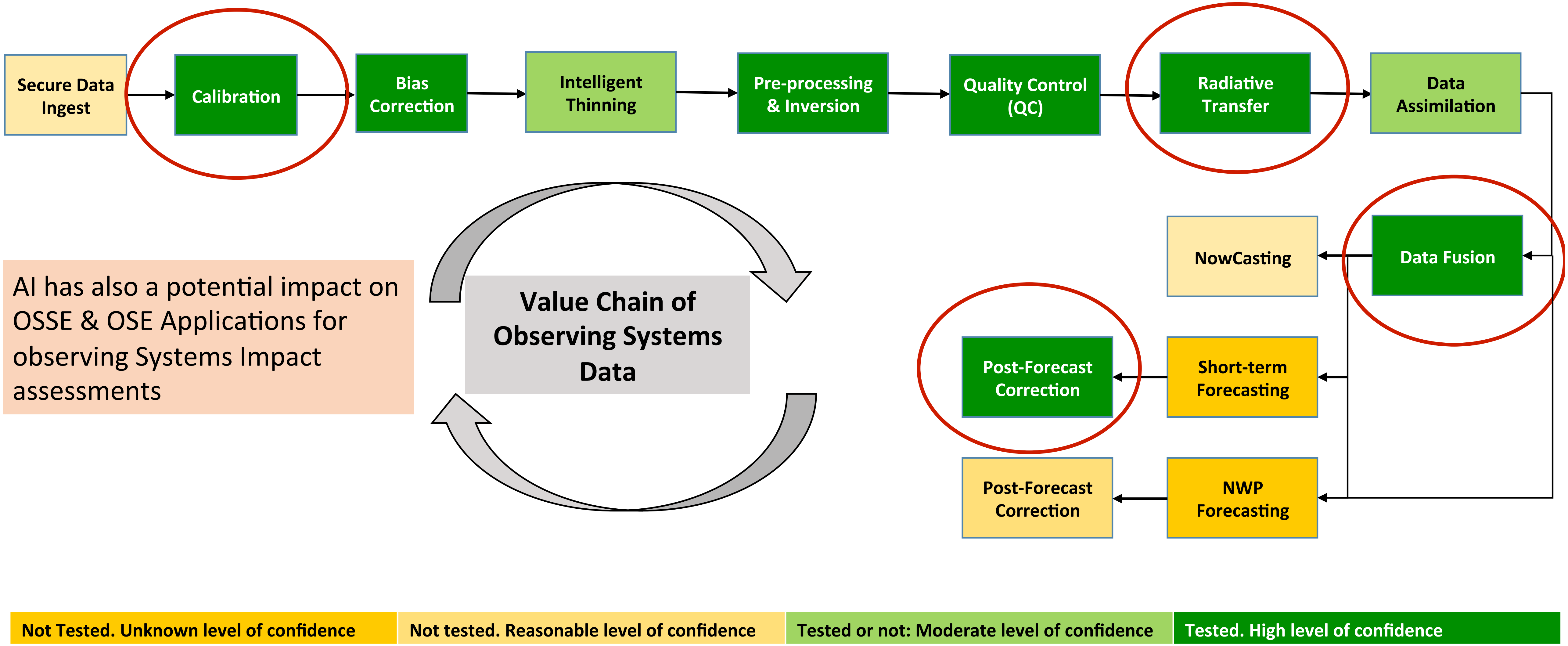
Why AI?

- **AI applied successfully in fields with similar traits as Environmental data & NWP/SA: (1) # obs. systems to analyze/assimilate/fuse and (2) predict behavior**
  - Medical field (Watson Project): Scan Image Analysis, Cancer detection, heart Sound analysis
  - In finance: Algorithmic Trading, market data analysis, portfolio management
  - In Music: Composing any style by learning from huge database & analyzing unique combinations.
  - Self-Driving Transportation Devices: Fusion of Multiple Observing Systems for situational awareness
  - …..
- **We believe Environmental data exploitation (remote sensing, data assimilation and perhaps forecasting), presents a viable candidate for AI application.**
- **This presentation is meant to present a few examples to convey that the potential is significant.**

Meta-Transfer Learning



Exploring AI for Remote Sensing, NWP & Situational Awareness (SA). Status



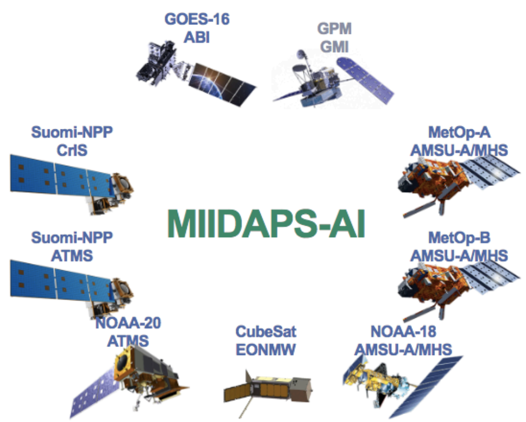
Methodology and Description
(baby steps)

- **Scope of the effort: Nowcasting/RS and Forecasting Adjustment**
  - focus on satellite-based analyses (RS), focusing on an enterprise algorithm used for inversion and assimilation pre-processing
  - but also assess capability of short term forecast correction
  - focus on atmosphere (T, Q, Wind) but highlight surface parameters and hydrometeors capability as well
- **Tools:** Google TensorFlow
- **Real data**
  - Focus on SNPP/ATMS and SNPP/CrIS

Training & Verification:

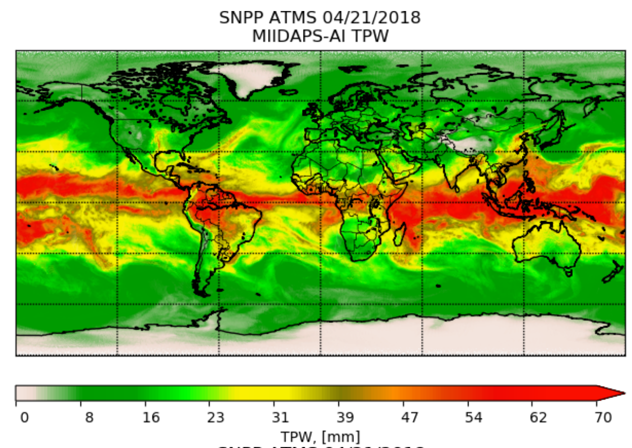
- Sets: ECMWF Analyses, G5NR fields, GDAS Analyses
- Noise addition: uncorrelated, Gaussian distributed noise with spread of (instrument noise*2) is added to simulations
- Sampling: Training data is randomly selected from a fixed set of ~5% of a days worth of data in each training epoch
- Simple training (sample over a day generally
- Independent sets used for verification, but still the same period

