# Hyperspectral Sounder Spectral Fingerprinting: Using Machine Learning Techniques to Enhance Model-Based Physical Inversion

Wan Wu[1], Xu Liu[1], Liqiao Lei[3], Xiaozhen Xiong[1], Qiguang Yang[3], Qing Yue[4], Sun Wong[4],Tao Wang[4], Lihang Zhou[2] , Daniel K. Zhou[1], and Allen M. Larar[1]

[1] NASA Langley Research Center, Hampton, VA 23681, USA

[2] NOAA JPSS Program Office, Lanham, MD 20706, USA

[3]ADNET Systems Inc., Bethesda, MD 20817, USA

[4]Jet Propulsion Lab, California Institute of Technology, CA, USA

# Outlines

- General overview of standard Level 2 data production algorithms of hyper-spectral IR sounders (e.g. IASI, CrIS, AIRS …)

- Performance criterion of the retrieval algorithms:
  - Radiometric consistency
  - Data procession speed
  - How to handle 'cloud', spatial resolution …

- Single Field-of-view Sounder Atmospheric Product (SiFSAP)

- Spectral fingerprinting methodology and applications

# Iterative Optimal Estimation vs Machine-Learning

| RTM based iterative method | Machine-Learning based method |
|---|---|
| Establishes radiometric consistency between retrieved results to sounder spectral radiances through radiative transfer models * | Performance is highly dependent on the quality and extent of the training data. Does not establish radiometric consistency. |
| Offers well-defined error estimates based on the physics of the retrieval process | Error estimation is based on statistics of the training data instead of the physics |
| | |
| Computationally expensive due to the iterative nature of solving the radiative transfer equations | Computationally fast by avoiding the iterative radiative transfer simulation |

**Climate trend analysis imposes stringent requirement on radiometric consistency and the need for error estimation;** The challenges in processing huge amount of satellite data necessitate a low-latency data production scheme.

**Weather applications benefits from the fast procession capability provided by the machine learning techniques;** However, the radiometric consistency-based quality control is critical for the accuracy.

# Review of legacy algorithms

- Optimal estimation based physical algorithms
  - NOAA Unique Combined Atmospheric Processing System (NUCAPS)
  - Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS)
    - use MERRA data to provide the FG and a priori constraint
  - Climate Heritage AIRS Retrieval Technique (CHART)
    - use results from the MIT Lincoln Lab neural network algorithm as the FG

- Neural network or regression based fast algorithm
  - IASI operational algorithm (Piece-Wise Linear Regression - PWLR)
  - University of Wisconsin-Madison Space Science and Engineering Center Dual-Regression algorithm
  - MIT Lincoln Laboratory Stochastic Cloud Clearing/Neural Network algorithm (SCCNN)

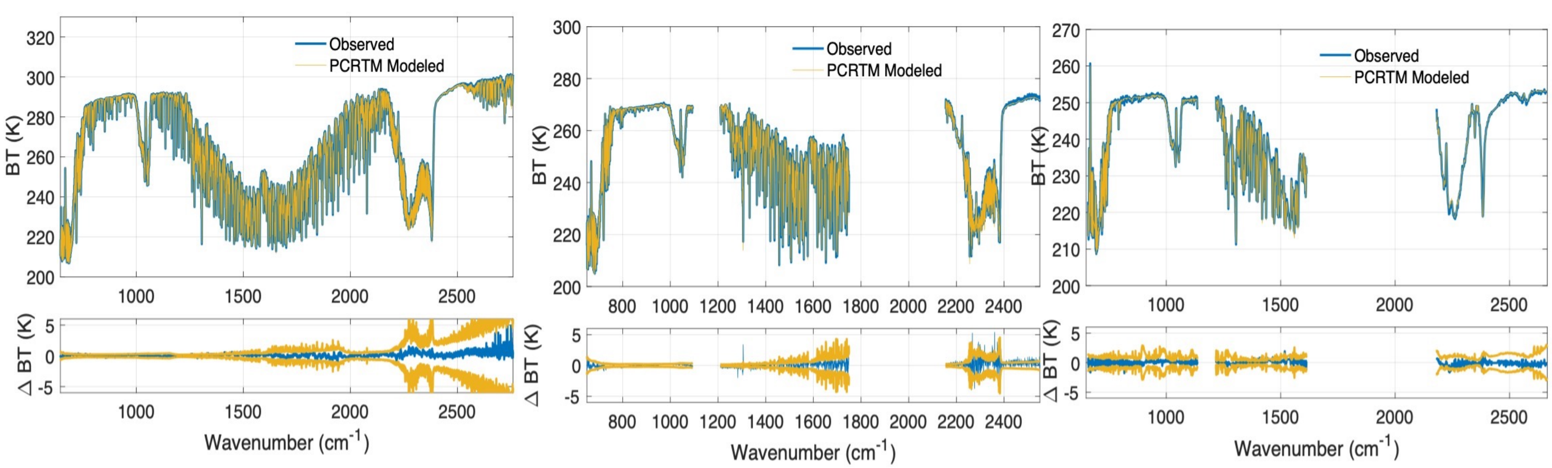
* NUCAPS, CLIMCAPS and CHART all use "**cloud cleared**" radiances , consequently

