Review comments on manuscript"Global TEC prediction performance assessment of IRI-2016 model based on EOF decomposition" by Li et al., 2019; submitted to Annales Geophysicae

The manuscript compares total electron content from Global Ionospheric Maps products and International Reference Ionosphere during 2013. Empirical Orthogonal Functions are employed to detail the differences between the two datasets.

Seasonal average analysis was performed which showed that the IRI model reproduces the equatorial ionisation anomaly distinctively while GIM TEC does show enhancement of TEC over equatorial/low latitude regions, but does not necessarily show the different bands of enhancement at the EIA crests. While related studies exists, I think this work is relevant especially if it clearly shows by how much the IRI under-predicts GIM TEC (in terms of TECu) in different latitude regions. However this is not clearly shown in the current paper.

Answer: According to the reviewer's suggestion, we have added some analysis and discussion about the discrepancies between GIM-TEC and IRI-TEC at different latitudes as follows in the revised manuscript:

Considering the different levels of ionospheric activities at different latitudes, mean and RMS values of the discrepancies between seasonal averages of GIM-TEC and IRI-TEC over different latitudinal regions in 2013 were calculated. Results are shown in Figure 2. From Figure 2, the mean and RMS values over the area near the equator generally exhibit peak values. GIM-TEC values over the equator and low latitudes are much larger than IRI-TEC values, especially over the ionospheric trough near the magnetic equator shown in Figure 1. The mean and RMS values over Southern Hemisphere during the December solstice are significantly large, and also very large over Northern Hemisphere during the June solstice. Therefore, there are large discrepancies between GIM-TEC and IRI-TEC over the summer Hemisphere.



Figure 2. Mean and RMS values of the discrepancies between GIM-TEC and IRI-TEC at different latitudes during four seasons.

Additionally, as the authors know, the IRI model provides TEC up to an altitude of 2000 km while GIM TEC products are based on GNSS observations (at about 20000

km). Assuming that the IRI model was 'accurate' at its specified height, it would be missing some plasmaspheric contribution. The authors have missed to point out this important aspect early in the paper. I believe it is related to line 10, page 6, and Figure 1. Information about this is later presented on page 15, line 10.

Answer: According to the reviewer's instruction, We advanced the relevant paragraph as follows on page 15 to page 6 in the revised manuscript:

"The IRI-2016 model provides ionospheric parameters of up to 2000 km and will inaccurately predict the TEC up to GNSS satellites located at an altitude of approximately 20,000 km. The IRI-TEC may be smaller than GIM-TEC because of the missing plasmaspheric content."

On page 15, we changed the statement as follows :

"Although the IRI-TEC will be smaller than the GIM-TEC because of the missing

plasmaspheric content,  $A_{11}$  of IRI-TEC in Figure 10(b) shows a quite large

underestimation compared with that of GIM-TEC."

Below are comments which may assist in improving the paper.

Page 3, line 35: I thought that the GIM TEC products are provided at time resolution of 2 hours. Please cross-check that they are also available for 15 minutes.

Answer: According to the reviewer's suggestion, we have checked the temporal resolution of IGS GIMs. In terms of temporal resolution, the GIM generated by each IAAC and IGS is different. Final GIMs produced by CODE, ESA, JPL and UPC are provided with a 2h temporal resolution, whereas the CODE-procuded IONEX maps are in 1-hour temporal resolution. The temporal resolution of CASG GIMs are 0.5-hour.

In order to describe it more accurately, we changed the expression as follows in the revised manuscript:

"The GIM TEC used in this study is the official IGS combined final product provided by the Crustal Dynamic Data Information System (ftp://cddis.gsfc.nasa.gov). Final GIMs are regular products of the International GNSS Service (IGS) since 1998. These GIMs are provided in the ionosphere exchange format with a spatial resolution of  $2.5^{\circ}\times5^{\circ}$  in geographic latitude and longitude and a temporal resolution of 2 h."

Page 4, line 5, please include original references for the hmF2 model options included within the IRI 2016 model. One is based on COSMIC observations (Shubin) and the other one on ionosonde measurements and spheric harmonic method (Altadil).

