Earth and Space Science

RESEARCH ARTICLE
10.1029/2018EA000447

Key Points:
- Data assimilation experiments with RO TEC from FORMOSAT-3/COSMIC and FORMOST-7/COSMIC-2 observing systems are presented
- The capability of the GSI Ionosphere data assimilation system to reproduce realistic ionospheric features is demonstrated
- The potential of RO TEC data from the upcoming FORMOSAT-7/COSMIC-2 to improve ionospheric weather specification is evaluated

Correspondence to:
C.-T. Hsu,
chihting.hsu@coloardo.edu

Citation:
Hsu, C.-T., Matsuo, T., & Liu, J.-Y. (2018). Impact of assimilating the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO data on the midlatitude and low-latitude ionospheric specification. *Earth and Space Science*, *5*, 875–890. https://doi.org/10.1029/2018EA000447

Received 20 AUG 2018
Accepted 30 OCT 2018
Accepted article online 8 NOV 2018
Published online 10 DEC 2018

Impact of Assimilating the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO Data on the Midlatitude and Low-Latitude Ionospheric Specification

C.-T. Hsu[1], T. Matsuo[1], and J.-Y. Liu[2]

[1]Ann and H. J. Smead Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, CO, USA, [2]Graduate Institute of Space Science, National Central University, Taoyuan, Taiwan

Abstract Global Navigation Satellite System (GNSS) Radio Occultation (RO) missions, such as the Formosa Satellite-3/Constellation Observing System for Meteorology, Ionosphere, and Climate (FORMOSAT-3/COSMIC) and the upcoming FORMOSAT-7/COSMIC-2, provide valuable profiling of the ionized atmosphere for the monitoring of space weather. This study shows that the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 missions' ability to monitor highly variable ionospheric weather can be considerably extended with the help of data assimilation. The Gridpoint Statistical Interpolation (GSI) Ionosphere is a new data assimilation system designed specifically for the low-latitude and midlatitude ionosphere. The capability of the GSI Ionosphere is first demonstrated with actual FORMOSAT-3/COSMIC RO total electron content (TEC) data for January 2013. Features of the ionospheric equatorial ionization anomaly in a coupled plasmasphere ionosphere thermosphere model become more consistent with the TEC maps created with independent ground-based GPS data. The consistency has improved by assimilation of FORMOSAT-3/COSMIC RO data up to about 50% in comparison to the control simulation case without data assimilation. To evaluate the impact of future RO missions on ionospheric weather specification, comparative Observing System Simulation Experiments (OSSEs) are carried out with synthetic RO TEC data. An OSSE of FORMOSAT-7/COSMIC-2 shows that the GSI Ionosphere can improve the ionospheric specification within ±30° geomagnetic latitude by 67% over the control case, which is comparable to the improvement yielded by FORMOSAT-3/COSMIC for 2009 (61%). These results indicate a great potential for improving the monitoring of realistic ionospheric weather with the help of FORMOSAT-7/COSMIC-2 RO TEC data.

1. Introduction

The refraction of radio signals due to plasma in the ionosphere enables remotely probing the electron density distribution in this region through the use of Radio Occultation (RO) technique. Data from the Global Navigation Satellite System (GNSS), including the United States' Global Positioning System (GPS) and the Russia's GLObal NAvigation Satellite System (GLONASS), can be used to probe the Earth's atmosphere by using RO technique. When GNSS radio signals at the L1 and L2 frequencies are occulted by the Earth's atmosphere and measured by GNSS receivers on Low Earth Orbit (LEO) satellites, the total electron content (TEC) along a given radio wave transmission path in the ionosphere can be derived.

The RO technique was first applied to probe the Earth's atmosphere using the GPS/Meteorology (GPS/MET) experiment. Following the successful GPS/MET experiment, the RO technique has been applied to many similar satellite missions in past decades. The Formosa Satellite-3/Constellation Observing System for Meteorology, Ionosphere, and Climate (FORMOSAT-3/COSMIC), launched on 15 April 2006, is the first constellation of satellites dedicated primarily to RO sounding of the Earth's atmosphere (Anthes et al., 2008). The FORMOSAT-3/COSMIC consists of six microsatellites. The main payload on the FORMOSAT-3/COSMIC, the integrated GPS occultation receiver, was able to track GPS frequencies and to provide roughly 1500–2000 GPS RO soundings per day once satellites entered their final mission orbits at 800-km altitude at the end of 2007 (Anthes, 2011). Although the satellites have suffered degradation, five of the six FORMOSAT-3/COSMIC constellation satellites were still operating and continuing to provide RO sounding data 6 years after launch (Schreiner & COSMIC Team, 2014; Yue et al., 2014). Following the successful FORMOSAT-3/COMSIC mission, six low-inclination-angle orbit (24–28.5°) satellites of the FORMOSAT-7/COSMIC-2 mission will be launched in 2018. The main payload of the FORMOSAT-7/COSMIC-2, the TriG

GNSS RO system, is capable of receiving both GPS and GLONASS signals. The FORMOSAT-7/COSMIC-2 low-inclination-angle satellites are expected to provide dense spatial and temporal coverage of high-quality data evenly distributed in low latitude and midlatitude (Yue et al., 2014).

The RO TEC data from FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 observing systems are very valuable for monitoring and prediction of ionospheric weather, especially when they are combined with an empirical or first-principles model of the ionosphere in a process of data assimilation. Yue, Schreiner, and Kuo (2012) and Yue, Schreiner, Kuo, Hunt, et al. (2012) have assimilated FORMOSAT-3/COSMIC and/or FORMOSAT-7/COSMIC-2 RO TEC data into the International Reference Ionosphere (IRI) and the NeQuick. Their results show that data assimilation of FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO TEC can help reproduce ionospheric electron density distribution, especially below the F_2 region. However, both IRI and NeQuick are empirical models that lack forecasting capabilities, and the data assimilation scheme does not incorporate information acquired from the past observations into the data assimilation analysis, as usually achieved in an application of the Kalman filter, in their studies. Lin et al. (2017) have presented a new approach to overcome this issue. They assimilate satellite-based RO TEC from the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 low- and high-inclination-angle satellites and ground-based GPS TEC into the IRI by the Gauss-Markov Kalman filter. The forecast model error covariance in the Gauss-Markov Kalman filter is a linear combination of the analysis error covariance and the climatological background error covariance estimated from the IRI. Both the analysis mean and covariance from previous data assimilation time steps are longitudinally shifted in universal time (UT) to reflect a diurnal variation of the ionosphere with solar zenith angles. Moreover, this method is able to bring information from past observations into the current data assimilation time step. Since the model used in their study is still an empirical model, the model error covariance cannot account for the realistic dynamical evolution of the ionosphere from the previous data assimilation time step to the current one. Assimilation of the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO data into an empirical model is a robust approach to mapping the global ionosphere, and suitable for climatology study, but is not designed to predict the short-term changes in ionospheric weather.

For ionospheric weather prediction, ideally, a first-principles model need to be incorporated in data assimilation systems. By combining the observation information from the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 missions into a numerical model, both nowcasting and forecasting of the ionosphere can be improved. Two widely recognized numerical ionospheric data assimilation systems, the Global Assimilative Ionospheric Model joint effort of University of Southern California and Jet Propulsion Laboratory (USC-JPL GAIM) and the Global Assimilation of Ionospheric Measurement developed by Utah State University (USU GAIM), had presented remarkable improvements of the ionospheric specification by assimilating FORMOSAT-3/COSMIC data (Komjathy et al., 2010; Scherliess et al., 2009; Schunk et al., 2016). Komjathy et al. (2010) show that assimilating both ground-based GPS TEC and FORMOSAT-3/COSMIC RO TEC into the USC-JPL GAIM can improve vertical electron density distributions. Scherliess et al. (2009) have assimilated RO data from the FORMOSAT-3/COSMIC satellites into the USU GAIM system in order to specify the low-latitude and midlatitude ionosphere together with the ionospheric drivers and their temporal and spatial variability. Lomidze et al. (2016) have assimilated FORMOSAT-3/COSMIC RO TEC data into the USU GAIM system to investigate ionospheric evening anomalies. Since the neutral atmosphere of the forecast model used in these two systems is parameterized using empirical models, coupling of the neutral and plasma species is not accounted for in a self-consistent manner.

The predictability of the ionosphere can be considerably improved by systematically incorporating the ion-neutral coupling into a data assimilation scheme (Hsu et al., 2014). The DART/TIE-GCM, composed of the National Center for Atmospheric Research (NCAR) Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM; Richmond et al., 1992) and ensemble filters implemented in the NCAR Data Assimilation Research Testbed (DART; Anderson et al., 2009), can take the neutral-ion coupling into account (Hsu et al., 2014; Lee et al., 2012; Matsuo et al., 2013; Matsuo & Araujo-Pradere, 2011). Lee et al. (2012) assimilated synthetic electron density profiles from the FORMOSAT-7/COSMIC-2 high- and low-inclination-angle satellites into the TIE-GCM and showed that compared with the Observing System Simulation Experiments (OSSE), the FORMOSAT-3/COSMIC, monitoring of ionospheric weather can be improved. Since then, the FORMOSAT-7/COSMIC-2 high-inclination-angle satellite mission have been canceled; thus, the impact of FORMOSAT-7/COSMIC-2 RO TEC needs reevaluation.

A new coupled data assimilation system, the Gridpoint Statistical Interpolation (GSI) Ionosphere, has been developed with the National Oceanic and Atmospheric Administration's (NOAA's) GSI Ensemble Square Root Filter (EnSRF) and a coupled model of the Global Ionosphere Plasmasphere (GIP) and the TIE-GCM (the GIP/TIE-GCM; Pedatella et al., 2011). Unlike earlier studies with the DART/TIEGCM (Chartier et al., 2016; Chen et al., 2016), the GSI Ionosphere does not require extrapolation of the modeled topside electron density to assimilate TEC data and can extract more information from TEC data by considering the contributions of both ionospheric and plasmaspheric electron densities. It can also account for thermosphere-ionosphere-plasmasphere coupling in a self-consistent manner. Hsu et al. (2018) present a number of 12-hr OSSEs for FORMOSAT-7/COSMIC-2 RO TEC using the GSI Ionosphere and have shown that the error in low-latitude and midlatitude ionosphere can be reduced by 67.85% over a control GIP/TIEGCM simulation case with the help of assimilating. The behavior of the root-mean-square difference (RMSD) between these OSSEs and the *truth* simulation has underscored that the ionospheric specification continues evolving during forecast cycles after data assimilation updates due to the physical, chemical, and electrodynamic processes described in the GIP/TIEGCM.