- **The consistency is established between retrieval results and cloud-cleared radiances, not the direct sounder observations.**
- **The spatial resolution is coarser than the native spatial of resolution of the sounder instrument.**

# Single Field-of-view Sounder Atmospheric Product (SiFSAP)

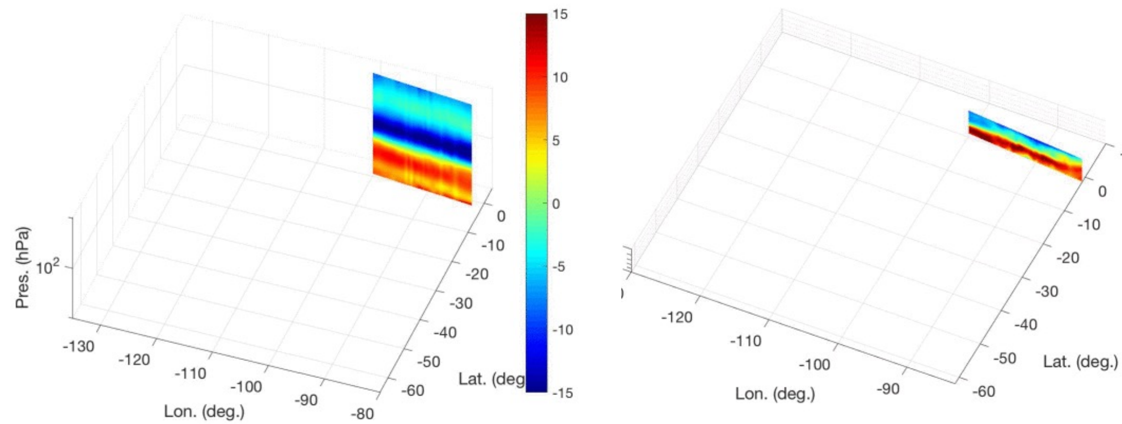
- The PCRTM based Single Field-of-view Sounder Atmospheric Product retrieval system bas been developed.
  - Include cloud scattering in forward simulation and the retrieval process to avoid using the cloud-clearing technique;
  - Directly fit the sounder measurements under all-sky condition to produce single FOV results
  - Minimize the constraints imposed by a priori by using climatology based a priori.
  - Unit data production speed on par with NUCAPS and CLIMCAPS, but produce **9 time more data** (spatial resolution matches to the native resolution of the sounder instrument).

IASI CrIS AIRS

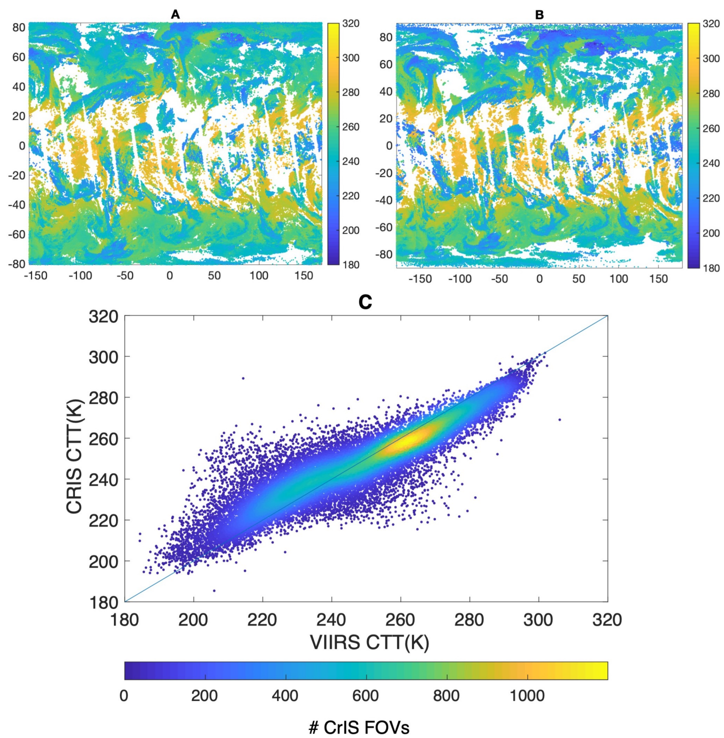


# SiFSAP of SNPP CrIS+ATMS

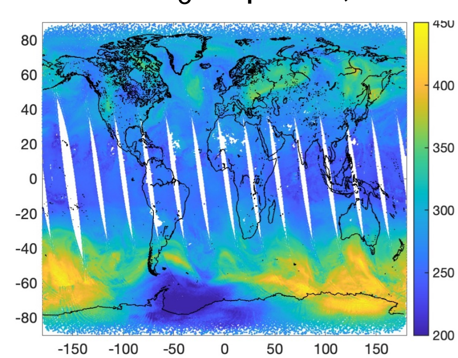
## 3-D atmospheric features revealed by SiFSAP CrIS T and Q



