https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

A New Approach to Inversion of Multi-spectral Data with Applications to FUV Remote Sensing

Matthew LeDuc[1], Tomoko Matsuo[2,1], and William Kleiber[1]

[1]Department of Applied Mathematics, University of Colorado Boulder, Boulder, CO, 80309
[2]Ann & H.J. Smead Department of Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, CO, 80309

Correspondence: Matthew LeDuc (matthew.leduc@colorado.edu)

Abstract. Many atmospheric measurement techniques involve inversion of photon counts detected by multi-spectral sensors spanning the X-ray to microwave regions of the electromagnetic spectrum. Although photon counts follow Poisson statistics, commonly used inversion techniques often rely on statistical assumptions that disregard the Poisson nature of the sensor data, limiting the scientific utility of datasets. Motivated to overcome this limiting assumption, this study focuses on retrieval techniques that involve the ratio of counts received in different sub-bands and introduces a new computationally efficient and robust approach to this type of inverse problem that respects the underlying count statistics. The method assumes that the received photon counts in each channel are a realization of a binned point process, allowing the ratio of the channel intensities to be modeled within a hierarchical Bayesian framework. This allows us to directly incorporate correlation between the bins via the prior that is modeled using a permanetal process. It further enables more accurate uncertainty quantification without costly sampling procedures common in Bayesian inversion methods. The method is verified and validated on thermospheric neutral temperature retrievals from simulated top-of-atmosphere FUV disk emission data corresponding to 2-8 November 2018, which includes a minor geomagnetic storm. It is also demonstrated on calibrated photon counts data from the NASA Global-scale Observations of the Limb and Disk (GOLD) mission from the same time period and from 11 May 2024 during a severe geomagnetic storm. The study demonstrates the method's ability to accurately recover neutral temperature in a variety of geomagnetic conditions, attesting to its potential to extend the fidelity of neutral temperature retrievals over broader solar zenith angles than currently possible with existing techniques.

1 Introduction

In recent decades, far-ultraviolet (FUV) remote sensing has played a key role in advancing space physics by providing valuable measurements of space plasma and neutral species. In the upper atmosphere, these emissions are primarily caused by various physical processes such as photoionization, photoexcitation, charge exchange and recombination that depend on thermospheric composition and temperature and ionospheric plasma density. The fact that these FUV emissions are observable from space without contending with significant background emissions makes FUV remote sensing attractive for upper atmosphere research (Paxton et al., 2017). In particular, the Lyman-Birge-Hopfield (LBH) band, one of the most prominent emission features, has been used to retrieve the column-averaged thermospheric temperature and composition, which are important parameters for

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

space weather research due to their relationship to aerodynamic drag experienced by objects in low Earth orbit (e.g., Aksnes et al., 2006; Zhang et al., 2019; Cantrall and Matsuo, 2021; Eastes et al., 2025). Current inverse modeling techniques, however, often make inappropriate assumptions on the statistical nature of photon count data that are especially problematic when photon counts are low, such as when solar zenith angle (SZA) is large. Even when photon counts are moderate to high, inversion techniques assuming Gaussian statistics have been observed to introduce bias into retrievals (e.g., Humphrey et al., 2009; Yamada et al., 2019; Nicolaou et al., 2024). Additionally, commonly used inversion techniques do not take into account the spatial structure of the photon count sensor data, limiting their ability to recover the underlying spatially correlated physical parameters.

One class of inversion methods widely used to retrieve neutral temperature and composition from LBH band emission is based on ratios of total observed counts in sub-bands of the received spectra. For example, portions of the LBH band corresponding to different transitions of molecular nitrogen are correlated with the changes in ambient temperature (e.g., Aksnes et al., 2006), and the ratios of these channels have been shown to have an approximately linear relationship with temperature (e.g., Cantrall and Matsuo, 2021; Zhang et al., 2019). Related techniques are widely used for $\Sigma O/N_2$ and QEUV FUV data products (e.g., Strickland et al., 1995; Zhang et al., 2004; Meier, 2021; Correira et al., 2021) as well as retrievals of ionospheric structure from FUV nightglow (Yin et al., 2023). Other examples include LiDAR profiling (Jia and Yi, 2014), spectrometry (Coath et al., 2013), x-ray astronomy (Park et al., 2006; Jin et al., 2006; Wang et al., 2024), astrochemistry (Boersma et al., 2012), and synthetic aperture radar (Gallardo i Peres et al., 2024). The relative simplicity of the two-channel ratio inversion techniques makes them amenable to integration of spatial models.

Following the work by Jin et al. (2006) and Park et al. (2006), we reformulate the problem of estimating the column-integrated neutral temperature from ratios of LBH channels as a hierarchical Bayesian inference problem. Specifically, the photon count data are used to infer the ratio of Poisson means, which we consider the geophysically relevant variable. This formulation allows us to rigorously characterize the effect of shot noise on the estimate of channel means, leading to a better understanding of the uncertainties in the estimated temperatures due to the shot noise. It is particularly relevant to upper atmosphere FUV remote sensing as shot noise effects are considered to be the largest source of uncertainty in the neutral temperature retrieval algorithm used for GOLD mission data product (Evans et al., 2024b). Additionally, this new formulation facilitates explicitly introducing spatial structure into the estimation by treating the received photon count data as a realization of a Poisson point process (Streit, 2010).

Among a variety of methods proposed to estimate the Poisson intensity from point process data, including the work by Nowak and Kolaczyk (2000); Antoniadis and Bigot (2006); De Oliveira (2013), we choose to model the intensity with a permanental process model (McCullagh and Møller, 2006), allowing us to use the theory of reproducing kernel Hilbert spaces to recover the intensity and perform uncertainty quantification (Flaxman et al., 2017; Walder and Bishop, 2017). Once the ratio of Poisson means Z is estimated, the effective neutral temperature is then recovered using the approximate linear relationship (Cantrall and Matsuo, 2021)

$$Z \approx mT_{eff} + z_0 \tag{1}$$

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

where T_{eff} is the effective neutral temperature and m and z_0 are slope and intercept parameters that depend on the specific LBH band and the observation geometry. These parameters are fitted using the Budzien vibrational-rotational model (Budzien et al., 2001). The method is verified and validated with simulated GOLD disk emission photon count data generated using an instrument simulator (Cantrall, 2022) and the NOAA Whole Atmosphere Model (WAM, Akmaev (2011)) for 2-8 November 2018. The method is furthermore demonstrated on calibrated, geolocated photon count data collected by GOLD for the same time period, as well as data from the Gannon storm in May 2024 (Grandin et al., 2024). These examples show that the proposed new approach is able to accurately retrieve the column-integrated temperature in a wide variety of geophysical conditions, both during geomagnetically calm and severely disrupted periods, and attest to the potential for extending temperature retrievals based on the LBH band emission to higher SZAs than currently considered. In addition, the method yields robust results with a computational cost that is feasible as an operational algorithm. Specifically, it achieves a full characterization of the posterior distribution of the neutral temperature without computationally intense sampling procedures that are common in Bayesian inversion methods. The method is implemented in an R package (R Core Team, 2025) that is publicly available.

In Sect. 2 we introduce the statistical modeling framework for temperature inference. In Sect. 3 and Sect. 4, we illustrate the new inversion procedures on simulated and actual GOLD disk emission data, respectively. Section 5 contains discussion of the computational performance, as well as directions for future research. Lastly, in Sect. 6 we provide the conclusion.

2 Hierarchical Bayesian Inverse Method

2.1 Building on Previous Work: The Forward Model

We seek to estimate the column-integrated thermospheric temperature from the top-of-atmosphere LBH band emission sensed by the GOLD instrument. Specifically, photon counts from radiometrically calibrated, geolocated L1C GOLD science data products are considered (McClintock et al., 2020). The column-integrated (effective) temperature T_{eff} can be given as a function of the wavelength λ as

$$T_{eff}(\lambda) = \frac{\int_0^\infty V(s,\lambda)\exp(-\tau(s,\lambda))T_n(s)\mathrm{d}s}{\int_0^\infty V(s,\lambda)\exp(-\tau(s,\lambda))\mathrm{d}s} \tag{2}$$

where s (cm) is the slant path distance, $V(\text{photons} \times \text{s}^{-1}\text{cm}^{-3})$ is the LBH volume emission rate, τ the optical depth, and T_{eff} (K) the neutral temperature (Evans et al., 2024b). This represents a weighted average of the neutral temperature along the line of sight, with the weight determined by the product $V(s,\lambda)\exp(-\tau(s,\lambda))$. As shown in Figure 4 of Evans et al. (2024b), the most weight is assigned to altitudes around 120-200 km. Since these observations are integrated quantities, they cannot be assigned to a specific altitude without a-priori knowledge of the temperature structure (e.g., Cantrall et al., 2019; Cantrall and Matsuo, 2021).

The forward modeling of the neutral temperature is based on the vibrational-rotational model (Budzien et al., 2001), which is essentially a lookup table of emission intensities for vibrational populations of molecular nitrogen as a function of temperature. This is convolved with the spectral point-spread function of the instrument to approximate the intensities as seen by the detector.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

Principal component analysis of simulated GOLD data has shown that the intensity in the upper portion of the LBH (2,0) band, from wavelength 138.56-139.2 nm, is positively correlated with temperature, while the intensity in the lower portion from 138-138.56 nm is negatively correlated with temperature (Cantrall and Matsuo, 2021). A similar structure is shown to hold in the (1,1) and (2,3) bands as well (Cantrall et al., 2019). So, as the column-integrated temperature increases, the observed photon counts in the upper portion of the band increases while the observed counts in the lower portion of the band decrease. This leads to the approximately linear relationship between the ratio of the long wavelength to short wavelength portions of the band and the neutral temperature (Cantrall and Matsuo, 2021), which is used in the study. As noted in Zhang et al. (2019) and Cantrall and Matsuo (2021), this two-channel ratio approach has several advantages. The most prominent is that determining the slope and intercept parameters in Eq. (1) does not require knowing the exact brightness of the band, meaning a fully calibrated instrument model is not required.

2.2 Modeling the Intensity Ratio

The first step is to model the two-channel intensity ratio as a random process. To motivate the choice of model for the process, we first consider the case where the inversion is done pointwise, as in Park et al. (2006) and Jin et al. (2006). If we assume that the count data from the channels are independently Poisson distributed with different mean parameters Λ_a and Λ_b, then the posterior distribution of the ratio $Z = \Lambda_a/\Lambda_b$ is a generalized Beta-Prime distribution $BP(\alpha, \beta, p, q)$, which has a density of the form

$$p_Z(z) \propto \frac{z^{\alpha p - 1}}{(1 + (z/q)^p)^{\alpha+\beta}} \tag{3}$$

if $z > 0$, and 0 elsewhere, where $\alpha > 0, \beta > 0, p > 0$ and $q > 0$ are parameters. A proof of this is given in Appendix A, and further discussion of this distribution along with applications can be found in McDonald (1987) and McDonald and Xu (1995).

We extend this to incorporate spatial information by utilizing Poisson point processes (Streit, 2010). A Poisson point process is a random measure defined on a space S, which for our application becomes a spherical cap domain, such that, for all measurable subsets R of S, we have that $N(R)$ the number of points in R is Poisson distributed with

$$\mathbb{E}\left(N(R)\right) = \Lambda(R) = \int_R \lambda(\boldsymbol{s}) \mathrm{d}w(\boldsymbol{s})$$

for some intensity $\lambda(\boldsymbol{s})$, where $w(\boldsymbol{s})$ is the ambient measure on S. These are extended by the Cox process model, where $\lambda(\boldsymbol{s})$ is itself a random function (Cox, 1955). We are interested in a special kind of Cox process known as the permanental process, where the assumption is that $\lambda(\boldsymbol{s}) = \frac{1}{2} c f(\boldsymbol{s})^2$, with $f(\boldsymbol{s})$ a Gaussian process (McCullagh and Møller, 2006; Møller, 2005). The estimation of $\lambda(\boldsymbol{s})$ then becomes a problem of estimating the latent process $f(\boldsymbol{s})$. In our specific case, the underlying point process has been binned, losing knowledge of the emission latitude and longitude. For this reason, instead of estimating $\lambda(\boldsymbol{s})$ we estimate $\Lambda(R_i)$ for each bin. However, our model can be generalized to the model of Walder and Bishop (2017) if, instead of binned data, the raw point process data are available.

