Radio Science

RESEARCH ARTICLE
10.1029/2023RS007907

Special Section:
URSI GASS 2023 Special Collection in Radio Science

Electron Density Specification in the Inner Magnetosphere From the Narrow Band Receiver Onboard DSX

Yi-Jiun Su[1], John A. Carilli[1], J. Brent Parham[2], Xiangning Chu[3], Ivan A. Galkin[4], and Gregory P. Ginet[2]

[1]Space Vehicles Directorate, Air Force Research Laboratory, Kirtland Air Force Base, NM, USA, [2]Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA, [3]Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, CO, USA, [4]University of Massachusetts Lowell, Lowell, MA, USA

Abstract Electron density plays an important role in the study of wave propagation and is known to be associated with the index of refraction and radiation belt diffusion coefficients. The primary objective of our investigation is to explore the possibility of implementing an onboard signal processing algorithm to automatically obtain electron densities from the upper hybrid resonance traces of wave spectrograms for future missions. U-Net, developed for biomedical image segmentation, has been adapted as our deep learning architecture with results being compared with those extracted from a more traditional semi-automated method. As a product, electron densities and cyclotron frequencies for the entire DSX mission between 2019 and 2021 are acquired for further analysis and applications. Due to limited space measurements, a synthetic image generator based on data statistics and randomization is proposed as an initial step toward the development of a generative adversarial network in hopes of providing unlimited realistic data sources for advanced machine learning.

Plain Language Summary Electron density is the most important fundamental plasma parameter, however, it is very difficult to directly measure in situ due to spacecraft potential. A convolutional neural network (CNN), developed to recognize features from biomedical images, has been adapted to pull out the resonance traces from space wave receivers automatically specifying densities along satellite orbits. The comparison between computer vision based on a CNN and human vision based on a semi-automated extraction is demonstrated in this paper. With additional development and refinement, our proof-of-concept study may be matured to a level suitable for incorporation into onboard signal processing units to reduce human labor and human-in-the-loop induced operational errors during future space missions.

Key Points:

- U-Net can automatically extract plasma frequencies from in situ wave receivers after it is trained from spectrograms with accurate labels
- Electron densities and cyclotron frequencies for the entire DSX mission are published for other research and applications

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y.-J. Su,
yijiunsu@gmail.com

Citation:

Su, Y.-J., Carilli, J. A., Parham, J. B., Chu, X., Galkin, I. A., & Ginet, G. P. (2024). Electron density specification in the inner magnetosphere from the narrow band receiver onboard DSX. *Radio Science*, *59*, e2023RS007907. https://doi.org/10.1029/2023RS007907

Received 6 NOV 2023
Accepted 20 JAN 2024

Author Contributions:

Conceptualization: Yi-Jiun Su
Data curation: Yi-Jiun Su, J. Brent Parham, Gregory P. Ginet
Formal analysis: Yi-Jiun Su
Investigation: Yi-Jiun Su
Methodology: Yi-Jiun Su, John A. Carilli
Project Administration: Gregory P. Ginet

1. Introduction and Motivation

The inner magnetosphere is regarded as the region surrounded by geomagnetic shells with an equatorial radius of ~7 Earth radii. It is composed of three populations of charged particles: (a) high-energy particles (≳MeV) trapped in the inner and outer radiation belts; (b) median-energy particles (~10–100s keV) drifting around the Earth to form the ring current; and (c) low-energy particles (<few eV) originating from the ionosphere and corotated with the Earth to form the plasmasphere. The investigations of space hazards have focused on high-energy electrons causing deep-dielectric charging when their fluxes are significantly enhanced for a substantial length of time and median-energy electrons depositing their charge on spacecraft surfaces, which can lead to discharges that damage electronics. Though not directly responsible for spacecraft hazards, low-energy particles provide the background medium for current flow and electromagnetic wave propagation, phenomena which affect the spatial, temporal and spectral distributions of damaging particles.

