# Data Assimilation for CEDAR Science

***Nick Pedatella***

*High Altitude Observatory, National Center for Atmospheric Research*

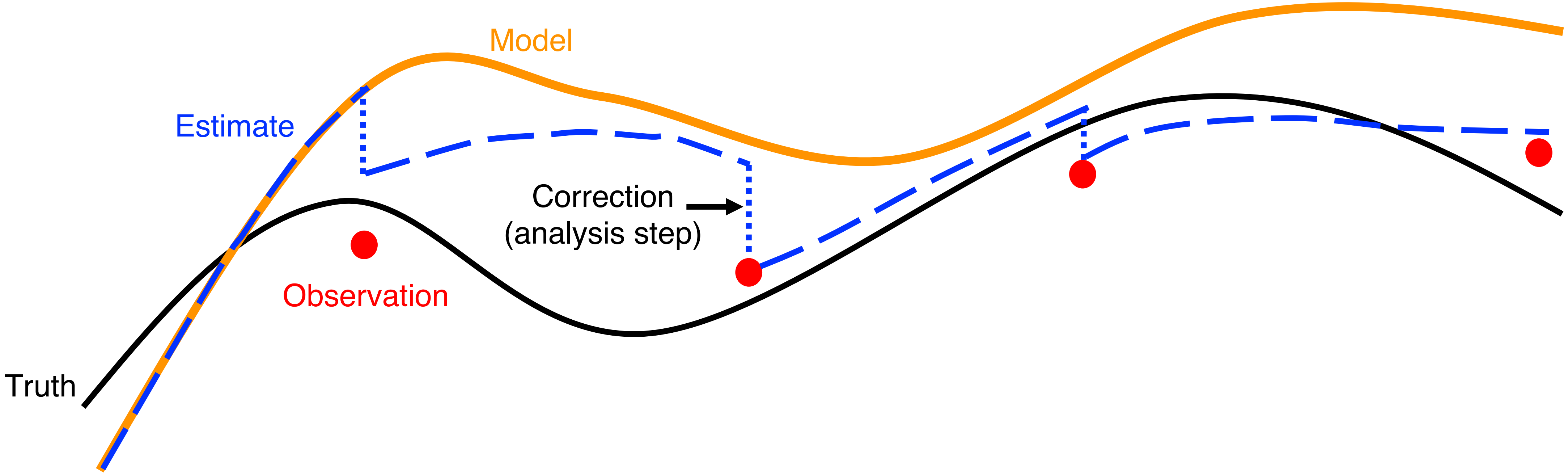
**June 25, 2023**

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

# Outline

- What is data assimilation?
- Why use data assimilation?
- Example applications
- Outstanding challenges

# Data assimilation is the process of statistically combining observations with a background model to determine the best estimate of the current state



# Data assimilation seeks to determine the best estimate of the current state based on the prior state and observations

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^f)$$

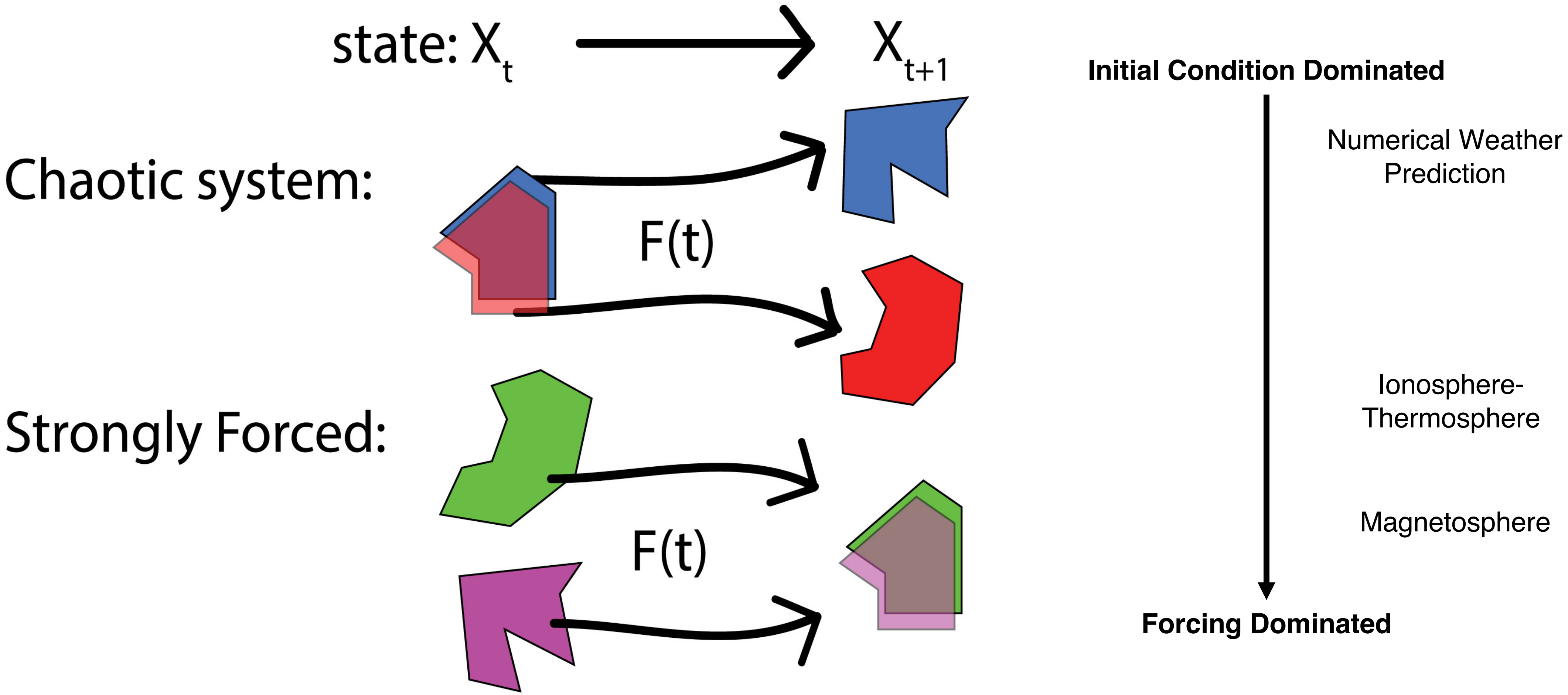
- $\mathbf{x}^a$: Analysis State
- $\mathbf{x}^f$: Forecast State
- $\mathbf{K}$: Weight or Gain Matrix
- $\mathbf{y}$: Observations
- $\mathbf{H}$: Observation or Forward Operator

