URSI GASS 2020, Rome, Italy 29 Aug. - 5 Sep. 2020

Probabilistic forecasts of ionospheric low-latitude scintillation using an ensemble data assimilation model

Luke D. Nugent* and Sean Elvidge
SERENE, University of Birmingham, UK

Abstract

A proxy for vertical plasma drift (PVPD) at the magnetic equator can be used as an indicator of whether strong ionospheric scintillation will occur at low-latitudes. Using this approach with a physics-based model has previously been shown to demonstrate forecasting skill. AENeAS is an ensemble data assimilation model with a physics-based background. In this work scintillation forecasting using the mean of the AENeAS ensemble is shown to increase forecasting skill when compared to forecasting with a deterministic physics-based model. AENeAS is also used to create probabilistic forecasts by generating an ensemble of PVPDs. Using kernel density estimation these PVPDs are combined to form a probability density function of the PVPD speed. The probability of strong scintillation occurring can then be determined. This method can be applied at a range of longitudes on the magnetic equator, thereby providing a global, probabilistic, early warning forecast of low-latitude scintillation.

1. Introduction

In the low-latitude ionosphere plumes of depleted plasma can rise into regions with higher plasma density during the evening. Signals passing through these equatorial plasma bubbles (EPBs) can be subjected to rapid fluctuations of amplitude or phase (scintillation) resulting in signal quality reduction or, in severe cases, total signal loss. As society's dependence on trans-ionospheric communications increases [1], the ability to mitigate the effects of EPBs becomes more important. Accurate forecasting of EPB generation will enable those reliant on trans-ionospheric transmissions to reduce the likelihood of being impacted by severe scintillation.

Anderson et al. [2] used the temporal rate of change of the virtual height of a 4 MHz ionosonde return signal between 1830 and 2000 local time (LT) as a proxy for vertical plasma drift (PVPD) speed at the magnetic equator. This corresponds to the rate of change of the altitude at which an electron density of 2×10^{11} e$^-$/m^3 occurs during this period. Anderson et al. demonstrated that this PVPD can be used as an indicator of whether strong scintillation would be observed within the region during the subsequent night. Limitations of this approach include that, as the observation period is from 1830 to 2000 LT, forecasts cannot be provided with significant antecedence. Furthermore, ionosondes would be required in every scintillation forecast region. To avoid these issues the PVPD forecasting approach can be adapted to use output from an ionospheric model. The greatest change in height of an electron density of 2×10^{11} e$^-$/m^3 between each model time step from 1830 to 2000 LT determines the PVPD speed so only electron densities and the corresponding altitudes are required for PVPD forecasting with a model. The adaptation of this approach to use an ionospheric model is, therefore, straightforward and provides a computationally cheap forecasting technique.

Using output from the physics-based Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM; [3]) PVPD forecasting demonstrated forecasting skill when compared to a more complex and computationally expensive forecasting approach [4] which uses TIE-GCM to calculate field-line integrated Rayleigh-Taylor growth rates [5]. PVPD forecasting skill was able to match or outperform the Rayleigh-Taylor growth rate method and persistence forecasting in almost all considered test cases [4] when thresholds used to determine whether a day is predicted/observed to have strong scintillation are not set too low or high (when most days in the test case would fall into one prediction or observation class).

The initial state of the upper atmosphere, from which a physics-based model is propagated forward in time, cannot be specified with unlimited accuracy. Furthermore, computational errors in the physics-based model solvers introduce further uncertainties during model propagation [6]. Therefore, rather than integrating a single path from a best guess of the initial state, it is more appropriate to consider the evolution of the uncertainty range of possible initial states. Such an approach can be implemented through use of an ensemble model [7]. In this work the Advanced Ensemble electron density (Ne) Assimilation System (AENeAS; [8]) will provide output for PVPD forecasts.

AENeAS is an ensemble model of the coupled ionosphere-thermosphere system which incorporates the assimilation of data from a variety of sources using the local ensemble transform Kalman filter. Each ensemble member uses an ionosphere-thermosphere model simulation to provide background conditions. In this work an independent TIE-GCM simulation is used for each of 32 ensemble members. Total electron content (TEC) observations are assimilated once in each 15 minute model time step.