Answer: According to the reviewer's instruction, we have added original references for the hmF2 model options included within the IRI 2016 model in the revised manuscript as follows:

"The recent version of this model is IRI-2016 (Bilitza et al., 2016; Bilitza et al., 2017). After IRI-2012, IRI-2016 exhibits the latest improvement in the model by introducing two new F2 peak height hmF2 modeling options with their data sources from ionosonde measurements (Altadill et al., 2013) and COSMIC radio occultations (Shubin, 2015)."

Altadill, D., Magdaleno, S., Torta, J. M., Blanch, E.: Global empirical models of the density peak height and of the equivalent scale height for quiet conditions, Adv. Space

Res., 52, 1756–1769, https://doi.org/10.1016/j.asr.2012.11.018, 2013.

Shubin, V. N.: Global median model of the F2-layer peak height based on ionospheric radio-occultation and ground-based Digisonde observations, Adv. Space Res., 56, 916–928, https://doi.org/10.1016/j.asr.2015.05.029, 2015.

Page 4, line 10: In the statement "The global TEC date calculated ...". The word 'date' should be data.

Answer: Yes, it is a mistake. we changed "date" to "**data**" in revised version. Thank you.

Page 5, line 5 is not clear. In the text "If the IRI TEC and GIM TEC are decomposed, then their EOF base functions and coefficients will exhibit poor comparability". Why would this be the case? And do you mean that this would be so, if they were decomposed separately? Assuming that they exhibit some similarities/differences, wouldn't such decomposition bring them out? May be not in magnitude of coefficients or base functions; but perhaps in the trend and identification of physical features?

Answer: This sentence is indeed unclear. As you understand, what we want to express is that if they are decomposed separately, it will be difficult to compare in magnitude. We changed the sentence to "If the IRI TEC and GIM TEC are decomposed separately, it is difficult to directly compare their EOF base functions and coefficients in magnitude." in revised version.

The spatial patterns and temporal variations of the global TEC data are separated by EOF decomposition and can be properly represented by the base functions and associated coefficients, respectively. For GIM-TEC data  $X_{GIM}$ , coefficients  $A_{k, GIM}$ 

and EOF base functions  $E_{k,GM}$  will be obtained by using EOF decomposition method.

For IRI-TEC data  $X_{IRI}$ , the coefficients  $A_{k,IRI}$  and EOF base functions  $E_{k,IRI}$  can be obtained:

$$X_{GIM} = \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k,GIM}$$
$$X_{IRI} = \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k,IRI}$$

EOF decomposition will extract main spatial patterns. The six main base functions  $E_k$  extracted by performing EOF decomposition on the global TEC from related

reference (Talaat and Zhu, 2016) and our study both include the following: the variation with the geomagnetic latitude reflecting the daily averaged solar forcing, the diurnal and semidiurnal periodic changes with longitude due to local time, and the interhemispheric asymmetry caused by the annual variation of the inclination angle of the Earth's orbit.

The spatiotemporal features extracted from IRI-TEC and GIM-TEC data have good consistencies, they are shown in Figs (2) and (3) of this document. Therefore, if GIM-TEC and IRI-TEC are decomposed separately, the results will exhibit obvious

similarities in trend and identification of physical features. However, it is not possible to make direct comparisons in magnitude, because  $A_{k, GIM}$  and  $A_{k,IRI}$  are different,

 $E_{k,GIM}$  and  $E_{k,IRI}$  are also different.

So we combined the data to form a whole data set for EOF decomposition and compared the two data sets.

$$\begin{bmatrix} X_{GIM} \\ X_{IRI} \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \end{bmatrix} \cdot A_k$$

Then, the GIM-TEC and IRI-TEC can be written and reconstruct as follows.

$$X_{GIM} = \sum_{k=1}^{N} E_{k,GIM} \cdot A_k$$
$$X_{IRI} = \sum_{k=1}^{N} E_{k,IRI} \cdot A_k$$

The same coefficients of the EOF base function  $A_k$  can be obtained, then  $E_{k,GIM}$  and

 $E_{k,RI}$  were compared to analyze the difference between GIM-TEC and IRI-TEC. We

think the conclusions obtained by the method of this paper are clearer.

Following on the previous comment, do you mean that IRI TEC and GIM TEC are combined to form one data file which is later used for decomposition?

