Assessment of the impact of FORMOSAT-7/COSMIC-2 GNSS RO observations on mid- and low-latitude ionosphere specification and forecasting: Observing system simulation experiments by ensemble square root filtering

Chih-Ting Hsu[1,2], Tomoko Matsuo[2], Xinan Yue[3], Tzu-Wei Fang[4], Timothy Fuller-Rowell[4], Kyo Ide[5], and Jann-Yenq Liu[1]
[1] Institute of Space Science, National Central University, Taoyuan, Taiwan [2] Aerospace Engineering Sciences, University of Colorado at Boulder, CO, USA
[3]Chinese Academy of Sciences, China [4] Space Weather Prediction Center, National Oceanic and Atmospheric Administration, Boulder, CO, USA
[5]Department of Atmospheric and Oceanic Science, University of Maryland, MD, USA

Motivations and Goals

The Formosa Satellite-7/Constellation Observing System for Meteorology, Ionosphere and Climate-2 (FORMOSAT-7/COSMIC-2) GNSS Radio Occultation (RO) payload can provide global observations of slant Total Electron Content (sTEC) with unprecedentedly high spatial and temporal resolution.

This presentation will demonstrate **(A) how the Ensemble Square Root Filter (EnSRF) [Whitaker and Hamill, 2001] can be used to assimilate sTEC observations effectively**, and **(B) impacts of FORMOSAT-7/COSMIC-2 GNSS RO data on low- and mid-latitude ionospheric specification and forecasting**.

Data assimilation system

Synthetic RO sTEC data are assimilated into a coupled model of thermosphere, ionosphere, and plasmasphere by using EnSRF.

Data - RO sTEC

- RO sTEC along a given radio path can be retrieved from signals received LEO GPS receiver
- RO path for a given sTEC can traverse through a large distance in the ionosphere and plasmasphere (up to 6000-7000 km).

Model –GIP/TIEGCM

Global-Ionosphere-Plasmasphere/Theremosphere-Ionosphere-Electrodynamics General Circulation Model (GIP/TIEGCM) [Pedatella et al, 2011] is made of following two models.

- TIEGCM – thermosphere ~ 400 - 800 km
- GIP – ionosphere and plasmasphere ~ 19000 km

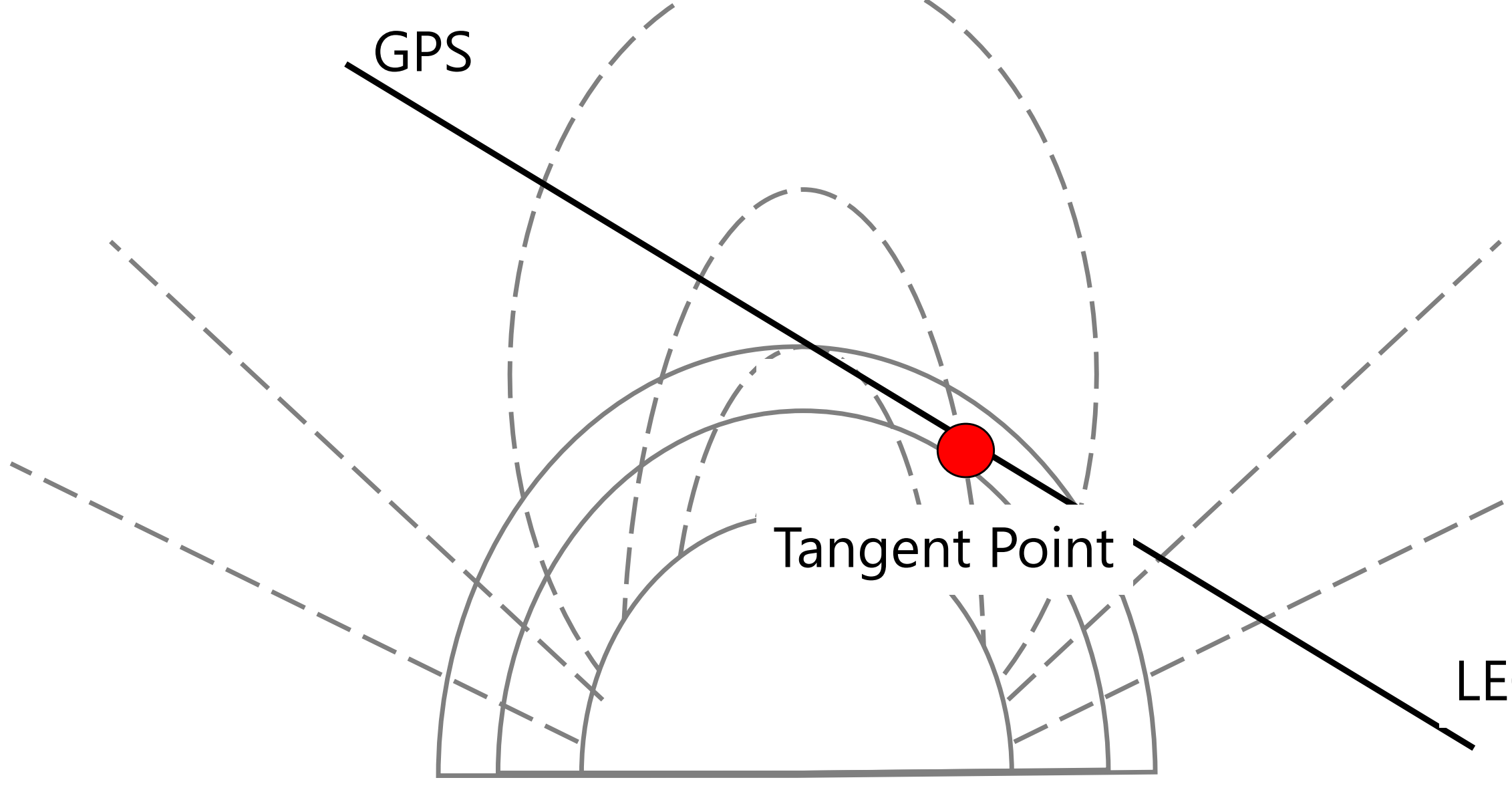


Figure 1. sTEC radio path between a LEO satellite and a GPS satellite and GIP/TIEGCM coordinates.