A commonly used indicator of amplitude scintillation is the S4 index:

$$S4 = \sqrt{\frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}}, \qquad (1)$$

where I is the signal intensity over a chosen period (usually 60 seconds). In this work the 90th percentile of S4 observations is determined for each hour between sunset and sunrise. The largest of these 90th percentile values ($S4_{90}$) is used to determine whether a day is classified as a strong scintillation day.

2. Deterministic forecasting: method and results

Nugent et al. [4] compared the forecasting skill of the Rayleigh-Taylor growth rate (RTGR; [5]) and proxy for vertical plasma drift (PVPD) approaches using TIE-GCM output. A direct comparison with PVPD forecasting skill using AENeAS can be obtained by using the mean electron densities and altitudes for the same latitude, longitude and pressure level across all ensemble members. This provides a single altitude for an electron density of 2×10^{11} e$^-$/m^3 at a specified location on the magnetic equator. The change in height of this electron density can then provide a maximum PVPD value as described in section 1.

Figure 1 demonstrates forecasting skill for a range of $S4_{90}$ thresholds (where the $S4_{90}$ threshold determines whether a day is classified as a strong scintillation day or not) at Vanimo, Papua New Guinea for 56 days in March and April 2000 (S4 data as used by Carter et al. [5]). Figure 1a shows the area under the receiver operating characteristic (ROC) curve which provides an indicator of forecasting skill when suitable PVPD/RTGR thresholds are not known. The area under the ROC curve (AUC) represents the probability that the PVPD/RTGR value for a randomly chosen strong scintillation day is greater than the PVPD/RTGR value for a randomly chosen weak scintillation day. Figure 1b shows the maximum Youden's Indices (YIs; also known as Peirce skill scores [9]) which provide an indicator of forecasting skill when suitable PVPD/RTGR thresholds are known. The maximum YI represents the model's greatest skill improvement over random chance. Further details are available in Nugent et al. [4].

Figure 1 clearly shows that in this test case, for all considered $S4_{90}$ thresholds using both AUCs and maximum YIs, AENeAS PVPD forecasting (brown) consistently outperforms TIE-GCM PVPD (blue) and RTGR (red) forecasting and persistence forecasting (green).

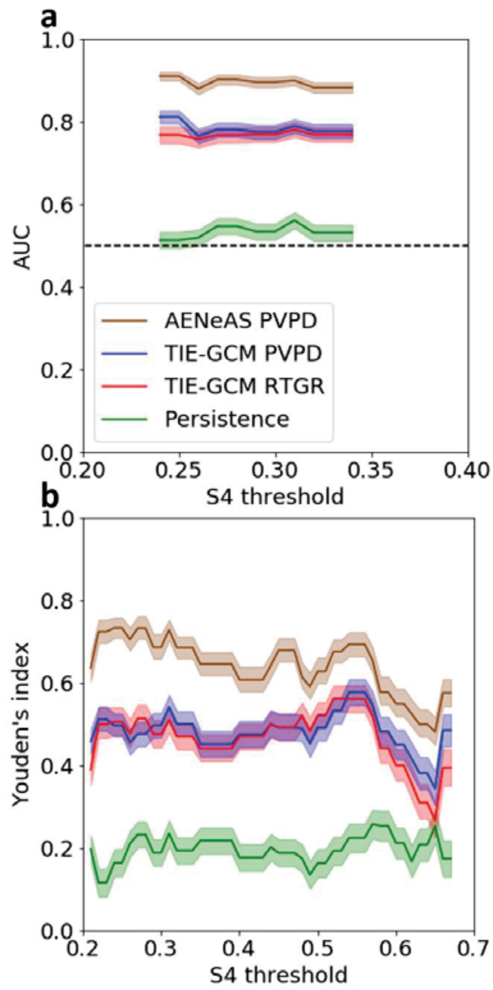

Figure 1. AUCs (a) and maximum YIs (b) for PVPD forecasting with AENeAS (brown) and TIE-GCM (blue), RTGR forecasting with TIE-GCM (red) and persistence forecasting (green) at Vanimo, Papua New Guinea in March and April 2000. Coloured shaded regions represent the range of skill values within the mean and two standard deviations from leave-one-out jackknifing (effectively a ~95% confidence interval of skill values subject to small changes in the data set). The horizontal dashed line in (a) represents a model with no skill.

