Towards Enhancements of Auroral Forecasting through Deep Learning

Part 1: A DAYSIDE AURORA DATASET FROM THE GOLD MISSION

Jordan Holmes[1], Scott L. England[1]

Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia USA[1]

Abstract

We present a comprehensive dataset of dayside auroral emissions observed by the Global-scale Observations of the Limb and Disk (GOLD) mission from October 2018 to June 2025. The dataset contains over 47,000 unique scans of the northern aurora in three far-ultraviolet spectral channels (OI 135.6 nm, NI 149.3 nm, and N_2 LBH), estimates of the background dayglow, binary masks of auroral locations, and other corresponding spatial and temporal metadata. The OI 135.6 nm, NI 149.3 nm, and N_2 LBH emissions are far-ultraviolet signatures of electron-impact excitation in the upper atmosphere and therefore serve as tracers of auroral electron precipitation. From this dataset, auroral pixels are directly available with no dayglow contamination of the emissions. Auroral signals are extracted through a multi-stage processing pipeline inspired by computer vision and machine learning techniques. This dataset provides a consistent view of the dayside aurora over the North American and Atlantic sectors, enabling studies of auroral dynamics with GOLD observations.

Introduction

The aurora is one of the most striking manifestations of space weather, driven by the precipitation of charged particles from various regions of the Earth's magnetosphere and solar wind into the upper atmosphere. As these particles enter the ionosphere–thermosphere (IT) system, they deposit energy that alters neutral and plasma dynamics[1, 2]. Their occurrence provides an indicator of magnetospheric coupling with the IT, and changes in the location, brightness and spectrum of the aurora is linked to geospace disturbances that can cause significant societal impacts including ground-based power outages, disruptions to radio wave propagation, and anomalies in spacecraft operation[3].

Although visible auroral displays are often appreciated as nighttime phenomena, they occur continuously throughout the day. Dayside auroral observations are particularly valuable for understanding solar wind-magnetosphere interactions and monitoring the dynamics of magnetic reconnection at the magnetopause[4]. However, dayside auroral observations have historically presented challenges due to the overwhelming solar induced emissions (dayglow) that can mask or obscure auroral emissions, especially at visible wavelengths.

Space-based far-ultraviolet (FUV) imaging has emerged as the primary technique for dayside auroral observations, as FUV emissions can be distinguished from background dayglow, and are not overpowered by reflected light from the surface of the Earth (which impacts visible wavelengths). Over the past two decades, instruments including the spectrographic imager (SI) and widefield imaging camera (WIC) onboard IMAGE[5, 6](2000-2005), the Global Ultraviolet Imager (GUVI) onboard TIMED[7, 8] (2002-present), and the Special Sensor Ultraviolet Spectrographic Imager (SSUSI) on multiple DMSP satellites (2003-present) have provided valuable UV auroral datasets[9]. These measurements have excelled at providing global coverage and high spatial resolution observations, offering crucial insights into auroral precipitation patterns.

The Global-scale Observations of the Limb and Disk (GOLD) mission provides a complementary observational capability that builds upon these datasets through a fundamentally different perspective. Launched in 2018 and positioned in geostationary orbit at 47.5°W, the GOLD instrument provides continuous monitoring of Earth's disk from a fixed vantage point[10]. The instrument features a FUV imaging spectrograph (132-164 nm) that captures key auroral emissions: atomic oxygen (135.6 nm), atomic nitrogen (149.3 nm), and the N_2 Lyman-Birge-Hopfield band

system. This geostationary perspective enables repeated imaging of the same geographic regions, making it uniquely suited for tracking temporal evolution and dynamics of the aurora. In 2020, Michell[11] demonstrated the instrument's capability to observe the aurora in two case studies. Inspired by this, but realizing the challenge in applying the prior methodology to the ~7 years of GOLD data, we developed our own scaled and automated approach to collect available dayside auroral observations from the GOLD mission.

GOLD Data

The dayside auroral observations can be found in the GOLD Level 1C (L1C) "DAY" data products. Each file corresponds to a single hemispheric scan with spatial dimensions of 52 x 92 non-zero pixels and a spectral dimension of 800 channels, spanning 132-162nm. Each image is produced by swathing the entrance slit over an entire hemisphere of the Earth as seen from GEO, gathering spectral information at each pixel. Full disk images of the Earth can be created by combining paired northern and southern hemispheric scans, with a 12 minute timing difference between the two hemispheric scans. Each full disk image covers geographic latitudes –74° to +74° and longitudes –127° to +32° with a nadir resolution 125 km × 125 km. During 2018–2021, GOLD produced 34 full disk scans per day with 30 minute cadence between subsequent full disk scans. After 2021, the cadence between subsequent full disk images decreased to 2 hours to extend the instrument's lifetime due to sensor degradation[12].