Step 1 convert model state variable on model space (e^-) to observed variable on observation space ($sTEC$)

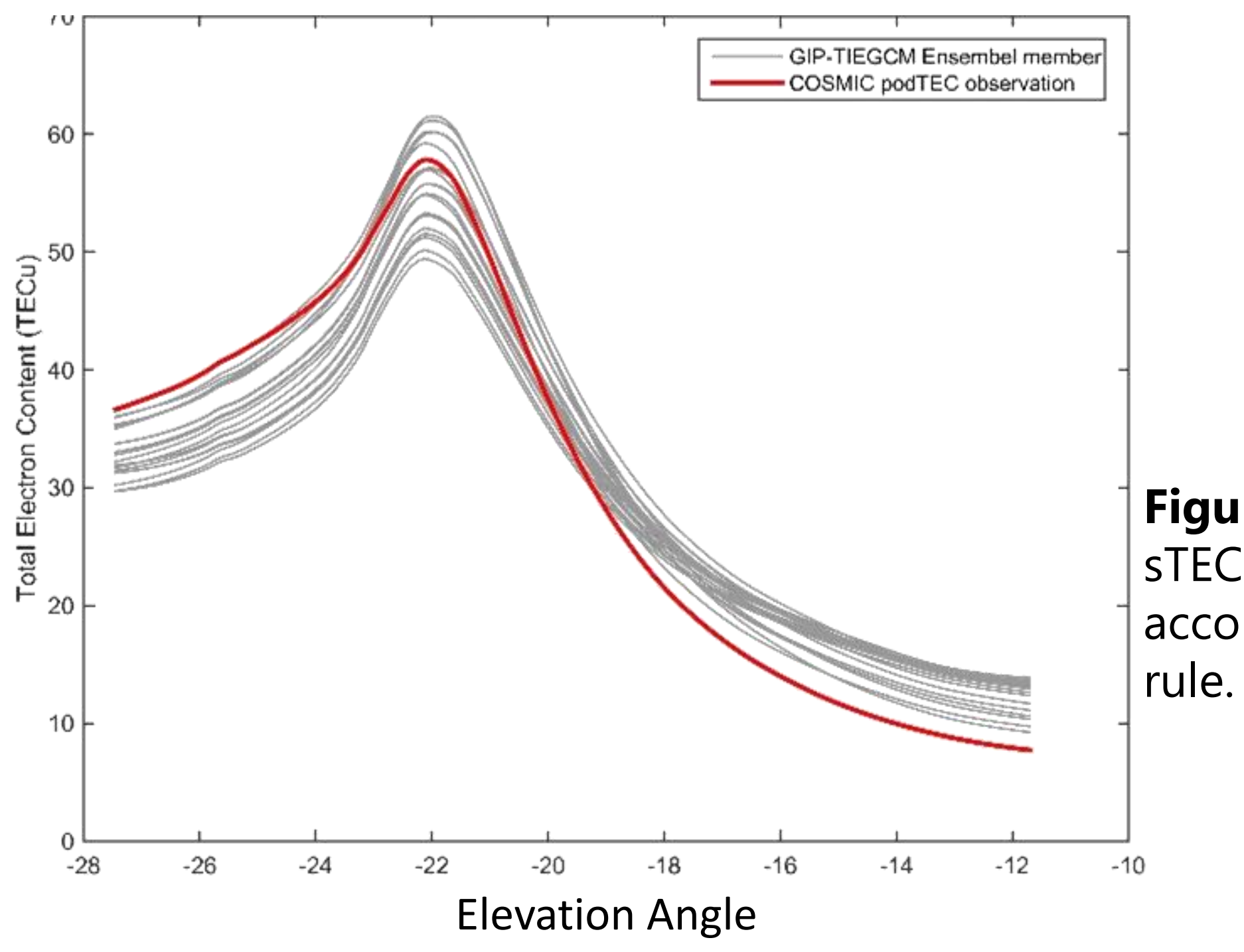


Figure 2. Comparison between observed sTEC from FORMOSAT-3/COSMIC (red line) and sTEC calculated from GIP/TIEGCM ensembles (grey lines).

Step 2 get the increment of observed variable ($\Delta sTEC$) according to Bayes rule