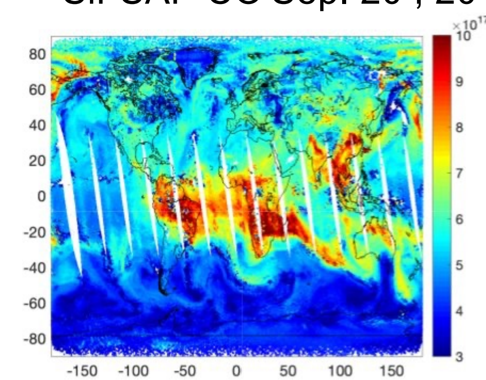

## SiFSAP CrIS vs VIIRS cloud effective Temp.




## SiFSAP $O_3$ Sep. 20 , 2019



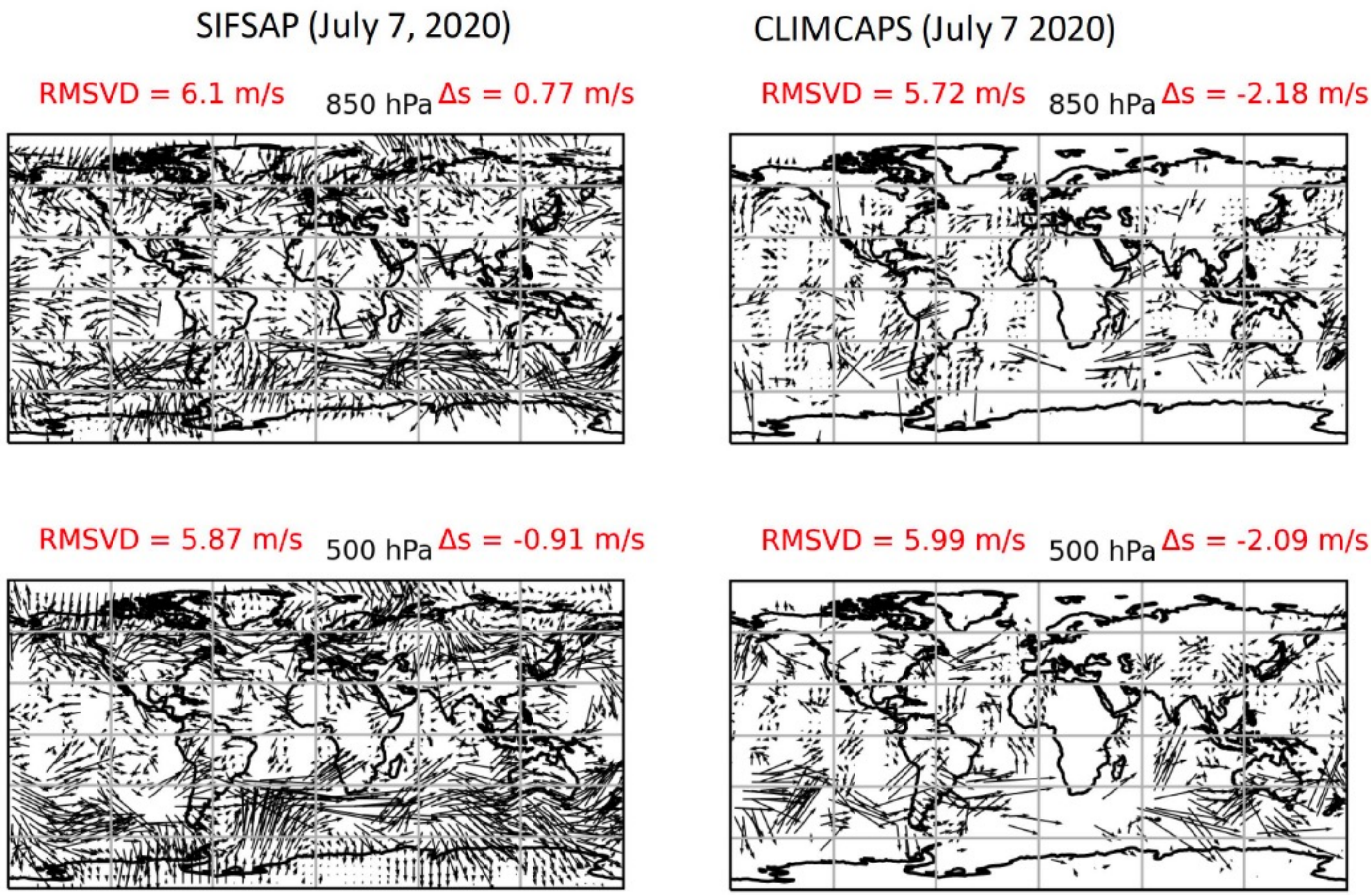
## SiFSAP CO Sep. 20 , 2019



# Example of 3-D Atmospheric Wind Vectors Derived from SNPP/CrIS and NOAA20/CrIS SiFSAP(Univ. Arizona)

Retrievals of 0.25 deg winds from SIFSAP (14 km 3D water vapor) are smoother and with lower speed bias (Δs) than CLIMCAPS winds (1 deg)