One of the fundamental space plasma parameters, plasma frequency, is proportional to the square root of electron density [$f_{pe}(\text{kHz}) = 8.98\sqrt{n_e\,(\text{cm}^{-3})}$]. Based on the kinetic plasma theory, n_e is obtained from the integration of the phase space density, in which cold plasmas with energy < few eV are the primary contributor. However, it is very difficult to measure in situ n_e directly due to the varying spacecraft potential that may distort the ambient plasma density count provided by onboard particle sensors (e.g., Pfaff et al., 1998). Radio wave instruments such as the low-power resonance relaxation sounder (e.g., Fung et al., 2022) that detect plasma responses at characteristic frequencies (plasma f_{pe}, cyclotron harmonics nf_{ce}, and the upper hybrid $f_{uh} = \sqrt{f_{pe}^2 + f_{ce}^2}$) are less susceptible to the spacecraft-induced effects. The long-duration responses are attributed to slowly propagating electrostatic

Software: Yi-Jiun Su, John A. Carilli, Ivan A. Galkin
Validation: Yi-Jiun Su, Xiangning Chu
Visualization: Yi-Jiun Su
Writing – original draft: Yi-Jiun Su
Writing – review & editing: Yi-Jiun Su, J. Brent Parham, Xiangning Chu, Ivan A. Galkin, Gregory P. Ginet

waves (Muldrew, 1972) reaching out 1,000+ meters into the surrounding plasma where instrument effects are negligible. The very-low-impact indirect n_e measurements are also provided by passive radio detectors of the distinct noise band near the upper hybrid resonance f_{uh} (e.g., LeDocq et al., 1994; Mosier et al., 1973) that, when observed, provides a few percent accurate f_{pe} in combination with the evaluated cyclotron frequency f_{ce} (Benson et al., 2004).

The Air Force Research Laboratory (AFRL) Demonstration and Science Experiments (DSX) flight mission was carried out from 25 June 2019 to 31 May 2021. The primary objective was to study the propagation of artificial space-based very low frequency (VLF) waves and their interaction with energetic particles in the radiation belts and slot region (McCollough et al., 2022). Ray tracing techniques (e.g., Reid et al., 2022) have been well established to study the propagation of electromagnetic waves; however, this type of model requires two fundamental plasma parameters to determine the ray path: plasma (f_{pe}) and cyclotron frequencies (f_{ce}). These parameters are required not just locally, but in a global sense to characterize wave propagation. Moreover, these two parameters are crucial for determining the index of refraction and the diffusion coefficient of radiation belt models (Albert, 2005). Cyclotron frequencies can be determined from the magnetic field magnitudes [$f_{ce}(\text{kHz}) = 2.7386 \times 10^{-2} B(\text{nT})$] measured from the Vector Fluxgate Magnetometer (VMAG) on DSX with high accuracy or obtained from modeled fields when in situ measurements are not available. In this work, passive spectrograms from the Narrow-Band Receiver (NBR, Reinisch et al., 2023) onboard DSX are automatically inspected for f_{uh} signatures of banded emissions to yield an accurate f_{pe} in combination with VMAG measurements. We note that the local f_{ce} and f_{pe} can also be determined from local resonance measurements of active spectrograms (e.g., Figure 28 in Reinisch et al., 2023). Due to limited resources and because it requires a case-by-case examination, the active spectrogram is not discussed here; however, the method developed from our study can be applied to extract resonance traces from the active spectrogram as well.

The severe scarcity of plasmaspheric observations that require sophisticated space-borne equipment flown to high-altitude orbits has been confining advances in empirical and semi-empirical modeling of the plasmasphere. Even nowadays, a widespread notion is that the accuracy of resulting models greatly depends on the availability of training data (e.g., Zhelavskaya et al., 2021). To illustrate the challenges, we will briefly review important discoveries, relevant model developments, and underlying sensor measurements of the plasmasphere.