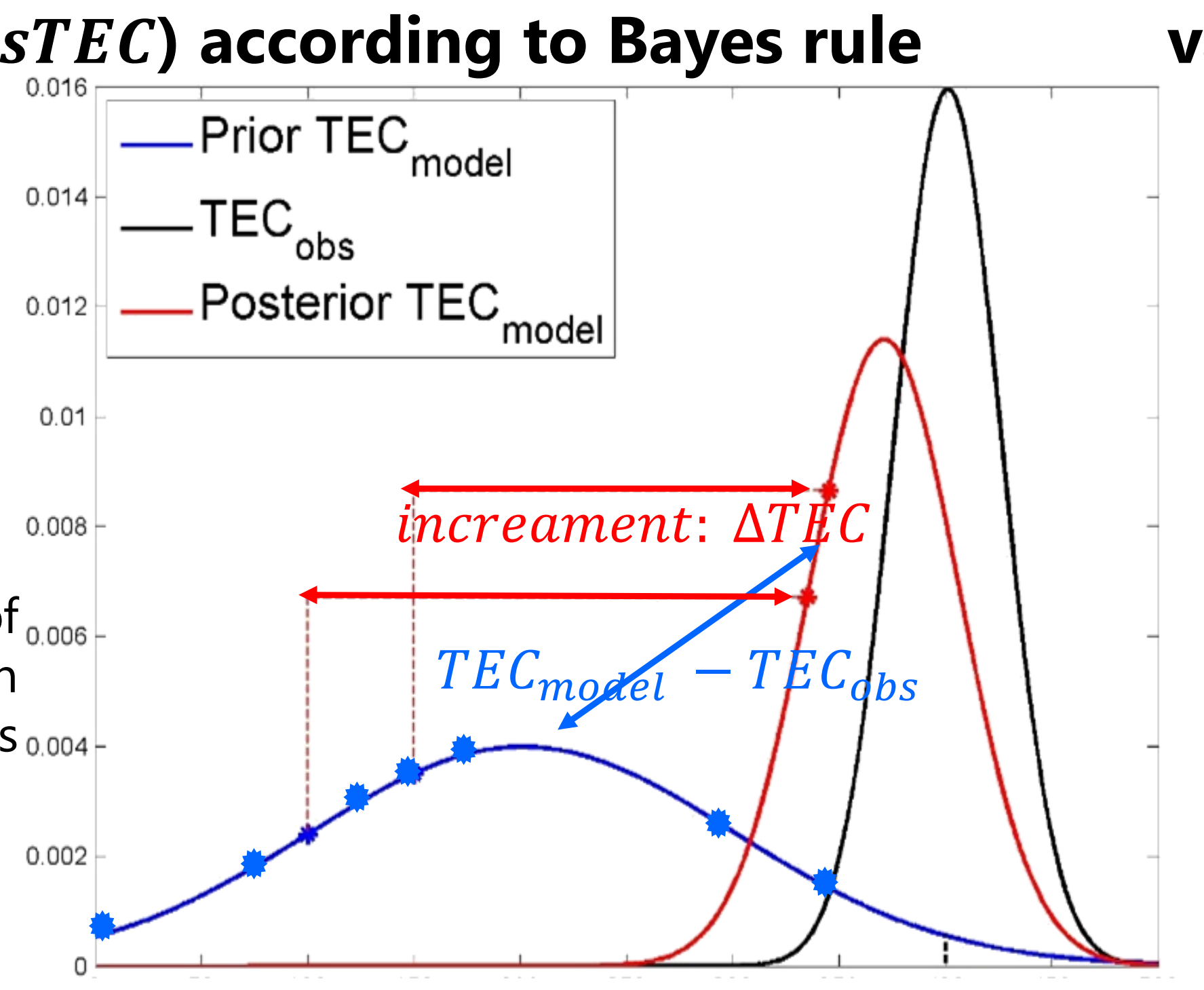
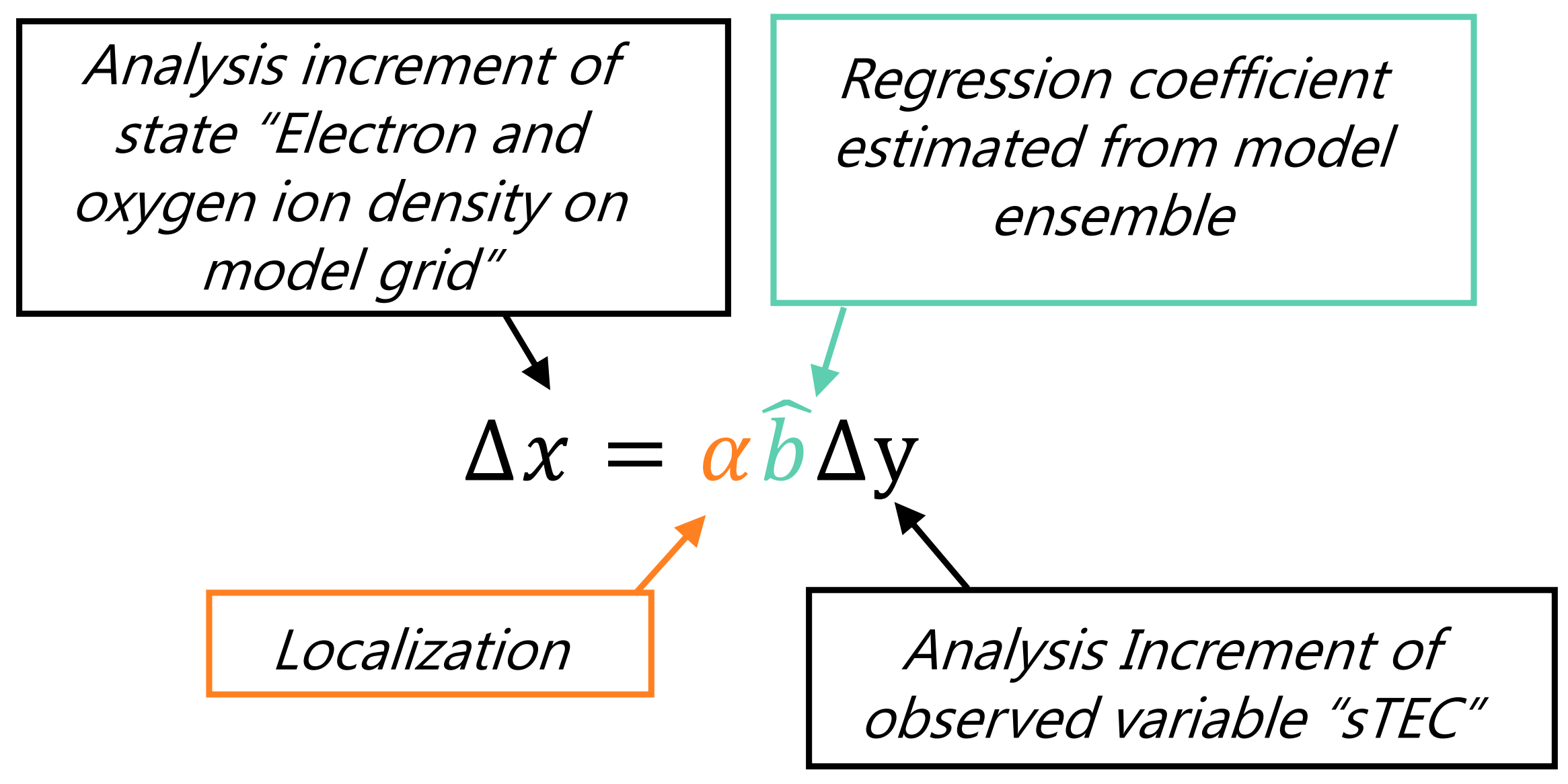


Figure 3. Basic idea of sTEC data assimilation according to Bayes rule.

Step 3 calculate the increment of model state variables (Δe^- and ΔO^+)



Analysis increment of state "Electron and oxygen ion density on model grid"

Regression coefficient estimated from model ensemble

$$\Delta x = \alpha \hat{b} \Delta y$$

Localization

Analysis Increment of observed variable "sTEC"

A. Experiments with Different Ensemble Sizes

Observing System Simulation Experiments (OSSEs) with 40, 70, and 100 GIP/TIEGCM ensemble members are carried out.

- Synthetic sTEC data sampled from a "true" state are assimilated into the model continuously from UT 0000 to UT 1200.
- Both e^- and O^+ density are updated by using EnSRF.
- GIP/TIEGCM ensembles are generated by perturbing model drivers according to a normal distribution specified below