Following Hsu et al. (2018), using actual RO TEC data, this study focuses on demonstrating the capabilities of the GSI Ionosphere to reproduce realistic features of the low-latitude and midlatitude ionosphere under geomagnetically quiet and low solar activity conditions and examining observation impacts of the RO TEC from FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 low-inclination-angle satellites on ionospheric weather specification and forecasting. Because the GSI Ionosphere is designed specifically for low-latitude and midlatitude regions with the upcoming FORMOSAT-7/COSMIC-2 mission in mind, in this study, only low-latitude and midlatitude (within ±60°) FORMOSAT-3/COSMIC RO TEC data are used to evaluate the capabilities of the GSI Ionosphere. Since atomic oxygen ion density ($[O^+]$) is one of the major prognostic variables that determine the ionospheric plasma density distribution in the GIP/TIE-GCM, the experiment results are discussed in terms of estimation errors of $[O^+]$. In addition, the Center of Orbit Determination in Europe (CODE) Global Ionospheric Maps (GIMs), generated from vertical TEC (vTEC) data of worldwide GNSS ground stations of the International GNSS Service (IGS) by using the spherical harmonic expansion (Schaer, 1999), are used as independent validation data.

2. Data Assimilation System

The GSI is an operational data assimilation system used at NOAA for numerical weather prediction. Recently, an ionospheric data assimilation capability, the GSI Ionosphere, has been added into the GSI for the purpose of ionospheric numerical weather prediction (Hsu et al., 2018).

The forecast model used in the GSI Ionosphere is the GIP/TIE-GCM, which is a coupled model of thermosphere, ionosphere, and plasmasphere. The neutral atmosphere and electrodynamics are simulated by the TIE-GCM in regular pressure coordinates, and the ionosphere and plasmasphere plasma is simulated by the GIP in geomagnetic fluxtube coordinates. The GIP consists of two model domains: a closed-field-line low-latitude and midlatitude domain and an open-field-line high-latitude domain. Since the plasma in the low-latitude and midlatitude domain is simulated along closed geomagnetic field lines, which are extended to about 19,000-km altitude, contributions of the electron density from the plasmasphere to TEC can be directly estimated using this model, which makes TEC data assimilation straightforward. By using the GIP/TIE-GCM, the GSI Ionosphere overcomes the upper boundary issue encountered in TEC assimilation by the DART/TIE-GCM. The main drivers of the GIP/TIE-GCM model include the F10.7 index (F107), cross-tail potential drop (CP), auroral hemispheric power (HP), and atmospheric tides. Note that the neutral atmosphere is an integral part of the model, and thus, ion-neutral coupling can systematically be taken into account by the assimilation system.

Specifically, an ensemble-based data assimilation method, the EnSRF, developed by Whitaker and Hamill (2002) employed in the GSI is used in this data assimilation system as described in Hsu et al. (2018). The EnSRF can be considered as a variant of the traditional Kalman filter (Kalman, 1960) and the ensemble Kalman filter (EnKF; Evensen, 1994). In Kalman filters, a data assimilation cycle consists of two steps: the analysis step, in which model state variables are updated by assimilation of observations, and a forecast step, in which updated state variables are used to initialize a forecast model and are integrated forward in time to the next analysis step by the model. The analysis increment is a function of the Kalman gain, which is estimated

from model ensembles in the EnKF and EnSRF, and thus, the finite ensemble size that is considerably smaller than the state dimension can lead to filter divergence and erroneous updates due to sampling errors. This issue can be usually corrected by covariance inflation and covariance localization auxiliary methods (Anderson & Anderson, 1999; Hamill et al., 2001; Houtekamer & Mitchell, 2001). The role of covariance localization and inflation in assimilation of FORMOSAT-7/COSMIC-2 RO TEC has been carefully investigated by Hsu et al. (2018).

3. OSSEs and Data Assimilation Experiments

Three sets of experiments are presented in this paper. Since the GIP/TEI-GCM's performance is poor under extremely low solar activity conditions, 1 January 2013 is chosen as a primary experiment period. An additional OSSE for the FORMOSAT-3/COSMIC mission on 1 January 2009 at the peak of FORMOSAT-3/COSMIC RO data volume is included for comparison.

In the first set of experiments (section 4.1), OSSEs designed for the FORMOSAT-3/COSMIC mission are executed to examine effects of the covariance localization scheme and ensemble size in the EnSRF on the $[O^+]$ analysis errors in the low-latitude and midlatitude region. A similar experiment for the FORMOSAT-7/COSMIC-2 has been presented in Hsu et al. (2018), but the effect of the EnSRF parameters for FORMOSAT-3/COSMIC have not been addressed before. Second, assimilation experiments of actual FORMOSAT-3/COSMIC data will be conducted to examine how the GSI Ionosphere can facilitate tracking realistic features of ionospheric weather (section 4.2). Finally, comparative OSSEs for low-latitude and midlatitude FORMOSAT-3/COSMIC data and FORMOSAT-7/COSMIC-2 low-inclination satellites data are presented in section 4.3. Note that FORMOSAT-3/COSMIC data with a tangent point above 45° latitude or below −45° latitude are excluded in the following experiments.

The model ensembles used in the EnSRF are initialized by perturbing the three main model drivers: F107, CP, and HP, according to Gaussian distributions. The mean values of F107, HP, and CP are set to 120 SFU, 16 GW, and 45 kV, and the standard deviations are set to 15 SFU, 2 GW, and 10 kV, respectively. Random draws of the F107 index are independent of both HP and CP random draws, but randomized HP values are set to correlate to CP. The spin-up period is composed of a 23-day stand-alone TIE-GCM model run and a 5-day GIP/TIE-GCM model run. For all OSSEs in this study, synthetic data sets of FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO TEC are sampled from a Nature Run (NR) of the GIP/TIE-GCM with F107, HP, and CP values of 140 SFU, 18 GW, and 55 kV, respectively.

In section 4.1, the impact of FORMOSAT-3/COSMIC RO TEC data assimilation on ionospheric specification is assessed by comparing two sets of OSSEs: Experiment A and Experiment B. In Experiment A, EnSRFs with model ensemble sizes of 10, 40, 70, and 100 are executed with an identical covariance localization setting. In Experiment B, the ensemble size is set to the one selected based on Experiment A, and the Gaspari and Cohn (GC) correlation function (Gaspari & Cohn, 1999) with different localization length scales are applied in the covariance localization schemes. Following the experiments presented by Hsu et al. (2018), the GC functions with horizontal localization length scales of 500, 1,000, 5,000, and 10,000 km and vertical localization length scales of 0.5, 1, 3, and 7 ln (mb) are employed in Experiment B. For comparison, OSSEs without horizontal and/or vertical covariance localization are also executed. Synthetic data used in Experiment A and Experiment B are sampled from the NR based on the geometry of the GPS and the FORMOSAT-3/COSMIC on 1 January 2013. These synthetic TEC data will be assimilated into the model hourly from 00 UT to 12 UT by using the GSI Ionosphere. Other experiment settings presented in the first part of this study are also the same as in Hsu et al. (2018).

A data assimilation experiment on actual FORMOSAT-3/COSMIC RO TEC on 1 January 2013 is presented in section 4.2 and is referred to as Experiment C. This period is under quiet geomagnetic conditions with solar activity in the low to moderate range. Figure 1 shows the Kp index, Dst index, and F107 index from 15 December 2012 to 15 January 2013. The overall geomagnetic and solar activity conditions are low. The F107 on 1 January 2013 is within one standard deviation from mean of the model ensembles used in the OSSEs. Therefore, the same set of model ensembles used for Experiments A and B is employed, and the covariance localization scheme and ensemble size are based on the results of Experiments A and B.

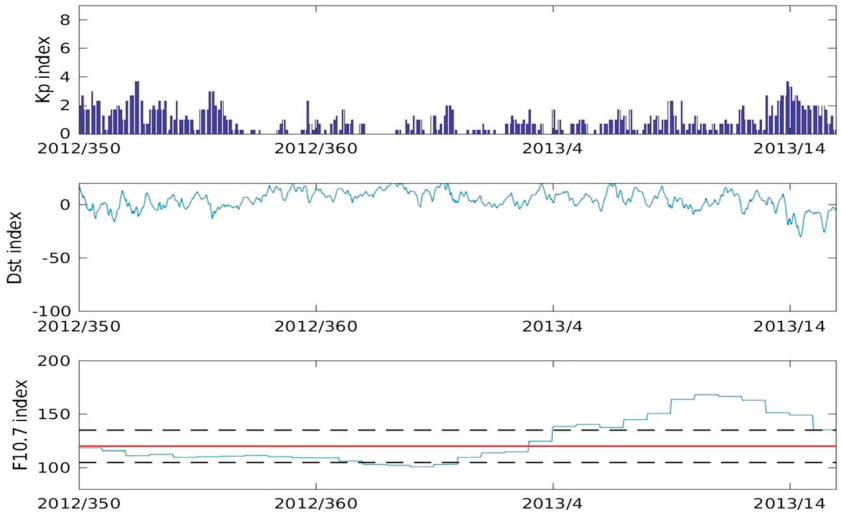

Figure 1. Kp index, Dst index, and F107 index from 15 December 2012 (day of year is 350) to 15 January 2013 (day of year is 15). In the bottom panel, the blue solid line shows the F10.7 index values and the red solid line and black dashed lines show value of mean and standard deviation of F10.7 index used to generate model ensemble.

In section 4.3, OSSEs are carried out from 00:00 to 12:00 UT on 1 January 2013 to compare the FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 observing systems in the low latitude and midlatitude and are referred to as Experiment D. Experiment D includes an additional OSSE for the FORMOSAT-3/COSMIC 00:00 to 12:00 UT on 1 January 2009. For these OSSEs, denoted as OSSE-F3C-09, OSSE-F3C-13, and OSSE-F7C2, different geometries of the LEO and GNSS satellites and RO event frequencies are used to generate synthetic RO TEC data. The covariance localization scheme and ensemble size are chosen for OSSE-F3C-09 and OSSE-F3C-13 based on the results of Experiments A and B. A horizontal localization length scale of the GC function is set to 5,000 km, and a 70-member EnSRF is chosen for OSSE-F7C2 based on Hsu et al. (2018). By comparing these OSSEs, observation impacts of the FORMOSAT-7/COSMIC-2 RO TEC data on ionospheric specification and forecasting will be evaluated.