Consider a region S divided into disjoint bins R_i such that $S = \bigcup_{i=1}^{d} R_i$. Given a point process on S, our data become pairs of counts a_i per bin R_i, $i = 1, \ldots, d$. We assume each a_i is Poisson distributed with intensity $\Lambda(R_i) = \frac{c}{2} f_i^2$ and attempt to

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

recover the latent vector $\boldsymbol{f} = (f_1, f_2, \ldots, f_d)^T$. The log-likelihood of $\boldsymbol{f}$ is given by

$$\ell(\boldsymbol{f}|\{a_i\}_{i=1}^d) = \sum_{i=1}^{d} a_i \log\left(\frac{c}{2} f_i^2\right) - \frac{c}{2}\langle \boldsymbol{f}, \boldsymbol{f}\rangle. \tag{4}$$

where $\langle\cdot,\cdot\rangle$ is the standard inner product on $\mathbb{R}^d$. Adding the prior $\boldsymbol{f} \sim \mathcal{N}(0, \frac{1}{\gamma}\mathbf{K})$ for some positive definite Hermitian matrix $\mathbf{K}$, we have that the log-posterior is

$$\ell(\boldsymbol{f}|\{a_i\}_{i=1}^d) = \sum_{i=1}^{d} a_i \log\left(\frac{c}{2} f_i^2\right) - \frac{c}{2}\langle \boldsymbol{f}, \boldsymbol{f}\rangle - \frac{\gamma}{2}\langle \boldsymbol{f}, \mathbf{K}^{-1}\boldsymbol{f}\rangle. \tag{5}$$

Since $\mathbf{K}$ is a covariance matrix, it can be written as $\mathbf{K} = \boldsymbol{\Phi}\mathbf{H}\boldsymbol{\Phi}^T$ with $\boldsymbol{\Phi}_{ij}$ the value of the j^{th} eigenvector at location i and $\mathbf{H} = \mathrm{diag}(\eta_i)$, with η_i the eigenvalues of $\mathbf{K}$. Then the inner products can be combined to generate an equivalent norm as

$$\frac{c}{2}\langle \boldsymbol{f}, \boldsymbol{f}\rangle + \frac{\gamma}{2}\langle \boldsymbol{f}, \mathbf{K}^{-1}\boldsymbol{f}\rangle = \frac{1}{2}\langle \boldsymbol{f}, \left(c\mathbf{I} + \gamma\mathbf{K}^{-1}\right)\boldsymbol{f}\rangle. \tag{6}$$

This new norm defines a norm on an equivalent kernel space with kernel $\tilde{\mathbf{K}}$ which we can write as

$$\begin{aligned}\tilde{\mathbf{K}} &= \boldsymbol{\Phi}\tilde{\mathbf{H}}\boldsymbol{\Phi}^T,\\ \tilde{\mathbf{H}} &= \mathrm{diag}\left(\frac{\eta_i}{c\eta_i + \gamma}\right).\end{aligned} \tag{7}$$

As in the infinite-dimensional case, we can say that the solution has the form $\hat{\boldsymbol{f}} = \tilde{\mathbf{K}}\hat{\boldsymbol{\alpha}}$ for some coefficients $\boldsymbol{\alpha}$[1]. This leads to the likelihood

$$\ell(\boldsymbol{\alpha}|\{a_i\}_{i=1}^d) = \sum_{i=1}^{d} a_i \log\left(\frac{c}{2}\left(\tilde{\mathbf{K}}\boldsymbol{\alpha}\right)_i^2\right) - \frac{1}{2}\boldsymbol{\alpha}^T\tilde{\mathbf{K}}\boldsymbol{\alpha} \tag{8}$$

with $\left(\tilde{\mathbf{K}}\boldsymbol{\alpha}\right)_i$ the i^{th} entry of $\tilde{\mathbf{K}}\boldsymbol{\alpha}$. This is the form of the log-posterior distribution of $\boldsymbol{\alpha}$ with log-likelihood function given by the first term in Eq. (8) and prior distribution $\boldsymbol{\alpha} \sim \mathcal{N}\left(0, \tilde{\mathbf{K}}^{-1}\right)$.

Using a Laplace approximation (Rasmussen and Williams, 2005) for the posterior, we find that the predictive mean is given by $\hat{\boldsymbol{f}} = \tilde{\mathbf{K}}\hat{\boldsymbol{\alpha}}$, where $\hat{\boldsymbol{\alpha}}$ is the maximum a-posteriori (MAP) estimate, and the predictive covariance of $\boldsymbol{f}$ is approximated by

$$\begin{aligned}\hat{\boldsymbol{\Sigma}} &= \tilde{\mathbf{K}} - \tilde{\mathbf{K}}\left(\mathbf{D}^{-1} + \tilde{\mathbf{K}}\right)^{-1}\tilde{\mathbf{K}},\\ \mathbf{D} &= \mathrm{diag}\left(\frac{\hat{\alpha}_i^2}{2a_i}\right).\end{aligned} \tag{9}$$

These results are derived in Appendix B. Then the distribution of $\hat{\Lambda}(R_i) = \frac{c}{2}\hat{f}_i^2$ given the data is well-approximated by

$$\hat{\Lambda}(R_i) \sim \Gamma(\hat{\alpha}, \hat{\beta}), \quad \text{with} \quad \hat{\alpha} = \frac{\left(\mu^2 + \sigma^2\right)^2}{2\sigma^2\left(2\mu^2 + \sigma^2\right)}, \quad \hat{\beta} = \frac{\mu^2 + \sigma^2}{\sigma^2 c\left(2\mu^2 + \sigma^2\right)} \tag{10}$$

[1]This is not strictly necessary for the problem we are solving of recovering $\boldsymbol{f}$ in a binned point process, however it is critical for the feasiblity of the method in recovering a random function, which is the case with unbinned point process data. This allows the estimation to be reduced to a finite dimensional problem.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

where μ and σ are the posterior mean and standard deviation of the estimate of $\hat{f}_i$ from the Laplace approximation and $\hat{\alpha}$ and $\hat{\beta}$ are the shape and rate parameters of the gamma distribution (analogous to Walder and Bishop (2017) Sect. 4.1.5, which uses the shape/scale parameterization).

With the above derivations, we can proceed as in the previous section to determine the parameters of the distribution of Z at a given spatial location. Once we know that $\hat{\Lambda}_a \sim \Gamma(\hat{\alpha}_a, \hat{\beta}_a)$ and $\hat{\Lambda}_b \sim \Gamma(\hat{\alpha}_b, \hat{\beta}_b)$ we have

$$Z = \frac{\hat{\Lambda}_a}{\hat{\Lambda}_b} \sim BP\left(\hat{\alpha}_a, \hat{\alpha}_b, 1, q\right), \tag{11}$$

where $q = \hat{\beta}_b / \hat{\beta}_a$, analogous to the pointwise model in Eq. (3). Due to the construction of the intensities as spatial processes, the parameters of the marginal posteriors are calculated using spatial information and thus endow spatial structure on the resulting temperature field.

2.3 Modeling the Temperature

From Eq. (1) we know that $Z \approx mT_{eff} + z_0$. Since $Z \sim BP\left(\hat{\alpha}_a, \hat{\alpha}_b, 1, q\right)$, the temperature $T_{eff} = \frac{1}{m}\left(Z - z_0\right)$ has the distribution and MAP estimate

$$\begin{aligned} T_{eff} &\sim -\frac{z_0}{m} + BP\left(\hat{\alpha}_a, \hat{\alpha}_b, 1, \frac{q}{m}\right) \\ \hat{T}_{eff} &= -\frac{z_0}{m} + \begin{cases} \frac{q}{m}\frac{\hat{\alpha}_a - 1}{\hat{\alpha}_b + 1}, & \hat{\alpha}_a > 1 \\ 0, & else \end{cases}. \end{aligned} \tag{12}$$

This allows estimation of the posterior of the neutral temperature without any sampling, leading to fast inference once the intensities are known. It also allows extension to problems where $Z \approx \left(mT_{eff} + z_0\right)^p$ for some exponent p, in which case $T_{eff} \sim -\frac{z_0}{m} + BP\left(\hat{\alpha}_a, \hat{\alpha}_b, p, \frac{1}{m} q^{1/p}\right)$.

3 Application to Simulated Data

To verify and validate the new approach, we apply the method developed in Sect. 2 to simulated GOLD LBH disk emission data generated using the NOAA National Weather Service Whole Atmosphere Model (WAM, Akmaev (2011)). The simulation study corresponds to Nov 2-8 2018. Synthetic GOLD emission data are generated as follows. WAM simulations conducted with realistic solar and magnetosphere forcing are used as input to the Global Airglow Model (GLOW, Solomon (2017)) which calculates the volume emission rates. These volume emission rates are passed into a GOLD instrument simulator developed in Cantrall (2022), which returns the slant column brightness (in units of counts per Angstrom) at each location on the detector. These brightnesses are convolved with the GOLD instrument point spread function to account for instrument effects and then used to simulate Poisson-distributed photon counts to include the effects of shot noise. Finally, these data are binned spatially, as done in the GOLD mission TDISK algorithm, to 250x250 km^2 resolution at satellite nadir, corresponding to about 1500 bins (Evans et al., 2024b). The emission location is assigned to the middle of the combined spatial bins. The column-integrated

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

effective temperatures directly calculated from WAM temperature fields using Eq. (2) in the same resolution serves as the ground truth.

Each retrieval takes approximately 40 seconds and uses less than 2GB of memory on a desktop computer with 16 GB RAM and an Intel i7 CPU. Most of the computing time is spent solving the optimization problem in Eq. (8) twice, although for large problems generating the equivalent kernel matrix is also costly. Because we have a closed form approximation to the posterior given by Eq. (12), we are able to avoid costly sampling algorithms that are common in other Bayesian retrievals, making the computations relatively efficient on these datasets. See Sect. 5.1 for more discussion on computational performance.

3.1 Choice of Kernel

Since the process we are interested in is observed on a section of a spherical shell by GOLD, we choose a kernel given by

$$k(\boldsymbol{s}_1, \boldsymbol{s}_2) = \sum_{l,m} \frac{1}{l^\nu (1+l)^\nu} V_{lm}(\boldsymbol{s}_1) V_{lm}(\boldsymbol{s}_2) \tag{13}$$

where the V_{lm} are spherical cap harmonics (SCHAs), which are eigenfunctions of the Laplacian on the spherical cap (Haines, 1985). Then the kernel matrix is given by $\mathbf{K}_{i,j} = k(\boldsymbol{s}_i, \boldsymbol{s}_j)$ and $\boldsymbol{s}_i$ is the reference latitude and longitude of R_i. These functions are well-studied in geophysics, especially in inverse problems that involve estimating the gradient of a process from incomplete measurements, where only a small portion of the globe can be observed at a time (e.g., Haines, 1988; Richmond and Kamide, 1988; Hwang and Chen, 1997). The parameter ν, which we call the smoothness parameter, controls the decay of the coefficients in the eigenvector expansion, leading to smoother fields with suppressed high-frequency variation when ν is large. In the context of random function estimation, for example when the data come from a realization of an unbinned point process, $f(\boldsymbol{s})$ is in $H^\nu(S)$ the Sobolev space of order ν when $\nu > 1$, and the Sobolev embedding theorem implies that $f(\boldsymbol{s})$ has n derivatives if $\nu > 1+n$ (Hunter and Nachtergaele, 2020). For this reason, we focus our analysis on the cases $\nu = 0.5, 1+10^{-8}, 2+10^{-8}$, with the latter two corresponding to a continuous and differentiable random field, respectively. Further details on the construction of the spherical cap harmonics are included in Haines (1985) and Hwang and Chen (1997), and a more in depth discussion of the kernel along with a comparison with other common choices is included in Appendix C.