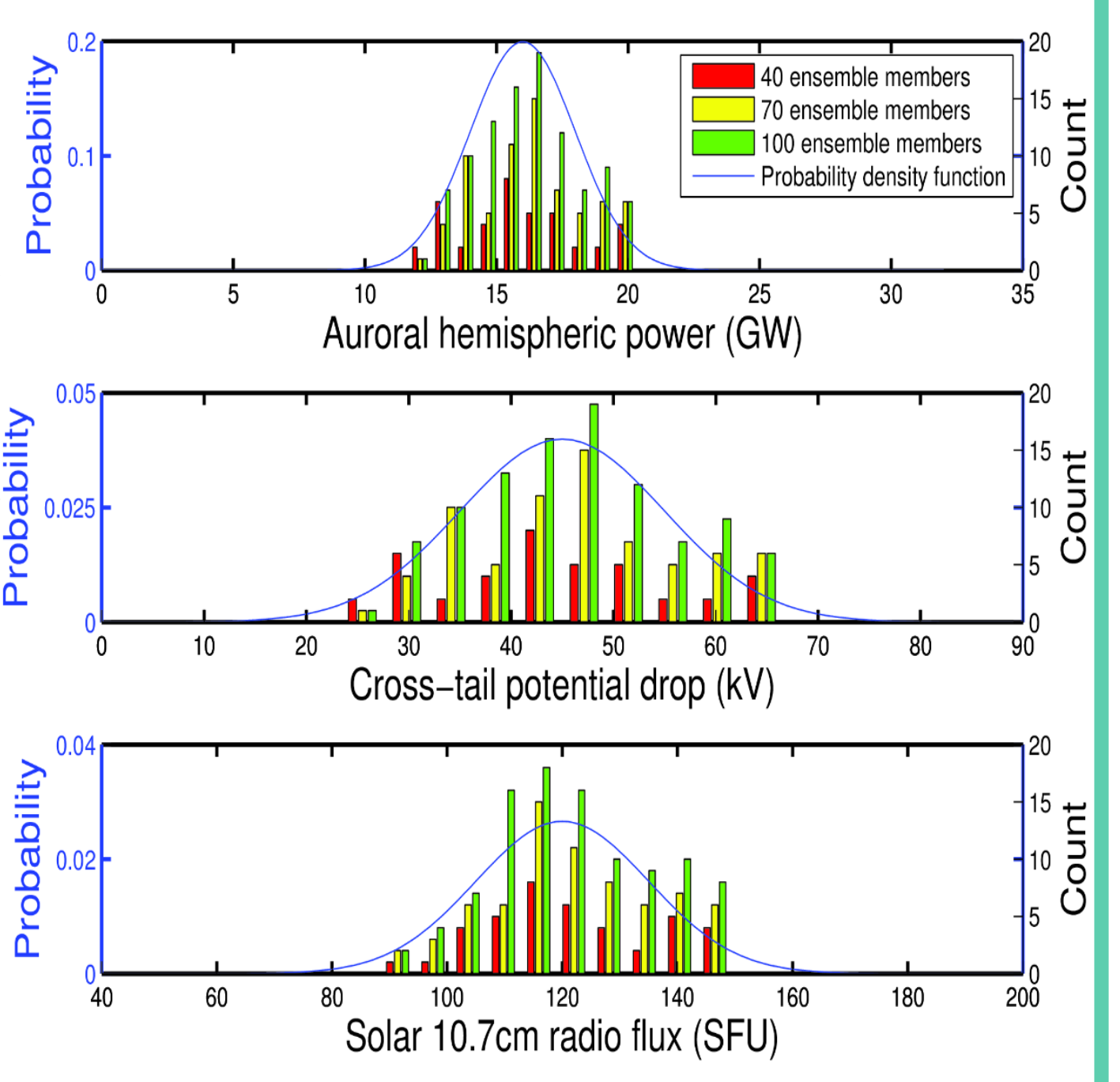
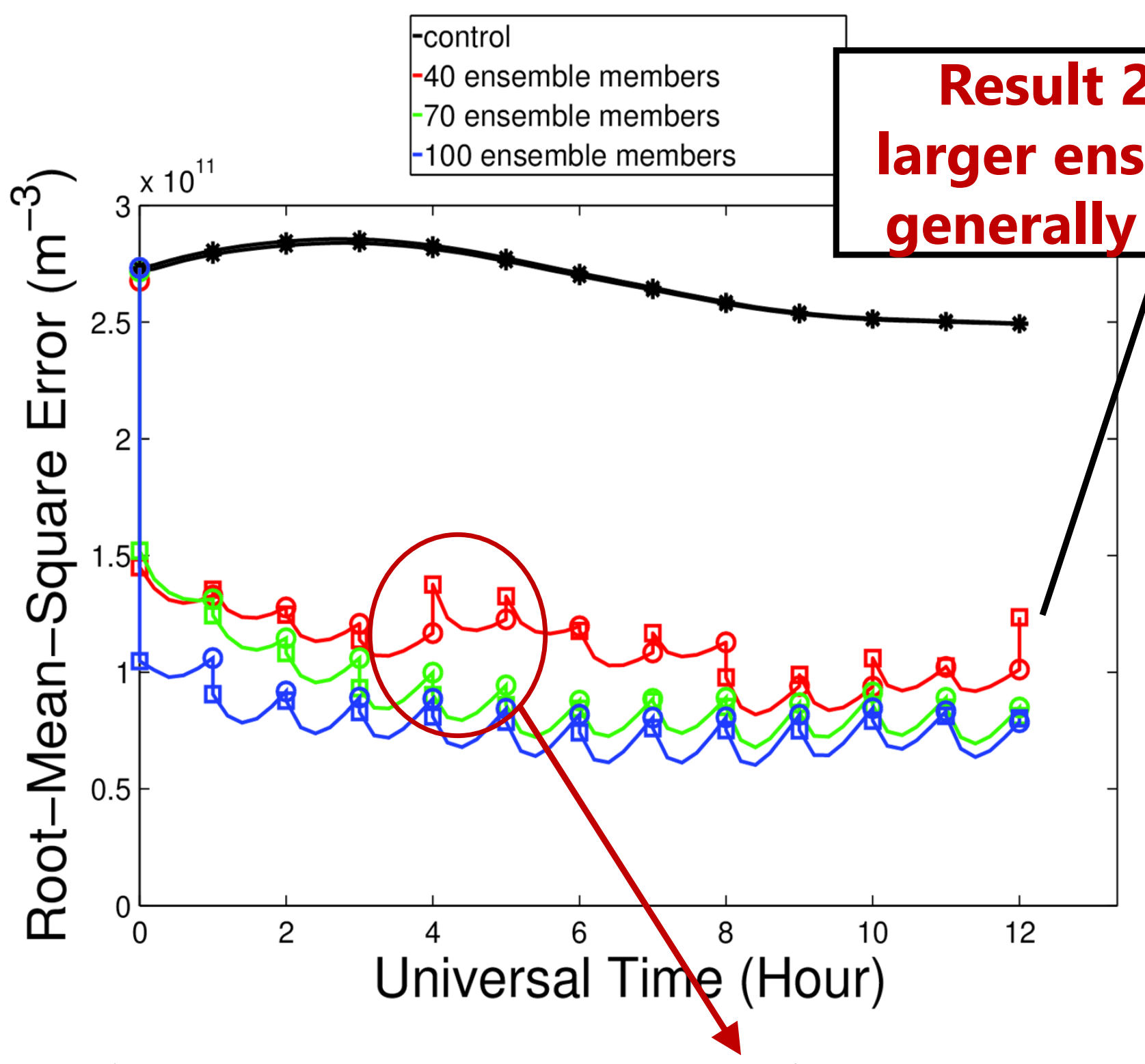


Figure 4. Input drivers that used to perturb the ensemble members



Result 2. EnSRF with a larger ensemble size (>70) generally performs better

Figure 5. Root-Mean-Square Error (RMSE) of O^+ density over mid- and low- latitude F-region during data assimilation cycle.