Courtesy of Amir Hassan Ouyed Hernandez and Xubin Zheng
Univ. Arizona



# A machine learning enhanced physical inversion scheme – spectral fingerprinting

$$\boldsymbol{r} = \mathrm{F}(\boldsymbol{x})$$

$$\mathbf{K} = {}^{\mathrm{dF}(\boldsymbol{x})}/_{\mathrm{d}\boldsymbol{x}}\Big|_{\boldsymbol{x}=\boldsymbol{x_0}}$$

$\mathbf{K}$ – spectral fingerprinting kernel

$$\triangle \boldsymbol{r} = \boldsymbol{r} - \boldsymbol{r_0}$$

$$\triangle \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x_0}$$

$$\boldsymbol{\Delta r} = \mathbf{K}\boldsymbol{\Delta x} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\Delta x} = (\mathbf{K}^T\boldsymbol{\Sigma}^{-1}\mathbf{K} + \boldsymbol{S_a^{-1}})^{-1}\mathbf{K}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Delta r}$$

Construct a comprehensive reference database consisting of $\boldsymbol{r_0}, \boldsymbol{x_0}$ and $\mathbf{K}$. Spectral fingerprinting kernels $\mathbf{K}$ are simulated using the state-of-art RTM (PCRTM) for each sample of $\boldsymbol{x_0}$.

For a given sounder observation $\boldsymbol{r}$, using supervise machine learning techniques (e.g. K nearest neighbor ) to identify spectral radiance reference state $\boldsymbol{r_0}$, geophysical reference state $\boldsymbol{x_0}$, and corresponding spectral fitting uncertainty $\boldsymbol{\Sigma}$ and a priori constraint $\boldsymbol{S_a}$.

Find the solution $\triangle \boldsymbol{x}$ using the liner inversion scheme within the optimal estimation framework.

- **Efficiency : Reduce iterative minimization steps which involves computationally intensive radiative transfer calculations;**
- **Accuracy : Achieve radiometric consistency by using model-based kernels and the optimal estimation scheme.**

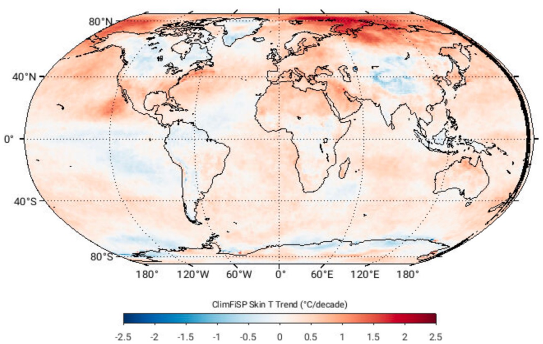
# Application Example

- **Updated version of Single Field-of-View Atmospheric Sounder Product (SiFSAP)**
  - Using machine learning based spectral classification to improve the spectral fitting constraint.
  - Reduce iterative minimization steps to accelerate the data processing speed**.**

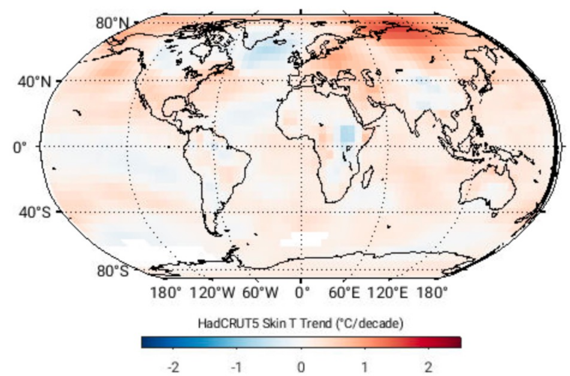
- **Climate Fingerprinting Sounder Products (ClimFiSP)**
  - Homogenization of $\Delta\boldsymbol{R}$ of AIRS and CrIS by using the **Climate Hyperspectral Infrared Radiance Product** (**CHIRP**). CHIRP combines AIRS and CrIS radiances using a common SRF and removes inter-instrument radiance biases.
  - Apply the constructed fingerprinting scheme to the complete hyper-spectral IR sounder data (AQUA+SNPP+J1) to derive climate data record (T, Q, Tskin, clouds, O3, CO …)
  - $\boldsymbol{r_0}$, $\boldsymbol{x_0}$ spectral radiance observations and corresponding geophysical variables of representative states can be constructed at different spatial-temporal level ( individual FOV, daily mean grid 0.5 × 0.5 lon. × lat., monthly mean grid … ).