3.2 Model Evaluation

We evaluate the model performance primarily using the continuous rank probability score (CRPS). The CRPS is given by

$$CRPS(F|y) = \int_{\mathbb{R}} \left(F(x) - H(x-y)\right)^2 \mathrm{d}x \tag{14}$$

where F is the predictive cumulative distribution function (CDF) of the neutral temperature, y the true neutral temperature, and H the Heaviside step function. This is a strictly proper scoring rule, meaning its unique minimizer is the deterministic CDF $H(x-y)$ (Gneiting and Rafferty, 2007). Since it incorporates the entire distribution rather than just a point estimate it allows us to assess the quality of the posterior beyond the error in estimation. CRPS is negatively-oriented in that smaller values indicate superior predictive models.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

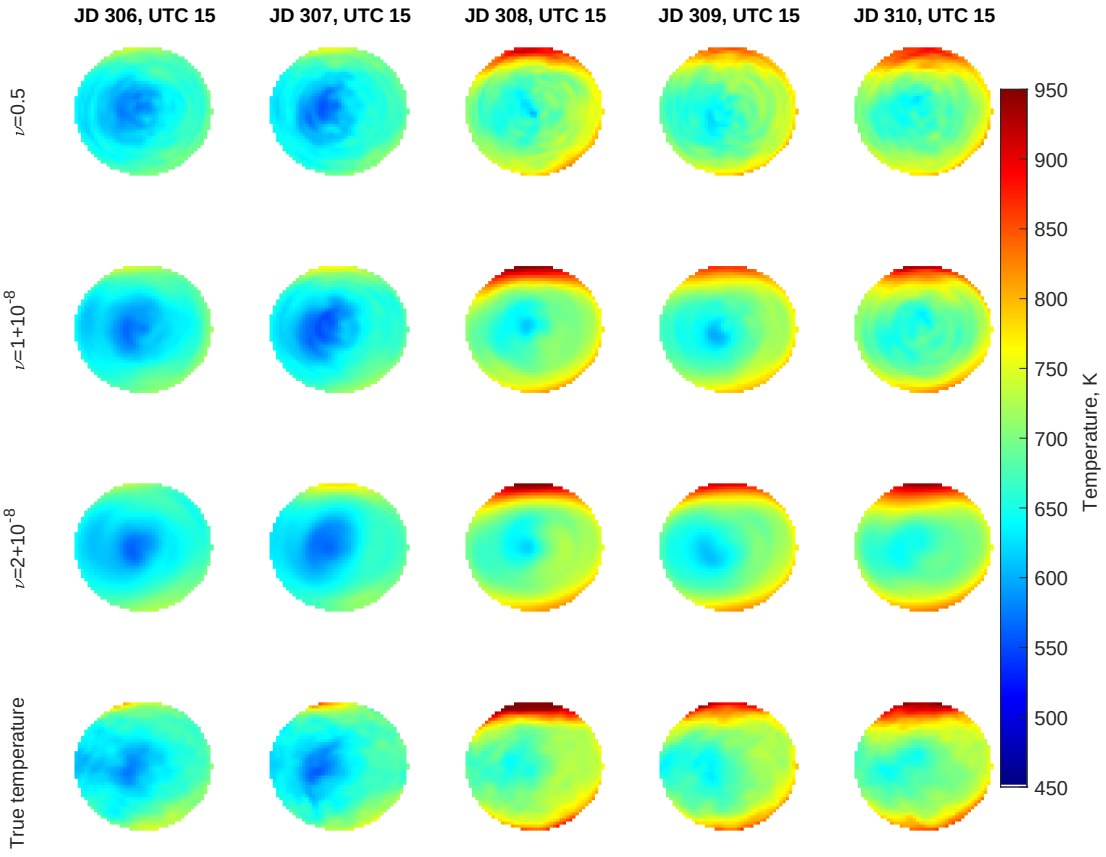

Figure 1. Estimated neutral temperature (first 3 rows) vs simulated field (last row), JD 306-310 at 15:00 UTC (Noon satellite local time) for smoothness parameters $\nu = 0.5$ (top row), $1 + 10^{-8}$ (2nd row), and $2 + 10^{-8}$ (3rd row). The algorithm is able to generate accurate estimates of the neutral temperature regardless of the smoothness parameter selected, although the parameter has a noticeable effect on the smoothness of the retrieved field.

As a first pass, we test how our model performs in estimating the neutral temperature at 15:00 UTC, or local noon at satellite nadir. At this time the largest portion of the instrument's viewing area is sunlit, meaning that the algorithm is expected to perform best in this situation. The retrieved fields from JD 306 - JD 310 are shown in Fig. 1, along with the "true" field from WAM simulations. The algorithm closely tracks the true temperature field for all cases, with variations in ν leading to smoother or rougher estimates of the underlying field.

In addition to performance at local noon, we are interested in the performance of our algorithm in low-light conditions. Currently, the GOLD mission discards data collected at an observing zenith angle (OZA) greater than 75° or SZA greater than 80° (Evans et al., 2024b). However, we expect that incorporating spatial information into the estimates can allow the limit on SZA to be pushed upward slightly. To investigate this, we processed data from 12:00 and 18:00 UTC as well. In this situation, a portion of the viewing area has a solar zenith angle higher than the nominal observation boundary of 80° used by the GOLD

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

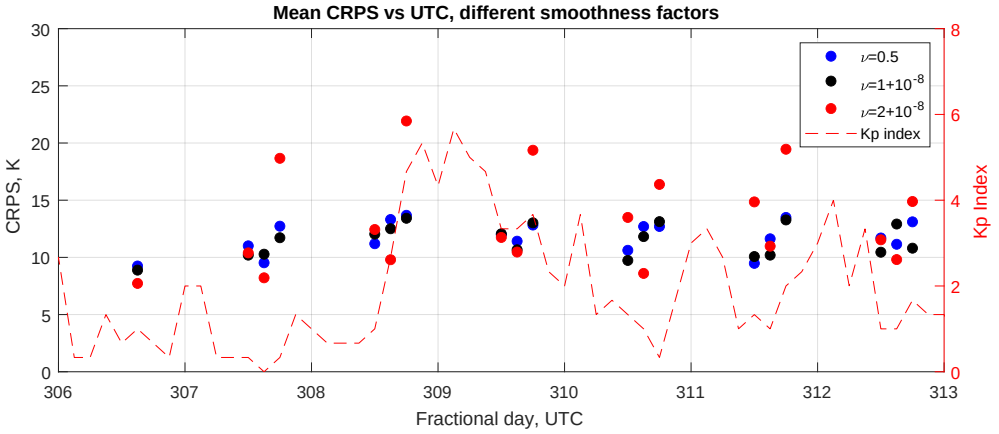

Figure 2. CRPS vs UTC for different kernel smoothness parameters, along with the K_p index from Matzka et al. (2021). The performance of the model depends strongly on local time and decreases slightly during storm time, especially in the case $\nu = 2 + 10^{-8}$.

mission, however there is still a large portion of the viewing area that is sunlit. Additionally, there are portions of the viewing area where the solar zenith angle exceeds 80° at local noon. These results for the average CRPS at each time are shown in Fig. 2 along with the K_p index. What we see is that the CRPS of the models for $\nu = 0.5$ and $\nu = 1 + 10^{-8}$ are roughly constant with time and show only a slight increase with the K_p index, never exceeding a CRPS of 15 K. However, the CRPS of the model with $\nu = 2 + 10^{-8}$ increases significantly, especially in the afternoon, and exceeds 20 K at 18:00 UTC on JD 308. This is because the onset of even this relatively minor storm generates structure in the field that the model is unable to capture.

Average CRPS over latitude, longitude, OZA, and SZA from the simulated data, is plotted in Figure 3. The CRPS is roughly constant in longitude, but varies with latitude, OZA, and SZA. To clarify whether SZA, OZA, or latitude is the primary driver of the increased CRPS, more case studies may be required. It is notable that these models maintain a CRPS less than 20 K up to solar zenith angles of roughly 90°, suggesting that the method can potentially help extend the scientific utility of FUV disk emission data at higher SZA than currently considered by the GOLD mission.

3.3 Uncertainty Quantification

Our model provides a full posterior distribution of the intensity ratio, and therefore the temperature. In this section, we examine how accurately our model captures uncertainty. This exercise is important to ensure that the model handles uncertainty induced by shot noise properly and that the estimates of the neutral temperature are reliable. We quantify the uncertainty captured by the model by determining the highest posterior density sets with coverage rate r ($r-$HPD set) for varying values of r. The $r-$HPD set is the set I_r such that $\int_{I_r} f(x)\mathrm{d}x = r$ and that $\forall \left(x \in I_r, y \in I_r^C\right), f(x) \geq f(y)$ (Amaral Turkman et al., 2019). If the posterior distribution accurately quantifies the uncertainty, then the true parameter lies in the $r-$HPD set with probability r. To determine how well the posterior HPD matches the nominal coverage for varying coverage parameters r, we compare the probability that the true value is within I_r for varying values of r from 0.05 to 0.95. These results are shown in Fig. 4 for the

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

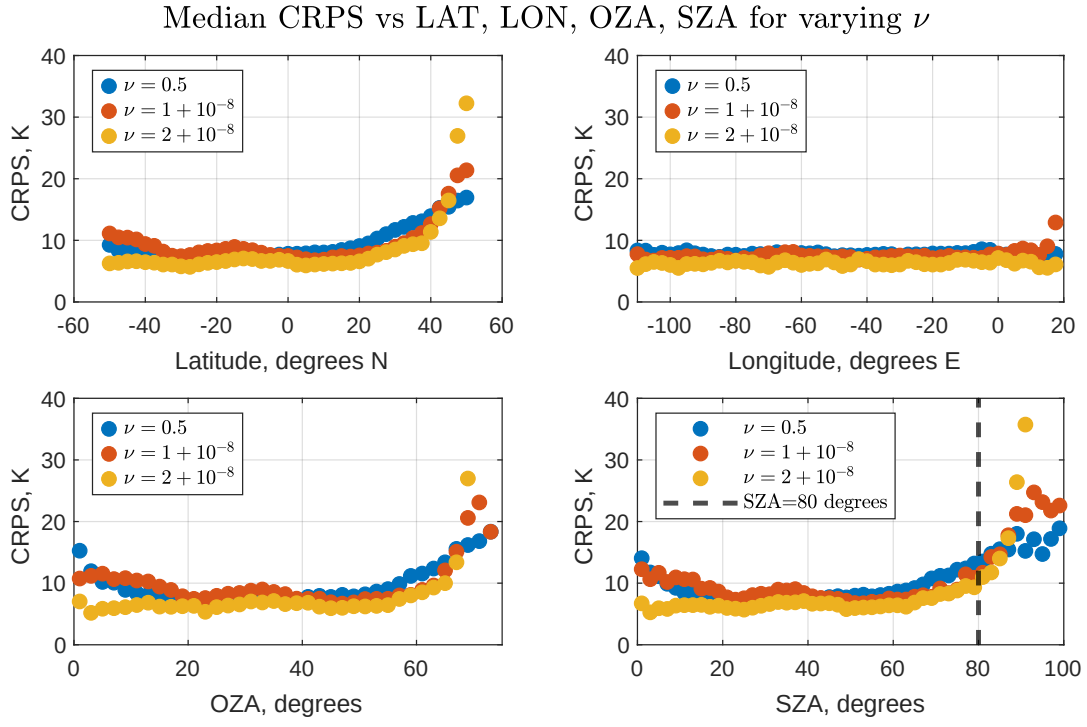

Figure 3. Mean CRPS for varying latitude, longituge, observing angle, and solar zenith angle for the simulated data study. All latitudes and longitudes are included when determining the mean CRPS as a function of latitude, observing angle, and solar zenith angle, while only latitudes between $\pm 10^{\circ}$ are considered when determining the CRPS as a function of longitude.

