Optimizing AENEAS: Enhancing the UK's Operational Model for Space Weather Forecasting

Jack W. R. Brown, David R. Themens, Sean Elvidge, and Matthew K. Brown

Abstract – An accurate representation of the coupled ionosphere-thermosphere system is essential for reliable space weather operations, as disturbances in these regions can disrupt modern technological systems. This has motivated the development of operational data assimilation models, such as the Advanced Ensemble NEtworked Assimilation System (AENEAS), which combine model forecasts with observations to improve upper-atmospheric state estimates. Although AENEAS offers considerable flexibility through configurable parameters governing ensemble behavior and observation handling, these settings had not previously been optimized. This work presents the first dedicated optimization framework for AENEAS, assessing a range of tuning parameters and quantifying their impact on model performance under both quiet and disturbed geomagnetic conditions. Validation against global ionosonde observations and comparison with established models demonstrate that targeted tuning consistently improves model behavior and statistical performance, while reducing the over-representation of G-condition within the model. G-condition occurrence decreases from 14.00% to 1.85%, compared with an observed occurrence rate of 1.48% for the test periods. The resulting improvements in model reliability support the use of the optimized configuration in operational settings.

Manuscript received 14 January 2026. S. Elvidge and AENEAS development are supported in part by the UK Space Weather Instrumentation, Measurement, Modelling and Risk (SWIMMR) Programme, NERC Grants NE/V002643/1 and NE/V002708/1. MB and DRT's contributions to the work are supported through the UK NERC FINESSE (NE/W003147/1) and DRIIVE (NE/W003368/1) grants. J.B. is supported through EPSRC doctoral training award #2931970.

GPS TEC data products and access through the Madrigal distributed data system are provided to the community by the Massachusetts Institute of Technology under support from US National Science Foundation grant AGS1242204. Data for the TEC processing is provided from the following organizations: UNAVCO, Scripps Orbit and Permanent Array Center, Institut Geographique National, France, International GNSS Service, The Crustal Dynamics Data Information System (CDDIS), National Geodetic Survey, Instituto Brasileiro de Geografia e Estatística, RAMSAC CORS of Instituto Geográfico Nacional de la República Argentina, Arecibo Observatory, Low-Latitude Ionospheric Sensor Network (LISN), Topcon Positioning Systems, Inc., Canadian High Arctic Ionospheric Network, Institute of Geology and Geophysics, Chinese Academy of Sciences, China Meteorology Administration, Centro di Ricerche Sismologiche, Système d'Observation du Niveau des Eaux Littorales (SONEL), RENAG: Reseau NAtional GPS permanent, GeoNet - the official source of geological hazard information for New Zealand, GNSS Reference Networks, Finnish Meteorological Institute, SWEPOS - Sweden, Hartebeesthoek Radio Astronomy Observatory, Crustal Dynamics Data Information System (CDDIS), Astronomical Institute of the University of Bern, TrigNet Web Application, South Africa, Australian Space Weather Services, RETE INTEGRATA NAZIONALE GPS, Estonian Land Board, and Virginia Tech Center for Space Science and Engineering Research.

Jack W. R. Brown, David R. Themens, Sean Elvidge, and Matthew K. Brown are with the Space Environment and Radio Engineering Group (SERENE), University of Birmingham, Birmingham, UK; e-mail: j.brown.14@pgr.bham.ac.uk, d.r.themens@bham.ac.uk, s.elvidge@bham.ac.uk, m.brown.12@bham.ac.uk.

1. Introduction

The thermosphere forms the layer of Earth's upper atmosphere above approximately 100 km, where temperatures increase because of the absorption of high-energy solar radiation by neutral species. Extreme ultraviolet radiation photoionizes these neutrals, forming the ionosphere, which is embedded within and strongly coupled to the thermosphere. The ionosphere-thermosphere (I-T) system plays a major role in the operation of modern space-based technologies: ionospheric plasma density affects Global Navigation Satellite Systems (GNSS) and high-frequency radio propagation. In contrast, thermospheric neutral density impacts satellite drag and orbit prediction in low Earth orbit [1]. As reliance on these systems increases, so too does the need for accurate monitoring and forecasting of upper-atmospheric variability. However, modeling this region remains challenging due to its complex composition, rapid variability, and strong coupling to both the lower atmosphere and Earth's magnetosphere. Existing approaches are commonly classified into empirical, physics-based, or data-assimilation models. The Advanced Ensemble Networked Assimilation System (AENEAS) belongs to the latter category, combining the physics-based Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM) [2] with a Local Ensemble Transform Kalman Filter (LETKF) to produce physically consistent, observation-constrained estimates of the I-T system [3].