# ClimFiSP surface skin temperature anomalies and trends Sep. 2002 – Jun. 2022

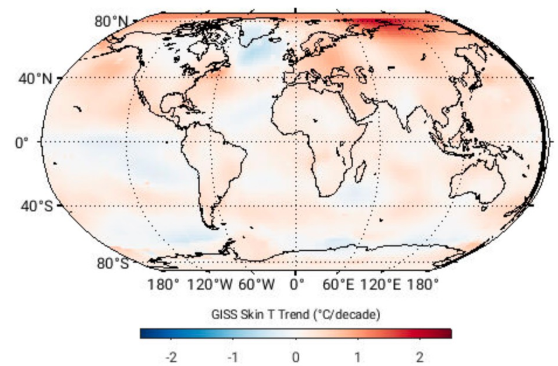
ClimFiSP



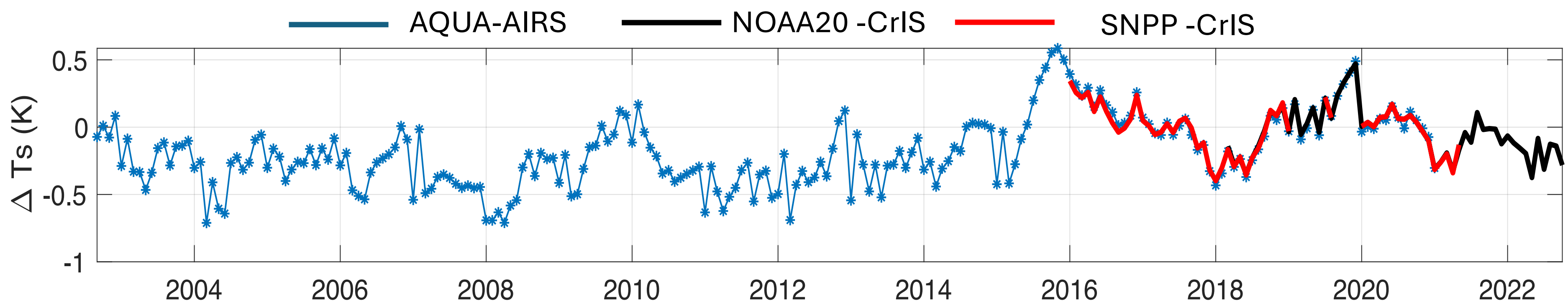

HadCRUT5




GISS

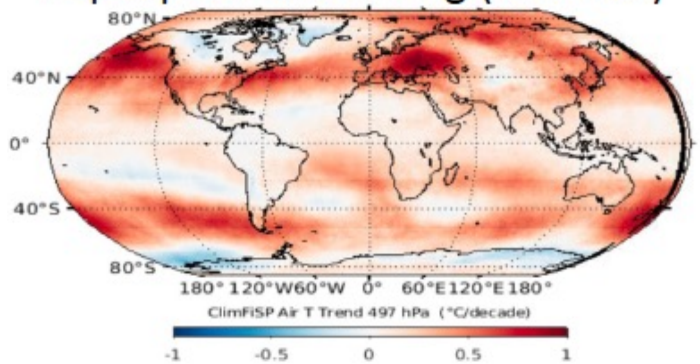



Tropical Region Skin Temp. Anomaly

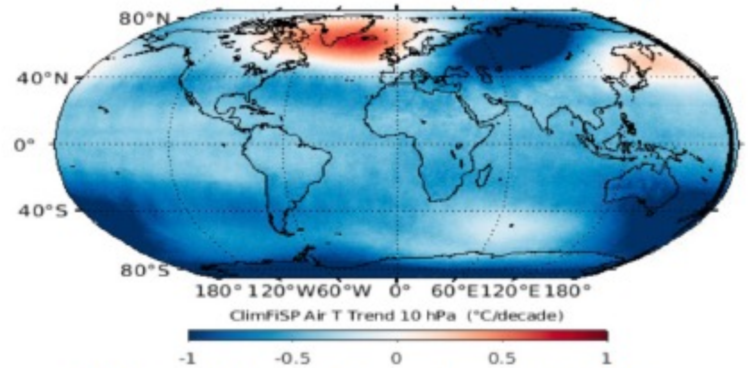


# ClimFiSP T & Q trends (K/decade) Sep. 2002 – June 2022

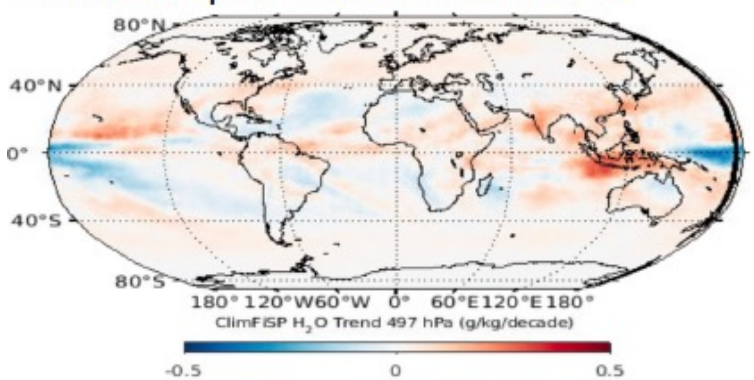
Tropospheric warming (500 hPa)