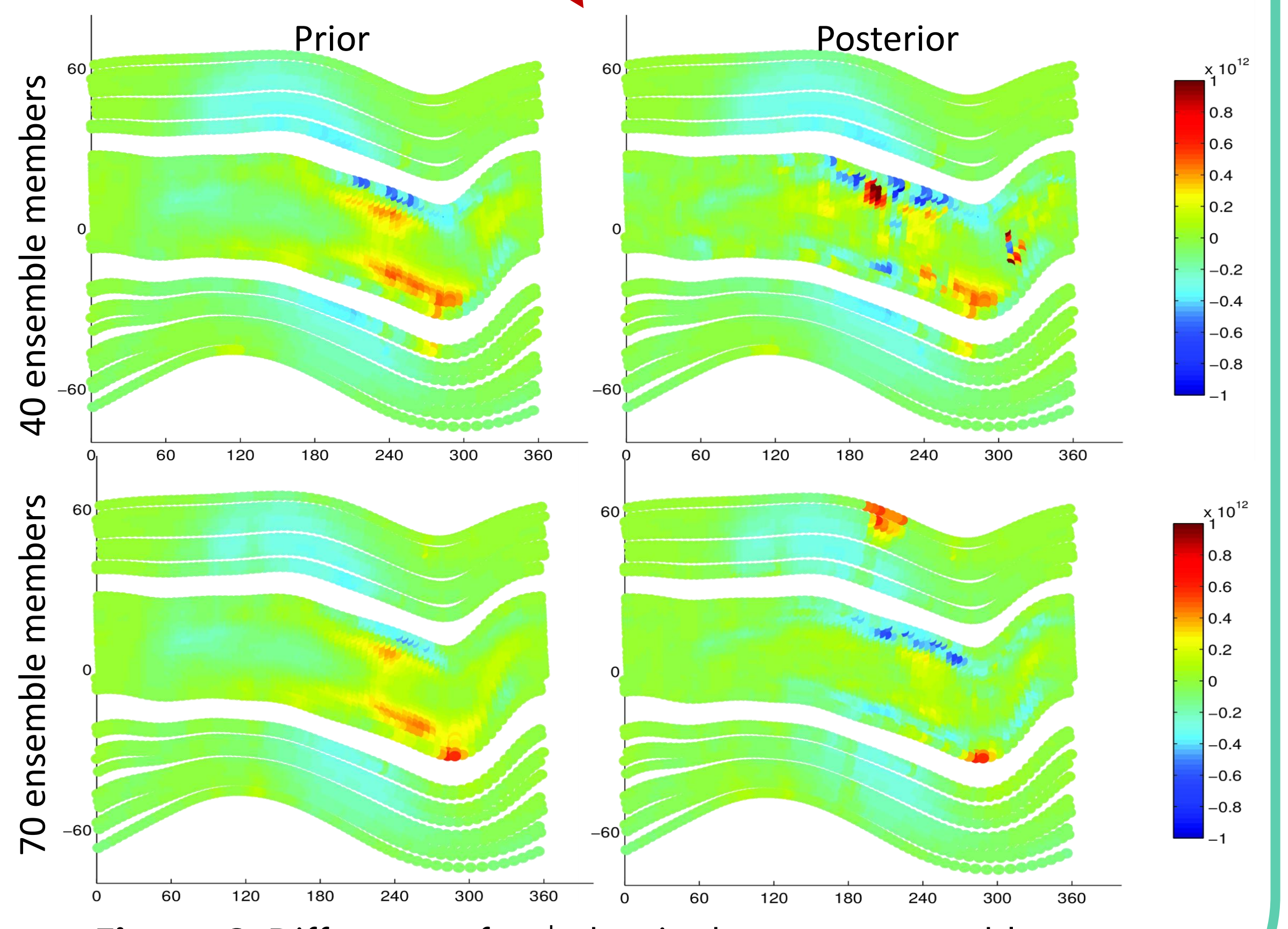
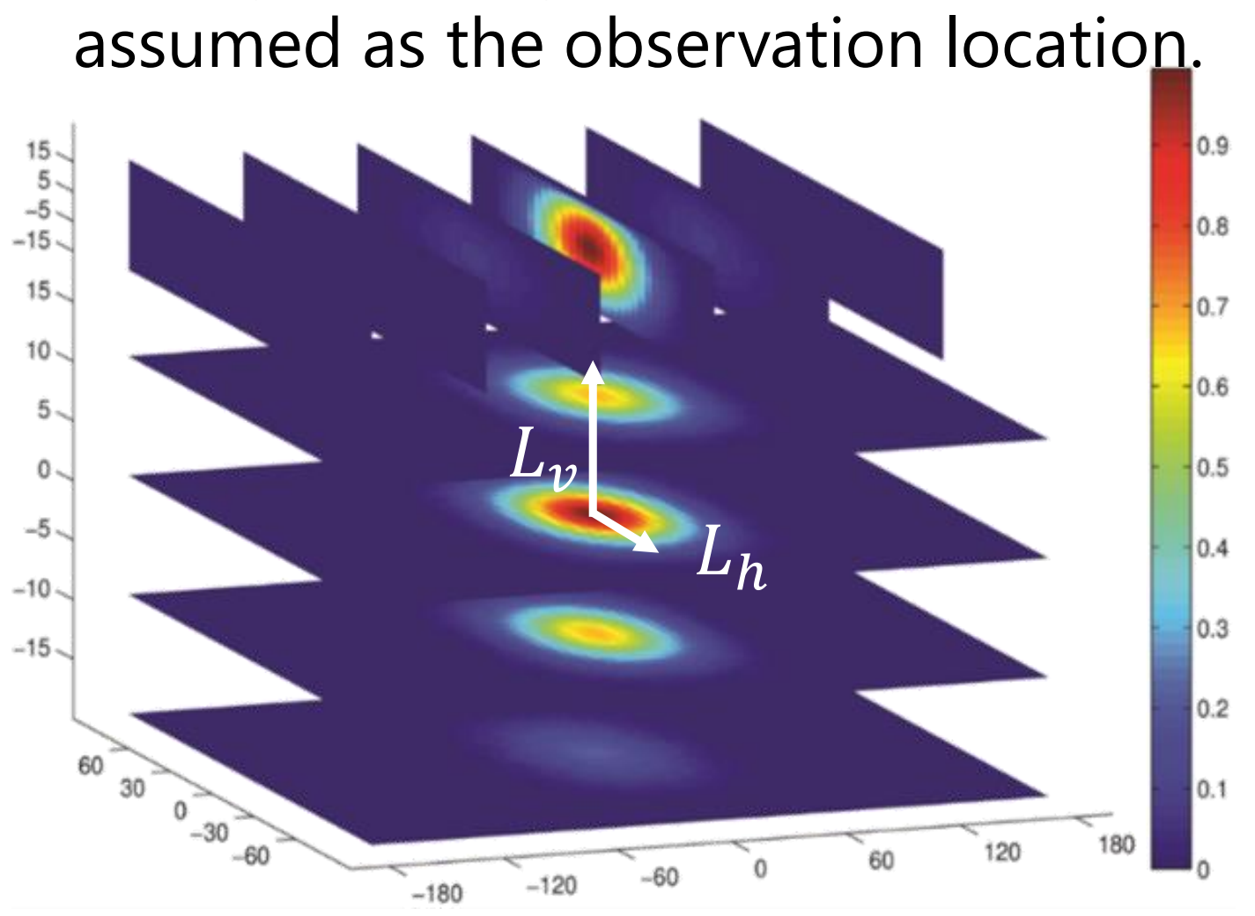