In this work, AENEAS outputs modeled variables on a global three-dimensional grid with a 5.0° × 5.0° horizontal resolution and a fixed-altitude vertical grid, interpolated from the scale-height-dependent TIE-GCM pressure levels. The TIE-GCM background extends from approximately 97 to 600 km (the upper boundary is dependent on solar activity), with finer resolution at the bottom side of the ionosphere and coarser spacing above approximately 500 km. Electron density is extended into the topside using NeQuick [4], providing coverage to approximately

20,000 km. The gridded upper-atmosphere variables form the state vector, which is updated by the LETKF using slant Total Electron Content (sTEC) observations within 15-minute assimilation windows. Model uncertainty is represented using an ensemble of 32 TIE-GCM members generated by perturbing key geophysical drivers, including the solar F10.7 flux and two Kp-derived parameters: hemispheric power and cross-tail potential. Driver values are updated every three hours using up-to-date observations, ensuring the ensemble evolves with changing physical conditions.

Recent developments in AENEAS have led to a new model version (v2.0) and its operational deployment at the UK Met Office for upper-atmospheric nowcasting and forecasting, motivating a comprehensive assessment of the factors influencing model performance. Although previous work has evaluated AENEAS against its background models and considered aspects of parameter configuration [3], the influence of its tunable parameters has not yet been examined systematically. The parameters considered in this study include observation-handling settings that govern the influence of assimilated data, such as outlier rejection, observation error inflation, and spatial localization (area of influence); as well as ensemble control settings, including the internal TIE-GCM timestep, the amplitude of perturbations applied to model drivers, covariance inflation, and change limits that cap the fractional adjustment applied to variables at each assimilation step. This study aims to identify an optimal parameter configuration that improves both model performance and overall model behavior.

2. Model Evaluation Framework

To identify an improved AENEAS configuration, a two-stage approach was adopted, consisting of parameter optimization followed by initial validation. In all testing, AENEAS assimilated GNSS sTEC observations from the Madrigal database, a web-based archive of upper-atmosphere measurements [5]. Although AENEAS can assimilate multiple observation types, including in-situ and radio occultation data, the current operational configuration is limited to GNSS sTEC due to data latency constraints; this restriction is therefore retained in this study. Pre-processing was applied to ensure data quality and computational efficiency. Raw 30-second sTEC measurements from all global stations were too dense for the 15-minute assimilation cycle and were temporally correlated along individual satellite-receiver paths, violating the LETKF assumption of uncorrelated errors. To address this, observations below 15° elevation were excluded to reduce multipath effects, and spatial thinning was applied by selecting a single receiver per 1.0° × 1.0° region for assimilation. The remaining data were averaged into 5-minute intervals to reduce both volume and temporal correlation. Each averaged value was assigned an uncertainty, ε_{TEC}, as:

$$\varepsilon_{TEC} = \sqrt{1.5^2 + (10\sigma_{TEC})^2}, \quad (1)$$

where σ_{TEC} is the standard deviation of TEC within each 5-minute window. The factor of 10 ensured consistency with AENEAS's operational observation-error statistics, while the 1.5 TECU term set the baseline error.