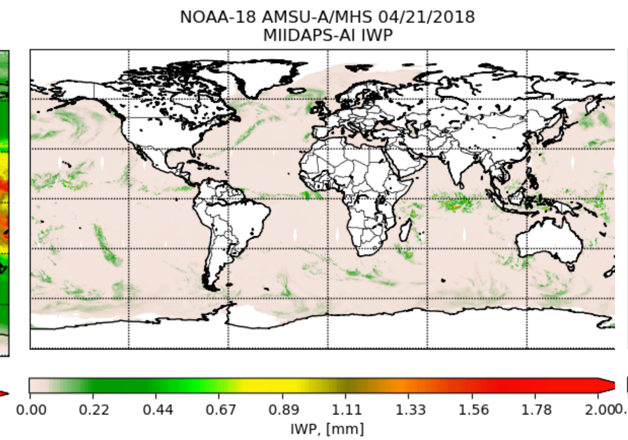
MIIDAPS-AI Product Examples – Real Polar Geo IR and MW Observations

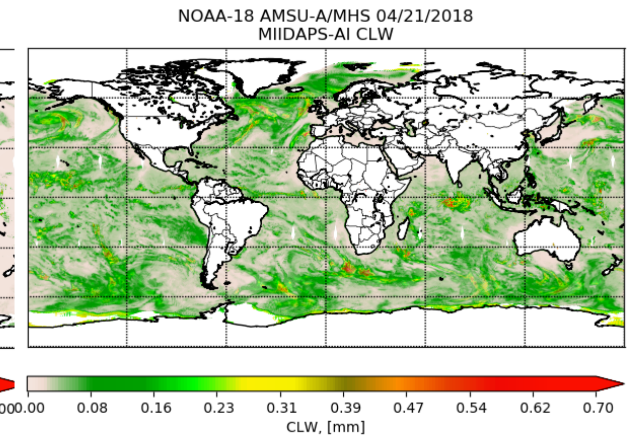


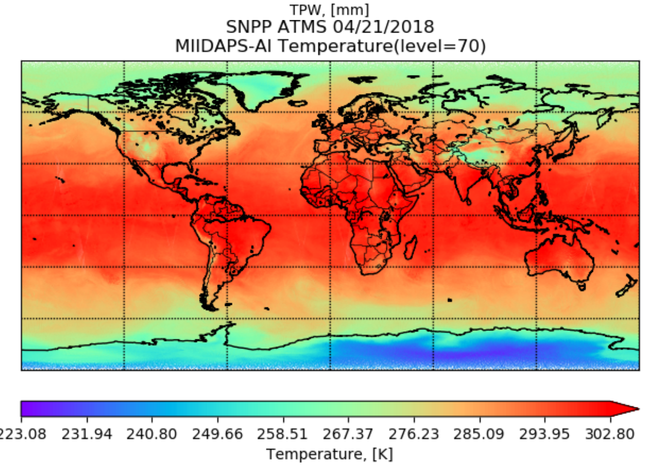
MIIDAPS-AI Remote Sounding Algorithm

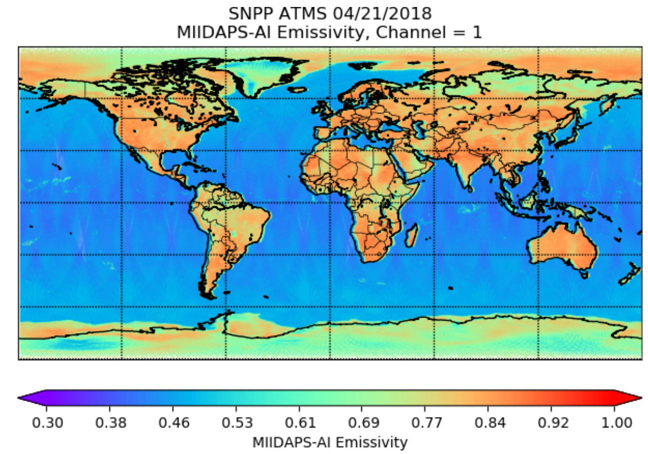
- *Algorithms are deep feed forward (and locally-connected) neural networks trained in simulation and applied to real observations*
- *Network architecture and hyper parameters are tested and optimized using Google TensorFlow™ and Keras*



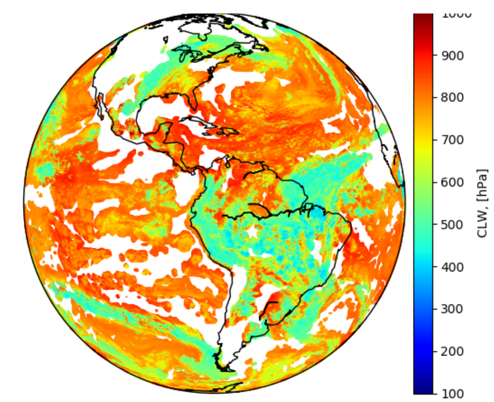








GOES-EAST ABI LIQUID CTP

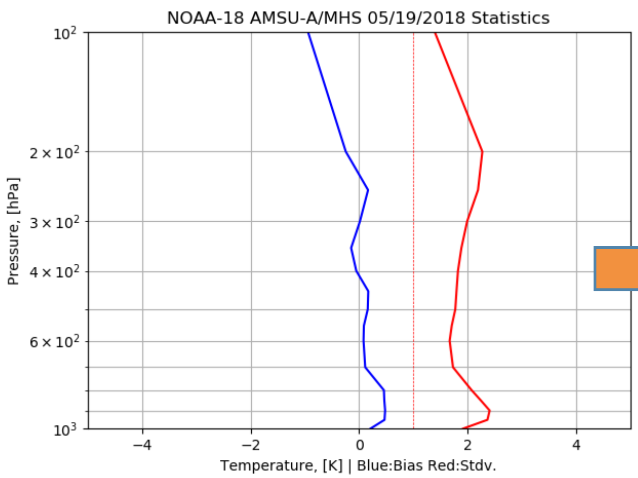


Can AI be used to perform bias correction of products/instrument observations?