Answer: Yes, IRI TEC and GIM TEC are combined to form one data file which is later used for decomposition.

The two sets of data are arranged in rows or columns as needed.

In our study, We analyzed the global TEC over a 1 year time period (2013) with a 2 h temporal resolution and  $37 \times 36 = 1332$  spatial grids, the total epoch number is

12×365=4380. Before EOF analysis, GIM TEC data  $X_{GIM}$  should be arranged as

follows:



Coefficients  $A_{k, GIM}$  and EOF base functions  $E_{k,GIM}$  will be obtained by using EOF

decomposition method:

$$X_{GIM} = \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k,GIM}$$

The same coefficients of the EOF base function, that is, the same time-varying features, can be obtained by arranging IRI-TEC and GIM-TEC according to the same number of columns. That is:



Then, we will get EOF decomposition result:

$$\begin{bmatrix} X_{GIM} \\ X_{IRI} \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \end{bmatrix} \cdot A_{k}$$

Then, the GIM-TEC and IRI-TEC can be written as follows.

$$X_{GIM} = \sum_{k=1}^{N} E_{k,GIM} \cdot A_k$$
$$X_{IRI} = \sum_{k=1}^{N} E_{k,IRI} \cdot A_k$$

It seems like that we extract common temporal variation factors  $A_k$ , and we can

therefore directly compare the spatial characteristics  $E_{k,GIM}$  and  $E_{k,IRI}$ .

Page 7, Table 1, indicate the units of some parameters; maximum, minimum and mean bias; e.g mean bias (TECU).

Answer: According to the reviewer's suggestion, we added the units of Maximum bias, Minimum bias and Mean bias in Table 1 in the revised manuscript.

Page 6, just after line 15: Bias values are computed using IRI TEC and GIM TEC? It is not clear how daily RMS values in 2013 displayed in Figure 2 are computed. Are they just average of the bias values calculated using IRI TEC and GIM TEC?

Answer: The sentence about how to calculate RMS in line 16, Page 6 is not clear. We changed it as follows in the revised manuscript:

"The gridded values of the global IRI-TEC and GIM-TEC at different UTs for each day of the year 2013 were used to calculate the daily RMS."

The expression of equation (10) in our manuscript is also not clear, so we changed

"
$$RMS = \left[\sum_{i}^{n} (Y_i - Y_i)^2 / n\right]^{1/2}$$
" to "RMS =  $\sqrt{\frac{1}{n} \sum_{i}^{n} (Y_i - Y_i)^2}$ " in the revised manuscript.

During a day, the global TEC data has 1322 grid points and 12 epochs. Therefore, both the GIM-TEC data and IRI-TEC data for one day have 1332\*12=15984 values. Daily RMS value in 2013 displayed in Figure 2 is computed by using equation (10):

$$RMS = \sqrt{\frac{1}{n} \sum_{i}^{n} (Y_i - Y_i')^2}$$

Where, n=15984,  $Y_i$  and  $Y'_i$  are data for GIM-TEC and IRI-TEC respectively.

Page 5, equation 10: Shouldn't RMS be RMSE? This seems to be what is plotted in Figure 2(a). RMSE values of IRI 2016, how are they computed?

Answer: Yes, here is a mistake. RMS and RMSE should be unified. We examined the entire manuscript and used "RMS" in all equations and figures in the revised manuscript.

Under subsection 3.2: the authors state "We combined the IRI TEC and GIM TEC data ...". If these datasets are combined, how do you obtain Figure 4? Answer: We combined the data as follows:



٦.

After performing EOF decomposition, we will get:

$$\begin{bmatrix} X_{GIM} \\ X_{IRI} \\ 2664 \times 4380 \end{bmatrix} = \sum_{k=1}^{N} \underbrace{E_{k} \cdot A_{k}}_{2664 \times 1} = \sum_{k=1}^{N} \begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \\ 2664 \times 1 \end{bmatrix} \cdot \underbrace{A_{k}}_{4380 \times 1} = \begin{bmatrix} \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k} \\ \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k} \\ \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k} \\ 1332 \times 4380 \end{bmatrix}$$
$$X_{IRI} = \sum_{k=1}^{N} \underbrace{E_{k,IRI} \cdot A_{k}}_{1332 \times 1} \underbrace{A_{380 \times 1}}_{4380 \times 1}$$