The optimization phase was conducted for a geomagnetically quiet day (8 February 2025) to isolate model behavior under stable conditions. A sequential tuning approach was implemented in which multiple candidate values of each parameter were tested at each step. Performance was assessed using statistical metrics, including root mean square error (RMSE), continuous ranked probability score (CRPS) [6], mean error (ME), Pearson correlation (R), and χ^2 to evaluate ensemble spread. At each stage, configurations were ranked across all metrics, and the highest-performing setup was retained for use in the subsequent stage. Promising parameter combinations outside the main sequence were also assessed, while configurations that degraded performance or led to ensemble collapse were rejected. The first parameter examined was outlier rejection, which removes observations lying outside 3σ from the ensemble mean before assimilation (on/off). The internal TIE-GCM timestep was then varied (30, 20, and 10 s) to test the trade-off between numerical stability and temporal resolution. Observation error was refined by inflating GNSS sTEC uncertainties by factors of ×1, ×6, ×10, and ×12 to reduce the relative weight of observations within the assimilation. Ensemble driver perturbations were subsequently explored. Kp-driven inputs were not modified, as even small perturbations resulted in ensemble collapse. Whereas F10.7 forcing, implemented via a random-walk process, was adjusted by increasing the allowable step size from ±3.0 to ±6.0 and ±12.0 sfu to expand the range of sampled solar flux conditions. Covariance inflation values of 1.0, 1.1, 1.5, and 3.0 were also tested to fine-tune the relative influence of the background model. Spatial localization angles of 10°, 15°, and 20° were tested to assess the effect of observational area of influence on model output. Finally, change-limit parameters for ion and electron temperatures (Ti & Te), electric potential and vertical transport (EP & VT), and neutral winds were adjusted. The baseline values of 0.00 were found to overly constrain physical variability, so values of 0.05, 0.10, and 0.20 were tested to increase model freedom while maintaining stability. The optimized settings are shown in Table 1.

To assess robustness and avoid overfitting to quiet-time conditions, the optimized configuration was then applied to the 14 to 17 July 2017 geomagnetic storm period. During this interval, optimized AENEAS performance was benchmarked against the default configuration and several established upper atmosphere models, including TIE-GCM, NeQuick, the IRI [7], and WACCM-X [8]. The July interval had originally been

Table 1. Summary of default and optimized settings.

Tuning parameter	Default	Optimized
Outlier Rejection	On	Off
TIE-GCM Timestep, s	30	30
Obs. Error Inflation	1×	10×
F10.7 Spread Scale	3.0	6.0
Covariance Inflation	1.1	1.5
Localization Angle	15°	10°
Change Limits: (Ti & Te, EP & VT, Winds)	0.0	0.2, 0.2, 0.1

selected for storm-time evaluation under the baseline configuration; however, initial tests revealed that AENEAS frequently entered unphysical states, further motivating the optimization presented here.

For both periods, model outputs were compared against global ionosonde observations of the F2-region critical frequency (foF2) and peak height (hmF2) using the stations shown in Figure 1.

3. Optimization Results

Figure 2 provides a representative example of the optimization effects for Grahamstown on 8 February 2025, comparing the default AENEAS with the optimized settings.

In the default configuration, the ensemble members remain tightly clustered around the ensemble mean, indicating a restricted ensemble spread and overconfident model uncertainty. The model also exhibits unstable behavior, particularly after 12:00 UT, where it departs from the observed foF2 trend. These characteristics likely arise from a combination of limited ensemble variability, poorly characterized observation errors, and outlier rejection. The tight ensemble restricts the range of I-T states the model can represent and makes the 3σ outlier threshold overly restrictive, leading to observations that could help correct the model state being rejected as outliers. Additionally, poorly characterized observation errors allow assimilated data to exert overly strong corrections, producing spike-like behavior and unrealistic variability. In contrast, the optimized configuration exhibits a broader and more

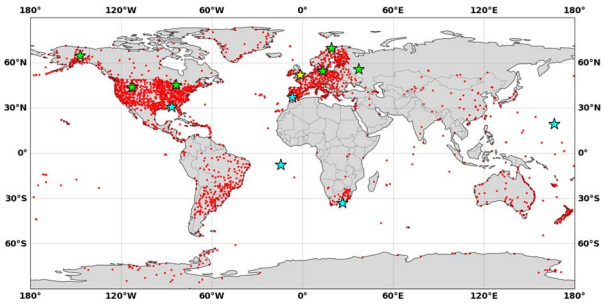


Figure 1. Global distribution of ionosonde stations used for model comparison (stars): blue for February 2025, green for July 2017, and yellow for stations used in both periods. Red circles indicate the 2240 GNSS receivers contributing assimilated observations.