A historical review of the early plasmasphere research can be found in Lemaire and Gringauz (1998). In 1966, the terms plasmasphere and plasmapause were introduced by Carpenter (1966) and Angerami and Carpenter (1966). Various empirical models have been developed based on data collected from different satellite missions. For example, Berchem and Etcheto (1981) deduced from ISEE-1 observations that the radial distribution of equatorial n_e varies as a power law of the radial distance from the Earth. The equatorial n_e inside and outside the plasmapause, as well as the location of plasmapause, based on the ISEE 1 satellite has been reported by Carpenter and Anderson (1992). The global core plasma (GCP) model, an amalgam of density models in the inner magnetosphere, was released by Gallagher et al. (2000). Based on CRRES observations, empirical plasmasphere and trough density models were published by Sheeley et al. (2001), and plasmapause models were reported by Moldwin et al. (2002) and O'Brien and Moldwin (2003). Denton et al. (2004) reported an electron density model utilizing data collected from the Polar spacecraft. Based on the Radio Plasma Imager (RPI) measurements on the IMAGE satellite, a field-aligned distribution of the plasmaspheric electron density was published by Ozhogin et al. (2012). Most recently, the statistics and empirical models of plasmapause from the Van Allen Probes (VAP) mission were published by Ripoll et al. (2022). Hartley et al. (2023) published an empirical plasma density model in the inner zone ($1 < L < 3$) based on VAP and Arase observations.

Comprehensive physics-based plasmaspheric models coupled with the ionosphere began to blossom in the 1980s. Due to the dense ionosphere, fluid approaches are most common when computational power is limited. For example, the Field Line Interhemispheric Plasma (FLIP) model solves the continuity and momentum equations for major species (Richards & Torr, 1988). The Dynamic Global Core Plasma Model (DGCPM) developed by Ober et al. (1997) solved a continuity equation for the flux tube content of cold ions with a simple convection field model. The time-dependent 3-dimensional plasmaspheric model, SAMI3 (Huba & Krall, 2013), solved for three major ion species and electron temperature equation with an empirical electric field model. Moreover, a 3-dimensional dynamic kinetic plasmaspheric model was developed by Pierrard and Stegen (2008). The kinetic development of plasmaspheric modeling was reviewed by Pierrard et al. (2009).

Due to the popularity of the Kalman filter, the data assimilative models of the plasmasphere have been developed. For example, Jorgensen et al. (2011) adopted DGCPM (Ober et al., 1997) with a simple data assimilation

approach driven by K_P to produce time-dependent plasmaspheric density profiles. Driven by GPS total electron content (TEC) observations, Nikoukar et al. (2015) published a data assimilation technique to specify plasmaspheric density based on an empirical GCP model (Gallagher et al., 2000) and an ionospheric data assimilation four-dimensional (IDA4D) algorithm (Bust et al., 2004). However, this type of model is not mature enough for real-time prediction because space-based measurements are limited for ingestion and model-produced outputs do not have sufficient data for comprehensive validation.

Recently, the fast expansion of neural networks has enabled the development of machine learning-based plasmaspheric models (Ace, 2021; Bianco et al., 2023; Chu et al., 2017a, 2017b; Huang et al., 2022; Zhelavskaya et al., 2017, 2018, 2021) to predict global plasma density in the inner magnetosphere. The accuracy of these models was found to be greatly dependent on the training data quality and availability. The conventional feed-forward neural network (NN) architectures could faithfully represent the training data set and capture the large-scale plasmasphere features such as plume dynamics and nightside erosion. However, their inductive reasoning over no-coverage areas and during geomagnetic storm events was subpar, leading to the "grey box" hybrid approach (Kroll, 2000) that employs feed-forward NNs to initialize and configure a physics-based model (Zhelavskaya et al., 2021).

Given the importance of the data collection for building and validating the plasmaspheric models, we concentrated on the task of reducing the DSX NBR passive spectrograms to the actionable information of the extracted plasma frequency readings. This task requires a human vision model to interpret the spectrogram images automatically. Although semi-automated tools have been developed to extract upper hybrid resonance traces from wave spectrograms for many satellite missions (e.g., Benson et al., 2004; Denton et al., 2012; Goldstein et al., 2014; Kurth et al., 2015; Obana et al., 2021; Trotignon et al., 2010) after the data were downloaded and collected, they require significant attention from trained operators or scientists as a human-in-the-loop requiring impractical investments of time and effort. Additionally, should a reliable fully automated f_{pe} evaluation become feasible, it would become a viable candidate for onboard implementation to provide plasma density monitoring in real-time during future flight missions. Such operational capability is instrumental to adaptive mission planning and spacecraft situation awareness.