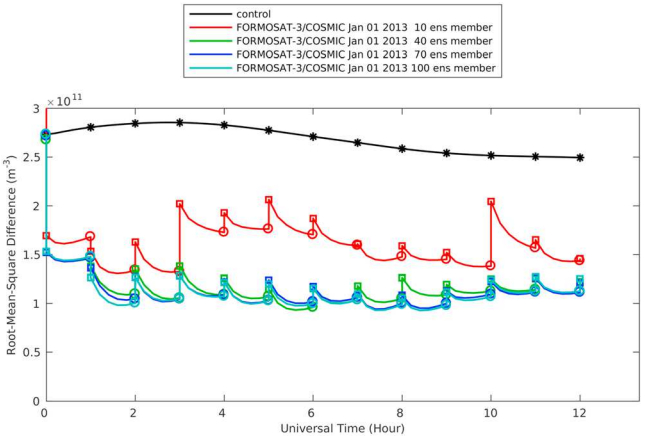

Figure 2. Root-mean-square difference (RMSD) of [O^+], between Nature Run and the ensemble mean of FORMOSAT-3/COSMIC Observing System Simulation Experiments (OSSEs) with the 10-, 40-, 70-, and 100-member Ensemble Square Root Filter, computed over the middle- and low-geomagnetic latitude region from 200- to 500-km altitude, from 00:00 to 12:00 UT. The circles and squares indicate the prior and posterior RMSDs at analysis steps, respectively. The black line is for the RMSD of the 100-member control ensemble forecast simulation. Data used here are synthetic total electron content data generated based on the geometry of FORMOSAT-3/COSMIC orbits during 1 January 2013.

4. Results

4.1. OSSEs of FORMOSAT-3/COSMIC GNSS RO TEC

4.1.1. Experiment A: Impact of Ensemble Size on Quality of Ensemble Data Assimilation

The objective of Experiment A is to determine the impact of ensemble sizes ranging from 10 to 100, with an interval of 30, on the quality of the assimilation analysis. The RMSD of [O^+] between OSSEs and NR over the middle- and low-geomagnetic-latitude region from 200- to 500-km altitudes is used as a measure of quality of the global data assimilation analysis. Figure 2 shows the RMSD from the OSSEs with the 10-, 40-, 70-, and 100-member EnSRFs. The GC localization function is employed to localize the covariance in the horizontal direction with a scale length of 10,000 km. No localization is applied in the vertical direction. For comparison, the RMSD of the 100-member control ensemble forecast experiment is also shown in Figure 2. The control ensemble forecast experiment is essentially a collection of all GIP/TIE-GCM ensemble simulations without data assimilation during the same period.

Overall, the RMSD of OSSEs are generally smaller than that of the 100-member control ensemble forecast experiment over the 12-hr experiment period, indicating that the monitoring of ionospheric specification becomes better with the help of assimilation of FORMOSAT-3/COSMIC RO TEC data. The performance of the EnSRF improves with an increasing

number of ensemble members. The OSSE with the 10-member EnSRF shows the worst performance among four filters. The OSSEs with the 40-, 70-, and 100-member EnSRFs result in a roughly comparable level of the RMSD over the entire data assimilation cycles. At the end of 12-hr experiments, the posterior RMSDs are 14.60×10^{10}, 12.24×10^{10}, 12.11×10^{10}, and 12.53×10^{10} m^{-3} for the OSSEs with 10-, 40-, 70-, and 100-member EnSRFs. The ratios of the posterior RMSDs to the RMSD of control ensemble forecast experiment are roughly 0.59, 0.50, 0.49, and 0.50 for the OSSEs with 10-, 40-, 70-, and 100-member EnSRFs. Generally, a smaller ratio means a smaller difference compared with the control ensemble forecast experiment, indicating a better OSSE result. In the following experiments shown in this study, the 70-member EnSRF is applied to all experiments of FORMOSAT-3/COSMIC RO TEC as well as FORMOSAT-7/COSMIC-2 RO TEC.

The GIP/TIE-GCM ensemble forecast within the EnSRFs cycling starts to become unstable, with the posterior RMSDs larger than prior RMSDs after the third analysis step at 02:00 UT. Unlike the result shown in the OSSEs of FORMOSAT-7/COSMIC-2 in (Hsu et al., 2018), this poor EnSRF performance cannot be corrected by increasing the ensemble size. Moreover, the model error becomes smaller over the course of given forecast steps. These behaviors will be discussed in detail later.

4.1.2. Experiment B: Impact of Localization Scheme on Quality of Ensemble Data Assimilation

In Experiment B, the impact of localization schemes on the quality of the data assimilation analysis is examined. Figure 3 shows the posterior RMSD from the OSSEs with the 70-member EnSRF with and without covariance localization at 00:00, 02:00, 04:00, 06:00, 08:00, 10:00, and 12:00 UT. The GC functions with four different vertical localization length scales and four horizontal localization length scales are applied. In addition, the EnSRFs with no vertical and/or horizontal localization are also executed.

The overall RMSDs that are shown in Figure 3 are smaller than the RMSD of the control ensemble forecast experiment that is shown in Figure 2, indicating positive impacts of the FORMOSAT-3/COSMIC RO TEC data on the quality of the assimilation analysis. Similarly to the OSSEs of FORMOSAT-7/COSMIC-2 shown in Hsu et al. (2018), the results are more sensitive to a choice of the horizontal localization length scales than that of the vertical localization length scales. The RMSDs of the OSSEs with localization with the smallest horizontal localization length scale are considerably larger than the rest of OSSEs. At the end of the data assimilation period at 12:00 UT, the EnSRF with horizontal localization with a length scale of 10,000 km and with no vertical localization leads to the smallest RMSD. In the following FORMOSAT-3/COSMIC RO TEC ensemble data assimilation experiments, the GC function will be applied to localize the covariance in horizontal direction with a horizontal length scale of 10,000 km. Note that, as mentioned earlier, a horizontal length scale of 5,000 km is used to localize the covariance in the horizontal direction for FORMOSAT-7/COSMIC-2 RO TEC ensemble data assimilation experiments according to Hsu et al. (2018).

4.2. Experiment C: FORMOSAT-3/COSMIC GNSS RO TEC Data Assimilation

In Experiment C, the GSI Ionosphere is applied to a case study. FORMOSAT-3/COSMIC RO TEC data obtained over 1–2 January 2013 are assimilated into the GSI Ionosphere hourly to reproduce realistic features of the ionosphere under low solar activity and geomagnetically quiet conditions. The 70-member EnSRF is applied with the covariance localized by the GC function in horizontal direction with a 10,000-km localization length scale. The same set of model ensemble simulation used in previous OSSEs is employed here. Model drivers for the GIP/TIE-GCM are held unchanged for each ensemble over the course of data assimilation experiment.

The GIMs of vTEC obtained from both the 70-member control ensemble forecast experiment and data assimilation experiments are here compared with the CODE GIMs. Starting at the end of 2002, the CODE GIMs constructed from the ground-based GPS data are available at a 2-hr interval, and there might not be sufficient data over the Pacific and Atlantic Oceans for mapping of the TEC. The GIMs of the control ensemble forecast experiment and data assimilation experiments are computed from the ensemble mean and by integrating electron densities into the model grid vertically from the bottom of the model domain at about 90 km to the top of the model boundary at about 19,000 km. In the following, GIMs of data assimilation experiments are referred to as the GSI GIMs, and the GIMs of control ensemble forecast experiment are referred to as the control GIMs. It is important to note that unlike Chartier et al. (2016), Chen et al. (2016), and Lin et al. (2017), since the GIP/TIE-GCM includes the plasmasphere, there is no additional correction or extrapolation when computing GIM from this system.

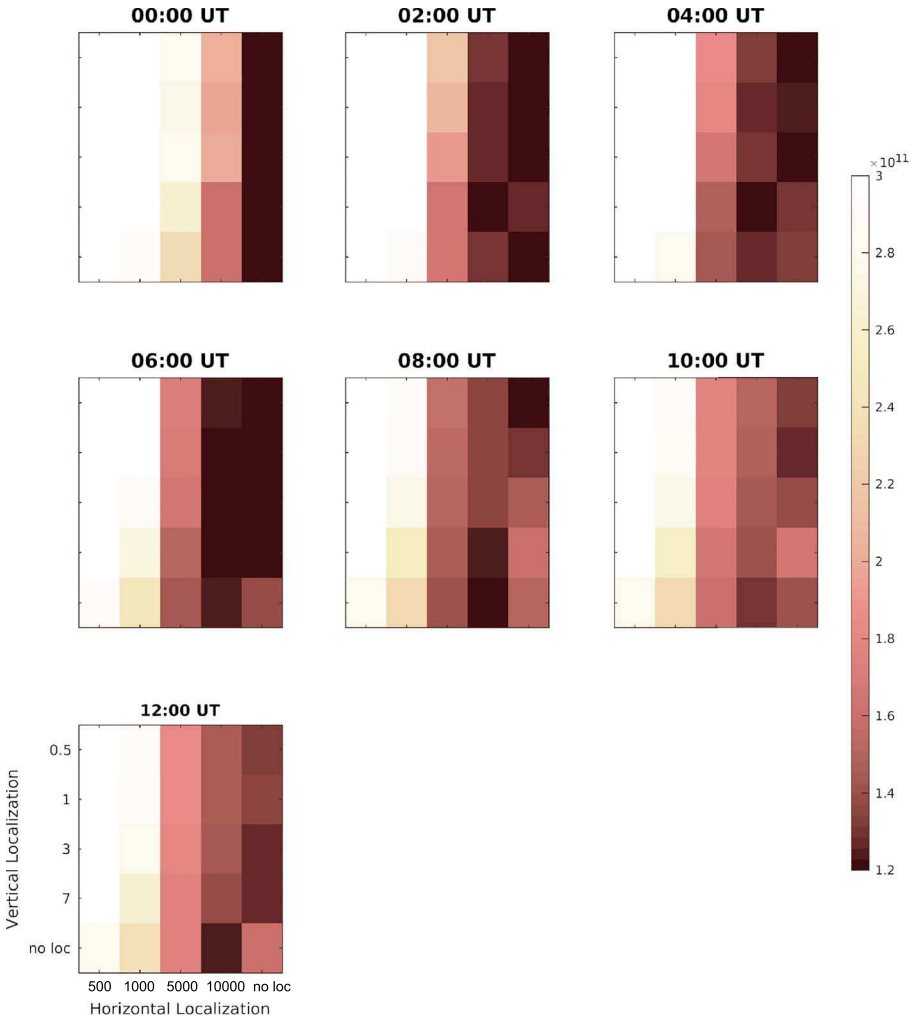

Figure 3. The posterior root-mean-square differences of Observing System Simulation Experiments (OSSEs), computed over the low- and middle-geomagnetic latitude region from 200- to 500-km altitude, are shown at 00:00, 02:00, 04:00, 06:00, 08:00, 10:00, and 12:00 UT for comparison of the Ensemble Square Root Filter (EnSRF) with different covariance localizations. The EnSRF with covariance localization using the GC function with four different vertical localization length scales (including 0.5, 1, 3, and 7 scale height) and four different horizontal localization length scales (including 500, 1000, 5,000, and 10,000 km) are applied in the OSSEs. The posterior for the EnSRF without covariance localization is shown for comparison.