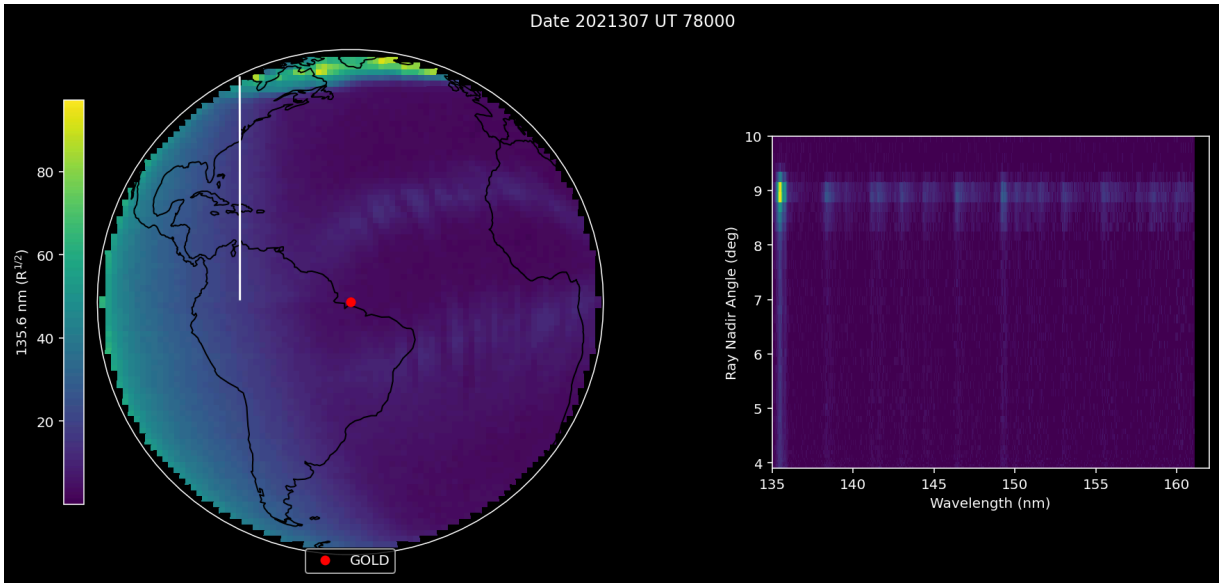

Figure 1: Shown here is an illustration of typical GOLD 135.6nm emission data from a single channel for a single slit position. On the left is an image of 135.6nm emission intensity displayed on a logarithmic scale in the range. Notice the slit only covers one hemisphere at a time. The right shows the spatial and spectral information along the slit. At the top of both figures, the aurora can be seen.

By virtue of GOLD's longitude and the location of the geomagnetic poles, the southern aurora is rarely visible except during major storms. While this might appear as a limitation, it provides an aurora-free reference that is dominated by dayglow. This allows southern scans to be used as a template of dayglow contamination that is present in the northern auroral observations.

Method

To construct a high-quality dataset of dayside auroral emissions from GOLD observations we developed a multi-stage processing pipeline that:

1. Removes dayglow contamination by modeling the dayglow using southern hemispheric observations.
2. Applies classical computer vision, clustering, and signal processing techniques to segment

auroral regions, producing an intermediate, automatically-labelled dataset.

3. Uses this intermediate dataset (3) to train a deep learning model (UNet[13]), improving segmentation across storm conditions and during periods of reduced GOLD measurement cadence in later years.

Below, we describe each stage in detail.