Γw

That is to say, the two sets of data are arranged together, and after the common coefficients are extracted, the base functions  $E_{k}_{2664\times 1}$  are separated to  $\begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \end{bmatrix}$ . So we can get two sets of base functions:  $E_{k}$  and  $E_{k}$  which are shown in Figure 4.

can get two sets of base functions:  $E_{k,GIM}$  and  $E_{k,RI}$ , which are shown in Figure 4. <sup>1332×1</sup> <sup>1332×1</sup>

In order to make the expression clearer, we revised equations (6) and (7) in page 5:

$$\binom{X_{GIM}}{X_{IRI}} = \sum_{k=1}^{N} \begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \end{bmatrix} \cdot A_{k} = \begin{bmatrix} \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k} \\ \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k} \end{bmatrix}$$
(6)

$$[X_{GIM} \quad X_{IRI}] = \sum_{k=1}^{N} E_k \cdot \left[ A_{k,GIM} \quad A_{k,IRI} \right] = \left[ \sum_{k=1}^{N} E_k \cdot A_{k,GIM} \quad \sum_{k=1}^{N} E_k \cdot A_{k,IRI} \right]$$
(7)

In Figure 3, is global data for 2013 used? How do you account for latitudinal

differences? Does this figure reflect only seasonal changes as indicated in the last statement on page 7?

Answer: Yes, global data for 2013 is used in Figure 3.

The EOF decomposition was conducted on GIM-TEC and IRI-TEC as follow,

$$\begin{bmatrix} X_{GIM} \\ X_{IRI} \\ 2664 \times 4380 \end{bmatrix} = \sum_{k=1}^{N} \begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \\ 2664 \times 1 \end{bmatrix} \cdot A_{k} = \begin{bmatrix} \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k} \\ 1330 \times 1 \\ \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k} \\ \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k} \\ 1330 \times 1 \\ 330 \times 1 \end{bmatrix}$$

The spatial patterns and temporal variations of the TEC are separated by EOF decomposition and can be properly represented by the base functions and associated

coefficients, respectively.  $\begin{bmatrix} E_{k,GIM} \\ E_{k,IRI} \\ _{2664\times 1} \end{bmatrix}$  represent the TEC's spatial distribution modes,

and they are base functions. Six main base functions are shown in Figure 4. And  $A_k$ 

represents the magnitude of the influence of the *k* th base function component at different epoch. Six coefficients of main base function  $A_k$  are shown in Figure 3.

Therefore, Figure 3 reflects only seasonal changes, while Figure 4 represents the spatial distribution characteristics.

Equation 7 and Figure 7: I am not sure of the physical significance and justification of combining IRI TEC and GIM TEC. Afterall, they have different inherent errors. What can be derived from this combination taken at same grid points can as well be determined from one dataset either GIM TEC or IRI TEC. Otherwise combining these datasets removes the differences/similarities that the authors would want to study? Provide a scientific justification for combining both datasets and what additional features or interpretations are obtained. I don't think that the text in line 15, page 15 is sufficient to justify this inclusion. This has already been discussed.

Answer: After performing EOF decomposition on GIM-TEC and IRI-TEC by using

equation (7), we will get base functions  $E_k$  and coefficients  $A_k$ .

$$\begin{bmatrix} X_{GIM} & X_{IRI} \end{bmatrix} = \sum_{k=1}^{N} E_k \cdot \begin{bmatrix} A_{k,GIM} & A_{k,IRI} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{N} A_{k,GIM} \cdot E_k & \sum_{k=1}^{N} A_{k,IRI} \cdot E_k \end{bmatrix}$$
(7)

However, the original TEC data can be reconstructed by using  $E_k$  and  $A_k$  as follow:

$$\begin{split} X_{GIM} = & \sum_{k=1}^{N} A_{k,GIM} \cdot E_k \\ X_{IRI} = & \sum_{k=1}^{N} A_{k,IRI} \cdot E_k \end{split}$$

In other words, the EOF decomposition process of equation (7) is reversible. Therefore, decomposition after combining the two sets of data does not lead to errors.