Figure 4 shows the prior and posterior GSI GIMs, CODE GIMs, and control GIMs over the low-latitude and mid-latitude region from 00:00 UT of 1 January 2013 to 00:00 UT 2 January 2013 at a 4-hr interval. At the beginning of the data assimilation period, the vTEC magnitude on both peaks of the Equatorial Ionization Anomaly (EIA) in the GSI GIM is much larger than that in the CODE GIM. In the GSI GIM, the EIA features from local daytime and sunset time are symmetric on the both sides of the geomagnetic equator, which is a typical vTEC distribution in a steady-state theoretical model simulation under geomagnetically quiet conditions. On the other hand, the EIA features in the CODE GIM are considerably asymmetric. After a few data assimilation cycles, as more data are assimilated into the model by the GSI Ionosphere, the vTEC magnitude on both peaks of the EIA in the GSI GIM becomes smaller and the location of EIA features in the GSI GIM becomes closer to that in the CODE GIM. Although the general ionospheric features in the GSI GIM and the CODE GIM appear to become closer over time, the vTEC magnitude of EIA peaks in the GSI GIM is considerably larger in comparison to the other two GIMs at 20:00 UT. A close examination of the posterior and prior GSI GIMs reveals that in the analysis step, the EnSRF overcorrects the state, causing a large peak vTEC in the southern hemispheric EIA region.

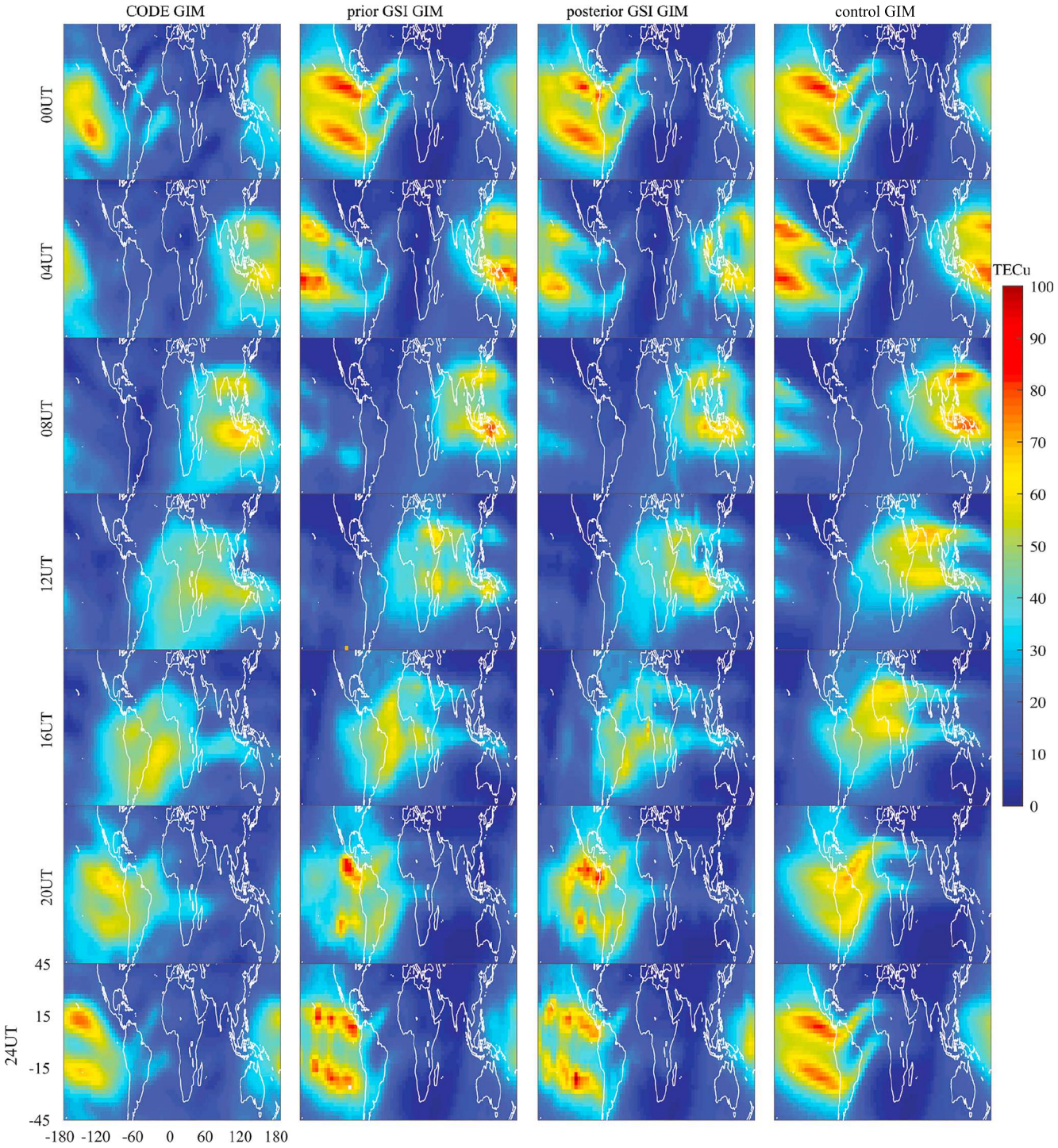

Figure 4. Low-latitude and midlatitude (left column) Center of Orbit Determination in Europe Global Ionospheric Map (CODE GIM), (middle column) Gridpoint Statistical Interpolation (GSI) GIM, and (right column) control GIM over the 1–2 January 2013 with a 4-hr interval. The *x* and *y* axes represent geographic longitude and latitude in degrees.

With the purpose of qualifying the difference in the CODE GIMs, GSI GIMs, and control GIMs, a Multivariate Alteration Detection (MAD) developed by Nielsen and Conradsen (1997) for remote sensing imagery and chi-square test is applied to our GIM images. The MAD is an algorithm for detecting the change of two multivariate images based on the canonical correlation analysis. Here the MAD is applied to quantify the difference

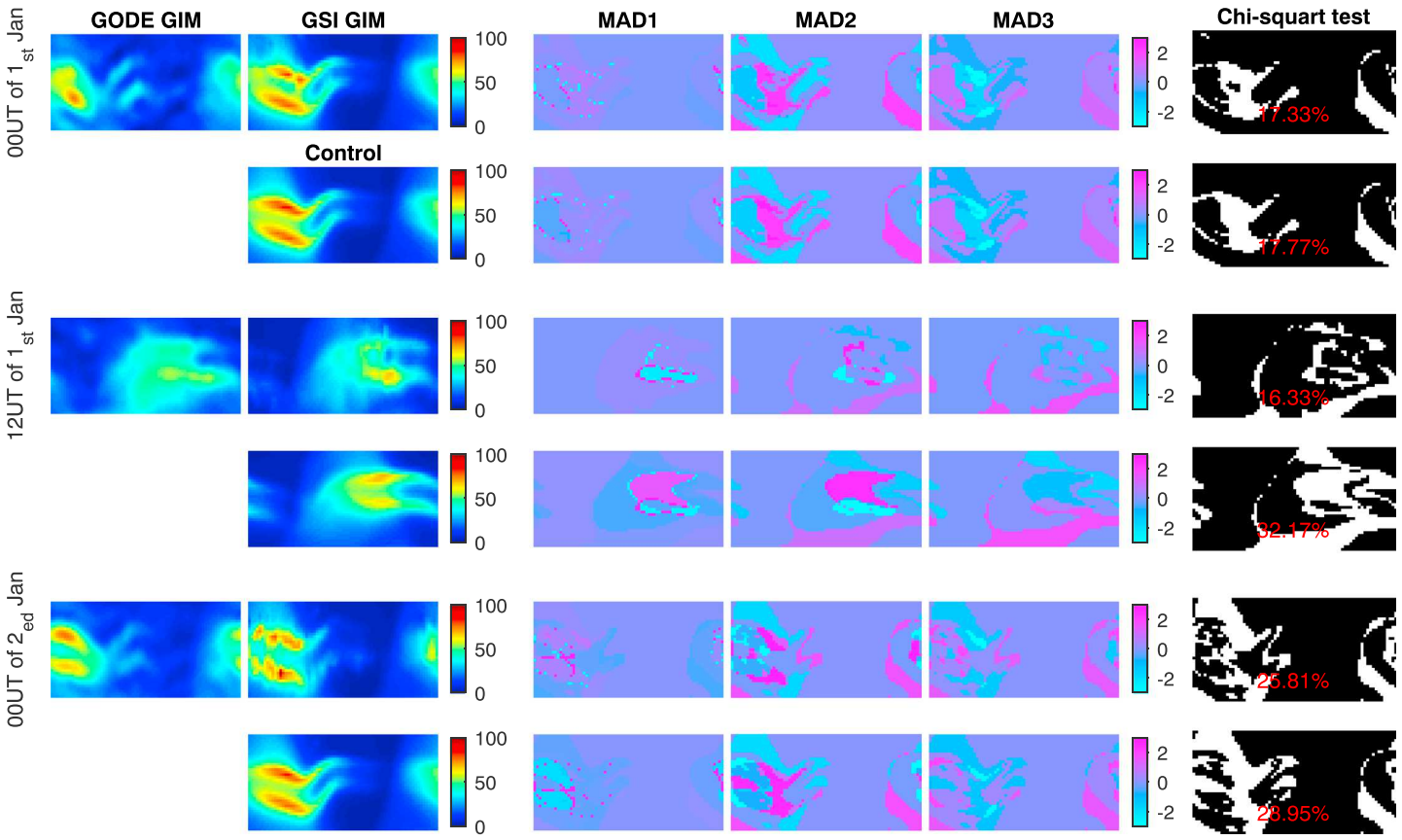

Figure 5. Global Ionospheric Maps (GIMs) and the Multivariate Alteration Detection (MAD) analysis results at 00:00 and 12:00 UT of 1 January and 00:00 UT of 2 January. The first column lists the GODE GIMs at three different UT. The second column lists the GSI GIM (first, third, and fifth panels from the top to bottom) and the GIMs of the control ensemble forecast experiment (second, fourth, and sixth panels from the top to bottom). The third to fifth columns lists the MAD1, MAD2, and MAD3 of the GSI GIMs and the GIMs of control ensemble forecast experiment. The final column is the result of chi-square statistics. The black regions indicate regions with no significant difference from the CODE GIMs, and the white regions indicate regions with the significant difference as defined in the main text. Percentage in red is percentage of grid point defined as significant difference.