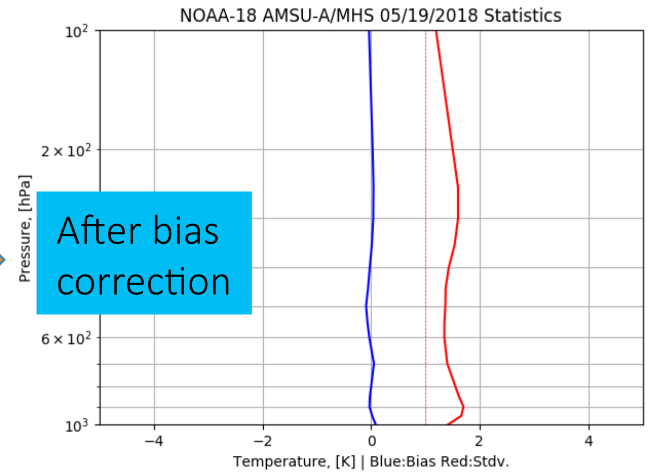
Predicted Temperature and Water Vapor profiles vs. ECMWF

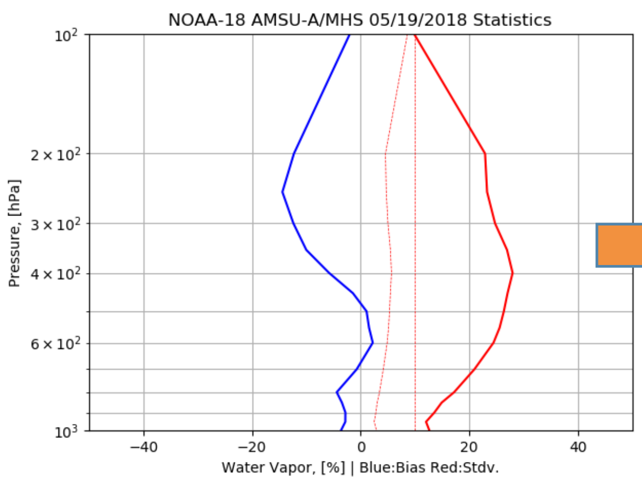
Inputs: AMSUA/MHS TB
Outputs: T, Tskin, Q

Inputs: AMSUA/MHS TB
Outputs: T, Tskin, Q, BC(TB_{sim}-Tb_{real})

