Dear Editor,

Please find below the review of the manuscript “Modeling Total Electron Content derived from radio occultation measurements by COSMIC satellites over the African Region” submitted to Annales Geophysicae by Patrick Mungufeni et al.

The manuscript presents an empirical model describing ionosphere total electron content over African region. Authors use experimental TEC data obtained using dual frequency GNSS RO receivers onboard of COSMIC satellites to construct the model. They validate the model using same type of data that was used to construct the model but for a different period.

General comments:

General impression is that the present work has no contribution to the current understanding of the low-latitude ionospheric physics/modelling. The work brings a little science and the newly created model could hardly be used in any real-life application. Authors are making too many assumptions and mistakes, sometimes trying to deliberately present performance results better than they are. Moreover, the performance of the model has not been compared to any other well-known model, leaving a room for doubts. Therefore, I recommend the manuscript (in its present form) is **rejected**. At the same time, the work might be improved and worth publication after substantial modifications. Please find below a list of critical issues along with possible improvements/corrections for a potential future re-submission.

Critical/Major comments:

P.1 L.27: Replace “good” with “applied”. Otherwise, provide a proof of the model “goodness”

P.2. L.35-38: Not all GNSS systems support ionospheric corrections. E.g. GLONASS does not broadcast any ionospheric model parameters. Correct the sentence accordingly.

P.2 L.40: Provide a reference to the original description of Klobuchar model:

“Klobuchar JA (1987) Ionospheric time-delay algorithm for single frequency GPS users. IEEE Trans Aerosp Electron Syst 23(3):325–331. https://doi.org/10.1109/TAES.1987.310829”

P.2 L.41-42: NeQuick G model is based on the NeQuick model, but not NeQuick 2. Correct the statement and the reference accordingly, e.g.

“EC (2016) European GNSS (Galileo) Open Service—Ionospheric correction algorithm for Galileo single frequency users, Issue 1.2, Sept. 2016, European Commission”

P.2. L.42: Change “The NeQuick is” to “The NeQuick and its subsequent modifications (NeQuick G and NeQuick 2) are”

P.2 L.53: IRI model does not provide information about “electron and ion velocities”. It only provides information about equatorial vertical ion drift. Correct the sentence accordingly.

P.2 L.55-56: Change “The model is primarily” to “IRI is an empirical model primarily”

P.3 L.74: Change “GIM” to “global ionosphere model”, as GIM is already defined to be Global Ionosphere Map.

P.3 L.76: Change “GIM model” to “global ionosphere model”

P.3 L.80-82: The high values of RMS in low latitude region provided by CODE is, primarily, due to the inability of the selected model function (spherical harmonics) to describe ionospheric structure in low latitude. Modify the sentence accordingly.

P.3 L.84: Change “the GIM model” to “global models”

P.4 L.115: Author use TEC integrated up to COSMIC satellite heights (800 km) to construct the model (*“integration being done up to the altitudes of the COSMIC satellites”*). However, the topside TEC values (according to numerous studies, e.g. by Bilitza 2009, Yizengaw 2008 etc.) can reach from 10% to 80% of the total electron content (from ground to GNSS satellite heights). This fact significantly reduces the scientific value and application of the developed model. Essentially, the model is useless for GNSS applications.

P.5 L.124-126: This statement “*Since the magnitudes of the TEC obtained from COSMIC occultation 124 measurements are close to ground based GNSS TEC*”, is not consistent with the previous statement and studies by Mungufeni et al. 2019. Where they show that, depending on the location, the RMS error can vary from 2 to 8 TECU and error distribution plots show values from -24 to 20 TECU. Such large errors cannot be considered “*close to ground-based GNSS TEC*”. Authors, at least, are expected to provide information about relative TEC errors (in %, rather than TECU) to claim that errors can be tolerated (if so).