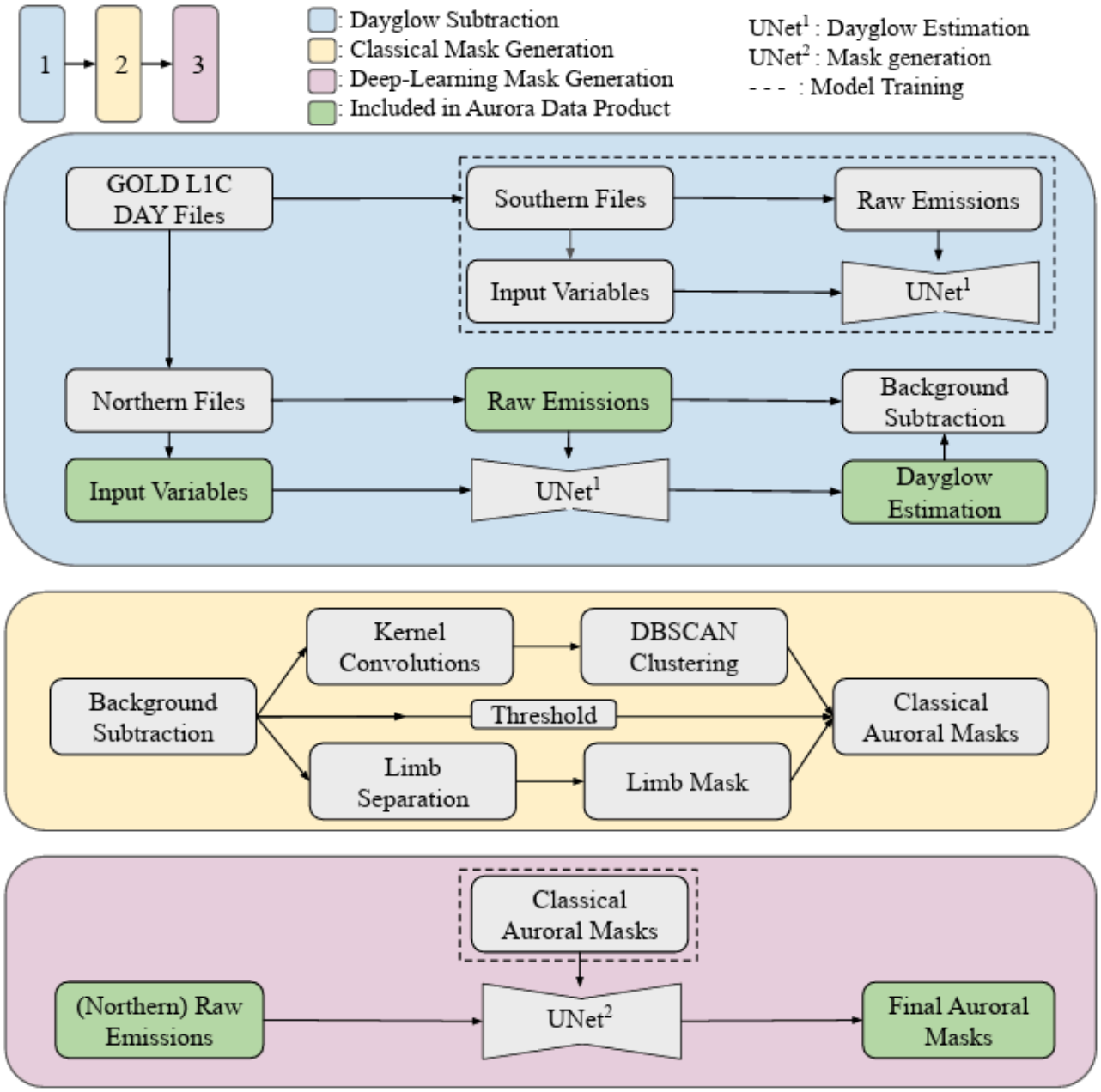

Figure 2: Schematic overview for generating the dayside auroral dataset. **Step 1** trains a UNet model to remove dayglow contamination from northern scans. **Step 2** applies signal processing techniques to produce binary masks of

the aurora. **Step 3** uses the classical masks to train a secondary UNet model to produce the final dayside auroral pixel locations. Green boxes indicate data that is kept in the final product.