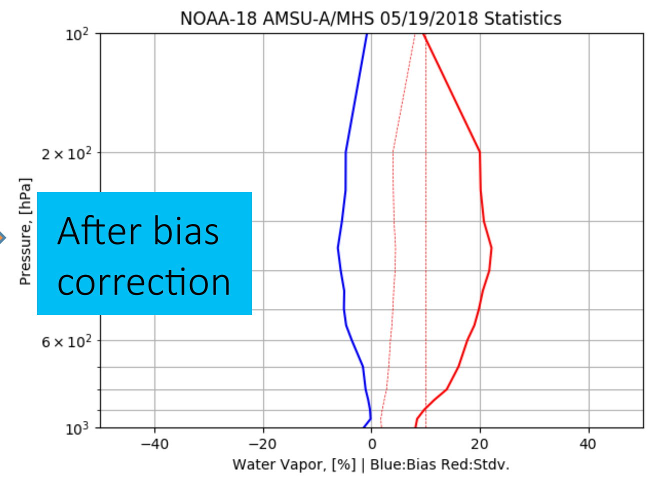


After bias correction



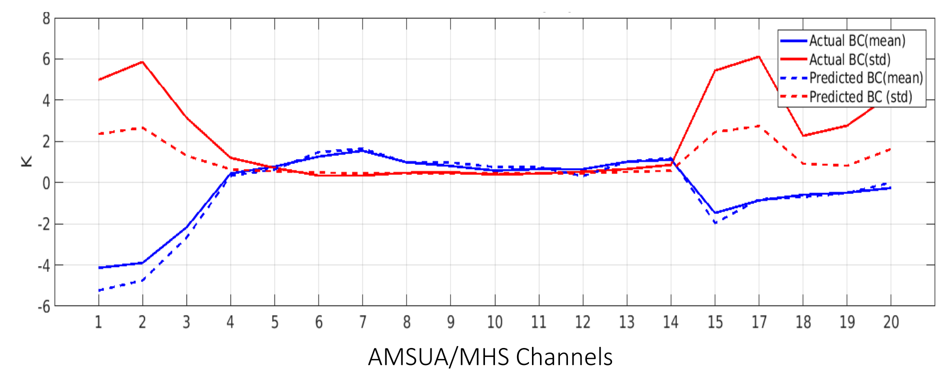


After bias correction



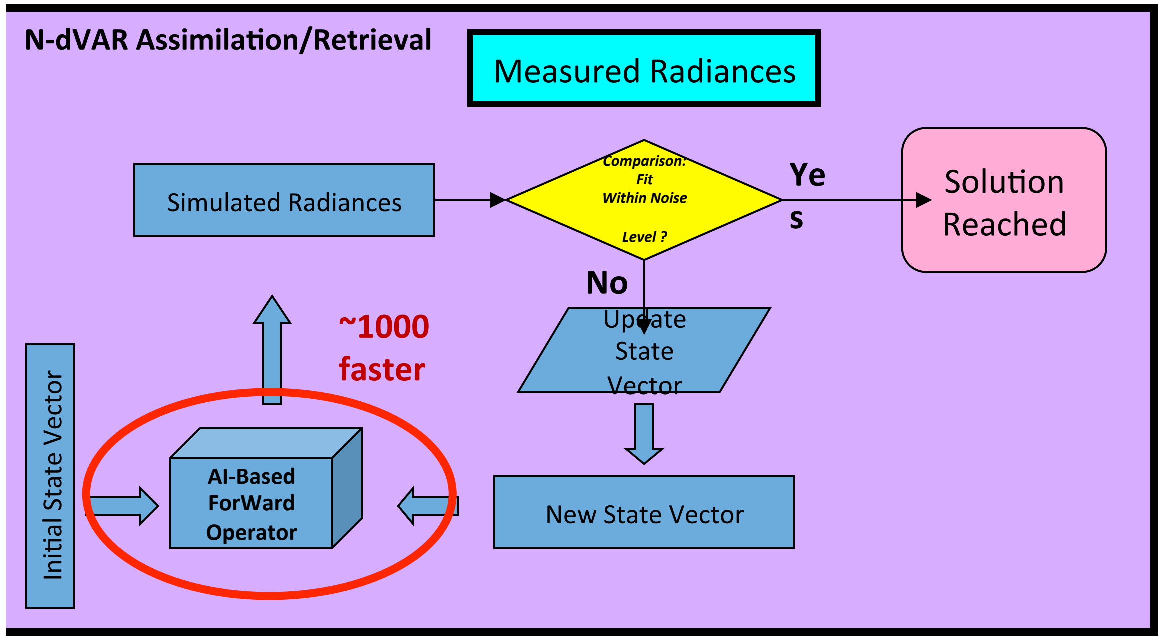
Adding ***BC*** into the training set improves the bias for both temperature and water vapor profiles

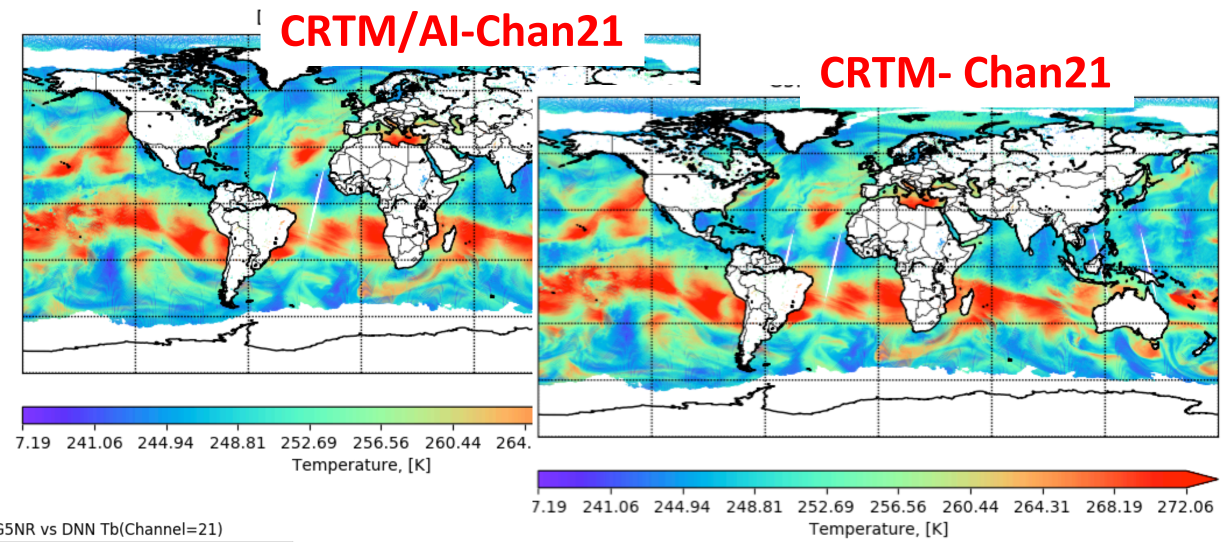
Predicted BC vs. Actual BC(TB_{sim}-Tb_{real})

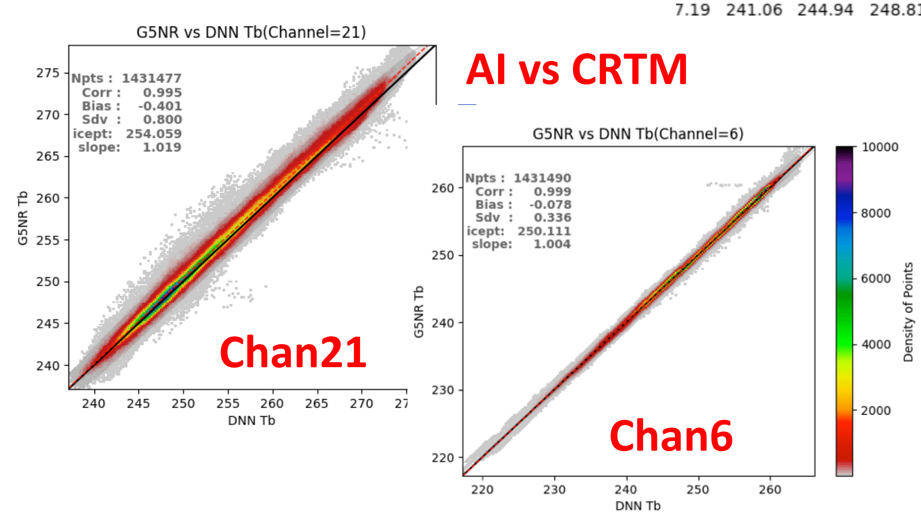


Can AI Be Used as Forward Operator?