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

- $\mathbf{B}$: background error covariance matrix
- $\mathbf{R}$: observation error covariance matrix

# Driver versus state estimation

Ionosphere-thermosphere is a partly forced system and the future state is thus due to a combination of the initial state and the external drivers



# Benefits of data assimilation

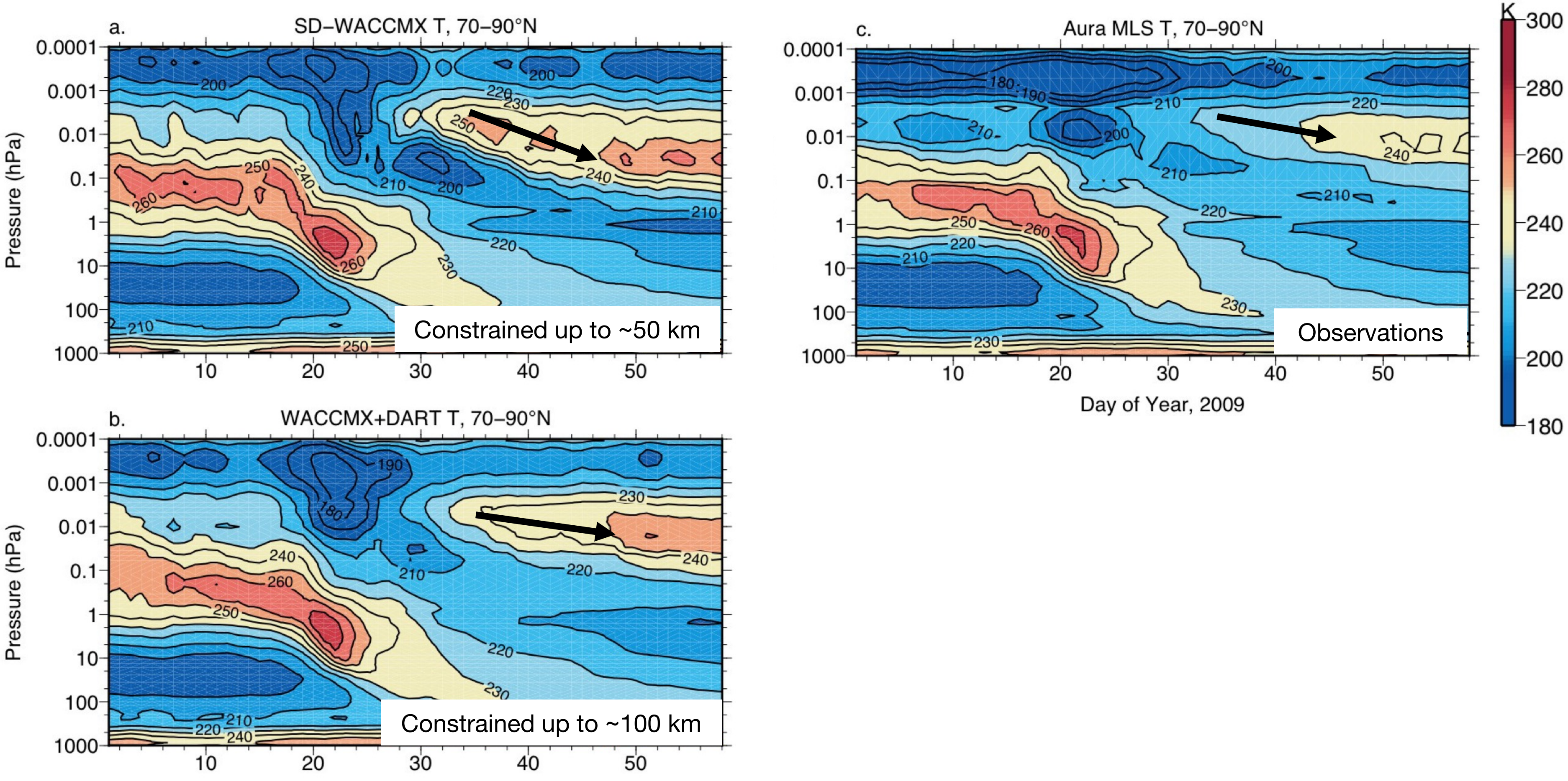
- **Observations are limited in scope**, and thus only provide limited knowledge about the variability in the atmosphere, including the sources of variability.