3. Probabilistic forecasting: methods

Deterministic forecasting provides users with a binary forecast of whether strong scintillation will or will not occur. Probabilistic forecasts, however, include the uncertainty of a predicted event. This can provide users with a clearer understanding of whether strong scintillation will occur.

Using the AENeAS ensemble mean to generate PVPD forecasts has demonstrated an improvement in forecasting skill when compared to PVPD forecasting with a single TIE-GCM simulation (Figure 1). However, the availability of output from each ensemble member provides an opportunity to generate a probabilistic scintillation forecast which may be more accurate and useful for users. Using the approach discussed in section 1 the maximum PVPD can be determined for each ensemble member to produce an ensemble of PVPD values. These values can be used to estimate a probability density function (PDF) for the PVPD speed. In this work PDFs are produced using kernel density estimation (KDE). The kernel has been chosen to be Gaussian and for each PVPD ensemble member a PDF is produced with mean equal to the PVPD value and a fixed standard deviation (SD). The estimated PVPD PDF is the normalised sum of these Gaussian distributions.

The choice of SD for the set of Gaussian distributions is an important factor as it can significantly affect the final PVPD PDF. If the SD is too small the resulting PDF will be too dependent on each ensemble member output value (effectively a spike at each output value which does not take into account that the ensemble member PVPDs are only samples from the overall PDF). If the SD is too large then information from the PVPD values about the underlying distribution will be lost. To find a SD which avoids these issues, j evenly spaced SD values between (e.g.) zero and ten (zero not included) are tested. Leave-one-out cross-validation determines the best SD to use in the following way:

1) In this work AENeAS uses 32 ensemble members thereby producing 32 PVPD values. Select one PVPD value to be the test set. The remaining 31 PVPD values are the training set.
2) From the set of j SDs, $\mathrm{SD}_i(i \in 1, \dots, j)$ is used as the KDE SD, thereby generating j PVPD PDFs using the 31 PVPD values in the training set.
3) Find the probability density of the test set value in each of the PVPD PDFs. These densities provide a likelihood function of SDs which would produce the test set value.
4) Repeat steps one to three 31 times so each PVPD value has been used as the test set value. 32 likelihood functions have now been produced.
5) The product of the 32 likelihoods associated with $\mathrm{SD}_i(i \in 1, \dots, j)$ provide the likelihood that the PVPD PDF generated using SD_i would produce the 32 PVPD values. The SD with the greatest of these combined likelihoods is chosen as the SD to be used for the KDE of the final PVPD PDF with all 32 PVPD values.

The PVPD PDF can then be used to estimate the probability that a PVPD is greater than a chosen threshold, $P(PVPD > PVPD_{thresh})$. The probability of strong scintillation occurring during the subsequent night, $P(S4 > S4_{thresh})$, can be determined using the law of total probability:

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^C) \\ &= P(A \mid B)P(B) + P(A \mid B^C)P(B^C), \end{aligned} \tag{2}$$

where B^C is the complement of B (i.e. $P(B^C)$ is the probability that event B does not occur) and $P(A \mid B)$ is the conditional probability that event A will occur given that event B has occurred. Therefore,

$$\begin{aligned} &P(S4 > S4_{thresh}) = \\ &\quad P(S4 > S4_{thresh} \mid PVPD > PVPD_{thresh}) \\ &\qquad \times P(PVPD > PVPD_{thresh}) \\ &+ P(S4 > S4_{thresh} \mid PVPD \leq PVPD_{thresh}) \\ &\qquad \times P(PVPD \leq PVPD_{thresh}). \end{aligned} \tag{3}$$