SCHA kernel with $\nu = 0.5, 1+10^{-8}, 2+10^{-8}$, and have been divided into times where $K_p < 3$, indicating calmer conditions, and $K_p \geq 3$, indicating higher levels of geomagnetic disturbance.

The coverage probabilities are estimated from 10 retrievals for each simulated field, and then compiled according to K_p index. The plots in Fig. 4 show the median coverage probability over all applicable times and spatial locations from the simulations, and error bars show the 10th and 90th percentiles of coverages. Generally, we see that the $\nu = 1+10^{-8}$ and $\nu = 2+10^{-8}$ models appear to do best at quantifying uncertainty, especially in geomagnetically calm conditions, and their performance degrades in geomagnetically active time periods. The $\nu = 2+10^{-8}$ model appears especially affected by this transition. This is due to the fact that setting ν higher makes the retrieved field smoother, so it cannot properly recover the smaller scale structures that appear during a storm. Regardless of the conditions, the model with $\nu = 0.5$ appears significantly overdispersed and unable to accurately quantify uncertainty. So, in addition to providing us with good predictive accuracy as seen in the previous section, selecting $\nu = 1+10^{-8}$ provides us with the best uncertainty quantification in all conditions.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

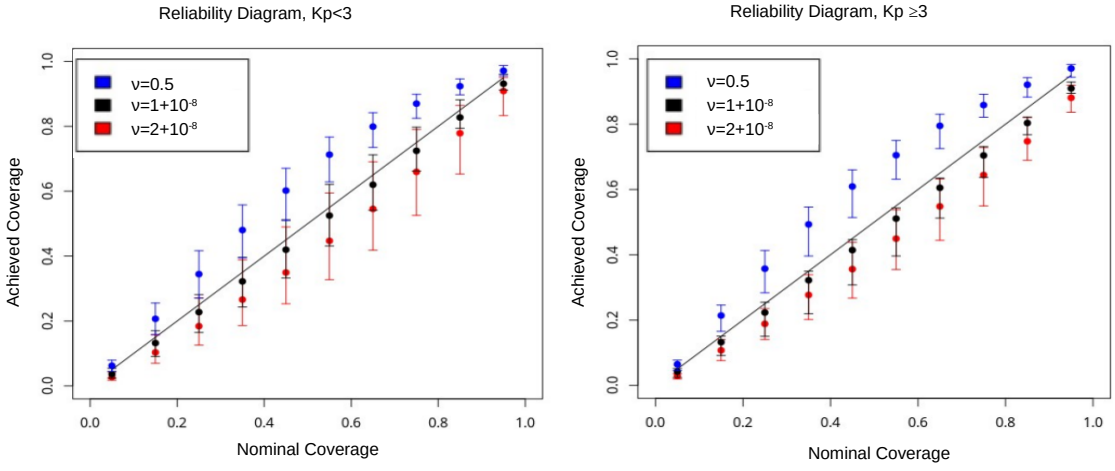

Figure 4. Reliability diagram for the model with varying smoothness parameters and K_p indices. Dots show median coverage probability of a highest-posterior density interval with nominal coverage probability along the x axis, error bars are 10th and 90th percentiles. The models perform worse at uncertainty quantification as K_p increases, however for calm conditions the $\nu = 1 + 10^{-8}$ model captures the uncertainty very well.

4 Application to GOLD Disk Emission Data

4.1 Nov 2018 Geomagnetic Storm

As a first demonstration of the method on real GOLD data, we apply it to GOLD L1C data from 2-8 November 2018. This is the same time period studied using WAM simulations in Sect. 3, and includes a minor geomagnetic storm on 4-5 November.

Since the GOLD instrument scans each hemisphere independently, the inversion incorporates two consecutive scans into each retrieval. The start times of these scans are separated in time by 12 minutes, which is short enough relative to the nominal thermospheric timescales to ignore the time difference between the scans. Since the scans overlap, we average the data in the overlap region, which reduces (but does not completely eliminate) the bias in the retrievals caused by the varying sensitivity of the detector near the equator (Evans et al., 2024b).

The results of this inversion are shown for 5 November (JD 309) in Fig. 5. We can see that the main features present in the GOLD TDISK data products are accurately recovered. There is a visible discrepancy near the equator, corresponding to the lower portion of the scan of the northern hemisphere and the upper portion of the scan of the southern hemisphere. Upon examination of the L1C data, we can see a corresponding enhancement in the count ratio at the same location that is present in the scan of the northern hemisphere, but not the southern hemisphere. This, coupled with the fact that similar artifacts are not seen in the results on simulated data (Fig. 1) suggests that this variation is due to the varying sensitivity of the detector across scans. In order to use this method, therefore, it is important to have a complete understanding of the variations in the detector

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

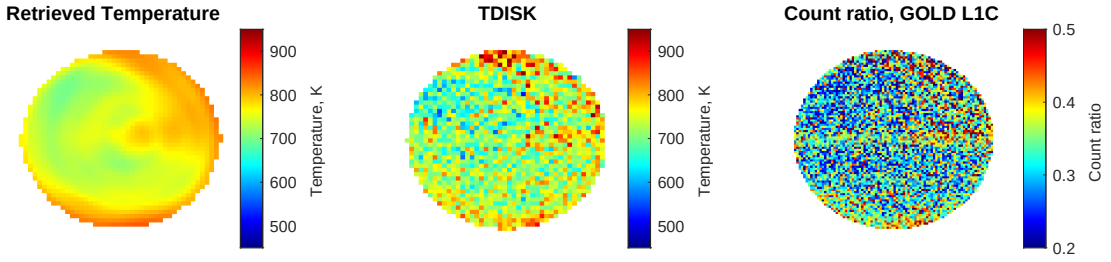

Figure 5. Retrieved temperature (left) compared to TDISK (center) and the ratio of counts in the upper and lower portion of the LBH (2,0) band (right) for JD 309, 2018 (5 November 2018) at around 1500 UTC ($K_p \approx 3.3$). The retrieved temperature shares the most prominent features of the TDISK retrieval, however it contains an additional feature along the equator. This feature appears to be inherited from the data, which shows elevated count ratios near the equator due to differences in the sensitivity of the instrument in the different scans of the equatorial region. Similar artifacts can be seen in older versions of the TDISK data.

sensitivity with look direction and eliminate artifacts in the retrievals. Similar artifacts can also be found in earlier versions of the TDISK data (e.g. Figure 7 of Cantrall and Matsuo (2021)).

4.2 May 2024 Geomagnetic Storm

The second demonstration example includes a major geomagnetic storm that occurred in May 2024 (also known as the Gannon Storm, Grandin et al. (2024)). With the Dst index reaching below -400 nT and a peak K_p index of 9, this storm resulted in a major expansion of the upper atmosphere causing the neutral temperature to elevate, leading to orbital decay of low Earth orbit satellites due to increased atmospheric drag.

The GOLD mission has detected previously unseen structure in the thermosphere during this event, simultaneously seen in both LBH and 135.6 nm radiances, as well as retrievals of neutral temperature, atmospheric composition, and total electron content. The detected equator-to-pole differences in neutral temperature of over 400K, with high-latitude temperatures in excess of 1400K (Evans et al., 2024a) are well beyond a typical FUV observation and retrieval scenario under nominal conditions. To investigate the performance of our algorithm with the corresponding GOLD L1C data under somewhat extreme conditions, we compare our results to the TDISK data product using both the SCHA kernel and a Wendland kernel function given by (Wendland, 1995)

$$K(d;r) = \begin{cases} \frac{1}{3}\left(1-\left|\frac{d}{r}\right|\right)^6\left(35\left|\frac{d}{r}\right|^2+18\left|\frac{d}{r}\right|+3\right), & |d|<r \\ 0, & |d|\geq r \end{cases} \tag{15}$$

Note that in the model model evaluation with simulated data for Novemeber 2018, the Wendland kernel with $r = 1$ is found to perform similarly to the SCHA kernel, albeit with slightly worse uncertainty quantification (see Appendix C for details). The

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

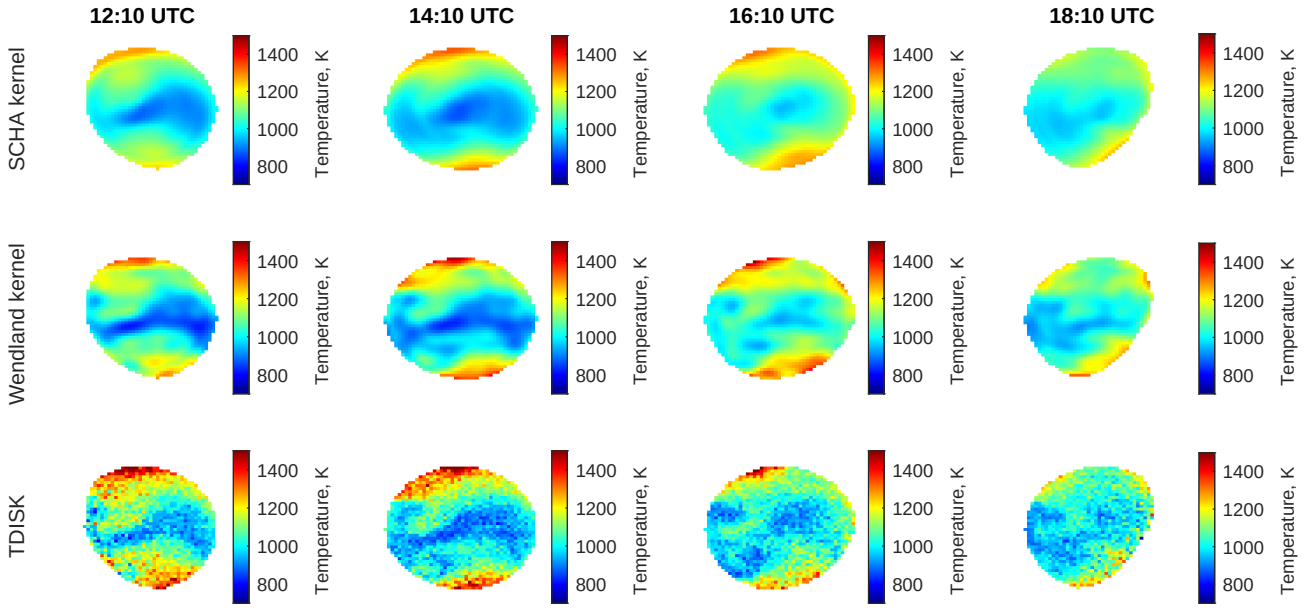

Figure 6. Retrieved temperatures for 12:10-18:10 UTC on 11 May 2024 compared with TDISK. The retrieved spatial structure generally resembles the TDISK feature, especially early in the day.

results of the retrieval for May 2024 are shown in Fig. 6 using $r = 1$ and $\nu = 1 + 10^{-8}$ in the Wendland and SCHA kernels respectively. We see that both models are able to recover most of the spatial structure in the field, especially earlier in the day. The retrievals can be improved by more careful selection of the parameters of the kernel, allowing capture of smaller scale structure such as the vortices that are apparent in TDISK at 16:10 and 18:10 UTC. These results also motivate future work for a retrieval method that does not need spatial binning. The spatial binning leads to a greatly increased SNR, however it ends up averaging out small-scale spatial structure generated by the storm that could be of scientific interest to the space physics community.

5 Discussion

5.1 Computational Performance

One of the drawbacks of Bayesian methods as practical retrieval algorithms is their computational load. Often, these methods require generating samples from the posterior distribution using an algorithm such as Markov Chain Monte Carlo, which can be computationally expensive. This work overcomes this problem with the use of a closed form posterior distribution, shifting the computational load to maximizing the likelihood in Eq. (8). Other options, such as a variational inference approach, are possible (e.g. Lloyd et al. (2015)); however, as shown in Walder and Bishop (2017), variational inference is much less efficient than the techniques described in this paper.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

The algorithm presented in the previous section is timed on a small runtime analysis study to recover the intensity function:

$$Z(x) = \frac{25\sin\left(\frac{\pi}{2}x\right)^2 + 10}{8\cos\left(\frac{\pi}{2}x\right)^2 + 10},\ -1 \leq x \leq 1 \tag{16}$$

a plot of which is shown in the upper portion of Fig. 7, using logarithmically spaced numbers of bins from 10 to 5600. A Wendland kernel with $r = 3/4$ as the kernel is used for the timing study. The minimization is performed with the implementation of conjugate gradient in the `optim` function in base `R` using the analytic gradient of Eq. (8) as an input.