- **Numerical simulations are uncertain**, though they provide global information about additional quantities of interest that are often unobservable.

- Data assimilation can provide the best estimate of the state of the atmosphere by statistically combining observations and a numerical model.

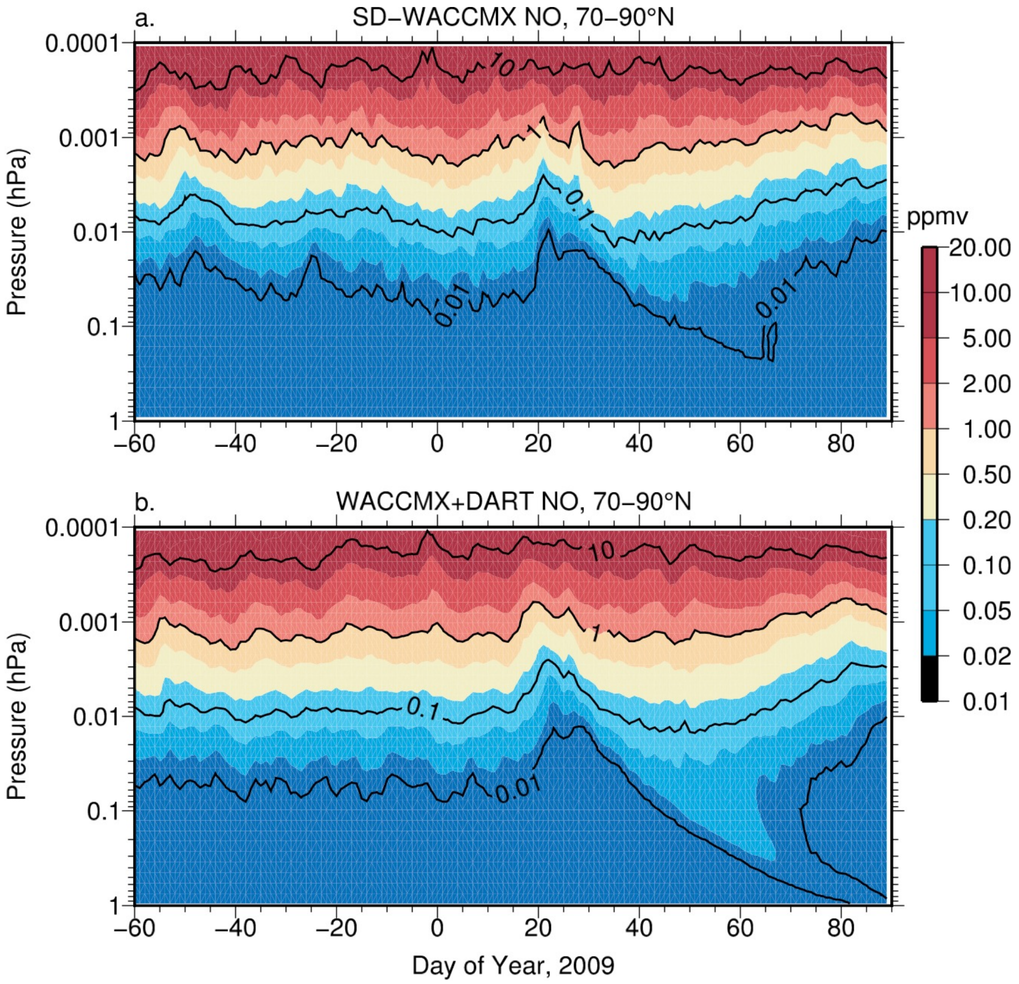
- Rationale for middle-upper atmosphere data assimilation:
  1. Improved estimate of the atmospheric state for process studies of the middle and upper atmosphere variability.
  2. Data assimilation increments provide insight into areas of large model errors
  3. Can be used to estimate uncertain model parameters and/or forcings
  4. Initial conditions for investigating the predictability of the middle and upper atmosphere.