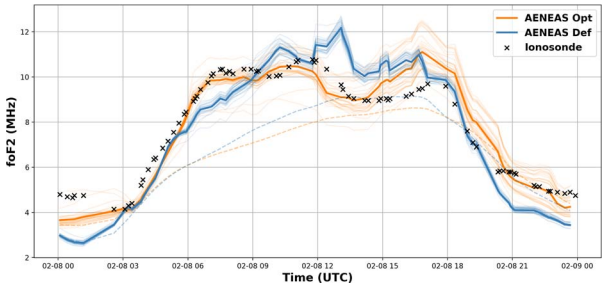


Figure 2. Ionosonde foF2 observations (black) at Grahamstown, South Africa (8 February 2025), compared with AENEAS default (Def) and optimized (Opt) configurations. Solid lines show ensemble means, faint lines indicate ensemble members, and dashed lines represent the corresponding background runs without data assimilation. Dates are in MM-DD HH format.

responsive ensemble spread that better represents the evolving uncertainty of the system. The ensemble remains compact during periods of good agreement with observations, while widening appropriately as the model departs from the observed trend, indicating a better balance between the background model and assimilated data. The optimized settings also provide greater freedom for the model to explore a wider range of I-T states, enabling the ensemble to more consistently encompass the observations. Consequently, the diurnal variation is reproduced more realistically, with the spike-like artifacts present in the baseline largely removed.

These improvements extend beyond the Grahamstown case and are seen in global performance metrics derived from all ionosonde stations for the period (Table 2). The optimized configuration shows consistent improvement across all metrics for both foF2 and hmF2. Reductions in CRPS and χ^2 indicate improved ensemble calibration and more realistic uncertainty representation. Increased correlation reflects the smoother, more physically consistent evolution of the optimized system relative to the erratic variability of the default. Reduced RMSE and bias confirm improved deterministic accuracy alongside the improved ensemble behavior.

4. Initial Validation Results

Pretesting of the July 2017 storm period revealed a major limitation in the baseline AENEAS configuration: the model frequently entered G-condition, in which the electron density peak occurred in the F1 region rather than the F2 region. Although G-condition is an observed ionospheric phenomenon, it is rare and

Table 2. Percentage change (% Δ) in performance metrics for the optimized settings during February 2025 calculated relative to the default. Positive values indicate improved performance in the optimized configuration.

% Δ	RMSE	CRPS	ME	R	χ^2
foF2	+18.6	+27.3	+19.1	+44.6	+42.4
hmF2	+26.6	+29.5	+22.1	+22.7	+39.7

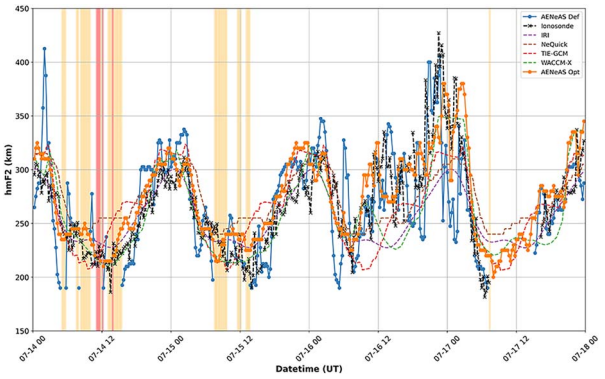


Figure 3. Modeled and observed hmF2 at Fairford, UK (14-17 July 2017). Ionosonde data are in black; optimized AENEAS (Opt) in orange; default AENEAS (Def) in blue; IRI in purple; TIE-GCM in red; NeQuick in brown; and WACCM-X in green. Background shading indicates periods of AENEAS-modeled G-condition: orange for default settings only and red for both. No shading is shown where ionosonde hmF2 observations are unavailable. Dates are given in MM-DD HH format.

