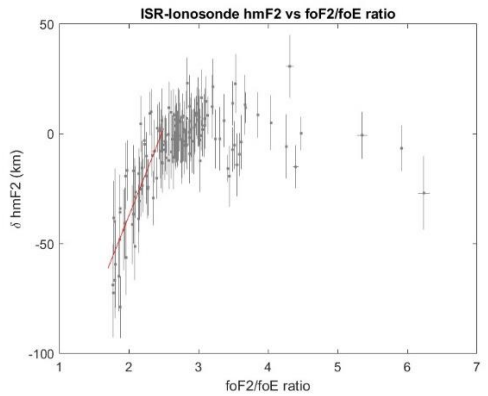
Response to reviewer Claudia Borries

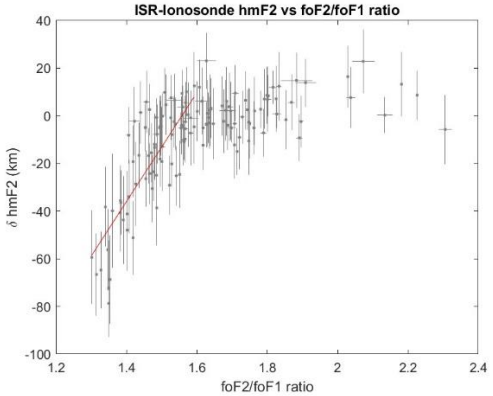
This manuscript written by C. Scott et al. addresses the limitations of using the height of the F2-layer maximum electron density (hmF2), which is a derived quantity from ionosonde data, for studying long-term changes in the ionosphere. This is an important topic because there have been already many publications using hmF2 for long-term studies and this work helps evaluating these results and using hmF2 more careful in future. From my point of view the manuscript is written excellently. It contains a very good overview of the state of the art at the beginning and the applied analysis and presentation of the results is adequate and well understandable. The authors detect a relation between the occurrence and strength of the F1 layer and the accuracy of the hmF2 layer (derived with one of the common approaches), which has not been described before. The results are discussed with respect to numerous related studies and the conclusions are logically derived from the results. The manuscript also contains some relevant results and discussions on the potential impact of climate change on the ionosphere. I evaluate the manuscript very good and I have some questions and remarks which may be considered before publication.

Questions and remarks:

- Section 3.5 provides corrections for the time delay in the ISR data. Why has this correction not been applied earlier in the study?

Thank you for this comment. This is a valid point and one we debated before we presented the work. We decided to present to analysis without the ISR data range correction (and simply quote the impacts on the various correlations when applied) because accounting for this correction required modelling the height of the F2 layer using the International Reference Ionosphere. We wanted to demonstrate that the result was not influenced by some implicit assumption in the model. The differences between the two analyses (with and without range correction applied) are small and the basic conclusions we draw from the analysis are the same in both cases. The equivalent plots for figures 4 and 7 in the revised manuscript but using the range-corrected data are shown below.





For transparency we would be happy make the range-corrected plots available as supplementary information.

We have modified the conclusions of this section to read;

'The results showed similar biases, with the coefficients of the linear fit exhibiting slight changes (gradient 79.6, offset -196.5 for the foF2/foE correction and a gradient of 46.6 with offset of -74.1 for the foF|2/foF1 correction). While it is not appropriate to apply these corrections to the individual points in the 35 year time series (since these have not been range-

corrected), as these are linear fits, the small changes to the coefficients will result in similarly small changes to the corrected values. The underlying conclusions concerning the impact of the foF2/foE and foF2/foF1 ratios on the empirical hmF2 formula are unaffected. The results showed similar biases, with the coefficients of the linear fit exhibiting slight changes (gradient 79.6, offset -196.5 for foF2/foE correction and a gradient of 46.6 with an offset of -74.1 for the foF|2/foF1 correction). While it is not appropriate to apply these corrections to the individual points in the 35 year time series (since these have not been range-corrected), as these are linear fits, the small changes to the coefficients will result in similarly small changes to the corrected values. The underlying conclusions concerning the impact of the foF2/foE and foF2/foF1 ratios on the empirical hmF2 formula are unaffected.'

- *line 482: Geomagnetic activity correlates to Joule heating driven by solar wind. How do the authors evaluate the potential that changes in thermosphere due to greenhouse effects may affect the magnitude of geomagnetic activity?*

This is an interesting question but one that is rather tangential to the subject covered in the current paper which concerns the calibration of ionosonde-based estimates of F2 layer heights (via an empirical formula) with those measured directly by Incoherent Scatter. While there may indeed be a modulation of the joule heating efficiency if the composition of the thermosphere is changed, the resulting changes in ionospheric layer height would be expected to affect both measures of F2 layer height and so would not introduce any bias into the calibration.