With 1500 bins, the average time to retrieve the ratio function is about 40 seconds on a desktop computer with 16 GB RAM and an Intel i7 CPU. For problems involving less than 500 bins, the algorithm converges in an average of about 3 seconds, and in about 30 minutes for problems involving 5600 bins. This suggests that the model can be made fast enough to run in an operational capacity if desired, especially for problems involving data sets in the low to mid thousands of points or less. Due to the eigenvalue decomposition necessary to determine $\tilde{\mathbf{K}}$ the algorithm scales asymptotically as $O(n^3)$, which can be mitigated by using a low-rank decomposition rather than a full eigendecomposition or sparse kernel matrices. Additionally, if the kernel has an explicit Mercer expansion, the need for the eigenvalue decomposition can be avoided altogether. The rest of the algorithm does not require any matrix inversions and scales approximately as $O(n^2)$. For problems with 1500 bins the peak usage is approximately 2 GB, making it easily small enough to run on commercially available laptops.

5.2 Future Work

The LBH (1,1) and (2,3) bands are not used in this study in part because they are not considered to have sufficient SNR for temperature retrievals (Cantrall, 2022). However, both of these bands may be beneficial to consider in the future. While the (2,0) band is not isolated from other LBH emissions, with the lower portion of the band overlapping with the (5,2) band, the (1,1) and (2,3) bands are. This means that, in principle, retrievals using these bands are independent of the specified populations of molecular nitrogen, removing a significant source of model specification error. Although several works, such as Aryal et al. (2022), Ajello et al. (2020), and Nicholls (1961), have attempted to determine these populations, any errors in the relative populations of the $v' = 2$ and $v' = 5$ species will lead to errors in the temperature estimation. For this reason, extending these techniques in a way to account for the lower SNR of the (1,1) and (2,3) bands or developing a technique that incorporates uncertainty in the vibrational populations may be of interest to practitioners.

One drawback of the two channel ratio method is that it is not robust to wavelength specification errors (Cantrall and Matsuo, 2021; Eastes et al., 2025). This can be partially overcome by using multiple bands, such as the (1,1) and/or (2,3) bands in conjunction with the (2,0) band, to do estimation and weighting observations by their inverse variance, as observed in Cantrall (2022). Under our model the posterior distribution no longer takes on a closed form. ~~form,~~ being given by the Lauricella D function (Pham-Gia and Turkkan, 2002). One avenue of future research is to determine whether there are suitable approximations to these distributions that would allow incorporation of additional spectral information into the retrieval approach while maintaining efficient computational performance.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

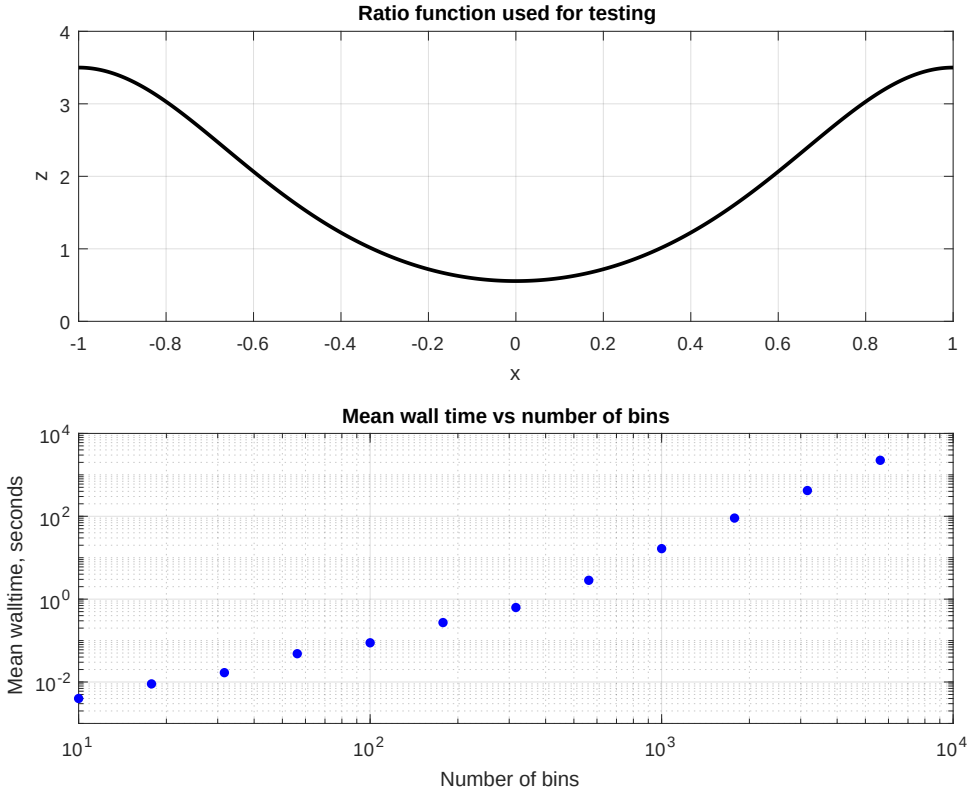

Figure 7. On the top: The ratio function $Z(x)$ estimated in the timing study. On the bottom: timing study results. Times shows are the mean wall times (in seconds) for the code to execute on a desktop computer over 5 repetitions.

While we have developed this method with the assumption that the data contain counts generated by a single process, that is not necessarily true. Photons from other sources inhibit retrieval, especially when the source counts are low, and the GOLD instrument's geostationary orbit means it can be subjected to high energy particles from outside sources (Evans et al., 2024b). While the GOLD mission L1C data product mitigates most of the effect of the background (McClintock et al., 2020), an algorithm that can account for it independently is desirable for further improvement. For instance, the model of Park et al. (2006) handled this by adding a separate term for the background count intensity and estimating both jointly. It is feasible to adopt such an approach to extend this work in the future.

6 Conclusions

While Poisson-distributed multi-spectral sensor data are common in many atmospheric remote sensing applications, current inversion methods used for retrieval of physical parameters often rely on statistical assumptions that disregard the real statistical properties of sensor data, limiting the scientific utility of datasets. To address this limitation, we introduce a retrieval approach

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

that is based on ratios of photon counts in non-overlapping spectral bands with a focus on FUV upper atmospheric remote sensing applications. This two-channel intensity ratio approach facilitates the development of a computationally efficient and robust method that respects the underlying statistics of the sensor data and naturally incorporates spatial structure. Specifically, the method uses a Poisson point process model with the intensity for each channel independently modeled with a squared Gaussian prior distribution. The new inverse modeling approach is shown to accurately recover spatial structure observed in the true fields and provide a complete posterior distribution of the physical parameter of interest that more faithfully captures variability of underlying geophysical process. The approach is demonstrated on thermospheric neutral temperature retrievals from simulated top-of-atmosphere FUV disk emission data during a minor geomagnetic storm period in November 2018, and from actual calibrated disk emission (L1C) data from the GOLD mission for the same period as well as a major geomagnetic storm that occurred in May 2024. We show that the retrieved column-integrated temperature generally agrees well with the GOLD mission (TDISK v5) data product, with the notable advantage of allowing retrieval of thermospheric neutral temperature at more extreme solar zenith angles than currently considered in the GOLD mission data product algorithms.

Code and data availability. The R package developed for this study is available on github at https://github.com/mfleduc/PoissonRatioUQ. GOLD data is available at https://gold.cs.ucf.edu/data/search/ courtesy of NASA/GOLD and the mission science team.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

Appendix A: Derivation of Eq. (3)

In Eq. (3), we present without proof the posterior distribution of the intensity ratio in the case of pointwise inversion with no background, as done in Jin et al. (2006) and Park et al. (2006), as a BP(α,β,p,q) distribution. This model contains as a special case priors of the form x^{-k_i}, equivalent to $\alpha_i = 1-k_i$ and $\beta_i = 0$. The proofs for the cases $n_a = n_b = 1$ with $k_a = k_b = 0$ and $k_a = k_b = 0.5$ are given in Jin et al. (2006), but the distributions are not named as the $BP(a+1-k_a, b+1-k_b, 1, 1)$ distribution.

Assume that the count data $\{a_i\}_{i=1}^{n_a}$ from the upper channel and $\{b_i\}_{i=1}^{n_b}$ from the lower channel are independent with

$$a_i \big| \Lambda_a \overset{\text{iid}}{\sim} \text{Poisson}(\Lambda_a)$$
$$b_i \big| \Lambda_b \overset{\text{iid}}{\sim} \text{Poisson}(\Lambda_b)$$

for all i, where n_a and n_b are the number of independent observations of the total counts in each channel. In the case of GOLD data where we do 2x2 spatial binning prior to processing, $n_a = n_b = 4$. Now consider the prior $\Lambda_i \sim \Gamma(\alpha_i, \beta_i)$ for each Λ_i. Let $a = \sum_{i=1}^{n_a} a_i$ and $b = \sum_{i=1}^{n_b} b_i$. Then since the Gamma distribution is the conjugate prior of Λ we have that $\Lambda_a \sim \Gamma(a+\alpha_a, n_a+\beta_a)$ and $\Lambda_b \sim \Gamma(b+\alpha_b, \beta_b+n_b)$.

So, the *intensity ratio* $Z = \frac{\Lambda_a}{\Lambda_b}$ has the distribution

$$\begin{aligned}
\mathbb{P}[Z \le z] = \mathbb{P}\left[\frac{\Lambda_a}{\Lambda_b} \le z\right] &= \int_0^\infty \mathbb{P}\left[\frac{\Lambda_a}{\Lambda_b} \le z | \Lambda_b = x\right] \mathrm{d}\mathbb{P}[\Lambda_b = x] \\
&= \int_0^\infty \mathbb{P}[\Lambda_a \le z\Lambda_b | \Lambda_b = x] \, \mathrm{d}\mathbb{P}[\Lambda_b = x] \\
= \frac{(\beta_a+n_a)^{a+\alpha_a}(\beta_b+n_b)^{b+\alpha_b}}{\Gamma(a+\alpha_a)\Gamma(b+\alpha_b)} &\int_0^\infty \left[\int_0^{zx} y^{a+\alpha_a-1} e^{-(\beta_a+n_a)y} \mathrm{d}y\right] x^{b+\alpha_b-1} e^{-(\beta_b+n_b)x} \mathrm{d}x
\end{aligned} \tag{A1}$$

The integrand is differentiable in z and is an $\mathcal{L}^1$ function in x, so by Lebesgue's Dominated Convergence Theorem we can pass derivatives into the integral. Then the probability density is given by

$$\begin{aligned}
p_Z(z) &= \frac{(\beta_a+n_a)^{a+\alpha_a}(\beta_b+n_b)^{b+\alpha_b}}{\Gamma(a+\alpha_a)\Gamma(b+\alpha_b)} \int_0^\infty \frac{\partial}{\partial z}\left[\int_0^{zx} y^{a+\alpha_a-1} e^{-(\beta_a+n_a)y} \mathrm{d}y\right] x^{b+\alpha_b-1} e^{-(\beta_b+n_b)x} \mathrm{d}x \\
&= \frac{(\beta_a+n_a)^{a+\alpha_a}(\beta_b+n_b)^{b+\alpha_b}}{\Gamma(a+\alpha_a)\Gamma(b+\alpha_b)} \int_0^\infty x^{a+b+\alpha_a+\alpha_b-1} e^{-(\beta_b+n_b+(\beta_a+n_a)z)x} \mathrm{d}x
\end{aligned} \tag{A2}$$

Letting $t = (\beta_b + n_b + (\beta_a+n_a)z)x$ we see that

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

$$p_Z(z) = \frac{(\beta_a+n_a)^{a+\alpha_a}(\beta_b+n_b)^{b+\alpha_b}}{\Gamma(a+\alpha_a)\Gamma(b+\alpha_b)} \int_0^\infty x^{a+b+\alpha_a+\alpha_b-1} e^{-(\beta_b+n_b+(\beta_a+n_a)z)x} \mathrm{d}x$$
$$= \frac{1}{B(a+\alpha_a, b+\beta_b)} \frac{\left(\frac{\beta_a+n_a}{\beta_b+n_b}\right)^{a+\alpha_a} z^{a+\alpha_a-1}}{\left(1+\frac{\beta_a+n_a}{\beta_b+n_b}z\right)^{a+b+\alpha_a+\alpha_b}} \tag{A3}$$

which is a $BP(a+\alpha_a, b+\alpha_b, 1, \frac{\beta_b+n_b}{\beta_a+n_a})$ distribution as desired.