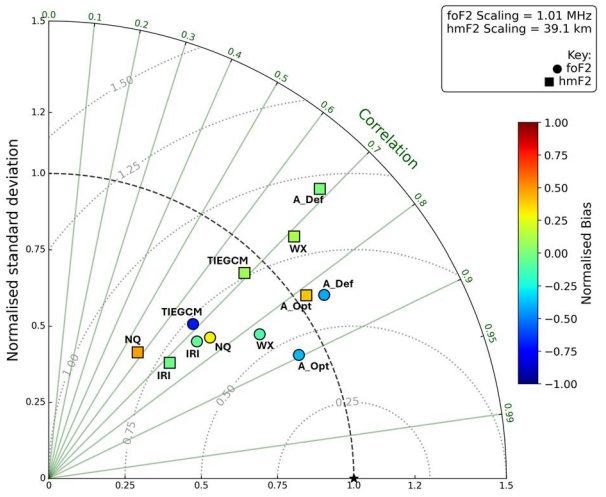


Figure 4. Modified Taylor diagram comparing upper-atmosphere model performance for 14 to 17 July 2017. A_Opt (AENEAS Optimized); A_Def (AENEAS Default Settings); NQ (NeQuick); WX (WACCM-X).

mainly associated with geomagnetically active periods [9]. Its repeated occurrence during the prestorm quiet phase, when the F2 layer is normally well defined, and its inconsistency with ionosonde observations indicated that the model was getting stuck in physically unrealistic states. This behavior suggested deficiencies in the modeled thermosphere, particularly in the representation of neutral composition and dynamics. Moreover, the baseline configuration lacked sufficient flexibility for the assimilated data to effectively correct these biases.

The optimization procedure significantly reduced the occurrence of G-condition, as shown in Figure 3. Although short intervals remain in the optimized run, they are less frequent and no longer produce the prolonged collapses present in the default, suggesting a more realistic background thermosphere that maintains the correct F-region structure. This improvement results from increased model flexibility, broader F10.7 perturbations generating a wider range of I-T states, and improved uncertainty characterization between the background model and assimilated data, reducing erratic behavior. Across all ionosonde observation times, simulated G-condition occurrence decreased from 14.00% (370/2642) in the default settings to 1.85% (49/2642) in the optimized settings, closely matching the observed occurrence rate of 1.48% (39/2642). Although the default configuration detected more observed cases (22 versus 7), it also produced many false positives, even in periods with a well-defined F2 region. Conversely, the optimized configuration significantly reduced these unrealistic occurrences, resulting in a more stable and consistent hmF2 representation.

Figure 4 summarizes relative model performance across all ionosonde stations for July 2017 using a modified Taylor diagram [10], comparing correlation, standard deviation, error standard deviation, and bias for foF2 and hmF2. Across both parameters, optimized AENEAS (A_Opt) outperformed the default (A_Def) in nearly all metrics, demonstrating the effectiveness of the optimization procedure. The only exception was hmF2 bias, where the neutral bias in A_Def resulted from compensating periods of over- and underestimation caused by large fluctuations; when absolute errors were considered, A_Opt showed a clear reduction in hmF2 bias. For foF2, the empirical models (IRI, NeQuick) performed well during quiet conditions but lacked storm-time responsiveness. WACCM-X captured storm-time structure more effectively, but A_Opt provided the most consistent representation of event-driven behavior, outperforming the other models in nearly all metrics. One remaining limitation is the moderate negative foF2 bias in A_Opt, inherited from the TIE-GCM background, which underestimated foF2 throughout. Although the optimized configuration reduced this bias relative to A_Def, a substantial portion of the assimilation effort was required to correct the background state, limiting AENEAS's ability to resolve smaller-scale structure.

For hmF2, the empirical models again performed well during quiet intervals but lacked sensitivity to storm-time dynamics. Physics-based models (TIE-GCM, WACCM-X) captured this variability more effectively but showed lower overall accuracy and correlation than the IRI. A_Opt achieved the highest correlation with observations and reduced hmF2 error relative to the other models but introduced a moderate positive bias absent in the standalone TIE-GCM run. This suggests that the bias was introduced through the assimilation process and likely reflects a limitation of TEC assimilation. Although TEC effectively constrains the magnitude of electron density and improves foF2 estimates, it is an integrated quantity and provides only limited information on vertical structure, especially in data-sparse regions with few intersecting ray paths. Consequently, the LETKF must

infer altitude-dependent corrections through ensemble covariance relationships, which may preferentially distribute electron density increments above the F2 peak, shifting the peak altitude upward and producing the observed positive hmF2 bias.