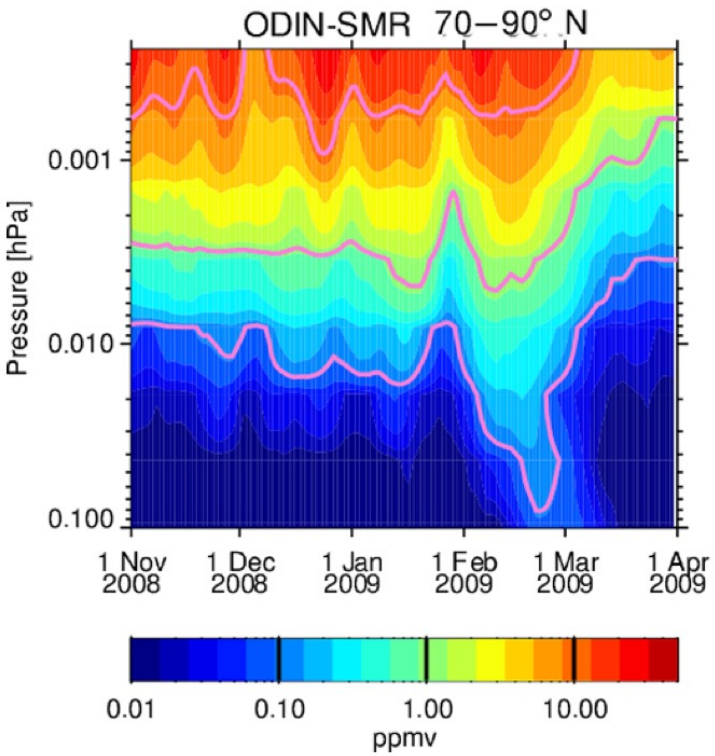
# Assimilation of mesospheric observations can alleviate model deficiencies leading to better representation of the dynamics and transport




(Pedatella et al., 2018)

# Assimilation of mesospheric observations can alleviate model deficiencies leading to better representation of the dynamics and transport