of the GSI GIM from the CODE GIM in comparison to the differences of the control from the CODE GIM by decomposing maps into three orthogonal components (MAD1, MAD2, and MAD3). The last MAD component (MAD3) indicates a multidimensional direction to which that largest difference information can be projected, and the second component (MAD2) has the largest difference information subject to MAD3. Detailed description of MAD is presented by Nielsen et al. (1998), Nielsen (2007), and Morton and Nielsen (2008). After that, the chi-square test is applied to these MAD components grid point by grid point to define the regions with significant difference between two maps (Morton & Nielsen, 2008). By counting the number of grid point that defined as regions with significant difference for GSI GIMs and control GIMs, we are able to qualify the impact of data assimilation. The confident interval of chi-square test used as a threshold value to define the region of the significant difference between CODE GIM and GSI GIM is 30%. Results of MAD and chi-square test are shown in Figure 5. A value of the MAD variate close to zero means no significate difference between two maps at the same location.

At the first assimilation cycle at 00:00 UT, the number of grid points with significant difference from the CODE GIM is roughly the same between the GSI GIM (468 grid points out of the total of 2,701 grid points, or 17.3%) and the control GIM (480 out of 2,701, or 17.8%). Data assimilation has little impact on making the GSI GIM appear closer to the CODE GIM in comparison to the control GIM.

At 12:00 UT, the number of grid points with significant difference is 16.3% for the data assimilation case and 32.2% for the control case. Data assimilation of the FORMOSAT-3/COSMIC RO TEC helps the GSI GIM become closer to the CODE GIM in comparison to the control ensemble forecast experiment by about 50%.

Table 1
Experiment Settings of OSSEs in Experiment D

Setting parameters	OSSE-F3C-09	OSSE-F3C-13	OSSE-F7C2
Data	FORMOSAT-3/COSMIC 1 January 2009 low-latitude and midlatitude data	FORMOSAT-3/COSMIC 1 January 2013 low-latitude and midlatitude data	FORMOSAT-7/COSMIC-2 low-inclination-angle satellites
Average number of soundings that are assimilated per cycle	96	68	368
Ensemble size	70	70	70
Horizontal localization scheme	GC function with 10,000-km localization length scale	GC function with 10,000-km localization length scale	GC function with 5,000-km localization length scale
Vertical localization scheme	No vertical localization	No vertical localization	No vertical localization

At the end of data assimilation period at 00:00 UT of 2 January 2013, the percentage of grid points with significant difference for the data assimilation case and control case are 25.4% and 29%, respectively. Comparing to 00:00 UT on the previous day, the percentage of grid points with significant difference for control case is increased by 11.2% (from 17.8% to 29%), and that for data assimilation case is increased by 8.1% (from 17.3% to 25.4%). The control GIM differs more from the CODE GIM at 00:00 UT of 2 January 2013, but the smaller percentage increase for the data assimilation case indicates the extent of observation impacts of the FORMOSAT-3/COSMIC RO TEC.

Experiment C demonstrates the improvement of the model's ability to track ionospheric weather and to reproduce general GIM features by assimilation of the FORMOSAT-3/COSMIC RO TEC using the GSI Ionosphere. However, partly because of a sparse data distribution of the FORMOSAT-3/COSMIC during 2013 in comparison to earlier periods of the FORMOSAT-3/COSMIC mission and the FORMOSAT-7/COSMIC-2, the GSI Ionosphere system overcorrects the magnitude of the EIA peaks from 10:00 to 22:00 UT. This issue will be addressed further later.

4.3. Experiment D: Comparison of FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 Observing Systems Simulation Experiment

OSSEs of the FORMOSAT-3/COSMIC mission (the OSSE-F3C-09 and the OSSE-F3C-13) and the FORMOSAT-7/COSMIC-2 mission (the OSSE-F7C2) from Experiment D are compared. The purpose of Experiment D is to compare the best ionospheric state estimation obtained for these OSSEs. Since FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 are different observing systems, as shown in Experiment B and Hsu et al. (2018), optimal EnSRF parameter settings, such as localization length scale and ensemble size, are likely different for OSSE. Table 1 summarizes the experiment settings of these three OSSEs. Note that the OSSE-F3C-13 shown here is same as the OSSE with a 70-member EnSRF shown in Figure 2 (blue line).

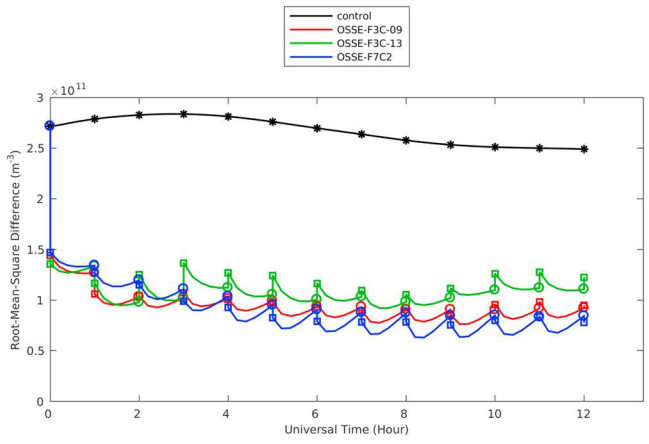

Figure 6. Similar to Figure 2. Root-mean-square difference of $[O^+]$, between the ensemble mean of Observing System Simulation Experiments (OSSEs) and Nature Run, over the middle- and low-geomagnetic latitude region from 200- to 500-km altitude, from 00:00 to 12:00 UT. The root-mean-square difference of the OSSE-F3C-09 (red), the OSSE-F3C-13 (green), and the OSSE-F7C2 (blue), and the 70-member control ensemble forecast experiment (black) are shown.

Figure 6 shows the RMSD of $[O^+]$ from OSSEs over the course of 12-hr data assimilation experiment. Overall, the RMSD of OSSEs are considerably smaller than that of the 70-member control ensemble forecast experiment over the entire data assimilation period, implying positive observation impact of both the FORMOSAT-3/COSMIC and FROMOSAT-7/COSMIC-2 mission on improving the numerical model's ability to monitor and forecast the low-latitude and midlatitude ionosphere. At the beginning of Experiment D, the RMSD of the OSSE-F3C-09 and OSSE-F3C-13 are slightly smaller than the RMSD of the OSSE-F7C2. This is because the horizontal localization length scale used in the OSSE-F3C-09 and the OSSE-F3C-13 is significantly larger than the one used in the OSSE-F7C2. The long-range covariance is more effective in correcting the global biases. Recall that the plasma density in the NR is biased, being set higher than the ensemble mean. Nevertheless, the RMSD of the OSSE-F7C2 continuously becomes smaller over the course of the data assimilation period and is the smallest among three OSSEs at the end. The RMSD of OSSEs after the final update at 12:00 UT are 9.43×10^{10}, 14.57×10^{10},and 8.16×10^{10} m^{-3} for the OSSE-F3C-09, the OSSE-F3C-13, and the OSSE-F7C2, respectively, and their