Appendix B: Derivation of the Posterior Mean and Covariance from Eq. (8)

From Eq. (8) we know that

$$\ell(\boldsymbol{\alpha}|\{a_i\}_{i=1}^d) = \sum_{i=1}^d a_i \log\left(\frac{c}{2}\left(\tilde{\mathbf{K}}\boldsymbol{\alpha}\right)_i^2\right) - \frac{1}{2}\boldsymbol{\alpha}^T\tilde{\mathbf{K}}\boldsymbol{\alpha}$$

In order to derive the posterior mean and covariance, we will first write $\boldsymbol{f} = \boldsymbol{\Phi}^T\boldsymbol{\beta}$. Then the likelihood becomes

$$\ell(\alpha|\{a_i\}_{i=1}^d) = \sum_{i=1}^d a_i \log\left(\frac{c}{2}\left(\boldsymbol{\Phi}^T\boldsymbol{\beta}\right)_i^2\right) - \frac{1}{2}\left\langle\boldsymbol{\beta}, \left(c\mathbf{I}+\gamma\mathbf{H}^{-1}\right)\boldsymbol{\beta}\right\rangle \tag{B1}$$

where $\mathbf{H} = \mathrm{diag}\,(\eta_i)$. From this we can see that the posterior mode estimate of $\boldsymbol{\beta}$ must satisfy

$$\hat{\boldsymbol{\beta}} = \left(c\mathbf{I}+\gamma\mathbf{H}^{-1}\right)^{-1}\nabla_{\boldsymbol{\beta}}\left[\sum_{i=1}^d a_i \log\left(\frac{c}{2}\left(\boldsymbol{\Phi}^T\hat{\boldsymbol{\beta}}\right)_i^2\right)\right] = \left(c\mathbf{I}+\gamma\mathbf{H}^{-1}\right)^{-1}\sum_{i=1}^d \frac{2a_i\boldsymbol{\Phi}_{:,i}}{\left(\boldsymbol{\Phi}^T\hat{\boldsymbol{\beta}}\right)_i} \tag{B2}$$

where $\boldsymbol{\Phi}_{:,i}$ is the i^{th} column of $\boldsymbol{\Phi}$. Then the posterior mode estimate of $\boldsymbol{f}$ is then given by

$$\hat{\boldsymbol{f}} = \boldsymbol{\Phi}^T\hat{\boldsymbol{\beta}} = \sum_{i=1}^d \left[\boldsymbol{\Phi}^T\left(c\mathbf{I}+\gamma\mathbf{H}^{-1}\right)^{-1}\boldsymbol{\Phi}_{:,i}\right]\frac{2a_i}{\left(\boldsymbol{\Phi}^T\hat{\boldsymbol{\beta}}\right)_i} \tag{B3}$$

From this, we can see the posterior mode has the form $\hat{\boldsymbol{f}} = \tilde{\mathbf{K}}\hat{\boldsymbol{\alpha}}$ by applying the eigenvector expansion of $\tilde{\mathbf{K}}$. This implies that, in our situation, $\hat{\alpha}_i = \frac{2a_i}{\hat{\boldsymbol{f}}_i}$, which is analogous to Eq. (11) in Walder and Bishop (2017). This posterior mode is used as the posterior mean in the Laplace approximation of the posterior distribution.

The inverse of the posterior covariance of $\boldsymbol{\beta}$ is given by the Hessian matrix at the posterior mode, given by

$$\mathbf{Q}^{-1} = -\frac{\partial^2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T}L(\boldsymbol{\beta}|\{a_j\}_{j=1}^d)|_{\hat{\boldsymbol{\beta}}} = \left(c\mathbf{I}+\gamma\mathbf{H}^{-1}\right) + \sum_{i=1}^d \frac{2a_i}{\left(\boldsymbol{\Phi}^T\hat{\boldsymbol{\beta}}\right)_i^2}\boldsymbol{\Phi}_{:,i}\boldsymbol{\Phi}_{:,i}^T \tag{B4}$$

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

Now letting $\mathbf{Z} = \left(c\mathbf{I} + \gamma\mathbf{H}^{-1}\right)$ and $\mathbf{D} = \text{diag}\left(\frac{\hat{\alpha}_j^2}{2a_j}\right)$ we have that $\mathbf{Q}^{-1} = \mathbf{Z} + \mathbf{\Phi}\mathbf{D}\mathbf{\Phi}^T$ and can apply the Woodbury-Morrison formula to write that

$$\mathbf{Q} = \mathbf{Z}^{-1} - \mathbf{Z}^{-1}\mathbf{\Phi}\left(\mathbf{D}^{-1} + \mathbf{\Phi}^T\mathbf{Z}^{-1}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^T\mathbf{Z}^{-1} \tag{B5}$$

The posterior covariance matrix of $\boldsymbol{f}$ is then given by

$$\hat{\mathbf{\Sigma}} = \mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi} \tag{B6}$$

$$= \tilde{\mathbf{K}} - \tilde{\mathbf{K}}\left(\mathbf{D}^{-1} + \tilde{\mathbf{K}}\right)^{-1}\tilde{\mathbf{K}} \tag{B7}$$

as desired.

Appendix C: The kernel

Since we wish to retrieve a field whose domain is the surface of the sphere, one option for constructing the eigenvector expansion of the kernel would be to use spherical harmonics. However, since we are only observing a portion of the sphere this would mean that the underlying intensity field is assumed to be 0 outside the viewing area and so is not continuous. Thus, using spherical harmonics, which are nonlocalized analytic functions, to represent the field can lead to nonphysical ringing in the retrieved fields. Instead, we choose the expansion of the kernel to be given in terms of spherical cap harmonics V_{lm} (SCHAs, Haines (1985)). These functions form an orthonormal basis for the Hilbert space of square integrable functions on the spherical cap, making them a natural basis for this problem. They have been used extensively in geoscience, for example in determining electrodynamics of the polar ionosphere (Richmond and Kamide, 1988), statistical analysis of TOPEX/POSEIDON data (Hwang and Chen, 1997), and mapping of total electron content over China and Iran (Liu et al., 2011; Razin and Voosoghi, 2017).

There are two families of spherical cap harmonics. One is the so-called even harmonics, which are calculated by solving the eigenvalue problem (Haines, 1985)

$$\begin{aligned}
\Delta V_{lm}(\theta,\phi) &= \frac{1}{\eta_{lm}} V_{lm}(\theta,\phi) \\
V_{lm}(\theta,\phi) &= V_{lm}(\theta,\phi+2\pi) \\
\frac{\partial}{\partial\phi} V_{lm}(\theta,\phi) &= \frac{\partial}{\partial\phi} V_{lm}(\theta,\phi+2\pi) \\
\frac{\partial}{\partial\theta} V_{lm}(\theta_0,\phi) &= 0 \\
V_{lm}(0,\phi) &= 0,\ m = 0 \\
\frac{\partial}{\partial\theta} V_{lm}(0,\phi) &= 0,\ m \neq 0
\end{aligned} \tag{C1}$$

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

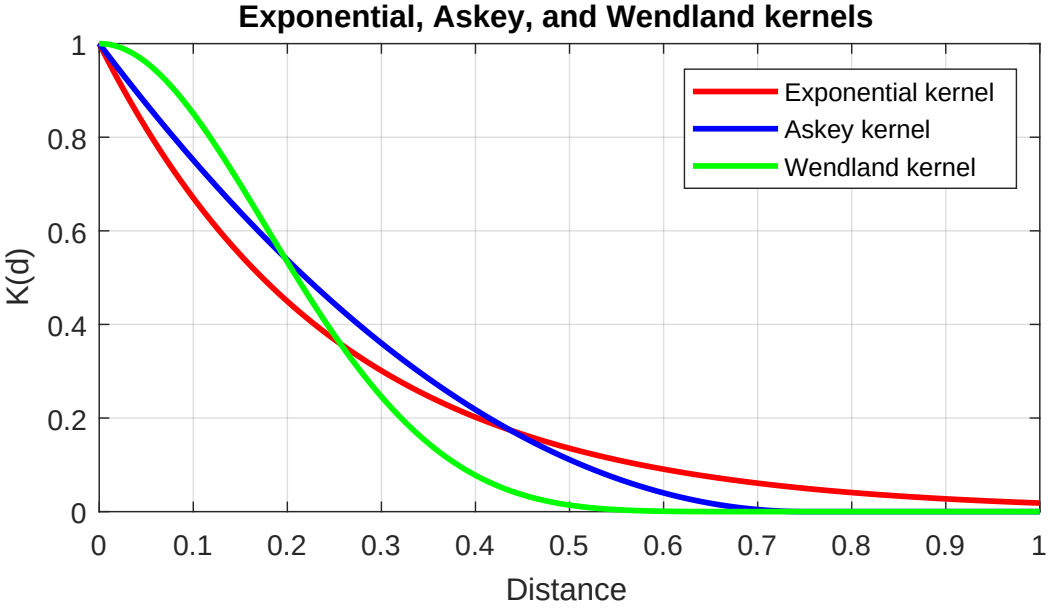

Figure C1. Askey (blue), exponential (red) and Wendland (green) kernels. The Askey and exponential kernels give rise to a rough process, while the Wendland kernel gives rise to a smoother process.

where θ_0 is the co-latitude of the cap boundary, taken for our purposes to be 64°. The odd harmonics are retrieved when the boundary condition $\frac{\partial}{\partial\theta}V_{lm}(\theta_0,\phi)=0$ is replaced by $V_{lm}(\theta_0,\phi)=0$. The even harmonics allow for the reconstruction of an arbitrary function on the boundary, and the odd harmonics allow reconstruction of a field with an arbitrary derivative at the boundary. Each of these sets of functions form an orthonormal basis of $\mathcal{L}^2(S)$, however they are not mutually orthogonal. Since we are concerned with reconstruction of the temperature field and not its derivatives, we choose to represent the field with the even harmonics. The solutions to this problem are the eigenvalue/eigenfunction pairs

$$V_{lm}(\theta,\phi)=e^{im\phi}P_{lm}(\cos(\theta))$$
$$\eta_{lm}=\frac{1}{l(l+1)} \tag{C2}$$

where P_{lm} are the associated Legendre functions, and l need not be an integer, but $m\in\mathbb{Z}$ and $l>|m|$. The associated Legendre functions are not polynomials, but instead are related to the hypergeometric function (Hwang and Chen, 1997). Due to the computational difficulty of solving the eigenvalue problem for various l,m we only calculate the SCHAs up to $m=20$, corresponding to 441 basis functions. Thus, the kernel $\mathbf{K}$ is given by Eq. (13), and the equivalent kernel is then given by

$$\tilde{\mathbf{K}}_{ij}=\sum_{l,m}\frac{1}{c+\gamma l^{\nu}(1+l)^{\nu}}V_{lm}(\theta_i,\phi_i)V_{lm}(\theta_j,\phi_j) \tag{C3}$$