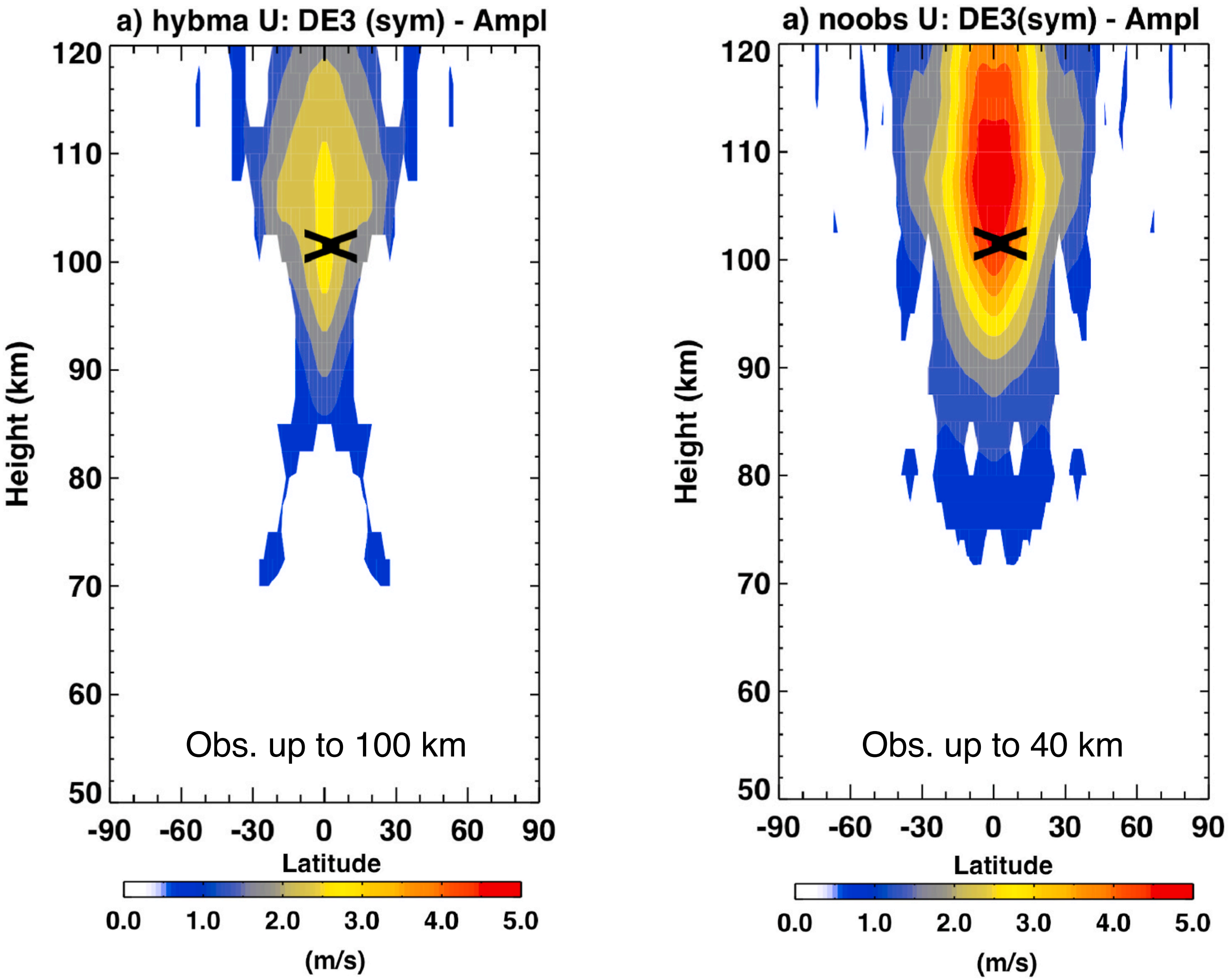



(Funke et al., 2017)

(Pedatella et al., 2018)

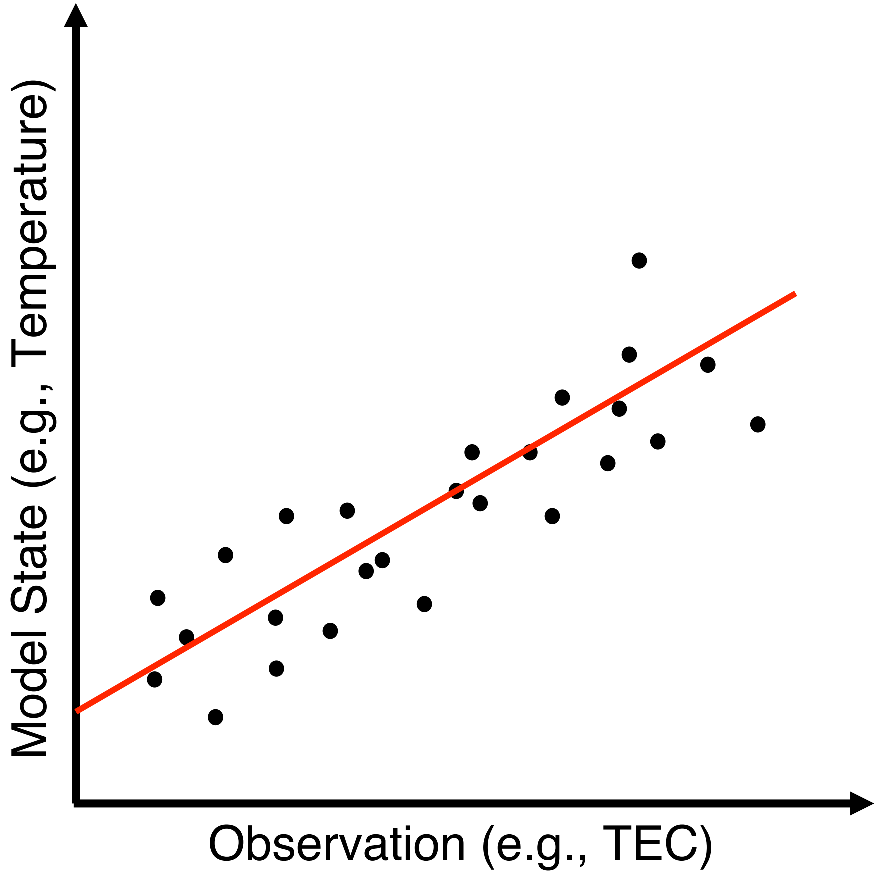
# Assimilation of middle atmosphere observations significantly impacts tidal amplitudes that ultimately impact the ionosphere-thermosphere