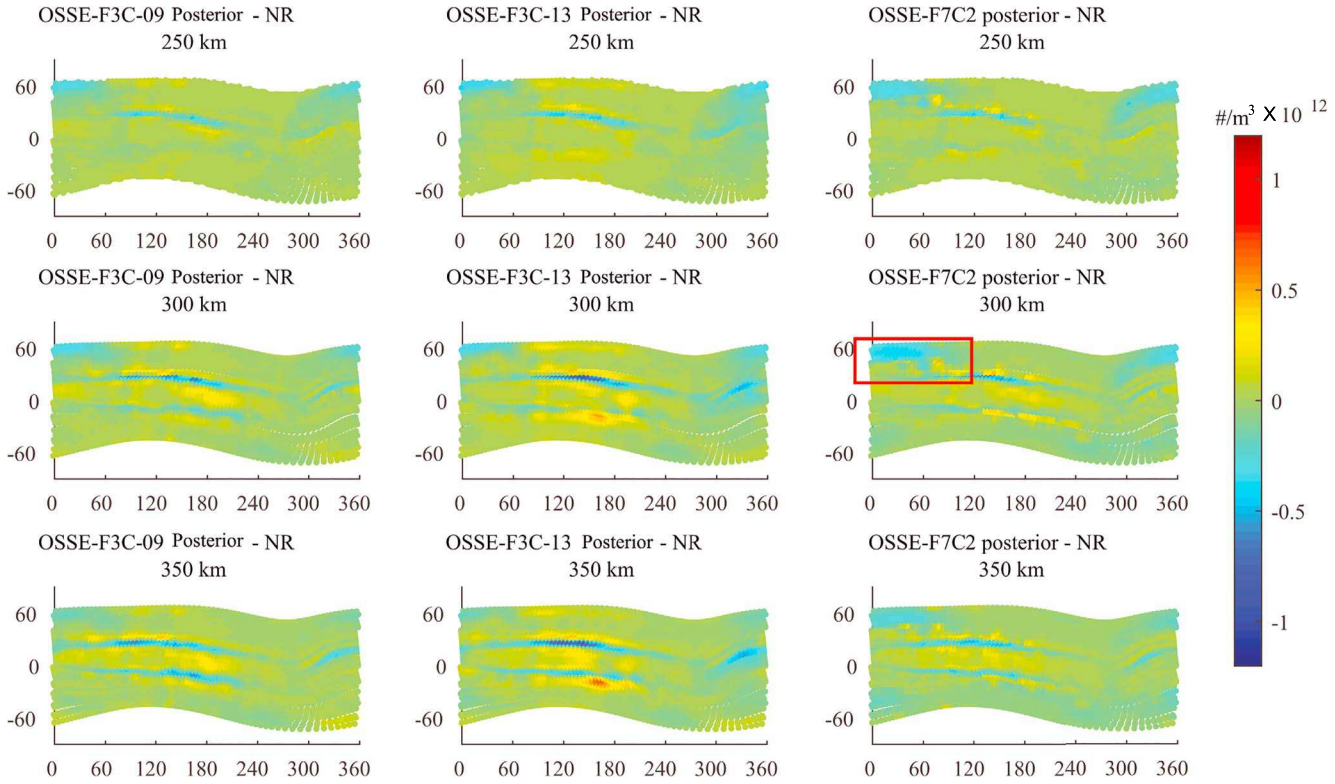

Figure 7. Differences of the $[O^+]$ between Observing System Simulation Experiments (OSSEs) and Nature Run at 12:00 UT after final update at 250-, 300-, and 350-km altitude, respectively. The red square denotes a region with significant difference between the OSSE-F7C2 and NR. The *x* and *y* axes represent geographic longitude and latitude in degrees.

ratios to the RMSD of their respective control ensemble forecast experiments are 0.44, 0.68, and 0.38. Figure 7 shows the difference of $[O^+]$ between OSSEs and NR at the end of experiments at different altitude. Although the OSSE-F7C2 shows a great improvement in the lower latitude ionospheric specification compared with the OSSE-F3C-09 and OSSE-F3C-13, the higher latitude improvement is less since the FORMOSAT-7/COSMIC-2 mission is mainly focusing on probing low-latitude and midlatitude regions. The ratios of the RMSD of OSSEs to the RMSD of control experiment within the ±30° geomagnetic latitude region and other regions are listed in Table 2. In the lower latitude region, the ratio for the OSSE-F7C2 is significantly smaller than that for the OSSE-F3C-09 and OSSE-F3C-13. Experiment D suggests a great potential of the FORMOSAT-7/COSMIC-2 low-inclination-angle satellites to improve the performance of ionospheric numerical weather prediction in lower latitudes through the RO TEC assimilation after all satellites enter their nominal mission orbits.

The OSSE-F3C-13 is the same OSSE shown in blue line in Figure 2, and thus, the same issues discussed earlier for Experiment A are present in Figure 6. At 02:00 UT and subsequent analysis times, the RMSD of the OSSE-F3C-13 increases in the analysis step and gradually decreases over the course of the one-hour forecast step. A comparison of Experiment A to Experiment D suggests some issues with the analysis step that cannot be resolved by increasing the ensemble size but can be corrected by increasing the data volume, indicating current limitations of the GSI Ionosphere for a small data volume case. In the forecast step of the OSSE-F3C-09 and the OSSE-F7C2, the RMSD decreases first before increasing, which indicates a forecast model error reduction. This phenomenon is termed the U-shape RMSD and examined in Hsu et al. (2018). Causes of these RMSD behaviors will be reviewed in section 5.

Table 2
The Ratio of the RMSD of OSSEs to the RMSD of Control Ensemble Forecast Experiment in Different Geomagnetic Latitude Regions

	OSSE-F3C-09	OSSE-F3C-13	OSSE-F7C2
$30^\circ \geq$ and $\geq -30^\circ$	0.39	0.51	0.33
$30^\circ <$ or $< -30^\circ$	0.52	0.63	0.74

Note. OSSEs, Observing System Simulation Experiments; RMSD, root-mean-square difference.

5. Discussion

The RMSD of the OSSE-F3C-13 shown in Figures 2 and 6 appears to increase in the analysis step and subsequently decrease during the

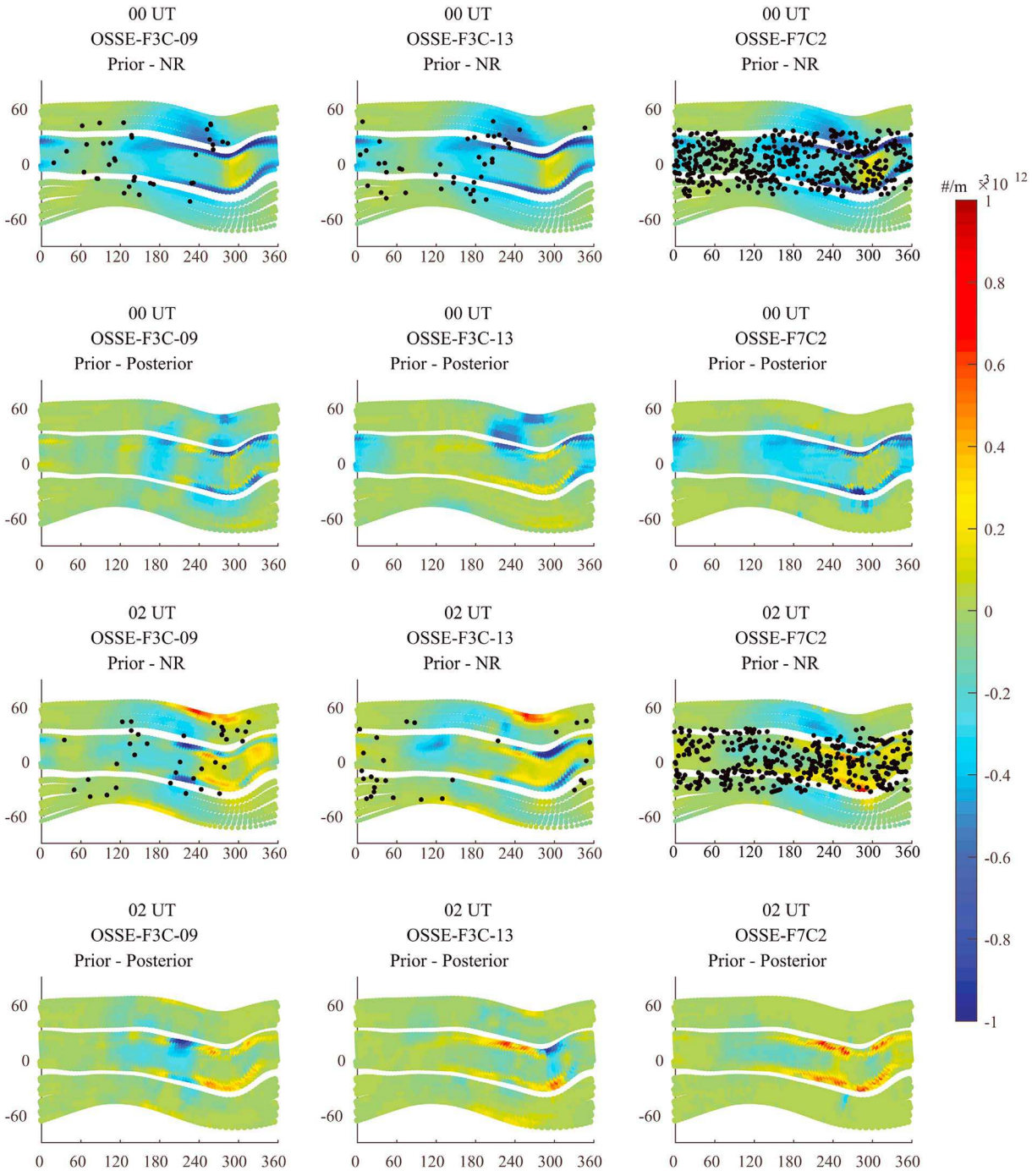

Figure 8. Differences of the $[O^+]$ at 330-km altitude at 00:00 and 02:00 UT. The first and third rows are the difference between prior ensemble mean of Observing System Simulation Experiments (OSSEs) and the Nature Run at 00:00 and 02:00 UT, respectively. The second and fourth rows are the difference between prior and posterior ensemble mean at 00:00 and 02:00 UT, respectively. The black dots denote the tangent points of RO data that are assimilated at 330-km altitude. The *x* axis and *y* axis represent geographic longitude and latitude in degrees.

forecast period. This issue is solved by increasing the data volume as suggested by Experiment D. To examine the filter behaviors in more detail, the differences of the $[O^+]$ at 330-km altitude between the prior ensemble mean and the NR and between the prior and posterior ensemble means are shown at 00:00 UT and 02:00 UT in Figure 8, for the OSSE-F3C-09, OSSE-F3C-13, and OSSE-F7C2, respectively. Note that 02:00 UT is when the RMSDs start increasing in the analysis step. While warm colors indicate that the oxygen ion density of

NR/posterior ensemble mean is higher than that of prior ensemble mean, cold colors indicate that the oxygen ion density of NR/posterior ensemble mean is lower than that of prior ensemble mean. For a successful data assimilation update, the differences of the prior ensemble mean and the NR should be consistent with the difference of the prior and posterior ensemble mean, because the EnSRF brings the posterior ensemble mean closer to the NR compared with the prior ensemble mean.

Figure 8 shows that distributions of RO TEC assimilated into the model at 00:00 and 02:00 UT are considerably uneven for the OSSE-F3C-09 and OSSE-F3C-13. The NR is biased to be higher than the prior model ensemble mean at 00:00 UT, so the EnSRF for the OSSE-F3C-13 is still able to center the posterior model ensemble around the NR after the assimilation update even with such a low data volume. The EnSRFs adjust the prior model ensemble in the direction that lead to a greater agreement with the NR at 00:00 UT for all experiments. However, at 02:00 UT, particularly in the OSSE-F3C-13, the $[O^+]$ between 180° and 320° longitudes in the low-latitude region where there is no data are adjusted in the opposite direction to the prior bias. It is difficult to extract sufficient information from observation for a detail correction particularly when data are unevenly distributed, so the EnSRF of OSSE-F3C-13 fails to adjust the $[O^+]$ in a correct direction locally. The data assimilation update becomes ineffective at 02:00 UT, leading to the increase in the RMSD as shown in Figure 6. The same issue is also seen from 10:00 UT to 22:00 UT in Experiment C, which has caused the magnitude of vTEC at the peak of EIA to be overcorrected in Figure 4. The OSSE-F7C2 of Experiment D suggests that a larger data amount with an even data distribution is likely to rectify this issue, and therefore, Experiment C result is expected to improve with the FORMOSAT-7/COSMIC-2 that can provide RO data at about 4 times larger volume.