Figure 6. Difference of O^+ density between ensemble mean and the "Truth" at 330 km altitude at 04:00 UT.

B. Experiments with Different Localization Length Scales

OSSEs with different localization length scales are carried out.

- Single sTEC data is assimilated into the model. The tangent point of this data is at local noon, 350 km, 0° longitude, and 0° latitude.
- Gaspari-Cohn (GC) function [Gaspari and Cohn, 1999] is used to specify α for a given normalized distance r. The tangent point is assumed as the observation location.



$$r = Dist/L$$

$$For\ r \le 1$$

$$\alpha(r) = GC(r)$$

$$For\ r > 1$$

$$\alpha(r) = 0$$

Figure 7. The 3-D structure of GC function.

Vertical localization length scale (ln(mb)) L_v	Horizontal localization length scale (km) L_h
0.5	500
1.0	1000
3.0	5000
7.0	10000
No localization	No localization

Table 1. List of vertical localization length scale used in this experiment. Vertical localization length scale is defined on scale height. One scale height equal to two TIE-GCM vertical level.

Table 2. List of horizontal localization length scale used in this experiment. Horizontal localization length scale is defined on km.

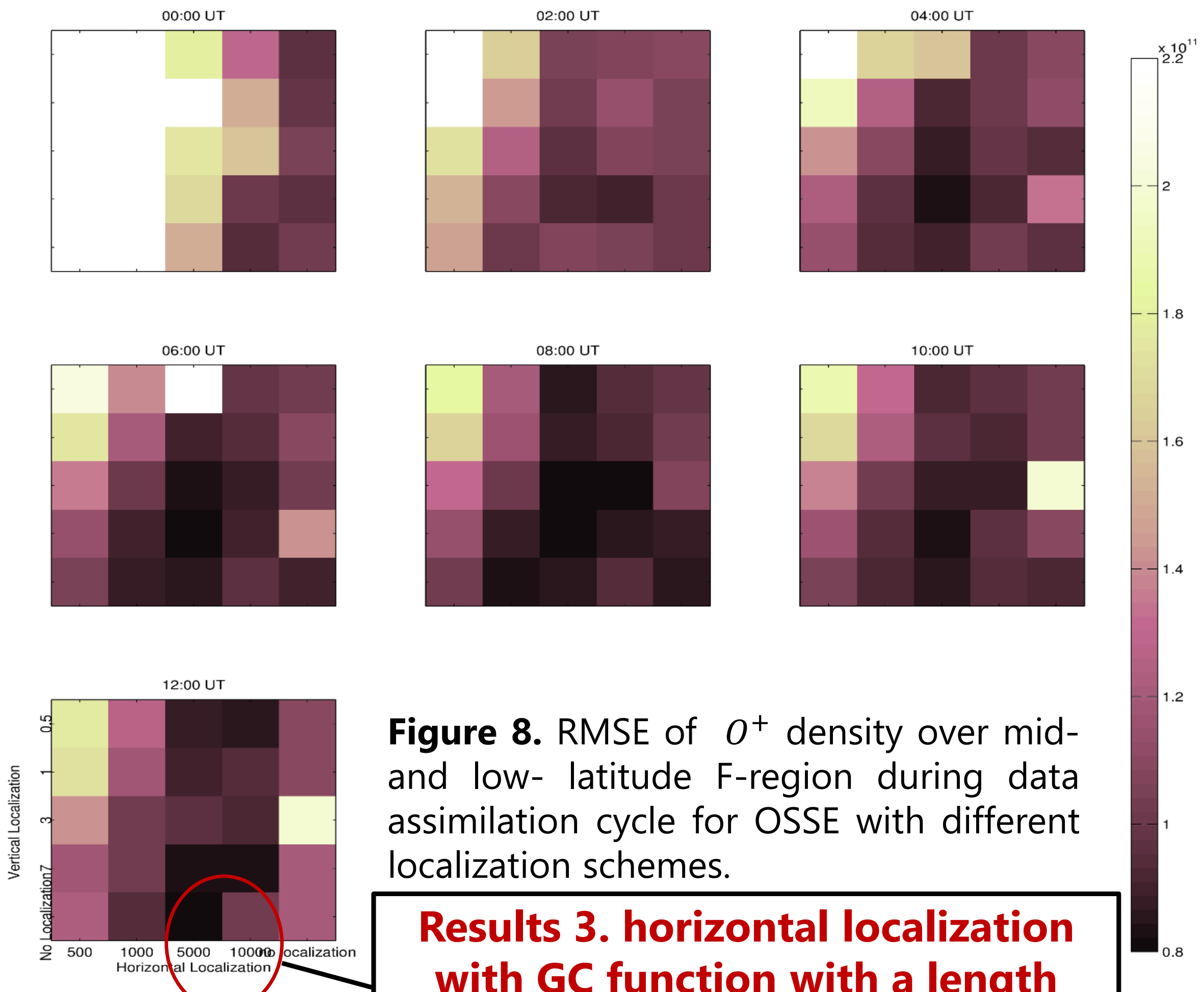


Figure 8. RMSE of O^+ density over mid- and low- latitude F-region during data assimilation cycle for OSSE with different localization schemes.