~2x larger amplitudes when only assimilating observations below 40 km

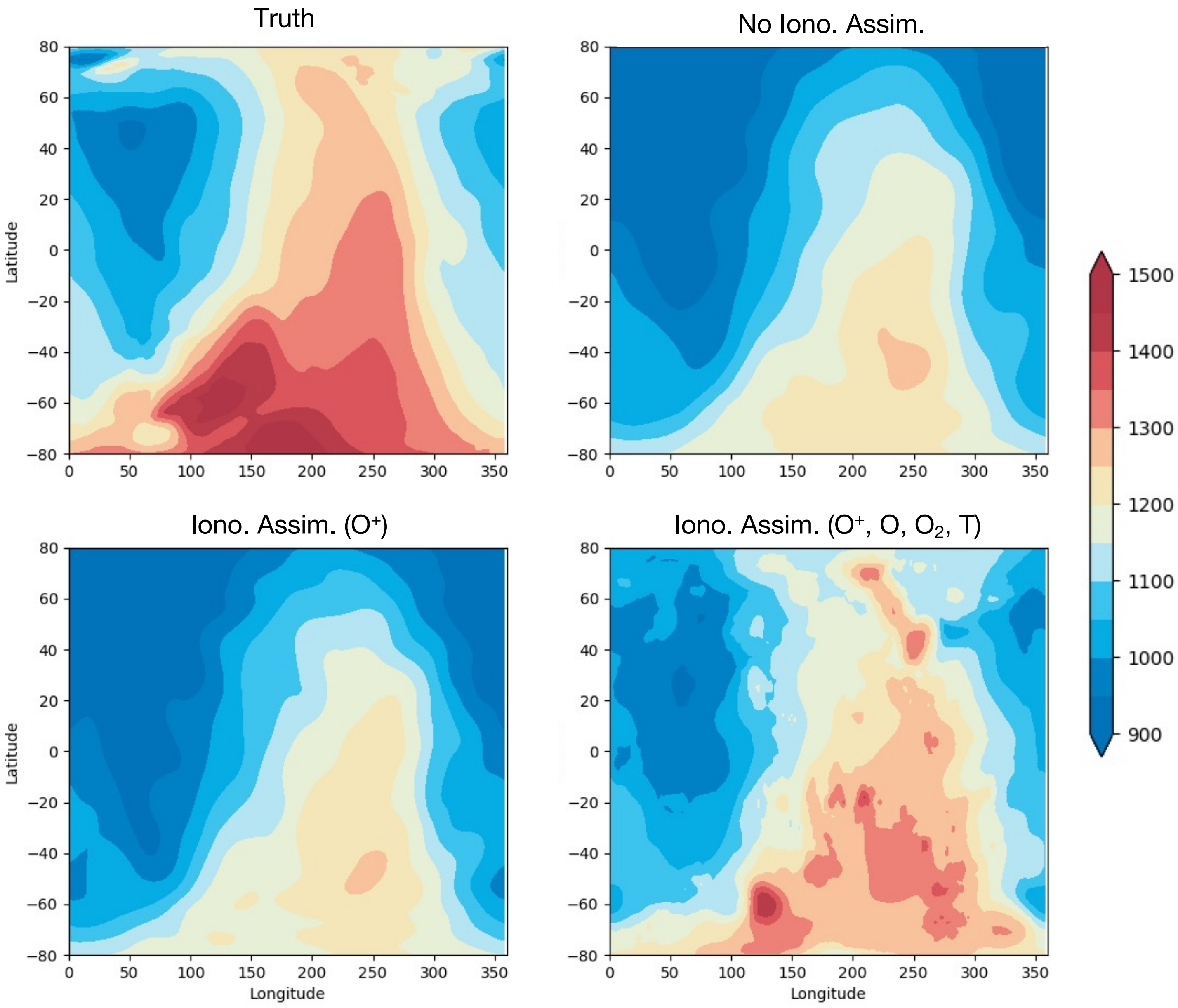
(Sassi et al., 2021)