During the forecast steps, the RMSD of the OSSE-F3C-09 and the OSSE-F7C2 shown in Figure 6 exhibit the U-shape RMSD, typically consisting of a 30-min error reduction followed by an error growth. In contrast, the RMSD of the OSSE-F3C-13 monotonically decreases over the entire duration of forecast periods. To better understand this behavior of RMSD of the OSSE-F3C-13, two additional experiments are carried out. After the same OSSE-F3C-13 and OSSE-F7C2 as previously done from 00:00 UT to 02:00 UT, the ensemble forecasts are launched at 02:00 UT and run through 12:00 UT. Figure 9 shows the RMSDs from these two additional forecasting experiments. The U-shape RMSD appears after the final analysis step at 02:00 UT for both experiments. The RMSD of experiment associated with OSSE-F3C-13 takes about 1 hr to reach to its minimum value and more than 2 hr to grow back to the level of RMSD at 02:00 UT. The RMSD of experiment associated with OSSE-F7C2 takes about half hour to reach to its minimum value and about 1 hr to grow back to the level of RMSD at 02:00 UT. Differences of the ensemble mean from the NR at 330-km altitude are shown for the prior and posterior ensemble at 02:00 UT and for the forecast ensemble from 02:00 to 04:00 UT with a 12-min interval in Figure 9. Warm colors represent that the ensemble mean is larger than NR, which indicates a positive bias, and cold colors represent that the ensemble mean is smaller than NR, which indicates a negative bias. From 02:00 UT to 04:00 UT, negative biases disappear and positive biases start appearing in the low latitudes, which is similar to the result presented by Hsu et al. (2018).

These additional experiments show that the U-shape RMSD appears in the forecast steps of OSSE-F3C-09, OSSE-F3C-13, and OSSE-F7C2, and the only difference is the timescale of the recovery from biased assimilation correction. Given a similarity of the RMSD behavior to the U-shaped RMSD previously examined in detail in Hsu et al. (2018), the error reduction and subsequent increase are likely caused by the same mechanism. An inadequate analysis update at 02:00 UT, partly due to an insufficient amount of the observation, results in the negative biases. Dynamical and chemical adjustments of the negative biases over the course of forecast periods involve (1) a correction of the original bias that appears as the initial RMSD reduction and (2) a formation of a new positive bias that leads to the RMSD growth. Hsu et al. (2018) have shown that the EnSRF adjustment of the prior model ensemble places the *F* region peak density height in the higher atmospheric region, where molecular species are less abundant or the atomic oxygen ion recombination rate is smaller, which results in plasma densities in the ensemble forecast simulation being larger than that in the NR in this region. This chemical adjustment process is also considered to amount to the formation of the positive bias shown here.

A further comparison of the RMSD from these two additional experiments from 02:00 to 12:00 UT suggests that the overall timescale of dynamical and chemical adjustment of the unbalanced ionospheric state back to the steady state is roughly the same between these two experiments in spite of differences in details of

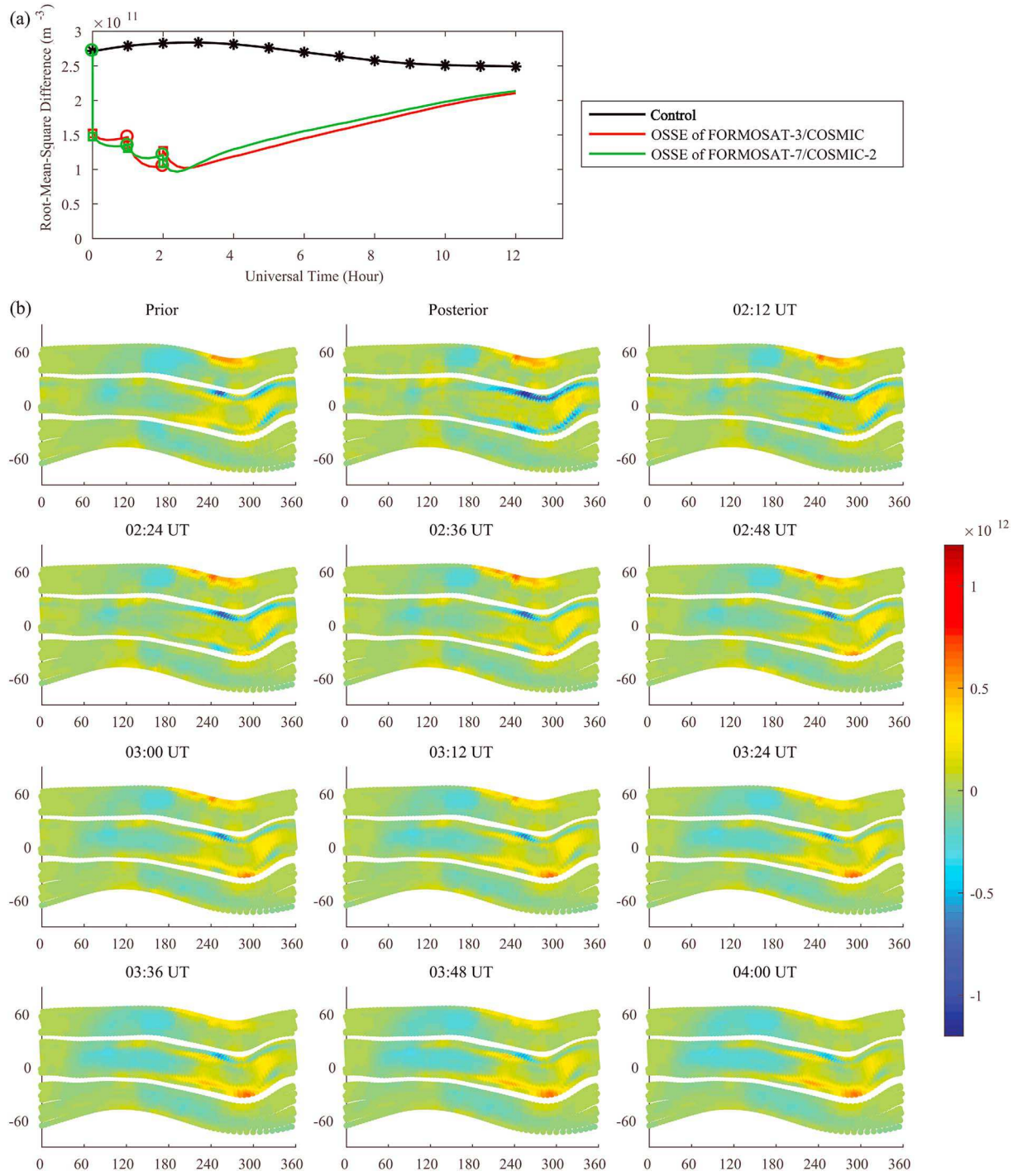

Figure 9. (a) Root-mean-square difference of [O^+] between the ensemble mean of Observing System Simulation Experiments (OSSEs) and Nature Run, computed over the middle- and low-geomagnetic latitude region from 200- to 500-km altitude, from 00:00 to 12:00 UT. Synthetic Radio Occultation total electron content data of FORMOSAT-3/COSMIC in 2013 (green line) and FORMOSAT-7/COSMIC-2 (red line) are assimilated into OSSEs from 00:00 UT to 02:00 UT with 70-member EnSRF. The 70-member control ensemble forecast experiment (black) are shown. (b) Differences of the [O^+] between the ensemble mean of OSSE of the FORMOSAT-3/COSMIC and NR at 330-km altitude from 02:00 to 04:00 UT with a 12-min interval. Prior and posterior difference at 02:00 UT are also shown. The *x* axis and *y* axis represent geographic longitude and latitude in degrees.

the RMSD behavior during the first hour of the forecast period. At 12:00 UT, the RMSD for the OSSE-F3C-13 and the OSSE-F7C2 are about 0.86 and 0.84 of the RMSD of the 70-member control ensemble forecast experiment.

6. Summary and Conclusions

This study evaluates the observation impacts of two RO missions, the FORMOSAT-3/COSMIC, and the upcoming FORMOSAT-7/COSMIC-2, on low-latitude and midlatitude ionospheric specifications and short-term forecasting, by using the GSI Ionosphere data assimilation system. By comparing the data assimilation experiment and OSSE results, the objective of this study is to evaluate the potential of the FORMOSAT-7COSMIC-2 to improve ionospheric monitoring after RO TEC in the near future. The actual low-latitude and midlatitude RO TEC data that are provided by the FORMOSAT-3/COSMIC mission are assimilated into the GIP/TIE-GCM to examine the capability of the GSI Ionosphere to reproduce realistic ionospheric features. Global vTEC maps produced by the GSI Ionosphere, here referred to as the GSI GIMs, are validated against the CODE GIMs. Cross comparison of OSSEs with synthetic data generated according to the actual 2009 and 2013 FORMOSAT-3/COSMIC and planned FORMOSAT-7/COSMIC-2 observing system parameters has identified considerable differences in the observations impact of these observing systems on the low-latitude and midlatitude ionospheric specification. These OSSEs are used to calibrate the ensemble size and covariance localization parameters in the GSI-EnSRF in order to optimize the performance of data assimilation.