P.6 L.150: The title of the reference Emmert et al. 2010 is incorrect:

Emmert, J. T., Richmond, A. D., and Drob, D. P.: Statistical analysis of the correlation 412 between the equatorial electrojet and the occurrence of the equatorial ionisation 413 anomaly over the East African sector, J. Geophys. Res., 15; A08322; 414 doi:10.1029/2010JA015326, 2010.

P.6 L.157-167: The selected spatial resolution of 15^{o} in longitude and 5-8^{o} in latitude is too coarse to describe the ionosphere reasonably, especially for the low latitude region, where TEC is changing dramatically from the crest down/up to two peaks of EIA. E.g. GIM maps (the source of the data for most of the empirical models discussed by the authors in the introductions section) use at least 5^{o} by 2.5^{o} resolution (lon and lat). Moreover, 15^{o} in longitude corresponds to 1 hour in LT. Gradients in TEC as a function of LT during sunrise and sunset hours may reach tens of TECU per hour (e.g. Mungufeni et al. 2019, Fig. 2). Therefore, such coarse spatial resolution in longitude will lead to big errors in the model description.

P.6 L.170: The whole solar cycle 24 has relatively low solar activity level compared to the two previous ones. Nevertheless, even if we look only at the 24^{th} solar cycle, 2011 and 2016 could hardly

be attributed as years of high solar activity level. Please, modify the statement accordingly (e.g. as it is done on P.7 L.182).

P.7 L.189: Please clarify, how 36 solar flux bins were obtained. From the description, it is only 3 solar flux ranges and 12 months, that gives 36 (3x12). But when listing by a variable, only number 3 has to be specified, as it is done, for example with the rest of the variable (hour, lat and lon). Indeed, if we take 60,480 TEC values indicated in L.189, this number can be obtained by multiplying 5x14x3x12x24, but not 5x14x36x12x24.

P.8 L.205: According to the definition of cubic spline, it is a spline constructed of piecewise third-order polynomials, meaning none of the B splines used in the model were cubic (order 2 and 4). Change the "cubic B spline" into "B spline of different orders" throughout the text and abstract.

P.9 L.218-220: Consider changing this sentence to something like "In order to assess the ability of the model to describe the data used to construct the model, modelled data were compared to the experimental one. The results of the self-consistency check are presented in Figure 1."

P.9 L.228-229: It is surprising that the authors compare the results of the climatological model (i.e. model where input data were averaged over time, e.g. one month) with GIM map for a single day of that month. Such a comparison is not correct. On top of that, by looking at TEC maps obtained from COSMIC and later by B spline model (columns 2 and 1), one can hardly see any separation between the peaks of the EIA, that can, taking into account averaging in all the bins (e.g. lat and lon) performed by authors, hardly be comprehended.

P.9 L.231-232: By looking at the color plots, a reader can hardly assess the performance of the model. It is suggested, in addition to the plots, to present/discuss the results of the mismodelling in terms of a bias and RMS of the error.

P.10 L.250-252: From Fig 1 it can not be clearly understood the secondary maximum if any, especially at -20 lat. Please, if you discuss a feature, try demonstrating it clearly to the reader. A separate figure, or at least, a dashed line at -20 and 4 in Fig 1 is needed to support the statement.

P.11 L.269-270: In row (b), Fig 1, none of the panel show peaks of the EIA. There is no clear separation of the crest and peaks of EIA. Nor in panels b1/b2 neither in b3. Modify the sentence accordingly.

P.277-279: The structure of the crest might differ based on various factors (including level of the geomagnetic disturbance). However, when taken as an average, a clear 2 peak structure is present in low latitudes, representing EIA.

P.12 L.298-299: The science question in this case is not how to model the observed data, but how to explain the data. What is the physical explanation for the absence of the EIA structure (two peaks and the crest) in TEC values calculated from the ground up to COSMIC satellite heights (~800km).