# Covariance information can be used to adjust unobserved states, extending the impact of observations

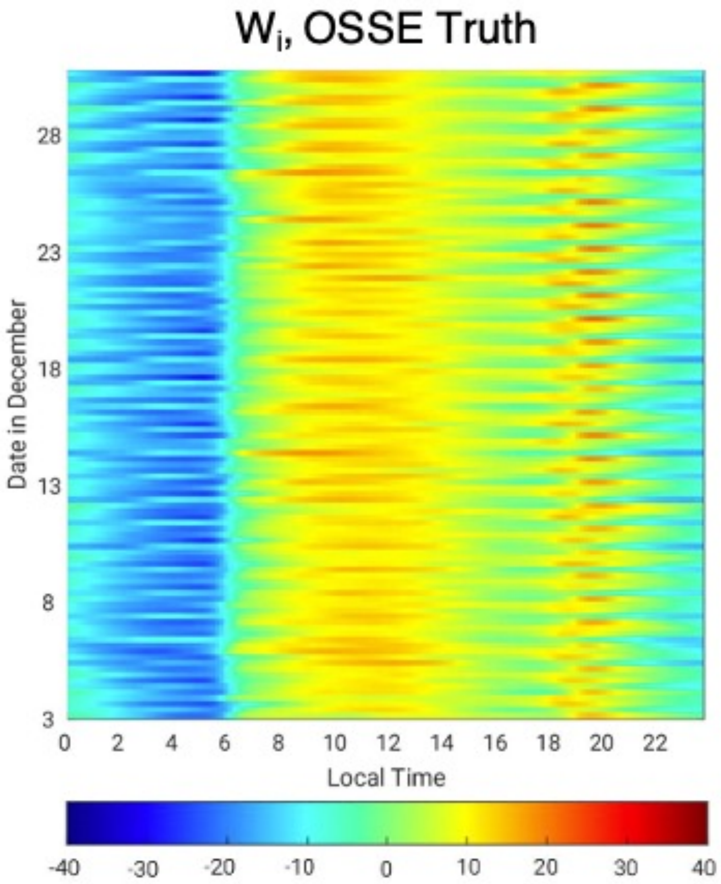



Sparsely observed neutral state can be constrained based on more dense ionosphere observations.

# Adjustment of thermospheric state from ionosphere observations leads to a reduction in thermosphere temperature errors

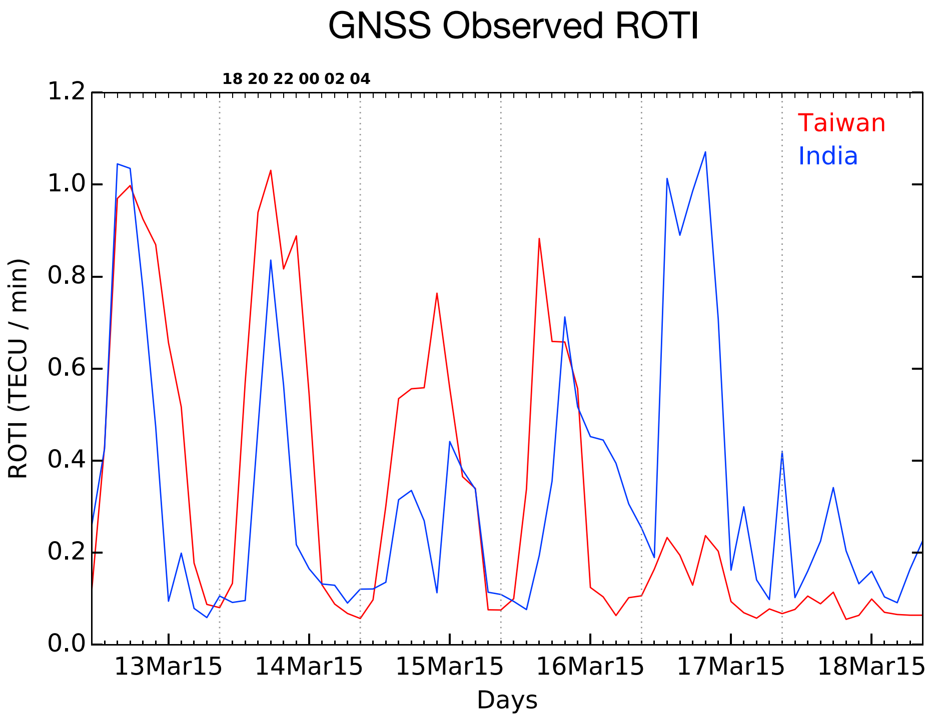


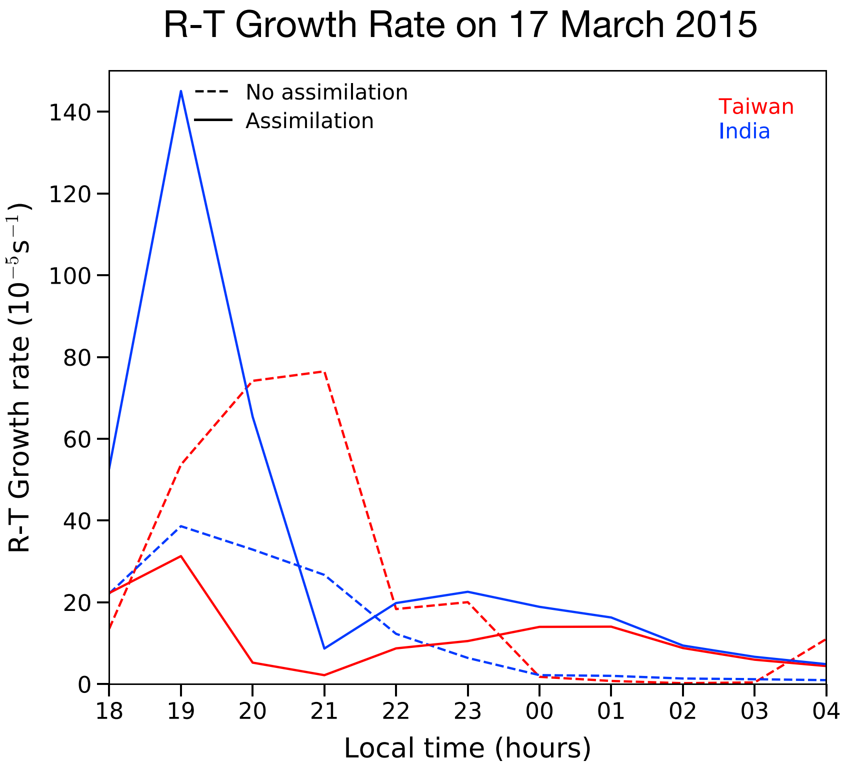
# WACCMX+DART equatorial vertical plasma drift velocity is improved when ICON/MIGHTI winds are assimilated