The main findings are as follows:

1. Generally, the assimilation of FORMOSAT-3/COSMIC and FORMOSAT-7/COSMIC-2 RO TEC data into the GIP/TIE-GCM using the GSI Ionosphere can improve the low-latitude and midlatitude ionospheric specification, during low solar activity and geomagnetic quiet conditions.
2. Comparative OSSEs for the FORMOSAT-3/COSMIC on 1 January 2009, 1 January 2013, and the FORMOSAT-7/COSMIC-2, show that FORMOSAT-7/COSMIC-2 has a great potential to improve the low-latitude ionospheric specification. The FORMOSAT-7/COSMIC-2 OSSE has a smaller RMSD within low geomagnetic latitude regions (33% of the RMSD for the control ensemble simulation without data assimilation) and a larger error in the high-geomagnetic latitude regions (74%) compared to FORMOSAT-3/COSMIC OSSEs. High-latitude data provided by other observing systems are necessary for improving global ionospheric monitoring.
3. The GSI Ionosphere is capable of assimilating real TEC data from the RO mission and reproducing global low-latitude and midlatitude ionospheric weather features even with a relatively low RO data volume provided by the FORMOSAT-3/COSMIC during 2013. Compared to the control GIM, the GSI GIM is 50% closer to the CODE GIM because of the assimilation of FORMOSAT-3/COSMIC data using the GSI Ionosphere.
4. OSSEs for the FORMOSAT-3/COSMIC on 1 January 2013 show that the GSI Ionosphere can effectively correct the overall distribution of $[O^+]$ but has issues with local adjustments especially in areas of low observation densities. This issue is identified as resulting from inadequate adjustments of the model ensemble in the EnSRF, which can be rectified by increasing the data volume rather than by an extensive calibration of the EnSRF parameters, such as the ensemble size and covariance localization scale. This underscores the limitation of the current implementation of the EnSRF in the GSI Ionosphere.

In conclusion, the above findings suggest that the GSI Ionosphere is able to track realistic features of the midlatitude and low-latitude ionosphere by assimilating actual RO TEC data into a coupled model of the plasmasphere, ionosphere, and thermosphere. It is expected to perform well with a large data volume provided by the upcoming FORMOSAT-7/COSMIC-2 mission to improve the GSI Ionosphere's capability as a low-latitude and midlatitude numerical ionospheric weather specification and forecasting system.

Acknowledgments

This study is supported by following grants: NASA award NNX14AI17G through Heliophysics Grand Challenges Research Program, NSF award AGS-1651469 through CEDAR program, and Taiwan Ministry of Science and Technology grant MOST 107-2119-M-008-018. The authors would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. This work is presented at the 2018 International Team meeting for Ionospheric Space Weather Studied by RO and Ground-based GPS TEC Observations, leaded by Jann-Yenq Liu, supported by International Space Science Institute, Bern, Switzerland. The authors sincerely thank NCAR COSMIC office and Dr. Xinan Yue for their great help with FORMOSAT-7/COMSIC-2 RO data. The authors also thank to Chi-Yen Lin for his valuable comments and suggestions. The main data assimilation result presented in this paper is publicly available from doi:10.17605/OSF.IO/AE7SP.

References

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, *127*(12), 2741–2758. https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. *American Meteorological Society*, *90*(9), 1283–1296. https://doi.org/10.1175/2009bams2618.1

Anthes, R. A. (2011). Exploring Earth's atmosphere with radio occultation: Contributions to weather, climate and space weather. *Atmospheric Measurement Techniques*, *4*(6), 1077–1103. https://doi.org/10.5194/amt-4-1077-2011

Anthes, R. A., Bernhardt, P. A., Chen, Y., Cucurull, L., Dymond, K. F., Ector, D., et al. (2008). The COSMIC/FORMOSAT-3 mission: Early results. *American Meteorological Society*, *89*(3), 313–334. https://doi.org/10.1175/BAMS-89-3-313

Chartier, A. T., Matsuo, T., Anderson, J. L., Collins, N., Hoar, T. J., Lu, G., et al. (2016). Ionospheric data assimilation and forecasting during storms. *Journal of Geophysical Research: Space Physics*, *121*, 764–778. https://doi.org/10.1002/2014JA020799

Chen, C. H., Lin, C. H., Matsuo, T., Chen, W. H., Lee, I. T., Liu, J. Y., et al. (2016). Ionospheric data assimilation with thermosphere-ionosphere-electrodynamics general circulation model and GPS-TEC during geomagnetic storm conditions. *Journal of Geophysical Research: Space Physics*, *121*, 5708–5722. https://doi.org/10.1002/2015JA021787

Evensen, G. (1994). Sequential data dssimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, *99*(C5), 10,143–10,162. https://doi.org/10.1029/94JC00572

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *125*(554), 723–757. https://doi.org/10.1002/qj.49712555417

Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, *129*(11), 2776–2790. https://doi.org/10.1175/1520-0493(2001)129%3C2776:DDFOBE%3E2.0.CO;2

Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *129*(1), 123–137. https://doi.org/10.1175/1520-0493(2001)129%3C0123:ASEKFF%3E2.0.CO;2

Hsu, C.-T., Matsuo, T., Wang, W. B., & Liu, J. Y. (2014). Effects of inferring unobserved thermospheric and ionospheric state variables by using an ensemble Kalman filter on global ionospheric specification and forecasting. *Journal of Geophysical Research: Space Physics*, *119*, 9256–9267. https://doi.org/10.1002/2014JA020390

Hsu, C.-T., Matsuo, T., Yue, X., Fang, T.-W., Fuller-Rowell, T., Ide, K., & Liu, J.-Y. (2018). Assessment of the impact of FORMOSAT-7/COSMIC-2 GNSS RO observations on midlatitude andlow-latitude ionosphere specification: Observing system simulationexperiments using Ensemble Square Root Filter. *Journal of Geophysical Research: Space Physics*, *123*, 2296–2314. https://doi.org/10.1002/2017JA025109

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Komjathy, A., Wilson, B., Pi, X., Akopian, V., Dumett, M., Iijima, B., et al. (2010). JPL/USC GAIM: On the impact of using COSMIC and ground-based GPS measurements to estimate ionospheric parameters. *Journal of Geophysical Research*, *115*, A02307. https://doi.org/10.1029/2009JA014420

Lee, I. T., Matsuo, T., Richmond, A. D., Liu, J. Y., Wang, W., Lin, C. H., et al. (2012). Assimilation of FORMOSAT-3/COSMIC electron density profiles into a coupled thermosphere/ionosphere model using ensemble Kalman filtering. *Journal of Geophysical Research*, *117*, A10318. https://doi.org/10.1029/2012JA017700

Lin, C. Y., Matsuo, T., Liu, J. Y., Lin, C. H., Huba, J. D., Tsai, H. F., & Chen, C. Y. (2017). Data assimilation of ground-based GPS and radio occultation totalelectron content for global ionosphericspecification. *Journal of Geophysical Research: Space Physics*, *122*, 10,876–10,886. https://doi.org/10.1002/2017JA024185

Lomidze, L., Scherliess, L., & Schunk, R. W. (2016). Modeling and analysis of ionospheric evening anomalies with a physics-based data assimilation model. *Journal of Atmospheric and Solar - Terrestrial Physics*, *140*, 65–78. https://doi.org/10.1016/j.jastp.2016.02.009

Matsuo, T., & Araujo-Pradere, E. A. (2011). Role of thermosphere-ionosphere coupling in a global ionospheric specification. *Radio Science*, *46*, RS0D23. https://doi.org/10.1029/2010RS004576

Matsuo, T., Lee, I. T., & Anderson, J. L. (2013). Thermospheric mass density specification using an ensemble Kalman filter. *Journal of Geophysical Research: Space Physics*, *118*, 1339–1350. https://doi.org/10.1002/jgra.50162

Morton, J. C., & Nielsen, A. A. (2008). Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sensing of Environment*, *112*(3), 1025–1036. https://doi.org/10.1016/j.rse.2007.07.013

Nielsen, A. A. (2007). The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing*, *16*(2), 463–478. https://doi.org/10.1109/TIP.2006.888195

Nielsen, A. A., & Conradsen, K. (1997). Multivariate alteration dection (MAD) in multispectral, bi-temporal image data: A new approach to change detction studies. Informatics and Mathematical Modelling, Technical University of Denmark, DTU..

Nielsen, A. A., Conradsen, K., & Simpson, J. J. (1998). Multivariate Alteration Detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, *64*(1), 1–19. https://doi.org/10.1016/S0034-4257(97)00162-4

Pedatella, N. M., Forbes, J. M., Maute, A., Richmond, A. D., Fang, T. W., Larson, K. M., & Millward, G. (2011). Longitudinal variations in the F region ionosphere and the topside ionosphere-plasmasphere: Observations and model simulations. *Journal of Geophysical Research*, *116*, A12309. https://doi.org/10.1029/2011JA016600

Richmond, A. D., Ridley, E. C., & Roble, R. G. (1992). A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophysical Research Letters*, *19*(6), 601–604. https://doi.org/10.1029/92GL00401

Schaer, S. (1999). Mappeing and perdicting the Earth's ionosphere using the Global Positioning System, PhD dissertation, Astron. Inst., University of Bern, Bern, Switzerland.

Scherliess, L., Thompson, D. C., & Schunk, R. W. (2009). Ionospheric dynamics and drivers obtained from a physics-based data assimilation model. *Radio Science*, *44*, RS0A32. https://doi.org/10.1029/2008RS004068

Schreiner, W., & COSMIC Team (2014). COSMIC Data Analysis and Archive Center (CDAAC)—Current status and future plans. Paper presented at *Eighth FormoSat-3/COSMIC Data Users' Workshop*, *National Science Foundation*, National Space Organization and Taiwan, National Science Council Taiwan, Boulder, CO.

Schunk, R. W., Scherliess, L., Eccles, V., Gardner, L. C., Sojka, J. J., Zhu, L., et al. (2016). Space weather forecasting with a Multimodel Eensemble Prediction System (MEPS). *Radio Science*, *51*, 1157–1165. https://doi.org/10.1002/2015RS005888

Whitaker, J. S., & Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, *130*(7), 1913–1924. https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2

Yue, X., Schreiner, W. S., & Kuo, Y.-H. (2012). A feasibility study of the radio occultation electron density retrieval aided by a global ionospheric data assimilation model. *Journal of Geophysical Research*, *117*, A08301. https://doi.org/10.1029/2011JA017446

Yue, X., Schreiner, W. S., Kuo, Y. H., Hunt, D. C., Wang, W. B., Solomon, S. C., et al. (2012). Global 3-D ionospheric electron density reanalysis based on multisource data assimilation. *Journal of Geophysical Research*, *117*, A09325. https://doi.org/10.1029/2012JA017968

Yue, X., Schreiner, W. S., Pedatella, N., Anthes, R. A., Mannucci, A. J., Straus, P. R., & Liu, J.-Y. (2014). Space weather observations by GNSS radio occultation: From FORMOSAT-3/COSMIC to FORMOSAT-7/COSMIC-2. *Space Weather*, *12*, 616–621. https://doi.org/10.1002/2014SW001133