And whether this phenomena is not a limitation of the technique applied to calculate TEC. Namely, TEC computed by integrating electron density profile, that by itself is a product of RO inversion, is subject to big errors, especially in places where big horizontal gradients exist (read, e.g. M.M Shaikh et al., Implementation of Ionospheric Asymmetry Index in TRANSMIT Prototype, DOI: 10.5772/58551). Without understanding the reasons of the observed behavior all the modelling efforts are meaningless.

P.13 L.313: One cannot see the "perfect match" of the observed and modelled data just by looking at the plots. At least a third row in form of difference map (error map) has to be presented to visually assess the error level. Moreover, statistical results (e.g. RMS and bias of the error) must be presented in order to make such a bold conclusion.

P.13 L.312-324: Authors do not discuss at all the TEC behavior observed in September at lat ~ -20, where its diurnal variation has a maximum during local night hours (21-03 LT). This maximum seems to exceed any other TEC values on this plot (row c, column 1 and 2) and looks like an error in the data processing. Such behavior seems to have no physical explanation.

P.14 Section 5: The authors fail to explain why they need yet another TEC model. Unless the performance of the newly created model is compared to existing models and it is demonstrated that its any better than the rest of the models present on the "ionosphere model market" (e.g. IRI, NeQuick, NTCM etc.), there is very little value in the study (both scientifically and application-wise).

P.15 L.350-353: Figure 4 does not show the full picture of the error distribution. It is clearly cut at -14 and 14 TECU. Iif one looks at Figure 3, errors in TEC can easily reach +-20 TECU (just draw a vertical line at any value of Observed TEC, e.g. at 30 TECU). It looks like the authors deliberately try to improve the results of their model performance.

Minor/Typo comments:

P.1 L.17: Change "derived" to "obtained"

P.1 L.19: Change "Geomagnetically quiet time (Kp < 3 and Dst > -20 nT) data during the years" to "Data during geomagnetically quiet time (Kp < 3 and Dst > -20 nT) for the years"

P.1 L.22 Change "to obtain the model" to "to obtain model coefficients"

P.1 L.26 Change "COSMIC TEC" to "COSMIC RO TEC"

P.2 L.31: Change "using Global Navigation Satellite Systems" to "in Global Navigation Satellite Systems"

P.2 L.30 Change "during day" to "during the day"

P.2. L.49: Space is missing between "European Geostationary"

P.2 L.50: Change "GPS And Geo-Augmented Navigation" to "GPS-aided Geo Augmented Navigation"

P.3 L.63: Space is missing in “analysis centers”

P.3 L.64: Space is missing in “using the”

P.3 L.64: Change “Global Ionospheric TEC data Map (GIM)” to “Global Ionosphere Maps (GIMs) containing vertical TEC data”

P.3 L.66: Change “Global Ionospheric TEC data Maps (GIMs)” to “GIMs”. It has been defined two lines above.

P.3 L.70: Space is missing in “the average”

P.3 L.71: Space is missing in “by CODE”

P.3 L.76: Space is missing in “constructed a”

P.3 L.77: Space is missing in ”GPS radio”

P.3 L.82: Space is missing in “related to”

P.4 L.87: Change “localized ionospheric structure” to “localized ionospheric structures”

P.4 L.88: Change “on a global scale model” to “in global models”

P.5 L.140: Space is missing in “during geomagnetically”

P.6 L.147: Change “solar activity” to “solar activity level”

P.6 L.164: Remove “15” in “reduced 15 to 5”

P.7 L.181 Space is missing in “the F10.7”

P.9 L.223: Change “Global Ionosphere Map (GIM) TEC (GIM-TEC)” to “GIM TEC”, as it was defined earlier, remove “Center for Orbit Determination in Europe” – it was defined earlier

P.9 L.225-226: Remove “The daily GIM-TEC values are derived using the GNSS data collected from over 200 tracking stations of IGS and other institutions”, as this information was given earlier in the text

P.10 L.238: Space is missing in “in turn”

P.14 L.336: Change “;” to “:”

P.14 L.337: Space is missing in “root mean squared”

P.17 L.373: Change “:” to “.” In “0.93”