Use of Deep Neural Network (DNN) for Radiative Transfer Modeling Purposes

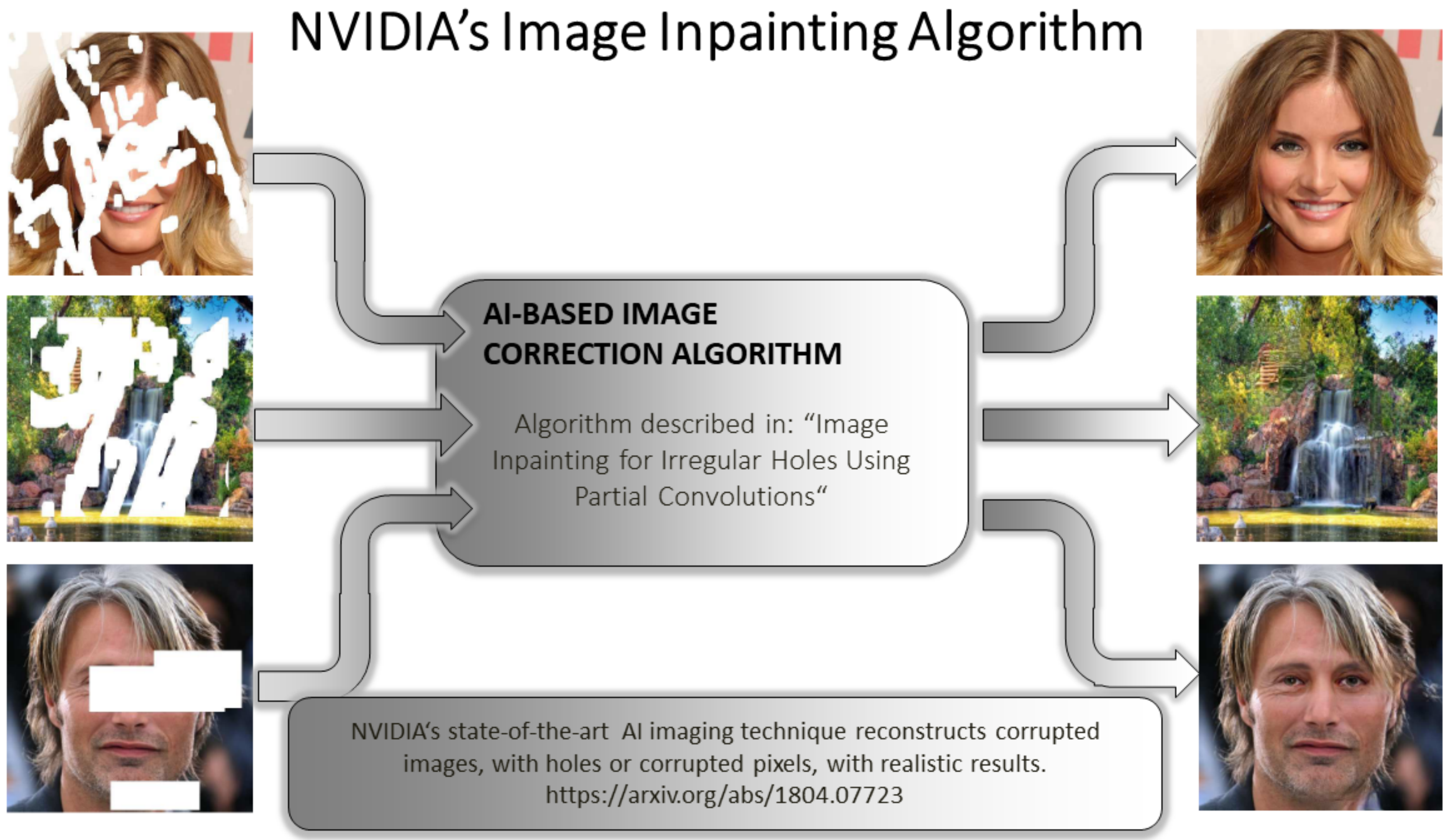




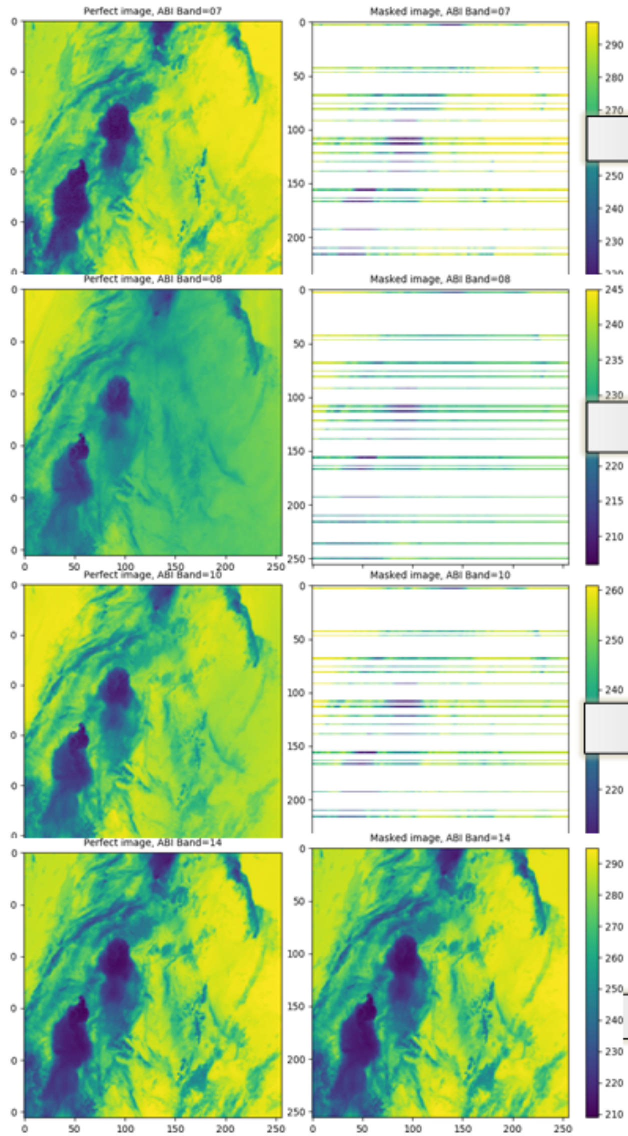


	CRTM-AI	CRTM
Processing Time for a full day data. A single sensor channel(ATMS). Excluding I/O	**<1 second**	**~ 1.3 hours**

Can AI be used to perform calibration correction for GOES-17 ABI thermal bands?



Can AI be used to perform calibration correction for GOES-17 ABI thermal bands?