5. Conclusion

This study demonstrated that targeted optimization of key AENEAS parameters significantly improved model behavior and performance. By refining observation handling and ensemble control parameters, the optimized configuration mitigates the overconfident and erratic behavior seen in the baseline system and significantly reduced the occurrence of unphysical G-condition profiles during the July 2017 interval, from 14.00% to 1.85%, approaching the observed rate of 1.48%. The improvements identified in the February 2025 optimization tests remained robust when applied to the disturbed July 2017 storm period. When evaluated against global ionosonde observations, the optimized AENEAS configuration outperforms both the default settings run and the standalone TIE-GCM run, reproducing storm-time variations in foF2 and hmF2 more consistently than these and other established upper atmosphere models. Some limitations remain, including a persistent quiet-time negative foF2 bias inherited from the TIE-GCM background, while a positive hmF2 bias arises from TEC-only assimilation. In addition, reliance solely on GNSS TEC observations leads to reduced coverage over oceanic and equatorial regions, where ground-based receiver density is sparse. These results therefore highlight the need to incorporate additional assimilated datasets, such as in-situ and radio occultation measurements, to improve spatial coverage, better constrain vertical structure, and enhance model accuracy. Overall, this work provides a strong foundation for broader validation of the optimized system across longer time periods, a wider range of solar and geomagnetic conditions, and further geographic locations. The results strengthen confidence in AENEAS for reanalysis applications and demonstrate clear potential for integration into the operational framework.

6. References

1. M. Hapgood, M. J. Angling, G. Attrill, M. Bisi, P. S. Cannon, et al., "Development of Space Weather Reasonable Worst-Case Scenarios for the UK National Risk Assessment," *Space Weather*, **19**, 4, 2021, p. e2020SW002593.
2. A. D. Richmond, E. C. Ridley, and R. G. Roble, "A Thermosphere/Ionosphere General Circulation Model with Coupled Electrodynamics," *Geophysical Research Letters*, **19**, 6, 1992, pp. 601-604.
3. S. Elvidge and M. J. Angling, "Using the Local Ensemble Transform Kalman Filter for Upper Atmospheric Modelling," *Journal of Space Weather and Space Climate*, **9**, 2019.
4. B. Nava, P. Coïsson, and S. M. Radicella, "A New Version of the NeQuick Ionosphere Electron Density Model," *Journal of Atmospheric and Solar-Terrestrial Physics*, **70**, 15, 2008, pp. 1856-1862.
5. A. Coster, "Data from the CEDAR Madrigal database. MIT/Haystack Observatory," http://millstonehill.haystack.mit.edu/list (Accessed: 30 May 2025).
6. J.E. Matheson and R. L. Winkler, "Scoring Rules for Continuous Probability Distributions," *Management Science*, **22**, 1976, pp.1087-1096.
7. K. Rawer, D. Bilitza, and S. Ramakrishnan, "Goals and Status of the International Reference Ionosphere," *Reviews of Geophysics*, **16**, 2, 1978, pp. 171-181.
8. H. L. Liu, C. G. Bardeen, B. T. Foster, P. Lauritzen, J. Liu, et al., "Development and Validation of the Whole Atmosphere Community Climate Model with Thermosphere and Ionosphere Extension (WACCM-X 2.0)," *Journal of Advances in Modeling Earth Systems*, **10**, 2, 2018, pp. 381-402.
9. V. Lobzin and A. Pavlov, "G Condition in the F2 Region Peak Electron Density: A Statistical Study," *In Annales Geophysicae*, **20**, 2002, pp. 523-537.
10. S. Elvidge, M. J. Angling, B. Nava, "On the Use of Modified Taylor Diagrams to Compare Ionospheric Assimilation Models," *Radio Science*, **49**, 9, 2014, pp. 737-745.