Although algorithms for extracting f_{uh} and f_{ce} signatures from passive spectrograms have been previously proposed and tested (e.g., Webb et al., 2007), the rapid growth of the convolution neural network (CNN) solutions in the deep learning domain of computer science has brought new possibilities for adopting well-established pattern recognition and image classification tools to extract signatures from complicated images. Recently, Hasegawa et al. (2019) and Matsuda et al. (2020) utilized residual networks (ResNet) to automatically determine and evaluate the electron density from the Plasma Wave Experiment (PWE) onboard Arase.

The model architecture used for our investigation is demonstrated in Section 3. In order to provide "truth" data for training and validation purposes, n_e are obtained from a traditional semi-automated extraction as described in Section 2. Our secondary objective is to release the fundamental plasma parameters (f_{pe} & f_{ce}) for the entire DSX mission to benefit model development and other scientific research. Section 4 compares the test results from our trained model with the truth data and other model outputs. Compared with other scientific satellite missions, the duration of the DSX flight mission is very short (<2 years) with many data gaps. Image classification models require a large amount of training images with labeled masks. We took advantage of the VAP mission from 2012 to 2019 to develop a synthetic image generator as described in Section 5. Finally, a summary with suggested future development is discussed in Section 6.

2. Data Description and Image Processing

2.1. DSX Data Description

After over a decade of delays, the DSX spacecraft was launched into a 6,000 × 12,000 km orbit with 42° inclination and 5.28 hr orbital period. Spacecraft orbit information in 10 s resolution for the entire mission utilized in this paper are calculated based on the Two-Line Element (TLE) set from Space-Track applying the Simplified General Perturbations 4 (SGP4) propagator adopted in the AE9/AP9/SPM software (Roth, 2022). The geomagnetic parameters, such as McIlwain's L-shell (L_m), magnetic local time (MLT), and magnetic field strength, are included in a standard output. The magnetic field is produced based on an internal field from the International Geomagnetic Reference Field (IGRF, Thébault et al., 2015) plus a quiet external field from Olson and Pfitzer (1974). Magnetic