Results 3. horizontal localization with GC function with a length scale of 5000-10000km helps the quality of sTEC data assimilation.

C. Final result

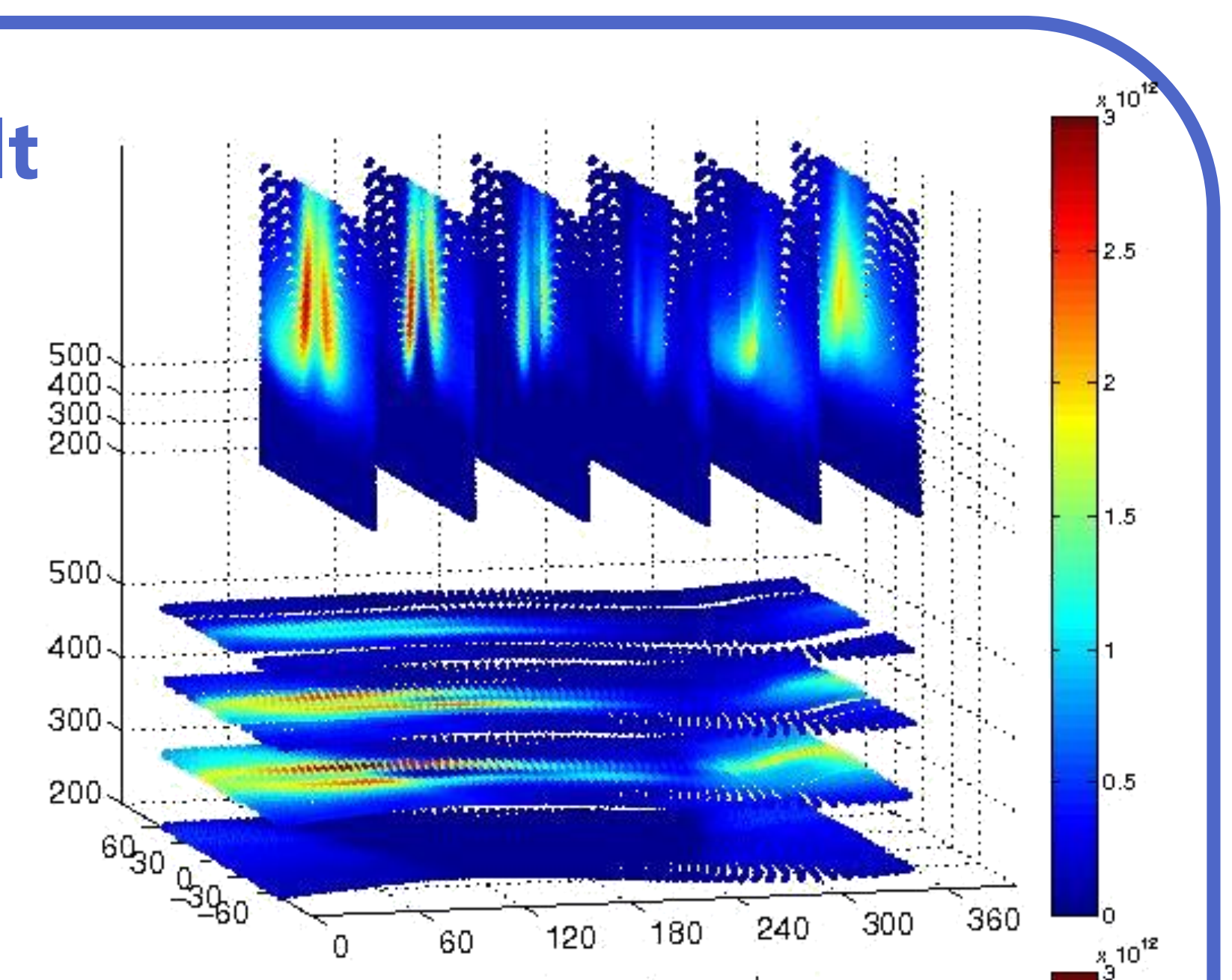


Figure 9. O^+ density distribution without data assimilation at 12:00 UT.