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

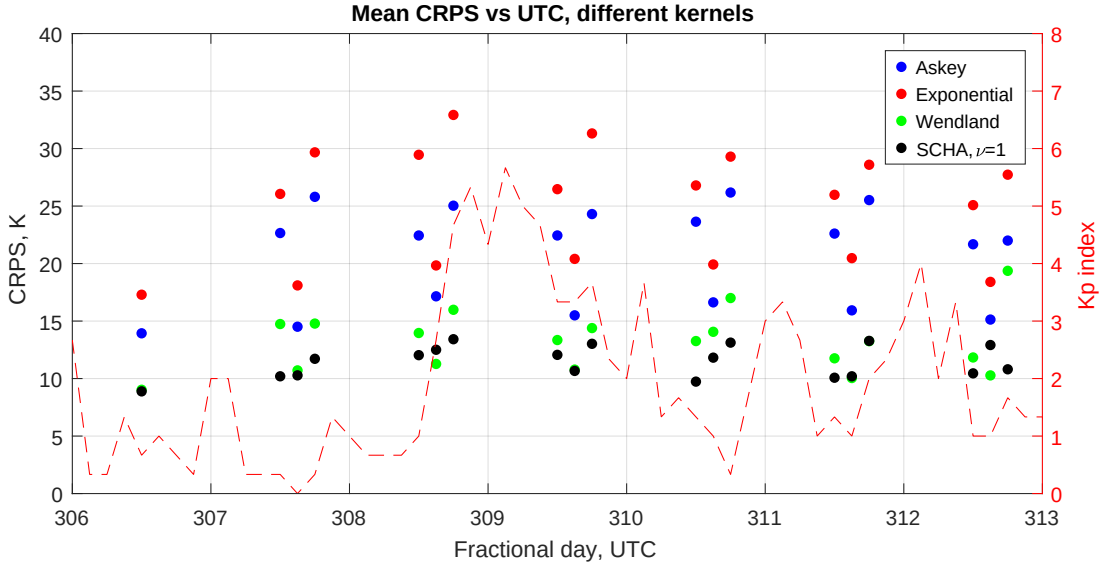

Figure C2. CRPS vs time for the Askey, exponential, Wendland, and SCHA kernels

As was shown in the body of the paper, the best of the SCHA kernels studied in terms of uncertainty quantification and retrieval accuracy was the one where $\nu = 1 + 10^{-8}$. We will compare this specific kernel with several other popular kernels that are positive definite on the sphere. A list of some possible kernels is included in Gneiting (2013). We choose to compare the SCHA kernel to the Askey kernel, given by

$$K(d;r) = \begin{cases} \left(1 - \left|\frac{d}{r}\right|\right)^2 & |d| < r \\ 0 & |d| \geq r \end{cases} \tag{C4}$$

, the exponential kernel, given by

$$K(d;r) = \exp\left(-\left|\frac{d}{r}\right|\right) \tag{C5}$$

, and the Wendland kernel given by Eq. (15). These kernels are plotted in Fig. C1. In all cases, the distance measure is the angular distance. From this figure, we see one feature that separates the Wendland from the other kernels: the number of derivatives that a kernel function has at $d = 0$ is related to the number of mean-square derivatives of the field (Stein, 1999), and the Askey and exponential kernels do not have any derivatives at $d = 0$, while the Wendland kernel does. This suggests that the Wendland kernel will generate smoother fields than the other two. Additionally, the Askey and Wendland kernels are compactly supported, meaning that they give rise to sparse kernel matrices. This property could make them useful in the analysis of large datasets, where correlations being zero outside a cutoff distance means that the algorithm can take advantage of sparsity in the resulting matrices. The exponential kernel is the covariance of an AR-1 process, leading to a sparse precision matrix for $\boldsymbol{f}$.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

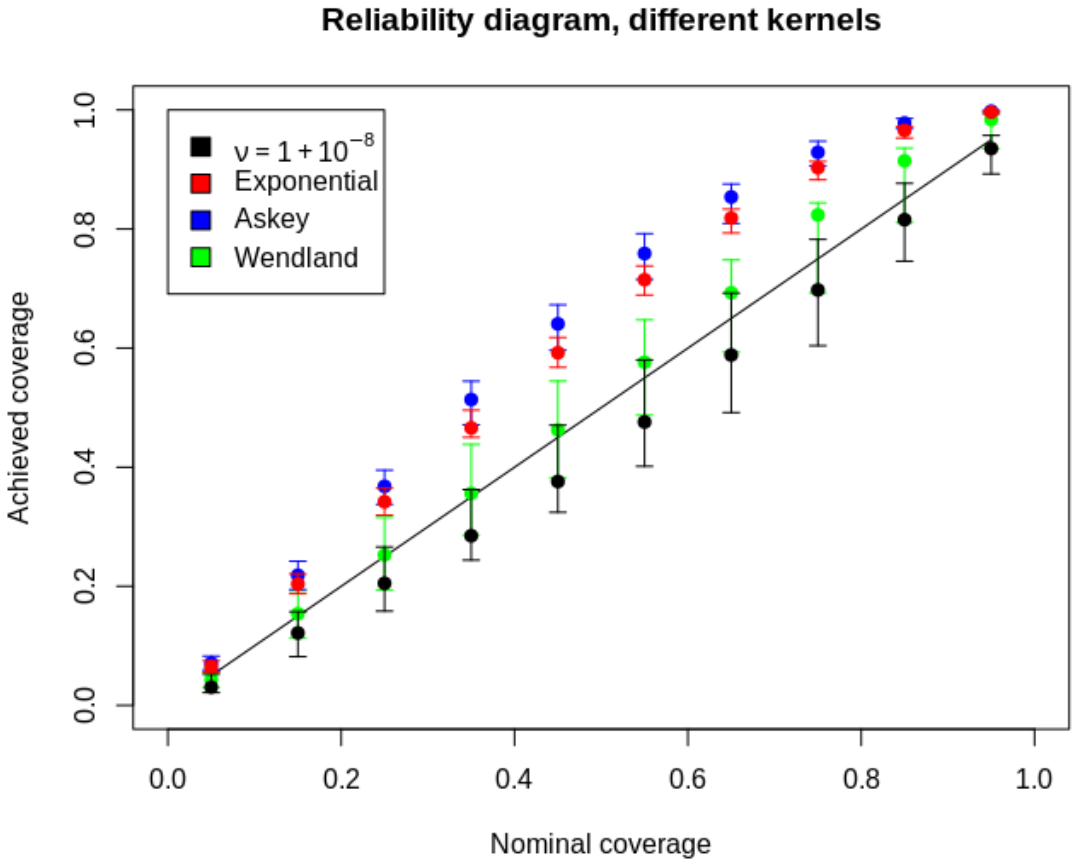

Figure C3. Reliability diagram for the different kernels, all K_p values.

We performed temperature retrievals on the simulated data examined in Sect. 3. In general, we found that the Wendland kernel had similar retrieval performance in terms of CRPS, with slightly worse uncertainty quantification as the nominal coverage increases. The other two kernels were not competitive in either metric. These results are shown in Figs. C2 and C3. We expect that this is due to the implied roughness of the fields from the Askey and exponential kernels. Since the retrieved field is an integrated quantity, we expect it to be relatively smooth, and from the data we see that in most conditions it is. This makes the Askey and exponential kernels less appropriate for the retrieval and less accurate in quantifying the uncertainty. The SCHA kernel with $\nu = 0.5$ showed similar levels of overdispersion in the posterior distribution, suggesting that incorrectly specifying the smoothness of the field is to blame.

While these results show that the SCHA kernel is the best option for our current analysis, we want to stress that any operational choice of kernel should involve careful cross-validation and other thorough testing before use. This is especially important because the application of the SCHA kernel to time periods where most of the disk is not illuminated could cause

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

problems akin to the ringing issue experienced with spherical harmonics. In a situation like this, a Wendland kernel or another one not investigated here may be preferable.

Author contributions. ML developed the technique and performed the analyses. TM and ML chose the datasets. TM and WK provided supervision and support. All authors provided interpretation and prepared the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. We would like to thank Dr. Clayton Cantrall for his help working with the curves of intensity ratio vs temperature generated for Cantrall and Matsuo (2021). We would like to acknowledge high-performance computing support from the Derecho system (doi:10.5065/qx9a-pg09) provided by the NSF National Center for Atmospheric Research (NCAR), sponsored by the National Science Foundation.

Financial support. This work was supported by NASA 80NSSC22K0175, UCAR SUBAWD006087, and NSF AGS-2231409 awards to the University of Colorado Boulder

.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

References

Ajello, J. M., Evans, J. S., Veibell, V., Malone, C. P., Holsclaw, G. M., Hoskins, A. C., Lee, R. A., McClintock, W. E., Aryal, S., Eastes, R. W., and Schneider, N.: The UV Spectrum of the Lyman-Birge-Hopfield Band System of N2 Induced by Cascading from Electron Impact, Journal of Geophysical Research: Space Physics, https://doi.org/10.1029/2019JA027546, 2020.

Akmaev, R. A.: Whole Atmosphere Modeling: Connecting Terrestrial and Space Weather, Reviews of Geophysics, 49, https://doi.org/10.1029/2011RG000364, 2011.

Aksnes, A., Eastes, R., Budzien, S., and Dymond, K.: Neutral temperatures in the lower thermosphere from N2 Lyman-Birge-Hopfield (LBH) band profiles, Geophys. Res. Lett., 33, https://doi.org/10.1029/2006GL026255, 2006.

Amaral Turkman, M. A., Paulino, C. D., and Müller, P.: Computational Bayesian Statistics: An Introduction, Institute of Mathematical Statistics Textbooks, Cambridge University Press, 2019.

Antoniadis, A. and Bigot, J.: Poisson inverse problems, The Annals of Statistics, 34, 2132 – 2158, https://doi.org/10.1214/009053606000000687, 2006.

Aryal, S., Evans, J. S., Ajello, J. M., Solomon, S. C., Burns, A. W., Eastes, R. W., and McClintock, W. E.: Constraining the Upper Level Vibrational Populations of the N2 Lyman-Birge-Hopfield Band System Using GOLD Mission's Dayglow Observations, Journal of Geophysical Research: Space Physics, 127, https://doi.org/10.1029/2021JA029869, 2022.

Boersma, C., Rubin, R., and Allamandola, L.: Spatial analysis of the polycyclic aromatic hydrocarbon features southeast of the Orion Bar, The Astrophysical Journal, 753, 168, https://doi.org/10.1088/0004-637X/753/2/168, 2012.

Budzien, S. A., Fortna, C. B., Dymond, K. F., Thonnard, S. E., Nicholas, A. C., McCoy, R. P., and Thomas, R. J.: Thermospheric temperature derived from ARGOS observations of N2 Lyman–Birge–Hopfield emission, Eos Trans. AGU, 82, Spring Meet. Suppl., 2001.

Cantrall, C.: New approaches for quantifying and understanding thermosphere temperature variability from far ultraviolet dayglow, Ph.D. thesis, University of Colorado-Boulder, 2022.

Cantrall, C. and Matsuo, T.: Deriving column-integrated thermospheric temperature with the N2 Lyman–Birge–Hopfield (2,0) band, Atmospheric Measurement Techniques, https://doi.org/10.5194/amt-14-6917-2021, 2021.

Cantrall, C. E., Matsuo, T., and Solomon, S. C.: Upper Atmosphere Radiance Data Assimilation: A Feasibility Study for GOLD Far Ultraviolet Observations, Journal of Geophysical Research: Space Physics, 124, 8154–8164, https://doi.org/10.1029/2019JA026910, 2019.

Coath, C. D., Steele, R. C. J., and Lunnon, W. F.: Statistical Bias in Isotope Ratios, Journal of Analytical Atomic Spectrometry, https://doi.org/10.1039/C2JA10205F, 2013.

Correira, J., Evans, J. S., Lumpe, J. D., Krywonos, A., Daniell, R., Veibell, V., McClintock, W. E., and Eastes, R. W.: Thermospheric Composition and Solar EUV Flux From the Global-Scale Observations of the Limb and Disk (GOLD) Mission, Journal of Geophysical Research: Space Physics, 126, https://doi.org/10.1029/2021JA029517, 2021.