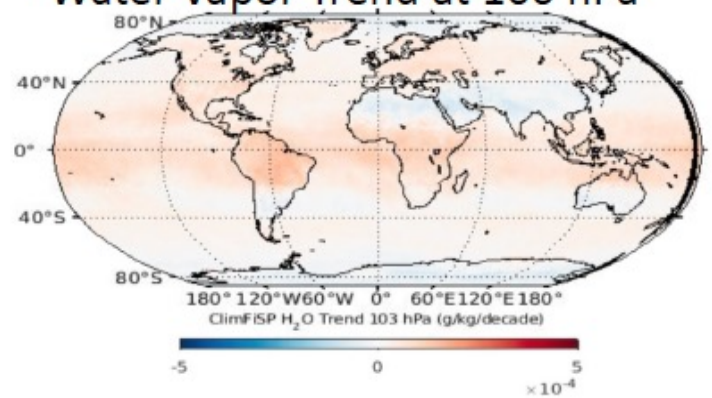

Stratospheric cooling (10 hPa)




Water Vapor Trend at 500 hPa

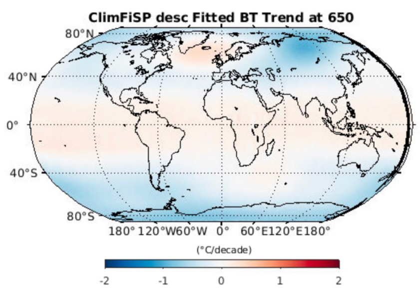


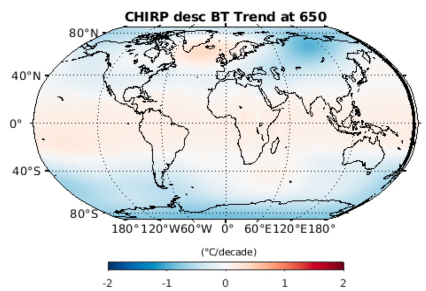

Water Vapor Trend at 100 hPa



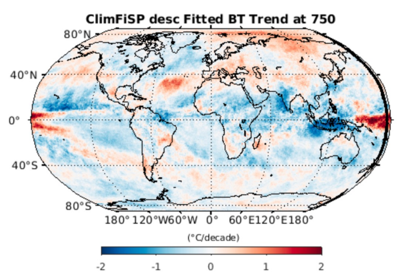
# ClimFiSP BT trends (K/decade)
# Meas. vs. Fitted (Sep. 2002 – Jun. 2022)