Step 1: The Dayglow Issue

Dayglow presents the main challenge in isolating dayside aurora. The perceived intensity varies strongly with solar zenith angle (SZA), emission angle (EMA), day of year (DOY), and geomagnetic activity (e.g. Kp). Physics-based models exist to describe this variability(e.g. GLOW[14]), but the outputs of the model may differ from experimental observations due to possible calibration issues. Instead, we use a data-driven approach in which southern scans form the empirical basis for dayglow estimation in northern scans. Direct subtraction of southern and northern scans is difficult because the 12-minute offset between them prevents complete cancellation. To address this, we trained a UNet[13] model to predict dayglow directly with inputs: SZA, EMA, DOY, & Kp. While many models can perform regression, UNet was selected for its strong generalization in both regression (dayglow estimation) and segmentation (relevant later). The architecture largely followed the standard design, with added padding to accommodate GOLD's rectangular image dimensions (52 × 92). After training, the final validation achieved $R^2 = 0.95$, indicating a well-trained dayglow model, with respect to the southern hemispheric data. For inference the variables of the northern hemispheric scans are used to gather an estimate on the dayglow. The predicted dayglow is then scaled via least-squares on low-latitude pixels to prevent oversubtraction of auroral pixels, with only positive residuals retained in the image.

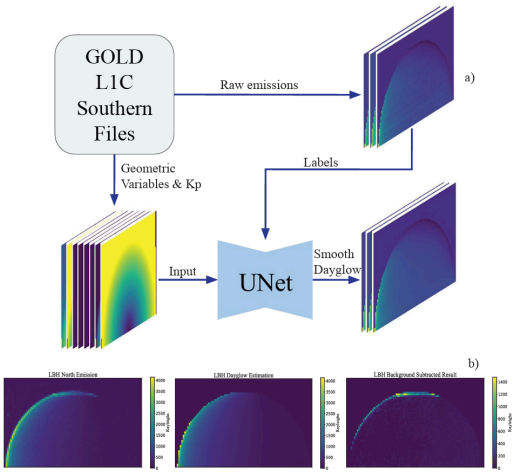

Figure 3: a) Dataflow for training the dayglow UNet model. Inputs include geometric variables (SZA, EMA, and DOY, encoded with sine–cosine functions) and geomagnetic activity level (Kp), which are paired with raw southern hemisphere emissions. The UNet outputs smoothed representations of the dayglow. b) Example of background subtraction applied to northern emissions. While positive residuals remain near the limb, most of the dayglow is effectively removed.

Step 2: **Classical Segmentation**

After subtraction, contamination of non-auroral pixels remains and requires further processing. In most scans from GOLD, the auroral pixels typically form a coherent and predictable structure near the edge of the Earth. This allows us to sequester these pixels. First, on the background subtracted image, we convolve it with a Laplacian-of-Gaussian kernel which acts as a sharpening filter, followed by a Gabor filter which acts as an extended edge detector. The response image from these operations are binarized, and the row-column pairs of 'on' pixels are fed into a clustering algorithm, DBSCAN[15]. This noise resistant clustering algorithm is useful since there shouldn't be high responses of non-auroral pixels to the previous kernel-convolution operations. Finally, since most of the auroral pixels lie on the limb of the Earth, some limb pixels that aren't aurora can get classified as such from the previous operations. Therefore, a sub-routine is necessary for determining which pixels are exclusively aurora. We now compare the 1-D intensity profile of the limb pixels of northern and southern scans. The difference between the northern and southern limb profiles typically contains two-peak structure. Pixels between these peaks are classified as auroral emissions. Boundaries of these peaks can be determined by zero crossings of the second derivative on the differenced limb signal. To reduce any potential noise in any one instance, we track limb-auroral trajectories over time and apply temporal smoothing, under the assumption that auroral morphology does not change drastically within the 30-minute interval between subsequent scans. The final 2-D auroral mask is then produced by combining results from both the 2-D feature-based, 1-D

limb-based algorithms, and pixels from the background subtracted images that are above an intensity threshold (50 Rayleighs).