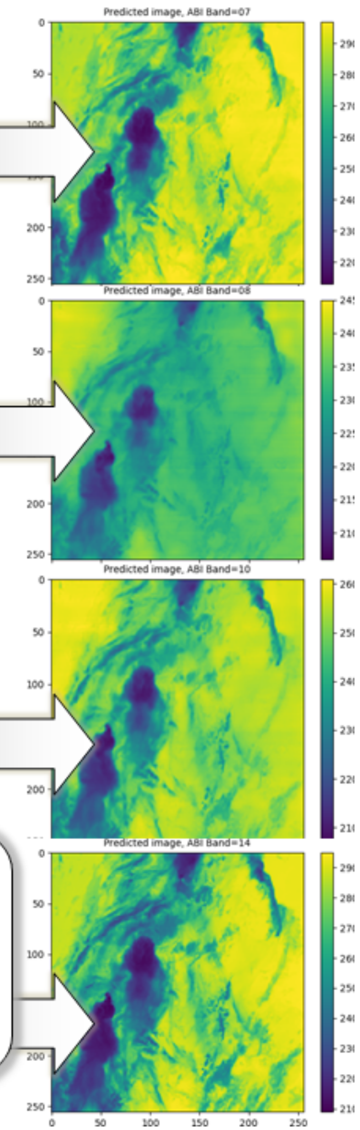
Proposed GOES-17 ABI Thermal Band Calibration Correction

AI-BASED IMAGE CORRECTION ALGORITHM

Correction based on NVIDIA's inpainting algorithm but applied for 256 x 256 pixel ABI thermal band images. Partial convolutions not implemented. Different loss fnc., architecture, hyperparams…all need tuned

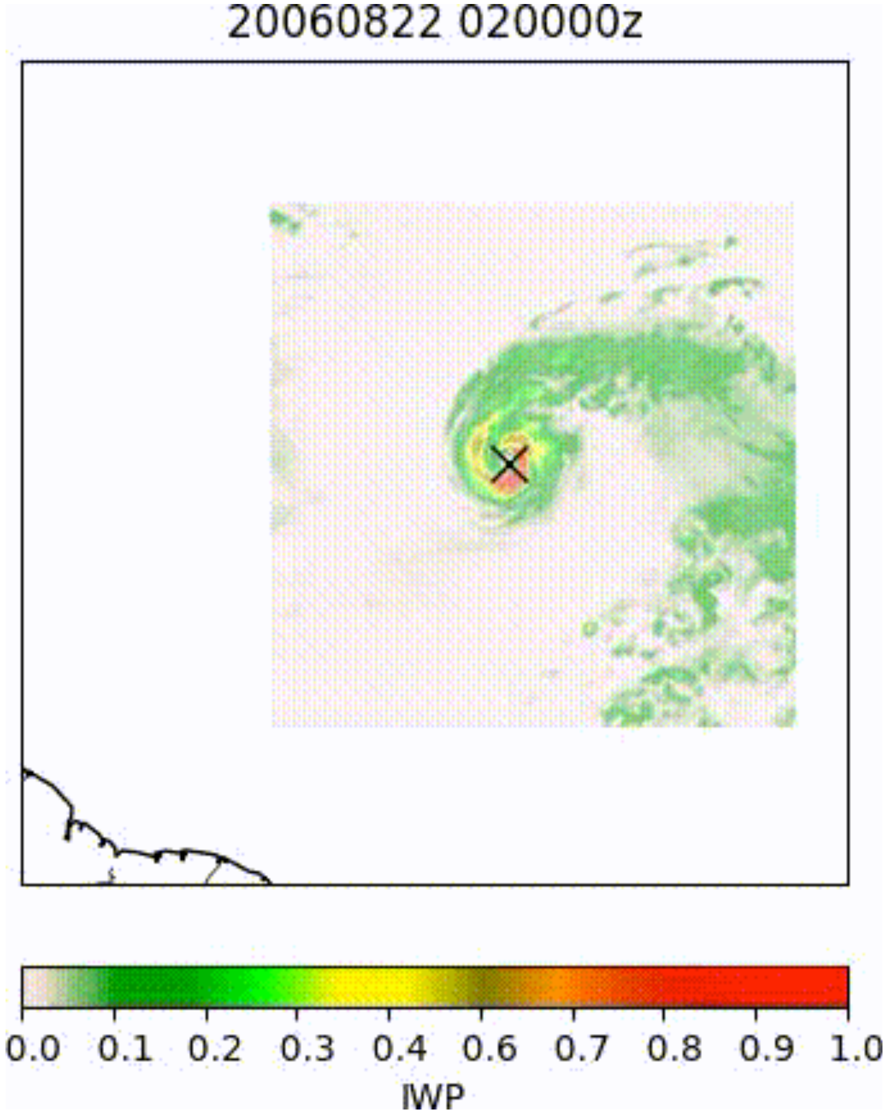
AI-based correction produces realistic images in corrupted bands.

Corrections can be applied when all thermal bands are corrupted or when only a subset of bands are corrupted. In the latter case, image reconstruction improved due to spatial/spectral redundancy.



Use of "morphing" AI Tool ("dogs" video morphing software) for Cloud/Precip morphing

Note the potential for morphing both the shape and color (i.e. equivalent of track and intensity)



- Used total integrated cloud ice from NASA GEOS-5 Nature Run (G5NR) "AL01" tropical cyclone at two time-steps (0200z and 0600z).
- Morphing software applied as a black box with some hand tuning of transformations between the two images.
  - Image at right sampled using 20 transformations between images