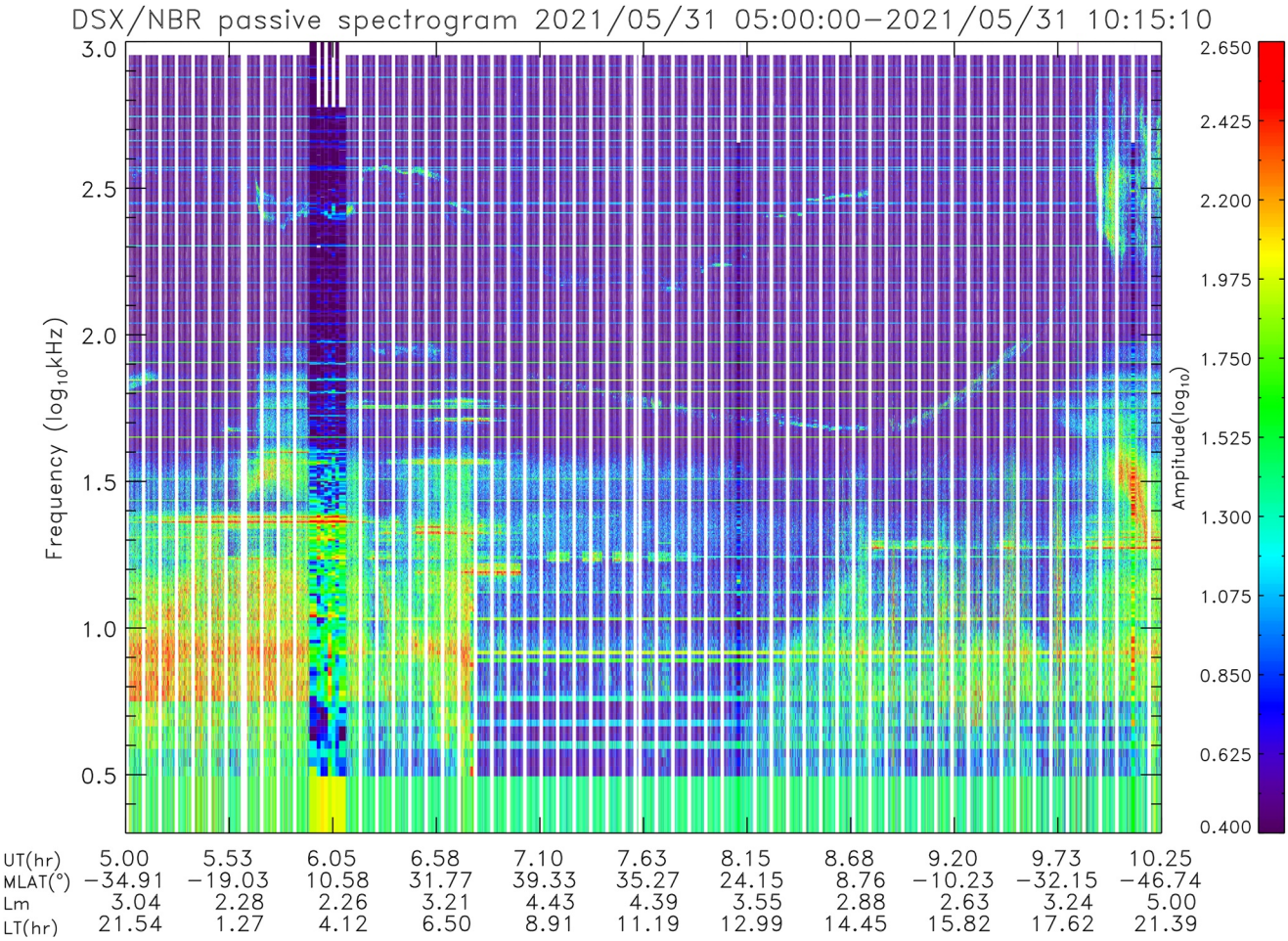

Figure 1. A NBR passive spectrogram based on the frequency sweep from 3 to 1,000 kHz. The wave amplitude in the logarithm is represented by color according to the color bar. The thin white stripes are not considered as data gaps because they do not exceed a critical time length, and are filled in using an interpolation algorithm during the image processing step in Section 2.2. Two discrete resonance traces are visible: (1) the upper hybrid resonance near 300 kHz (~2.5 in log scale); and (2) the gyroresonance near 60 kHz (~1.8 in log scale).

latitudes (MLAT) are calculated in Solar Magnetic (SM) coordinates. The data are organized orbit-by-orbit from perigee to the next perigee. We refer to this data set as the predicted magnetic ephemeris.

The primary data set was collected from NBR, a component of the Transmitter, Narrowband receiver, and Tuner circuitry (TNT) suite (Reinisch et al., 2023). NBR is a proven low-noise heritage analog receiver design from the Radio Plasma Imager (RPI, Reinisch et al., 2000) on the NASA IMAGE spacecraft. A fixed bandwidth of 300 Hz could be tuned at any frequency, scanning through the frequency range with desirable steps for environmental plasma wave measurements or in sync with the transmission frequency and steps for sounding measurements. In fact, we initially used passive spectrograms from RPI to test a simple convolution neural network (Su & Carilli, 2023) because the wave data with corresponding f_{pe} and f_{ce} are publicly available through the Coordinated Data Analysis Web (CDAWeb). Closer inspection of the CDAWeb data set revealed data noise due to the automatic analysis mishaps and notable time mismatches. Yet, after removal of ~1/3 of available f_{pe} and f_{ce} values, our initial test (Su & Carilli, 2023) was satisfactory, paving the way to adopt more sophisticated network architectures reported herein.

An original NBR passive spectrogram per orbit is displayed in Figure 1 as an example. Interested readers can refer to a collection of articles regarding the geospace electromagnetic waves and radiation by Labelle and Treumann (2006). The *y*-axis of Figure 1 represents the frequency sweep from 3 to 1,000 kHz in logarithm. The first spectrogram began on 8 August 2019. Ideally, there should be 3009 complete orbits by the end of the mission on 31 May 2021. Due to data gaps, only 1467 orbital spectrograms (~49%) are available including partial