Anderson et al. [2] reported that the maximum five minute average S4 between sunset and sunrise was found to be greater than 0.5 on 90% of days with observed PVPDs greater than 20 m/s and less than 0.5 on 85% of days with PVPDs less than 20 m/s. Using these values the probability that the maximum five minute average S4 is greater than 0.5 is given by

$$\begin{aligned} &P(S4 > 0.5) = \\ &\qquad P(S4 > 0.5 \mid PVPD > 20\, m/s) \\ &\qquad\quad \times P(PVPD > 20\, m/s) \\ &\qquad + P(S4 > 0.5 \mid PVPD \leq 20\, m/s) \\ &\qquad\quad \times P(PVPD \leq 20\, m/s) \\ &= 0.9\, P(PVPD > 20\, m/s) \\ &\quad + (1 - 0.85) \times [1 - P(PVPD > 20\, m/s)] \\ &= 0.15 + 0.75\, P(PVPD > 20\, m/s), \end{aligned} \tag{4}$$

where $P(PVPD > 20\, m/s)$ is determined from the PVPD PDF. However, these thresholds and forecast success rates are likely to vary under differing conditions such as location, season and solar activity (e.g. [10]).

4. Probabilistic forecasting: testing

A commonly used method to determine forecasting skill for probabilistic forecasts is the Brier Score (BS; [11]), given by

$$BS = \frac{1}{n}\sum_{t=1}^{n}(f_t - E_t)^2. \tag{5}$$

When used for a binary scintillation forecast n is the number of days in the test case, f_t is the forecasted probability of strong scintillation occurring during day t and E_t represents whether the event (strong scintillation) actually occurred. $E_t = 0$ if the event did not occur on day t and $E_t = 1$ if the event did occur. A model which provides a forecast of 100% certainty that each event will or will not occur and is correct on every occasion (i.e. a perfect model) will achieve $BS = 0$ whereas a model with 100% certainty forecasts which are wrong on every occasion will have $BS = 1$. A model with no skill (i.e. the model predicts 50% probability of the event occurring on every occasion) will have $BS = 0.25$. In this work the probabilistic forecasting approach will be compared to deterministic forecasting techniques by assigning a strong scintillation prediction forecast of 0% or 100% for the deterministic model dependent on whether the deterministic model predicts a value below or above the $S4_{90}$ threshold respectively (so the BS will simply be the proportion of days which were incorrectly predicted).

Probabilistic PVPD forecasting will also be compared to forecasts from the empirical WideBand MODel (WBMOD; [12]) which provides predictions of the proportion of time for which scintillation levels exceed a chosen threshold.

5. Summary

Determining a proxy for vertical plasma drift (PVPD) at the magnetic equator using the physics-based model TIE-GCM has previously been shown to provide an effective predictor for whether strong scintillation will occur in the subsequent night at equatorial latitudes [4]. Use of an ensemble data assimilation model (AENeAS) has been shown to increase forecasting skill (compared to TIE-GCM) when using the mean of ensemble output to generate PVPDs. A method to provide probabilistic forecasts of low-latitude scintillation using output from each AENeAS ensemble member has been proposed. PVPDs are calculated for each ensemble member and this ensemble of PVPD values is used to estimate the probability density function of PVPD speed using kernel density estimation. Using this probability density function and the law of total probability an estimation of the probability of strong scintillation occurring can be calculated if deterministic forecast success rates and suitable PVPD thresholds are known. The forecasting skill of this approach will be assessed using the Brier Score and compared to the forecasting skill of deterministic forecasting approaches and the WideBand MODel (WBMOD). By applying the probabilistic PVPD forecasting approach with AENeAS at a range of longitudes it will be possible to provide a global, probabilistic, early warning forecast of low-latitude ionospheric scintillation (Figure 2).