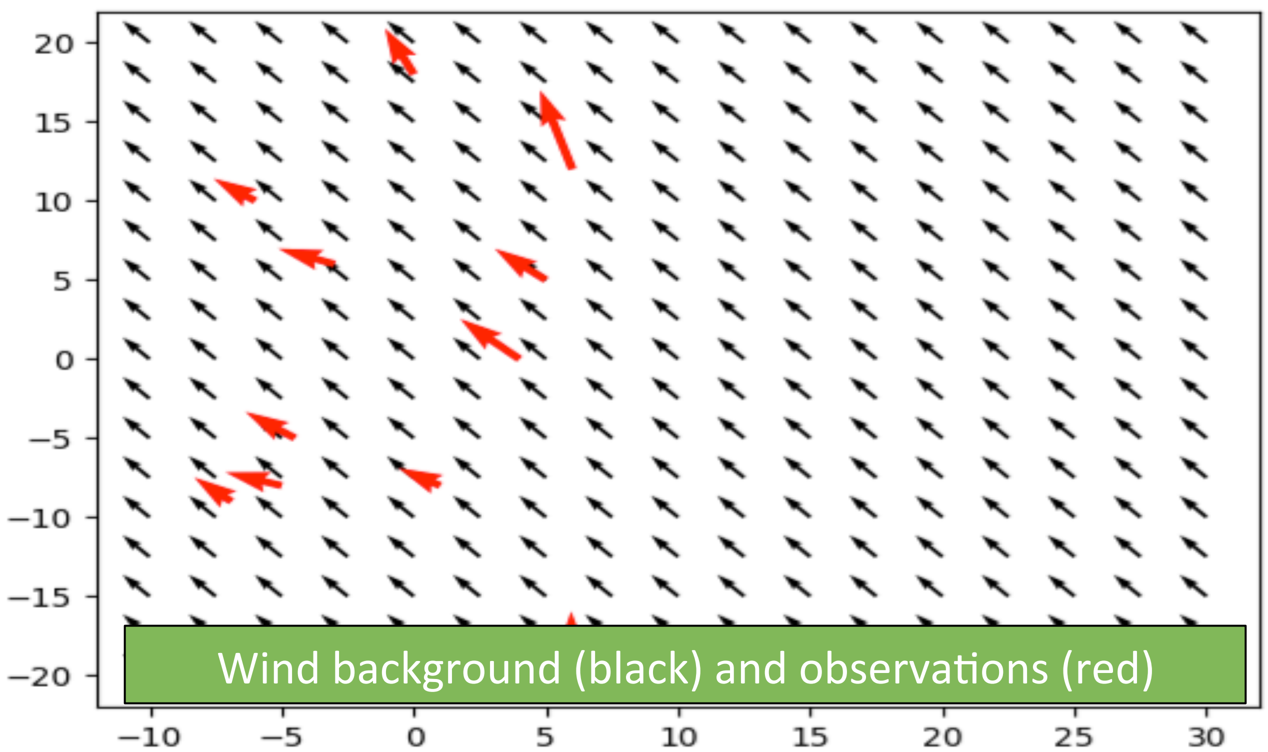
Credit: Example output and software from:
http://andrew.gibiansky.com/blog/image-processing/image-morphing/

Can AI Tools Be Used for Data Fusion & Data Assimilation?

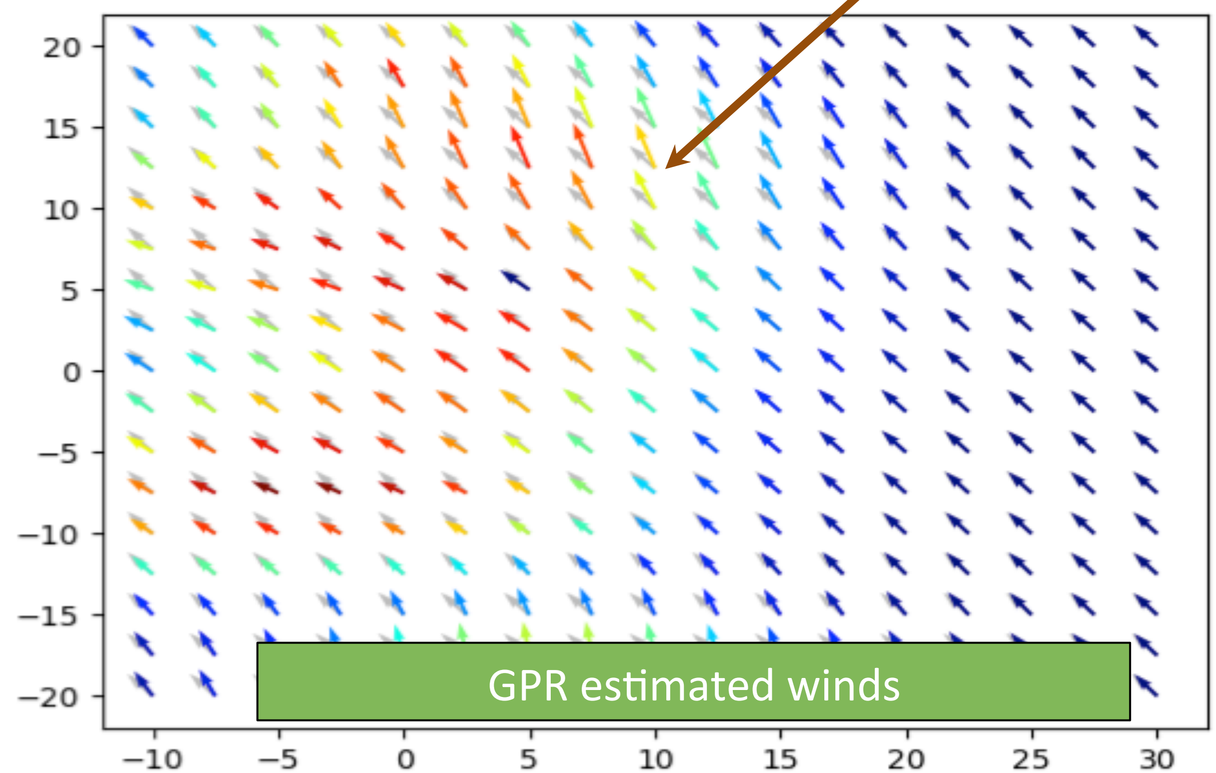
Use of GPR (Gaussian Process Regression) AI Model for Data Fusion/Assimilation (Case of AMV)