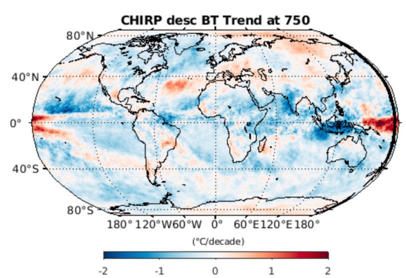
650 cm $^{-1}$



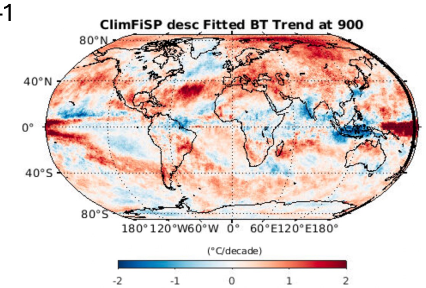



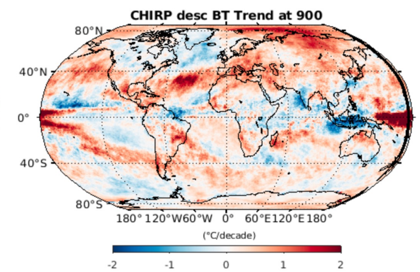

750 cm $^{-1}$



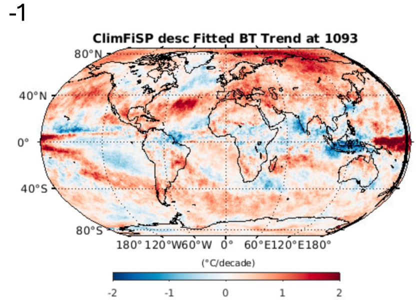



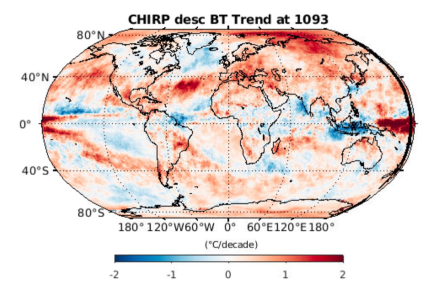

900 cm $^{-1}$



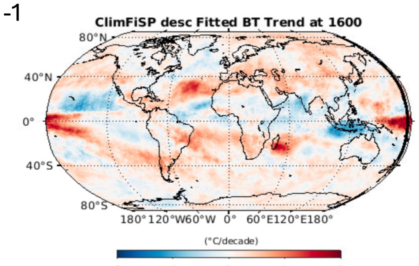



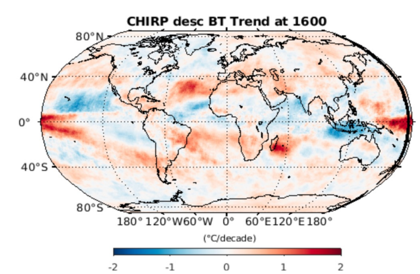

1093 cm $^{-1}$




1600 cm $^{-1}$



# Take home messages

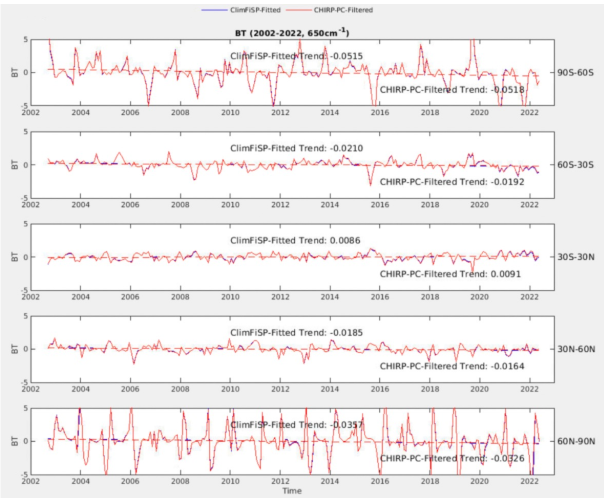
- New techniques have been developed to push the application limits of hyper-spectral IR sounders for weather and climate studies.

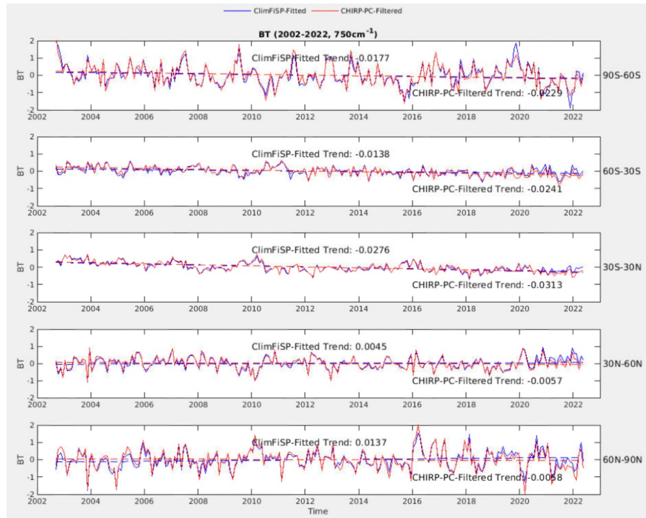
- The spectral fingerprinting method is a hybrid approach that enhances the OEM based physical inversion scheme using machine learning techniques. The method achieves a high efficiency data production while ensuring the radiometric consistency.

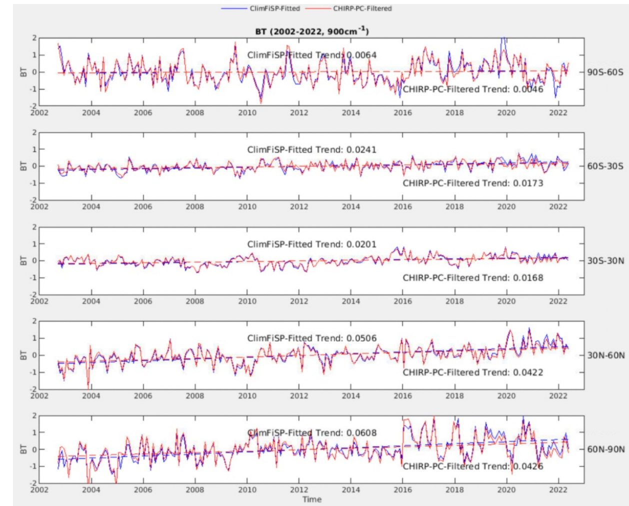
- SiFSAP which has been released at the NASA Goddard GES DISC will be updated by utilizing the spectral fingerprinting methodology.

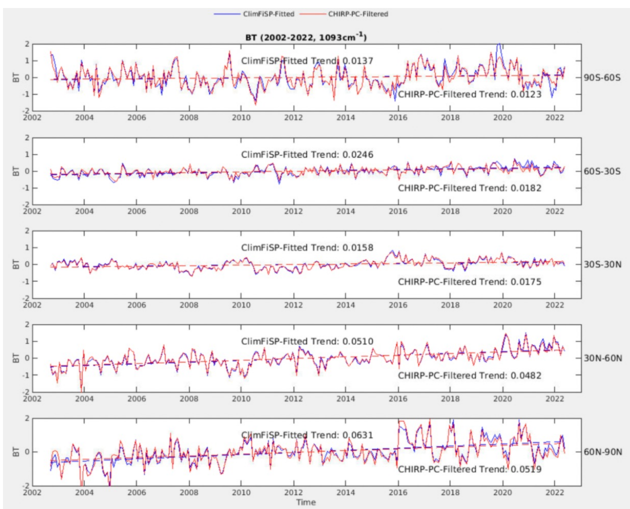
- ClimFiSP is going to be publicly available soon.