(Hsu, Pedatella, and Anderson, 2021)

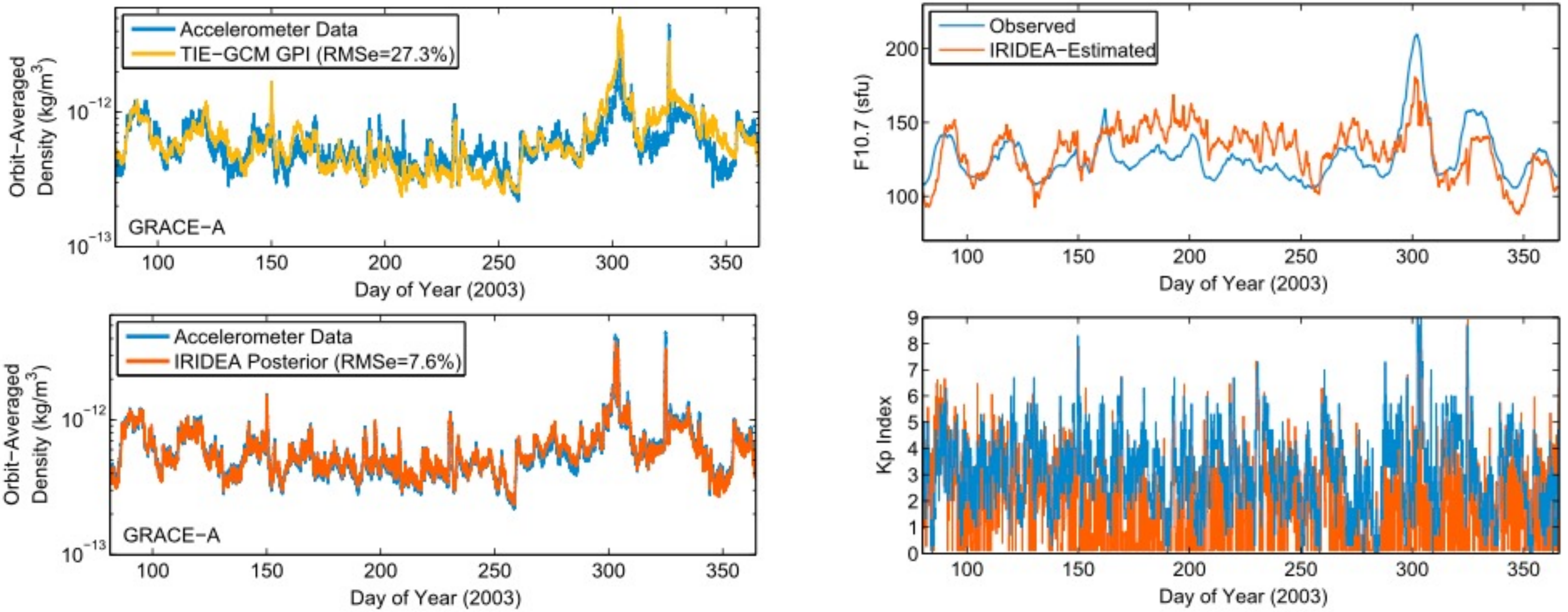
# Data assimilation (TIEGCM+DART) better reproduces the longitude variations in the R-T growth rate during the 2015 St. Patrick's Day storm, allowing investigation into the physical mechanisms responsible for these differences




(Rajesh et al., 2017)

# Estimation of forcing parameters through data assimilation leads to improvements in thermosphere mass density specification




(Sutton, 2016)

# Outstanding Challenges in Middle-Upper Atmosphere Data Assimilation

- Understanding the advantages/disadvantages of state vs. driver estimation an how they can be best applied for different scientific problems and operational goals.
- Make the most effective use of the sparse available observations in the mesosphere, thermosphere, and ionosphere.
- Determine best methods for generating ensemble spread for ensemble data assimilation
- Develop advanced data assimilation techniques (e.g., 4D-Var; hybrid) for the middle-upper atmosphere