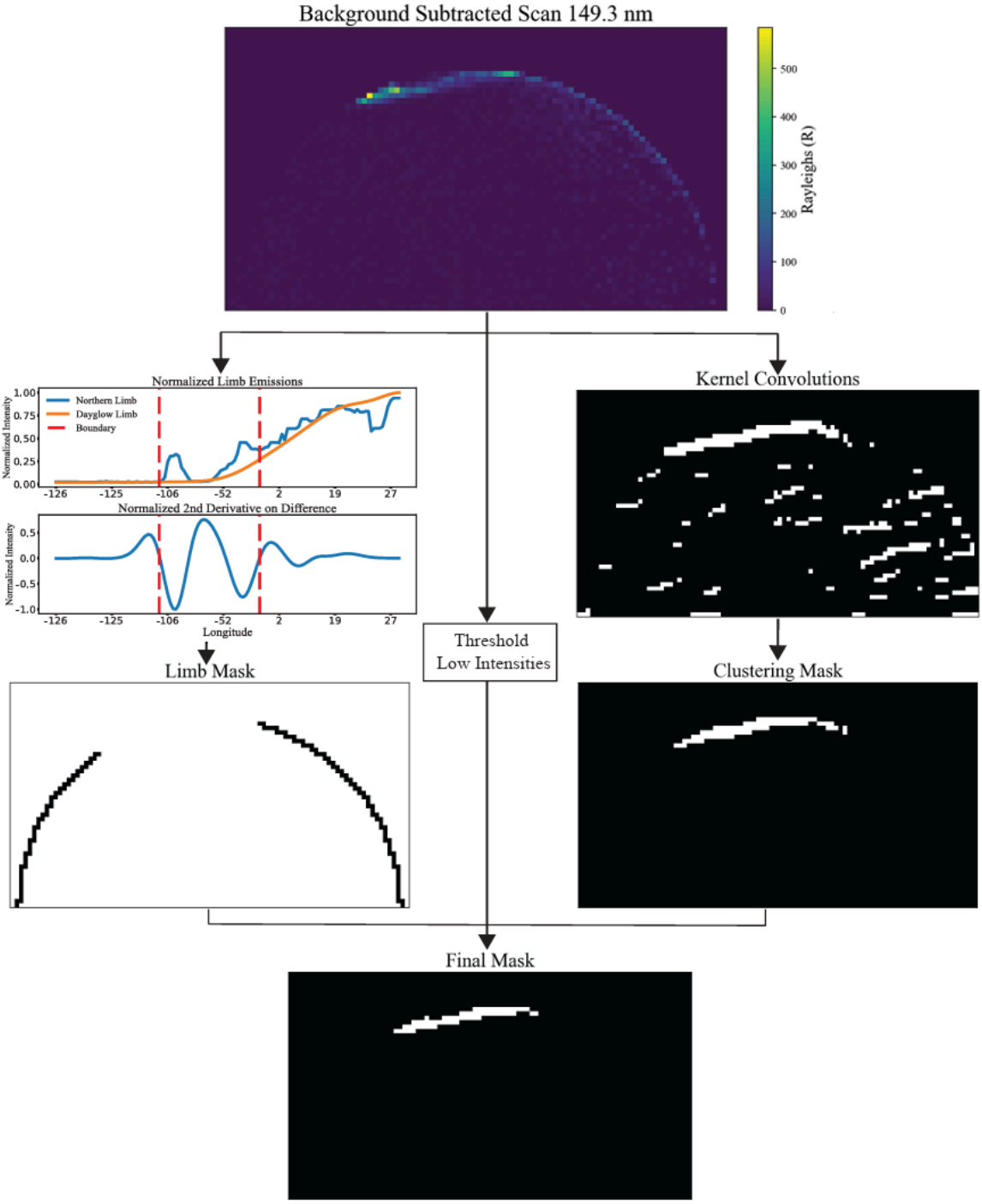

Figure 4: The flow for generating a mask in **one** instance begins with the differenced image. From this, two complementary routines are applied: a spatial routine (right column) identifies the auroral region, while the 1D signal-based routine (left column) separates the auroral signal from limb emissions. Combining the outputs of both

approaches produces a more accurate mask with fewer false positives. The temporal smoothing step of the limb mask is not shown here, but is required to ensure consistency from instance to instance.

Step 3: **Deep Learning-based Segmentation**

Accurate boundary smoothing is only possible during years with higher-cadence measurements (2018–2021). After 2021, the smoothing assumption breaks down and the previously mentioned mask generation process becomes unstable. To overcome this limitation, and generate masks for later years of the mission, we train a UNet model to directly produce the binary masks. The training labels are the 2020 auroral masks produced by the algorithm, and the inputs to the model are simply the northern scans. Validation yielded a Matthews Correlation Coefficient (MCC) of 0.893 against the classically labelled masks. The MCC is considered a more representative metric for evaluating the performance of a binary classifier, particularly on imbalanced datasets, as it accounts for all values in the confusion matrix[16].