- *The authors use the ratio of foF2 and F10.7 as a composition proxy and refer to Wright and Conkright (2001). Such a proxy sounds very favourable, but reading Wright and Conkright (2001) I cannot see a justification that the ration of foF2 and F10.7 is a proxy for thermosphere composition. Wright and Conkright (2001) worked with a sunrise extrapolation index SRCC, which is related indirectly to foF2. Wright and Conkright (2001) correlate the ratio log(SRCC/F10.7) with log(O/N2) and find a rather moderate correlation. The authors describe in their conclusion that they intended to provide a composition index, but the morphology of the proposed one differs from that of [O/N2]. If the ratio [foF2/10.7] is used by Scott et al. as a proxy for thermosphere composition, they need to provide better justification.*

Thank you for pointing this out. This was also queried by reviewer #2. We have added the following text to describe this in more detail;

"Similarly, changes in neutral composition affect the peak electron concentration of the F2 layer. At noon, where the F2-layer approaches a steady-state condition, production and loss are in equilibrium. If the loss rate of ionisation is enhanced by the presence of molecular ions, the peak electron concentration of the layer will be reduced. Since the electron concentration is proportional to the peak frequency squared, comparison can be made between measurements at similar dates but different times in the solar cycle by scaling these values by the ion production rate, q. A good proxy for q is the F10.7cm solar radio flux. Thus

$$N \propto \frac{f^2}{q} \equiv \frac{foF2^2}{F10.7}$$

Using noon values of foF2 to track qualitative changes in thermospheric composition was suggested by Rishbeth et al (1995) with Wright and Conkwright (2001) comparing the efficacy of this simple index (their FFD index, averaged over 5 hours around noon) with other more complex indices derived from the rate of change of the ionosphere at sunrise."

Minor issues

- line 75 “Data from a such …”

Corrected. Thank you.

- lines 92 and 96 error in citing

Corrected, thank you.

- line 243 “… function of The Earth …”

Corrected. Thank you

- line 296 “In the ionosphere, While the …”

Corrected, thank you.

- line 371 “… in to the …”

This text has been changed in response to reviewer #2.

- line 389 “… formulae tends introduces …”

This text has been changed in response to reviewer #2.

- line 413 “(\pm 25 km at 250)”. Add “km altitude” after the 250

Changed as suggested.

- line 416: Using am index is very reasonable. However, since it is not yet very popular to use, I recommend adding some justification, why this is used instead of the more frequently used kp/ ap indices.

Thanks for this useful comment. In order to do this we have introduced the following text;

“We here choose to use the am index rather than the Ap index used in previous studies. The response patterns of the individual magnetic observatories used to compile such indexes depend strongly on the level of geomagnetic activity. At low activity levels the effect of solar zenith angle on ionospheric conductivity dominates over the effect of station proximity to the midnight-sector auroral oval, whereas the converse applies at high activity levels. It has been shown (Lockwood et al, 2019) that these biases are far smaller for the am index than for Ap”

- line 417: “… between the two, …” it is not immediately clear what are the two parameters that are correlated. Accordingly, it is not clear for the correlation values in lines 423 and 426, too.

Text expanded in all cases to explicitly state that the correlation is between this parameter and the model error.

- line 430: What is meant with “longitude sector near to the geomagnetic pole (\approx 48-50N)”? How can a longitude sector be close to the pole and why does it have latitude (north) coordinates?

Thank you for pointing out this ambiguity. We have expanded the text to read;

“… is a mid-latitude station in a geographic longitude sector near to the geomagnetic pole (at a geomagnetic latitude during this epoch of ~48-50 N).

- line 438: I get confused with the description. Chilton does not have a semiannual variation in foF2?

Thank you for pointing out this ambiguity in the text, We have restructured this paragraph to read;

“In contrast, Stanley in the Falkland Islands (at a geomagnetic latitude of ~35- 39 S during this epoch) is a station that is far enough from the magnetic pole that compositional changes between equinox and winter months are relatively small compared with the associated change in solar zenith angle, resulting in a semiannual variation in foF2 (Millward et al,1996).”

- line 438-439 “such as is seen”: Where is it seen? Is there a figure or paper?

This point has also been addressed in the above restructuring

- line 437: What means far enough away? Does it just need to be outside the auroral oval or even further away?

The distance varies depending on geomagnetic activity levels. We have added the following text to support this qualitative statement.

“The relative magnitudes of the annual and semi-annual variations at a given station vary depending on geomagnetic activity, resulting in the long-term trends identified by Scott, Stamper and Rishbeth, (2014).”

- line 440 “at these stations”: here the two station are addressed and in the second part of the sentence only Chilton. This is confusing.

Thank you for pointing this out. We have split the discussion of Chilton into a separate sentence;

“Such differences are also likely to influence the relative values of the foF2/foE and foF2/foF1 ratios at these stations. For example, the ratios at Slough/Chilton will be lower during the summer when compositional change suppresses foF2 while foE and foF1 are at their peak.”