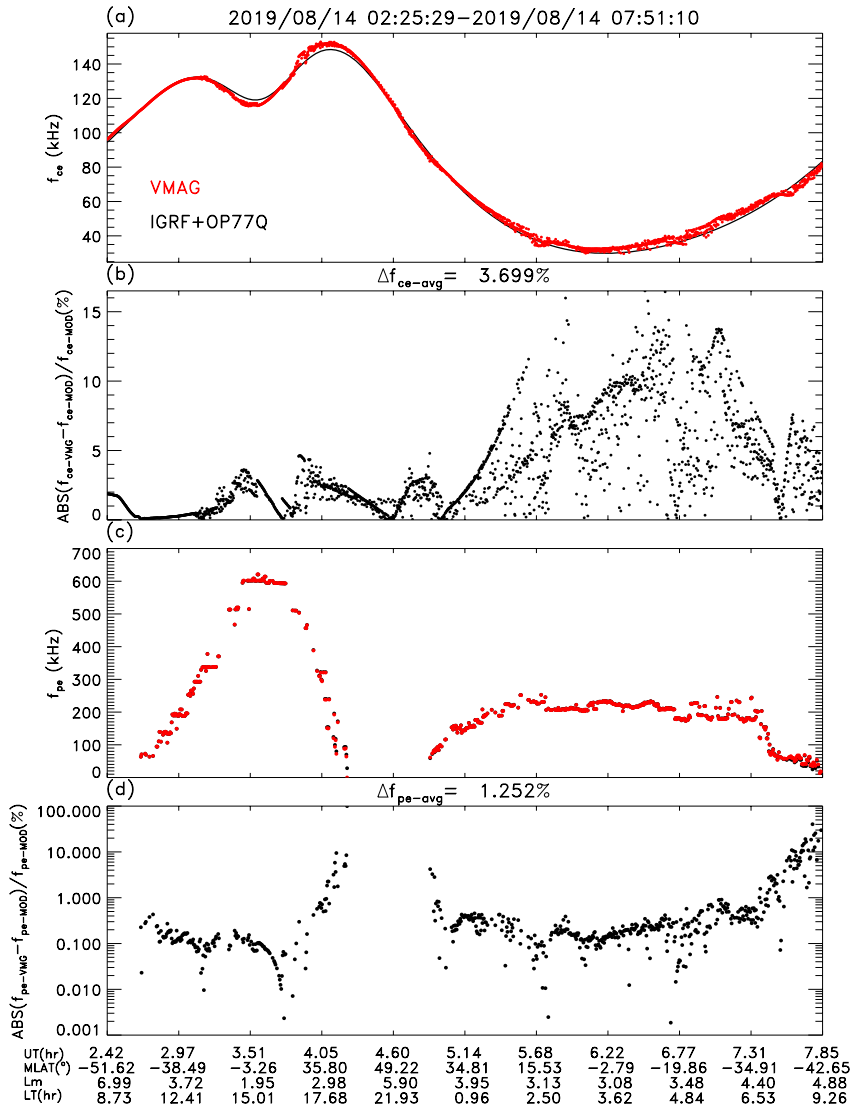

Figure 2. (a) A f_{ce} comparison between VMAG measurements (red) and IGRF + OP77Q modeled values (black) in 10 s resolution over one DSX orbit. (b) The percentage of f_{ce} difference between measurements and modeled values $\Delta f_{ce} = 100\% \times |f_{ceVMAG} - f_{ceIGRF+OP77Q}|/f_{ceIGRF+OP77Q}$. The averaged orbital difference is labeled on top of panel (b). (c) A f_{pe} comparison between values derived from VMAG (red) and IGRF + OP777Q (black). (d) The percentage difference between f_{pe} measurements and modeled values, where the *y*-axis is in log scale. The averaged orbital difference in linear scale is label on top of panel (d).

orbits. ~46% of available spectrograms have data gaps ranging from <10% to >90% of an orbit; however, we still include those partial spectrograms as long as the f_{uh} resonance traces are visible by human inspection to maximize the number of data points in our study.