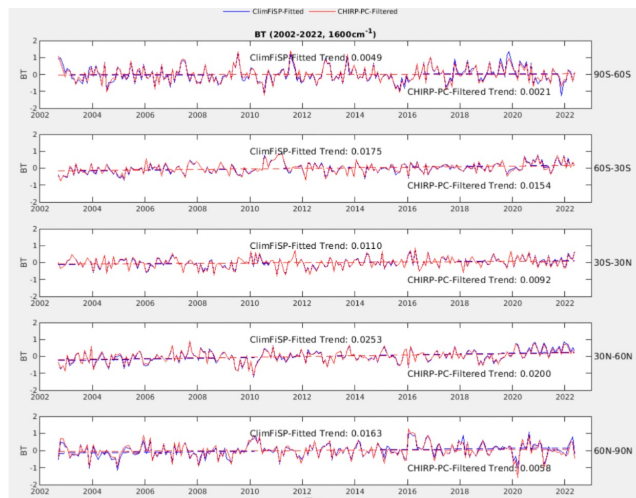
# ClimFiSP Spectral Fitting



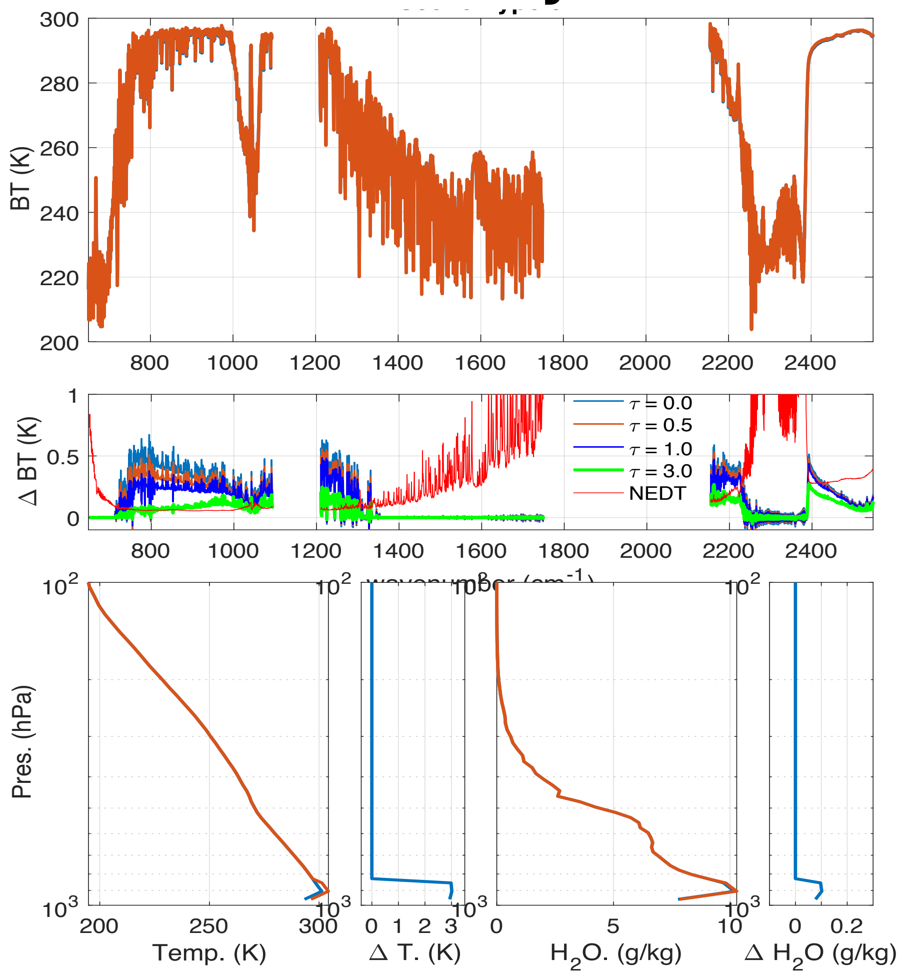



# The importance of establishing radiometric consistency to sounder observations




**A Planetary Boundary Layer study case**

Spectral fitting accuracy ultimately determines to what extend the information content from the measurement is utilized

$$\mathbf{A} = \left(\mathbf{K}^T \mathbf{S}_r^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}\mathbf{K}^T\mathbf{S}_r^{-1}\mathbf{K}$$

Spectral fitting uncertainty

Using averaging kernel to characterize vertical resolution is meaningless if spectral fitting uncertainty does not comply with the estimation based on *a prior*!