We showed the six main base functions  $E_k$  extracted from combined data of IRI

TEC and GIM TEC by using Equation (7) in Figure (1). And we also performed EOF decomposition on GIM-TEC and IRI-TEC separately by using Eqs (a) and (b), the base function  $E_{k,GIM}$  and  $E_{k,IRI}$  are shown in Figs (2) and (3).

$$X_{GIM} = \sum_{k=1}^{N} E_{k,GIM} \cdot A_{k,GIM}$$
(a)  
$$X_{IRI} = \sum_{k=1}^{N} E_{k,IRI} \cdot A_{k,IRI}$$
(b)

Although  $E_k$  in Figure (1) extracted from combined data is not as same as  $E_{k,GM}$  or

 $E_{k,IRI}$  in Figs (2) and (3), they do reflect consistent spatial distribution characteristics of global TEC.

Only if common base functions  $E_k$  of Equation (7) are used, we can compare  $A_{k,GIM}$ and  $A_{k,IRI}$  directly. The results will show the difference of the intensity of each base function between GIM-TEC and IRI-TEC.



Figure (1). Base function  $E_k$  extracted from combined data of GIM-TEC and IRI-TEC



Figure (2). Base function  $E_{k,GIM}$  extracted from GIM-TEC



Figure (3). Base function  $E_{k,IRI}$  extracted from IRI-TEC

Unless I am not understanding equation 7, how do you separately derive A1-A6 for GIM TEC and IRI TEC that you have plotted in Figure 8? Once again, is this necessary? What additional information do we get in Figure 8?

Answer: Maybe the equation (7) is not so clear, we have changed it as follows:

Then, GIM-TEC and IRI-TEC can be written:

$$X_{GIM} = \sum_{k=1}^{N} E_k \cdot A_{k,GIM}$$
$$X_{IRI} = \sum_{k=1}^{N} E_k \cdot A_{k,IRI}$$

Therefore, we can get  $A_{k,GIM}$  for GIM-TEC and  $A_{k,IRI}$  for IRI-TEC, which are shown

in Figure 8. From Figure 8, we can see that the variation of  $A_1$  is strongly correlated with solar activity.  $A_2$  and  $A_3$  have a diurnal variation with UT, and also have a semiannual cycle.  $A_4$  has a distinct annual cycle, and  $A_5$  and  $A_6$  exhibit a semidiurnal cycle and a semiannual cycle. GIM-TEC and IRI-TEC have good consistencies in the above period terms, but the comparison of specific difference in each periodic variation is difficult.

So, we conducted EOF decomposition on  $A_1 - A_6$  according to the equation (13) to divide diurnal variation with UT and seasonal variation characteristics of  $A_{k,GIM}$  and  $A_{k,IRI}$ . The results are shown in Figure 9. Therefore, the article continued to discuss the differences between t  $A_{k,GIM}$  and  $A_{k,IRI}$  based on Figure 9.

We have added some discussion about Figure 8 in the revised manuscript as follows: "The time-varying characteristics of the coefficients in Figure 8 are very consistent with the results shown in Figure 3. From Figs. 8(a) and (b), the variations of  $A_1$  are mainly related to solar activity, and solar activity is the primary determinant of the first base function  $E_1$  in Figure 7(a), which describe the overall average of global

TEC. From Figs. 8(c)–(f), the EOF coefficients  $A_2$  and  $A_3$  of GIM-TEC and IRI-TEC all obviously exhibit a diurnal period and a semiannual period. They reflect the diurnal variation of solar radiation change with longitude due to the LT.  $A_4$  in Figs. 8(g) and (h) indicate a strong annual cycle variation of the interhemispheric asymmetry of the TEC.  $A_5$  and  $A_6$  show a semiannual period of the base functions

 $E_5$  and  $E_6$ , which represent a longitudinal variation that changes with LT. The EOF coefficients of GIM-TEC and IRI-TEC have consistent annual, semiannual, diurnal, and semidiurnal variations. Therefore, Figure 8 manifests that GIM-TEC and IRI-TEC have highly consistent temporal variation characteristics based on the same spatial distribution modes  $E_k$  according to equation (7)."