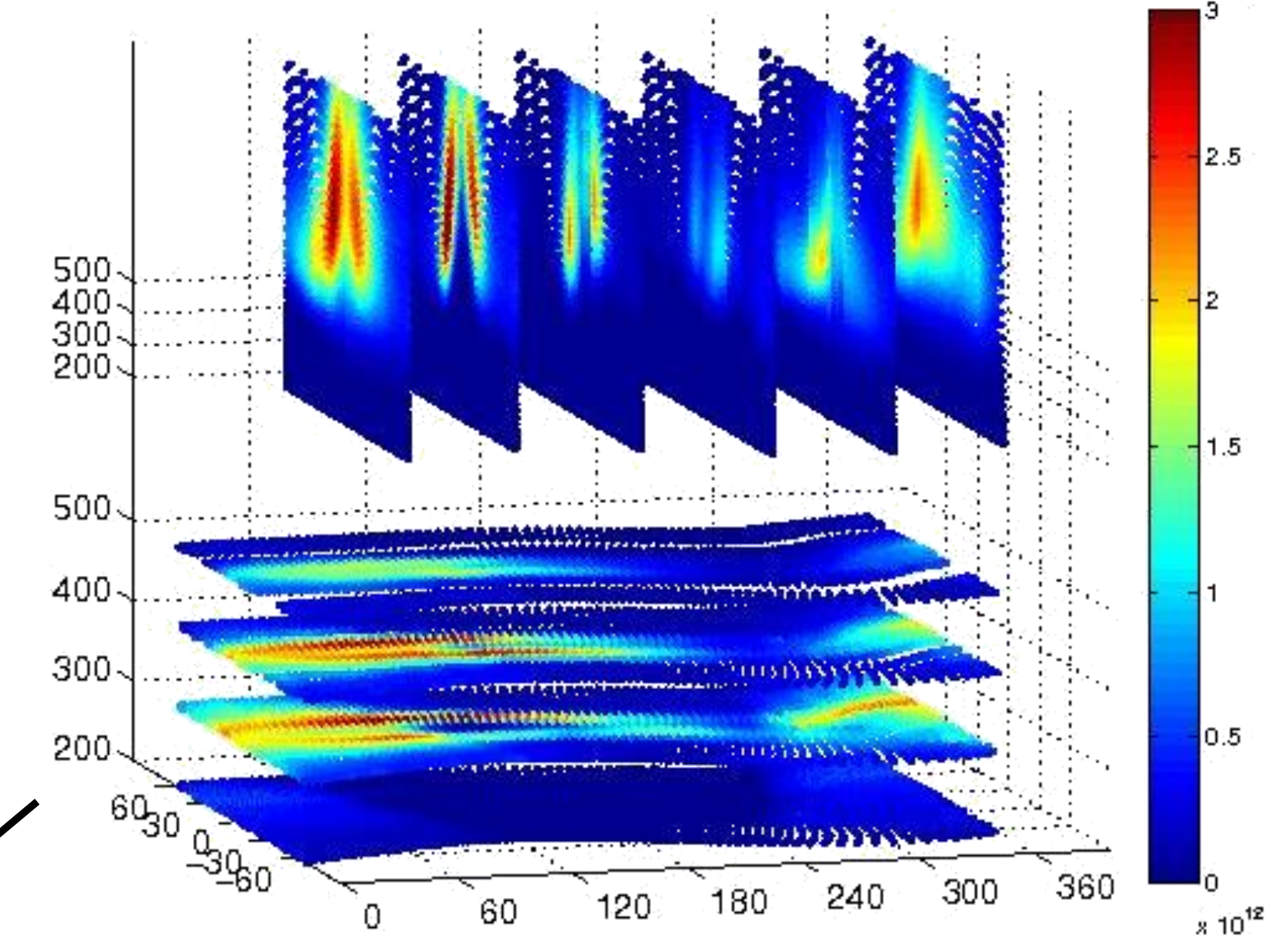


Figure 10. O^+ density distribution with data assimilation at 12:00 UT.

Result 1. sTEC data assimilation improve ionospheric specification

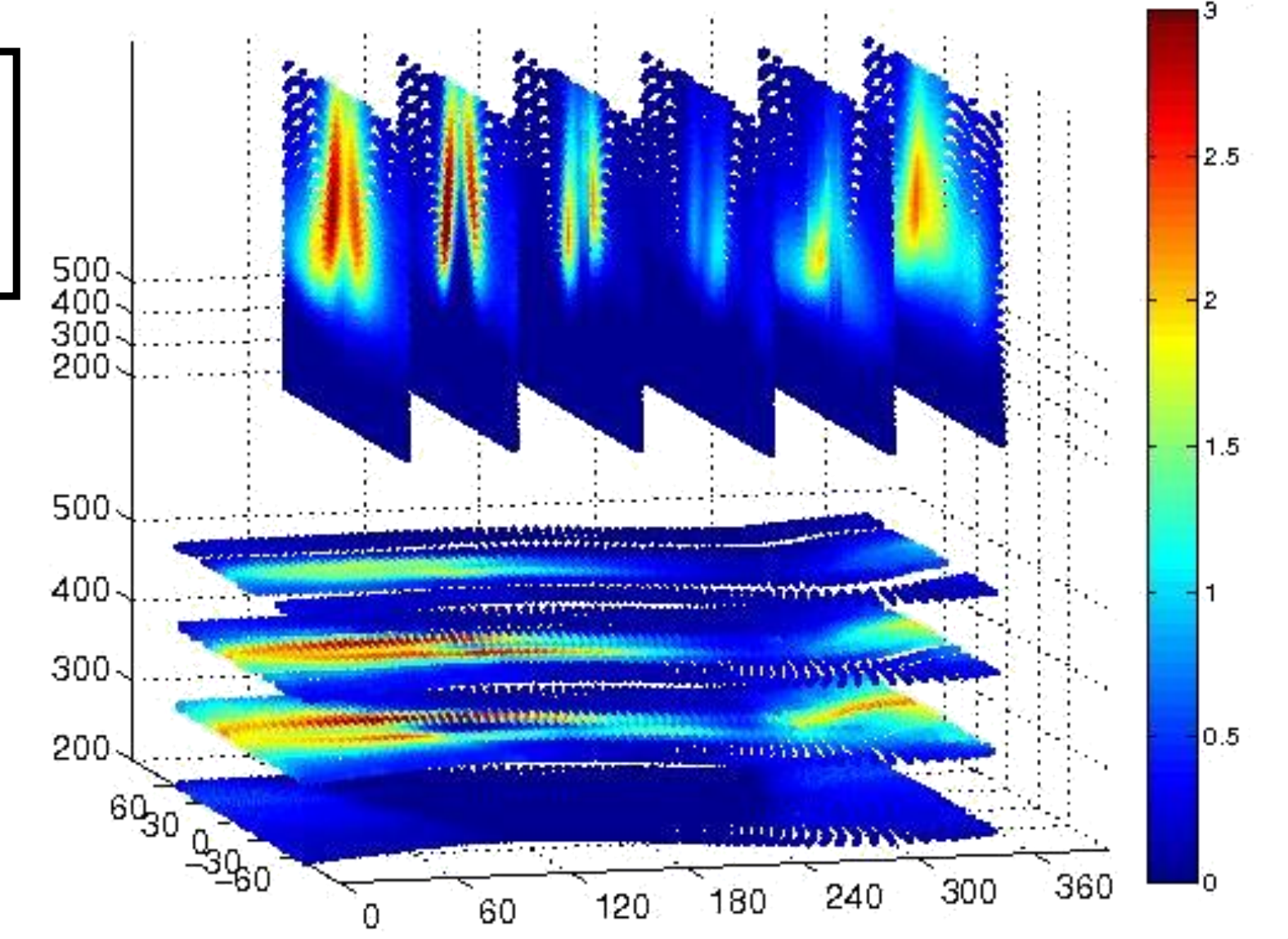


Figure 11. O^+ density distribution of the "truth" at 12:00 UT.

Conclusions

1. Overall, data assimilation of FORMOSAT-7/COSMIC-2 RO data can improve the mid- and low-latitude ionospheric specification.
2. For a given localization length scale, EnSRF with a larger ensemble size (>70) generally performs better for assimilation of RO sTEC. Considering the computational demand, we decide to use EnSRF with 70 ensemble members for GIP/TIE-GCM RO sTEC data assimilation.
3. The covariance localization with GC function with a length scale of 5000-10000km in horizontal direction helps the quality of sTEC data assimilation. On the other hand, the vertical localization appears to have a mixed effect.
4. Our future work includes: a) applying an empirical localization function that designed specifically for sTEC data assimilation. b) updating thermospheric variables during analysis step of EnSRF.