Validation

Direct validation of the produced GOLD auroral masks requires comparison with an independent dataset of auroral occurrence. Since there is no concurrent dayside aurora dataset that provides extended overlap in coverage with GOLD, validation is performed against the Zhang-Paxton (ZP) auroral oval model[17]. The ZP model was derived from nearly four years of FUV auroral images collected by the TIMED/GUVI instrument from 2002-2005, making it a suitable model to compare against GOLD. The ZP model is parameterized by Kp level and specifies precipitating electron energy flux as a function of magnetic latitude (MLAT) and magnetic local time (MLT). The ZP model poleward and equatorward boundaries are defined by the lowest MLAT of 0.25 ergs cm^{-2} s^{-1} flux, with all enclosed grid cells between these contours classified as aurora. Here we adopt this 0.25 ergs cm^{-2} s^{-1} flux boundary as our auroral threshold for comparison with GOLD.

Since the ZP model is defined in magnetic coordinates, all GOLD observations were transformed from geographic to apex quasi-dipole coordinates using the apexpy library[18,19], with the date of observation as the epoch and an altitude of 150 km.The GOLD auroral observations were then compared to the ZP model auroral classifications at each MLAT/MLT grid point. Since the resolution of GOLD becomes very coarse near the limb, a single GOLD pixel in this region may span a range of auroral and sub-auroral locations. To address this, we compute a finer grid of points within each GOLD pixel – a 5×5 array of subpixels – with each subpixel converted to quasi-dipole coordinates. A GOLD pixel is classified as containing an expected auroral signature if at least 3 of the 25 subpixels exceed the 0.25 erg cm^{-2} s^{-1} flux threshold according to the ZP model.

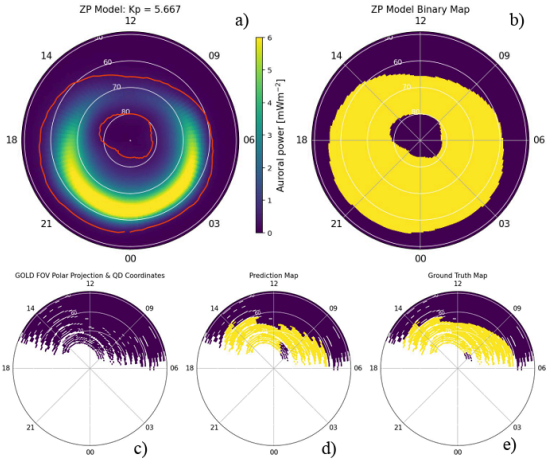

Figure 5: a) Example output prediction of the ZP model at moderate geomagnetic activity. b) All bins of the model above the selected threshold. c) Single instance of GOLD's viewing perspective in quasi-dipole coordinates. d) Instance of an auroral prediction output from the methods section. e) Ground truth map obtained by combining panels c) and d).

Comparing outputs for a high cadence year (2020), the classical and deep learning methods for segmentation achieved MCC scores of 0.82 and 0.86, respectively, when evaluated against the ZP. The classical method provided reliable identifications of auroral pixels, while the deep learning model, aided by regularization and data augmentation techniques, demonstrated a somewhat stronger agreement with the ZP model. Given this improved performance and its ability to generalize across different geomagnetic

conditions, the trained deep learning model was adopted to generate auroral masks for the complete dataset spanning October 2018 to June 2025.

Conclusions

This dataset provides large-scale daytime auroral maps with significant reuse potential for the broader space weather community, comprising over 47,000 image–label pairs spanning six years of GOLD mission data, and showing strong agreement with external FUV-image-based auroral precipitation models. Although not an original objective of the GOLD mission, the dataset enables researchers to leverage GOLD's unique observations – unavailable from other missions – for new studies. By removing data-processing barriers, it facilitates a wide range of physical and correlational analyses without requiring researchers to build their own preprocessing pipelines. Specific potential applications include: (1) statistical analysis of auroral boundary information and its relationship to solar wind parameters, (2) validation of magnetosphere-ionosphere coupling models, (3) development of a deep learning auroral forecasting model with multimodal data from many different missions.

References

[1] Lu, G., Richmond, A. D., Emery, B. A. & Roble, R. G. Magnetosphere–ionosphere–thermosphere coupling: Effect of neutral winds on energy transfer and field-aligned current. *J. Geophys. Res.* **100**, 19643–19659 (1995). https://doi.org/10.1029/95JA00766

[2] Richmond, A. D. & Lu, G. Upper-atmospheric effects of magnetic storms: a brief tutorial. *J. Atmos. Sol.-Terr. Phys.* **62**, 1115–1127 (2000).