Cox, D. R.: Some Statistical Methods Connected with Series of Events, Journal of the Royal Statistical Society: Series B (Methodological), 17, 129–157, https://doi.org/10.1111/j.2517-6161.1955.tb00188.x, 1955.

De Oliveira, V.: Hierarchical Poisson models for spatial count data, Journal of Multivariate Analysis, 122, 393–408, https://doi.org/10.1016/j.jmva.2013.08.015, 2013.

Eastes, R., Evans, J. S., Gan, Q., McClintock, W., and Lumpe, J.: Remote sensing of lower-middle-thermosphere temperatures using the N2 Lyman–Birge–Hopfield (LBH) bands, Atmospheric Measurement Techniques, https://doi.org/10.5194/amt-18-921-2025, 2025.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

Evans, J. S., Correira, J., Lumpe, J. D., Eastes, R. W., Gan, Q., Laskar, F. I., Aryal, S., Wang, W., Burns, A. G., Beland, S., Cai, X., Codrescu, M., England, S., Greer, K., Krywonos, A., McClintock, W. E., Plummer, T., and Veibell, V.: GOLD Observations of the Thermospheric Response to the 10–12 May 2024 Gannon Superstorm, Geophysical Research Letters, 51, https://doi.org/10.1029/2024GL110506, 2024a.

Evans, J. S., Lumpe, J. D., Eastes, R. W., Correira, J., Aryal, S., Laskar, F., Veibell, V., Krywonos, A., Plummer, T., and McClintock, W. E.: Disk Images of Neutral Temperature From the Global-Scale Observations of the Limb and Disk (GOLD) Mission, Journal of Geophysical Research: Space Physics, 129, https://doi.org/10.1029/2024JA032424, 2024b.

Flaxman, S., Teh, Y. W., and Sejdinovic, D.: Poisson intensity estimation with reproducing kernels, Electronic Journal of Statistics, 11, 5081–5104, https://doi.org/10.1214/17-EJS1339SI, 2017.

Gallardo i Peres, G., Dall, J., Mason, P. J., Ghail, R., and Hensley, S.: A Generalized Beta Prime Distribution as the Ratio Probability Density Function for Change Detection Between Two SAR Intensity Images With Different Number of Looks, IEEE Transactions on Geoscience and Remote Sensing, 62, 1–14, https://doi.org/10.1109/TGRS.2024.3369509, 2024.

Gneiting, T.: Strictly and non-strictly positive definite functions on spheres, Bernoulli, 19, 1327–1349, https://doi.org/10.3150/12-BEJSP06, 2013.

Gneiting, T. and Rafferty, A.: Strictly Proper Scoring Rules, Prediction, and Estimation, Journal of the American Statistical Association, https://doi.org/10.1198/016214506000001437, 2007.

Grandin, M., Bruus, E., Ledvina, V. E., Partamies, N., Barthelemy, M., Martinis, C., Dayton-Oxland, R., Gallardo-Lacourt, B., Nishimura, Y., Herlingshaw, K., Thomas, N., Karvinen, E., Lach, D., Spijkers, M., and Bergstrand, C.: The Gannon Storm: citizen science observations during the geomagnetic superstorm of 10 May 2024, Geoscience Communication, 7, 297–316, https://doi.org/10.5194/gc-7-297-2024, 2024.

Haines, G.: Computer programs for spherical cap harmonic analysis of potential and general fields, Computers & Geosciences, 14, 413–447, https://www.sciencedirect.com/science/article/pii/0098300488900271, 1988.

Haines, G. V.: Spherical cap harmonic analysis, Journal of Geophysical Research: Solid Earth, 90, 2583–2591, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB090iB03p02583, 1985.

Humphrey, P. J., Liu, W., and Buote, D. A.: χ^2 and Poissonian data: Biases even in the high-count regime and how to avoid them, The Astrophysical Journal, 693, https://doi.org/10.1088/0004-637X/693/1/822, 2009.

Hunter, J. K. and Nachtergaele, B.: Applied Analysis, World Scientific, 2020.

Hwang, C. and Chen, S.-K.: Fully normalized spherical cap harmonics: application to the analysis of sea-level data from TOPEX/POSEIDON and ERS-1, Geophysical Journal International, 129, 450–460, https://doi.org/10.1111/j.1365-246X.1997.tb01595.x, 1997.

Jia, J. and Yi, F.: Atmospheric temperature measurements at altitudes of 5-30km with a double-grating-based pure rotational Raman lidar, Appl. Opt., 53, 5330–5343, https://doi.org/10.1364/AO.53.005330, 2014.

Jin, Y., Zhang, S., and Wu, J.: Hardness Ratio estimation in Low Counting X-Ray Photometry, The Astrophysical Journal, https://doi.org/10.1086/508677, 2006.

Liu, J., Chen, R., Wang, Z., and Zhang, H.: Spherical cap harmonic model for mapping and predicting regional TEC, GPS Solutions, 15, 109–119, 10.1007/s10291-010-0174-8, 2011.

Lloyd, C., Gunter, T., Osborne, M., and Roberts, S.: Variational Inference for Gaussian Process Modulated Poisson Processes, in: Proceedings of the 32nd International Conference on Machine Learning, edited by Bach, F. and Blei, D., vol. 37 of *Proceedings of Machine Learning Research*, pp. 1814–1822, PMLR, https://proceedings.mlr.press/v37/lloyd15.html, 2015.

Matzka, J., Bronkalla, O., Tornow, K., Elger, K., and Stolle, C.: Geomagnetic Kp index. V. 1.0., https://doi.org/10.5880/Kp.0001, 2021.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

McClintock, W. E., Eastes, R. W., Beland, S., Bryant, K. B., Burns, A. G., Correira, J., Danielll, R. E., Evans, J. S., Harper, C. S., Karan, D. K., Krywonos, A., Lumpe, J. D., Plummer, T. M., Solomon, S. C., Vanier, B. A., and Veibel, V.: Global-Scale Observations of the Limb and Disk Mission Implementation: 2. Observations, Data Pipeline, and Level 1 Data Products, Journal of Geophysical Research: Space Physics, 125, https://doi.org/10.1029/2020JA027809, 2020.

McCullagh, P. and Møller, J.: The Permanental Process, Advances in Applied Probability, 38, 873–888, https://doi.org/10.1017/S0001867800001361., 2006.

McDonald, J. B.: Model selection: some generalized distributions., Communications in Statistics - Theory and Methods, 16, https://doi.org/10.1080/03610928708829422, 1987.

McDonald, J. B. and Xu, Y. J.: A generalization of the beta distribution with applications, Journal of Econometrics, https://doi.org/10.1016/0304-4076(94)01612-4, 1995.

Meier, R. R.: The Thermospheric Column O/N2 Ratio, Journal of Geophysical Research: Space Physics, 126, e2020JA029 059, https://doi.org/10.1029/2020JA029059, 2021.

Møller, J.: Properties of spatial Cox process models, Journal of Statistical Research of Iran, 2, 1–18, 2005.

Nicholls, R.: Franck-Condon factors to high vibrational quantum numbers I: N2 and N2+, Journal of Research of the National Bureau of Standards. Section A, Physics and Chemistry, 65, 451, https://doi.org/10.6028/jres.065A.047, 1961.

Nicolaou, G., Livadiotis, G., Sarlis, N., and Ioannou, C.: Resolving velocity distribution function parameters from observations with significant Poisson statistical uncertainty, RAS Techniques and Instruments, 3, 874–878, https://doi.org/10.1093/rasti/rzae059, 2024.

Nowak, R. D. and Kolaczyk, E. D.: A statistical multiscale framework for Poisson inverse problems., IEEE Transactions on Information Theory, 46, 1811– 1825, https://doi.org/10.1109/18.857793, 2000.

Park, T., Kashyap, V. L., Siemiginowska, A., van Dyk, D. A., Zezas, A., Heinke, C., and Wargelin, B. J.: Bayesian Estimation of Hardness Ratios: Modeling and Computations, The Astrophysical Journal, 652, 610, https://doi.org/10.1086/507406, 2006.

Paxton, L. J., Schaefer, R. K., Zhang, Y., and Kil, H.: Far ultraviolet instrument technology, Journal of Geophysical Research: Space Physics, 122, 2706–2733, https://doi.org/10.1002/2016JA023578, 2017.

Pham-Gia, T. and Turkkan, N.: Operations on the Generalized F Variables and Applications, Statistics, https://doi.org/10.1080/02331880212855, 2002.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2025.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, The MIT Press, ISBN 9780262256834, https://doi.org/10.7551/mitpress/3206.001.0001, 2005.

Razin, M. and Voosoghi, B.: Regional ionosphere modeling using spherical cap harmonics and empirical orthogonal functions over Iran, Acta Geodaetica et Geophysica, 52, 19–33, https://doi.org/10.1007/s40328-016-0162-8, 2017.

Richmond, A. D. and Kamide, Y.: Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, Journal of Geophysical Research: Space Physics, 93, 5741–5759, https://doi.org/10.1029/JA093iA06p05741, 1988.

Solomon, S. C.: Global modeling of thermospheric airglow in the far ultraviolet, JGR Space Physics, 122, 7834–7848, 2017.

Stein, M. L.: Interpolation of Spatial Data: Some Theory for Kriging, Springer Series in Statistics, Springer, 1999.

Streit, R. L.: Poisson Point Processes:Imaging, Tracking, and Sensing, Springer, https://doi.org/10.1007/978-1-4419-6923-1, 2010.

Strickland, D. J., Evans, J. S., and Paxton, L. J.: Satellite remote sensing of thermospheric O/N2 and solar EUV: 1. Theory, Journal of Geophysical Research: Space Physics, 100, 12 217–12 226, https://doi.org/10.1029/95JA00574, 1995.

https://doi.org/10.5194/egusphere-2025-5570
Preprint. Discussion started: 19 December 2025

Walder, C. J. and Bishop, A. N.: Fast Bayesian Intensity Estimation for the Permanental Process, in: Proceedings of the 34th International Conference on Machine Learning, edited by Precup, D. and Teh, Y. W., vol. 70 of *Proceedings of Machine Learning Research*, pp. 3579–3588, PMLR, https://proceedings.mlr.press/v70/walder17a.html, 2017.

Wang, C., Liao, J.-Y., Guan, J., Liu, Y., Li, C.-K., Sai, N., Jin, J., and Zhang, S.-N.: Analysis of bright source hardness ratios in the 4 yr Insight-HXMT galactic plane scanning survey catalog, Research in Astronomy and Astrophysics, 24, 025 013, https://doi.org/10.1088/1674-4527/ad18a3, 2024.

Wendland, H.: Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree., Adv Comput Math, 4, 389–396, 1995.

Yamada, S., Axelsson, M., Ishisaki, Y., Konami, S., Takemura, N., Kelley, R. L., Kilbourne, C. A., Leutenegger, M. A., Porter, F. S., Eckart, M. E., and Szymkowiak, A.: Poisson vs. Gaussian statistics for sparse X-ray data: Application to the soft X-ray spectrometer, Publications of the Astronomical Society of Japan, 71, https://doi.org/10.1093/pasj/psz053, 2019.

Yin, X., Qin, J., and Paxton, L. J.: A new strategy for ionospheric remote sensing using the 130.4/135.6 nm airglow intensity ratios, Earth and Planetary Physics, 7, 445–459, https://doi.org/10.26464/epp2023042, 2023.

Zhang, Y., Paxton, L. J., Morrison, D., Wolven, B., Kil, H., Meng, C.-I., Mende, S. B., and Immel, T. J.: O/N2 changes during 1–4 October 2002 storms: IMAGE SI-13 and TIMED/GUVI observations, Journal of Geophysical Research: Space Physics, 109, https://doi.org/https://doi.org/10.1029/2004JA010441, 2004.

Zhang, Y., Paxton, L. J., and Schaefer, R. K.: Deriving Thermospheric Temperature From Observations by the Global Ultraviolet Imager on the Thermosphere Ionosphere Mesosphere Energetics and Dynamics Satellite, Journal of Geophysical Research: Space Physics, 124, 5848–5856, 2019.