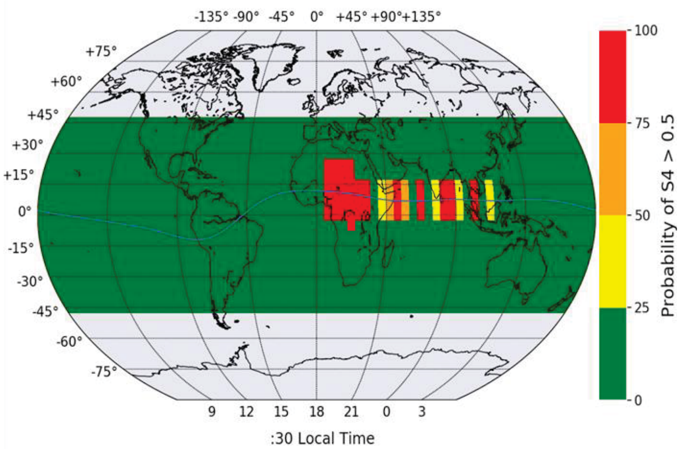

Figure 2. Example of low-latitude probabilistic scintillation forecast. White regions are outside the region of forecast.

6. Acknowledgements

This work has been partly funded by the Dstl Space programme and supported by the UK Met Office. We are grateful to the World Data Centre of the Australian Bureau of Meteorology, Space Weather Services for the provision of ionospheric scintillation data.

References

[1] M. Hapgood, "Satellite navigation—Amazing technology but insidious risk: Why everyone needs to understand space weather," *Sp. Weather*, vol. 15, no. 4, pp. 545–548, Apr. 2017.

[2] D. N. Anderson, B. Reinisch, C. Valladare, J. Chau, and O. Veliz, "Forecasting the occurrence of ionospheric scintillation activity in the equatorial ionosphere on a day-to-day basis," *J. Atmos. Solar-Terrestrial Phys.*, vol. 66, no. 17, pp. 1567–1572, Nov. 2004.

[3] A. D. Richmond, E. C. Ridley, and R. G. Roble, "A thermosphere/ionosphere general circulation model with coupled electrodynamics," *Geophys. Res. Lett.*, vol. 19, no. 6, pp. 601–604, Mar. 1992.

[4] L. D. Nugent, S. Elvidge, and M. J. Angling, "Comparison of low-latitude ionospheric scintillation forecasting techniques using a physics-based model," *Sp. Weather*, 2020 [*submitted*].

[5] B. A. Carter *et al.*, "An analysis of the quiet time day-to-day variability in the formation of postsunset equatorial plasma bubbles in the Southeast Asian region," *J. Geophys. Res. Sp. Phys.*, vol. 119, no. 4, pp. 3206–3223, Apr. 2014.

[6] J. R. Rozante, D. S. Moreira, R. C. M. Godoy, and A. A. Fernandes, "Multi-model ensemble: technique and validation," *Geosci. Model Dev. Discuss*, vol. 7, pp. 2933–2959, 2014.

[7] S. Elvidge, H. C. Godinez, and M. J. Angling, "Improved forecasting of thermospheric densities using multi-model ensembles," *Geosci. Model Dev*, vol. 9, pp. 2279–2292, 2016.

[8] S. Elvidge and M. J. Angling, "Using the local ensemble Transform Kalman Filter for upper atmospheric modelling," *J. Sp. Weather Sp. Clim.*, vol. 9, p. A30, Aug. 2019.

[9] C. S. Peirce, "The numerical measure of the success of predictions," *Science (80-.).*, vol. ns-4, no. 93, pp. 453–454, Nov. 1884.

[10] J. M. Smith, F. S. Rodrigues, B. G. Fejer, and M. A. Milla, "Coherent and incoherent scatter radar study of the climatology and day-to-day variability of mean F region vertical drifts and equatorial spread F," *J. Geophys. Res. A Sp. Phys.*, 2016.

[11] G. W. Brier, "Verifications of forecasts expressed in terms of probability," *Mon. Weather Rev.*, vol. 78, no. 1, pp. 1–3, Jan. 1950.

[12] J. A. Secan, R. M. Bussey, E. J. Fremouw, and S. Basu, "An improved model of equatorial scintillation," *Radio Sci.*, vol. 30, no. 3, pp. 607–617, 1995.