[3] Buzulukova, N. & Tsurutani, B. Space Weather: From solar origins to risks and hazards evolving in time. *Front. Astron. Space Sci.* **9**, 1017103 (2022). https://doi.org/10.3389/fspas.2022.1017103

[4] Frey, H. U., Han, D., Kataoka, R. *et al.* Dayside Aurora. *Space Sci. Rev.* **215**, 51 (2019). https://doi.org/10.1007/s11214-019-0617-7

[5] Mende, S., Heetderks, H., Frey, H. et al. Far ultraviolet imaging from the IMAGE spacecraft. 1. System design. Space Science Reviews 91, 243–270 (2000). https://doi.org/10.1023/A:1005271728567

[6] Mende, S., Heetderks, H., Frey, H. et al. Far ultraviolet imaging from the IMAGE spacecraft. 3. Spectral imaging of Lyman-α and OI 135.6 nm. Space Science Reviews 91, 287–318 (2000). https://doi.org/10.1023/A:1005292301251

[7] Christensen, A. B., et al. (2003), Initial observations with the Global Ultraviolet Imager (GUVI) in the NASA TIMED satellite mission, J. Geophys. Res., 108, 1451, doi:10.1029/2003JA009918, A12.

[8] Paxton, L. J. & Meng, C.-I. Auroral imaging and space-based optical remote sensing. *Johns Hopkins APL Tech. Dig.* **20**, 556 (1999).

[9] Larry J. Paxton, Daniel Morrison, Yongliang Zhang, Hyosub Kil, Brian Wolven, Bernard S. Ogorzalek, David C. Humm, Ching-I. Meng, "Validation of remote sensing products produced by the Special Sensor Ultraviolet Scanning Imager (SSUSI): a far UV-imaging spectrograph on DMSP F-16," Proc. SPIE 4485, Optical Spectroscopic Techniques, Remote Sensing, and Instrumentation for Atmospheric and Space Research IV, (30 January 2002); https://doi.org/10.1117/12.454268

[10] Eastes, R. W., McClintock, W. E., Burns, A. G., Anderson, D. N., Andersson, L., Aryal, S., et al. (2020). Initial observations by the GOLD mission. *Journal of Geophysical Research: Space Physics*, 125, e2020JA027823. https://doi.org/10.1029/2020JA027823

[11] Michell, R. G. (2020). Auroral structure and dynamics from GOLD. Journal of Geophysical Research: Space Physics, 125, e2019JA027650. https://doi.org/10.1029/2019JA027650

[12] Eastes, R. W., McClintock, W. E., Burns, A. G. *et al.* The Global-Scale Observations of the Limb and Disk (GOLD) Mission. *Space Sci. Rev.* **212**, 383–408 (2017). https://doi.org/10.1007/s11214-017-0392-2

[13] GOLD Science Team. *GOLD Data Release Notes – Rev. 5.9* (NASA GOLD Mission, 2025). Available at: https://gold.cs.ucf.edu/data_release_notes

[13] Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional networks for biomedical image segmentation. Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. 9351, 234–241 (2015).

[14] Solomon, S. C. (2017), Global modeling of thermospheric airglow in the far ultraviolet, J. Geophys. Res. Space Physics, 122, 7834–7848, doi:10.1002/2017JA024314.

[15] Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. Second Int. Conf. Knowledge Discovery and Data Mining (KDD-96)*, 226–231 (AAAI Press, 1996).

[16] Chicco, D. & Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **21**, 6 (2020). https://doi.org/10.1186/s12864-019-6413-7

[17] Zhang, Y. & Paxton, L. J. An empirical Kp-dependent global auroral model based on TIMED/GUVI FUV data. *J. Atmos. Terr. Phys.* **70**, 1231–1242 (2008). https://doi.org/10.1016/j.jastp.2008.03.008

[18] Emmert, J. T., Richmond, A. D. & Drob, D. P. A computationally compact representation of Magnetic-Apex and Quasi-Dipole coordinates with smooth base vectors. *J. Geophys. Res.* **115**, A08322 (2010). https://doi.org/10.1029/2010JA015326

[19] Richmond, A. D. Ionospheric Electrodynamics Using Magnetic Apex Coordinates. *J. Geomagn. Geoelectr.* **47**, 191–212 (1995). https://doi.org/10.5636/jgg.47.191