Color confidence/error estimates

Background and Measurements



GPR-Based Analysis

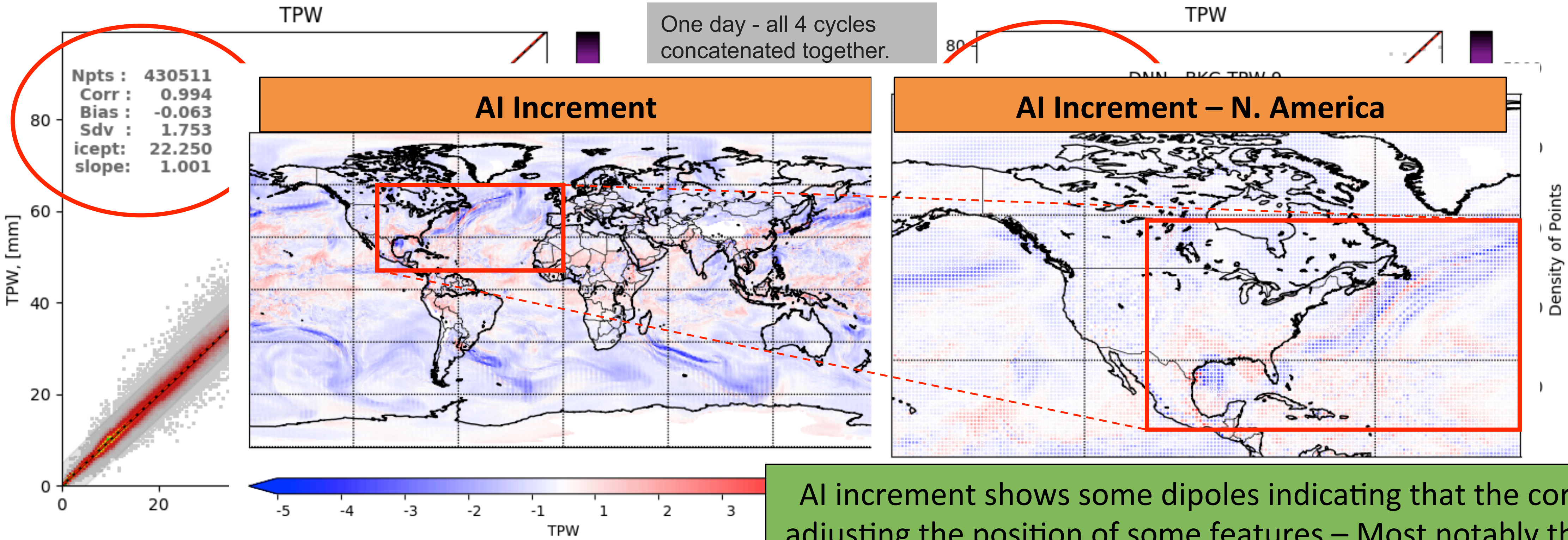


- Synthetic wind observations (red) are injected onto background (black) fields and GPR used to “fuse” the two.
- Color code corresponds GPR confidence – warmer colors reflect high confidence, while colder colors reflect low confidence estimates – and are consistent with observation locations.

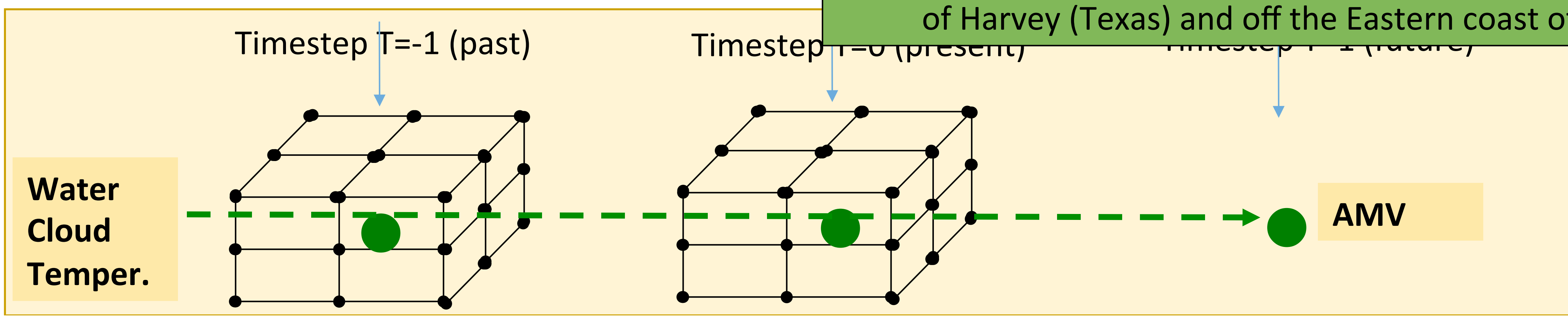
Correcting TPW Forecasting with AI?

ECMWF vs AI-corrected 6h fcst valid @ECMWF analysis time

ECMWF vs 6 hr frcst valid @ECMWF analysis time.



AI increment shows some dipoles indicating that the correction is adjusting the position of some features – Most notably the position of Harvey (Texas) and off the Eastern coast of N.America



Conclusions

- Increase in number, diversity and sources of global observing systems (GOS) including private sector. This presents unprecedented (and welcome) added resiliency and quality of the GOS. However this presents challenges: Cost and infrastructure to leverage/exploit them.

- Computing constraints, perhaps require us to explore new approaches for the future (not so distant). AI-Based Analyses (satellite-exlusive) are found to be radiometrically, spatially and geophysically consistent with traditional analyses.

- Goal of this study is not to show AI can do better, but that it can provide at least similar quality, much faster. It appears to be doing that.

- Different components can benefit from AI (Inversion, Data Assimilation, RT, QC, Data Fusion,..) for NWP and Situational Awareness SA.

- Encouraging results so far were found when assessing derivation of AMV using AI (not shown) and when assessing the feasibility of correcting GFS forecasts (using ECMWF as a target). Pointing to the potential for using AI for actual forecasting (at least short-term).

- Training is key for AI. Nature Run Datasets presents a good source for this.

- Pursuing AI applications, we believe, will allow us to :
  - (1) Reduce pressure on Infrastructure (ground systems), HPC and cost
  - (2) benefit from new environmental data (and face increased volume), including satellite data from all partners, including IoT
  - (3) Improve Latency
  - (4) Reduce cost of running legacy systems (remote sensing and data assimilation/fusion systems)
  - (5) Increase percentage of satellite data being assimilated (improved thinning, QC-ing, faster processing, etc)
  - (6) Reduce time to run OSE/OSSE and in general data assimilation/fusion systems, for decision making purposes
  - (7) <u>Perhaps</u> Improve forecast as a result of above and because AI can be exploited for forecast improvement