In order to extract n_e values from the f_{uh} resonance trace, the f_{ce} determination is required. The data used in this paper are based on level-1 VMAG records with 0.05 s resolution (20 Hz). A sliding window of 10 s is used to smooth high-resolution measurements prior to application. Due to data gaps or unrealistic values, ~7.6% of magnetic field data associated with available NBR spectrograms are filled with modeled values. A systematic comparison between VMAG and IGRF + OP77Q values was performed: the difference between measurements and modeled values is ~1.5% on average over all data. The largest average difference over an orbit (3.7%) is shown in Figure 2. The differences are typically larger when the field magnitudes are smaller, that is, further away from the Earth and/or closer to the magnetic equator. The plasma frequency derived from $\sqrt{f_{uh}^2 - f_{ce}^2}$ is displayed in Figure 2c, where f_{uh} was extracted based on the procedures described in Section 2.2. The percentage

of f_{pe} difference is shown in Figure 2d. Although DSX could reach large L-shells, its apogee is ~2.88 R_E where the Earth's internal magnetic field dominates. For missions without in situ magnetometers such as IMAGE, the modeled magnetic fields are reasonable alternatives with relatively good approximations in the plasmasphere based on our investigation.

2.2. Image Processing

As seen in Figure 1, the wave amplitudes from the original frequency sweeps are cluttered with other unrelated radio emissions. For image recognition, removing unnecessary information to bring out the pattern of interest is the first step. Based on the study shown in Figure 2, the minimum of f_{ce} is typically greater than 30 kHz along the DSX orbit. It is safe to remove man-made and natural waves below 30 kHz, because f_{uh} should always be above or equal to f_{ce}. A typical frequency set of 339 steps was chosen from 30 kHz to 1 MHz independent of time. The base-10 logarithmic frequency is defined as the image's y-axis.

The second step is removing or minimizing horizontal stripes attributed to the narrowband observatory noise and external fixed-frequency signals. A selective background intensity can be estimated by manually examining the passive spectrograms, focusing on locations without wave activity. If the 30th percentile of intensities for a frequency step over an orbit is above the background level, intensities of an entire frequency step are subtracted by the 30th percentile level with a lower limit of the background value. This is not the only way to remove constant noise lines (see, e.g., Galkin et al., 2005); however, it is quick to run through available spectrograms above 30 kHz for the entire mission without adjusting parameters after it was initially set.

The third step is to determine the resolution of the image's x-axis. For image segmentation, all training, validation, and test images should have the same size and dimension. However, the actual cadence of the passive-mode NBR measurements varied on orbit, depending on the mission science plan, with the highest 5-s resolution in the survey mode. In order to keep the computational training time reasonable on a 24 GB Nvidia Quadro RTX 6000 Graphics Processing Unit (GPU), a user defined resolution, 1,000 time steps per image ($\Delta T < 20$ s), is selected for each satellite orbit. In addition, the intensity range is reduced to a 0–254 byte table, while 255 is reserved for pixels without data.

The fourth step, the most labor-intensive step, is to label the image as accurately as possible for model training and internal validation processes. The labeled mask also provides the "truth" data for the final test process and the comparison with other model outputs. This step is similar to a traditional method, which extracts f_{uh} by human vision. LabelMe is a 3rd party software developed in Python for the purpose of image labeling. It is easy to install and operate. Users can draw arbitrary polygons around the features of interest and save the coordinates to a structured ASCII JSON data format. Because f_{uh} lines are often very faint, three sets of modeled densities described in Section 5 and the IGRF + OP77Q magnetic field model, as well as three sets of f_{uh}, are plotted for each orbit next to the image to guide human vision when marking the polygon. Human bias might result from this process, however, it is still more accurate than any model output. The polygon is not required to be precise, hence, it is much faster than clicking on the actual feature. An algorithm was then developed to extract the maximum intensity of each time step within the polygon to determine the feature of interest, also known as the mask in image segmentation. From our initial exercises, we learned that test results are better if finer features are labeled in the training process. Hence, the mask associated with each image is constructed based on one pixel per time step. The width of one vertical pixel is ~0.0043 in $\log_{10}$ kHz (i.e., ~1% spectral resolution, $\Delta f/f$) corresponding to ~0.0086 in $\log_{10}$ cm^{-3} (i.e., ~2% density resolution, $\Delta n/n$.) A criterion is also set that the maximum intensities within the polygon should be a factor of 5 or higher than those from nearby pixels in the frequency axis for each time step. There is no interpolation between marked pixels to prevent an artificial smooth plasmapause. Finally, we made use of f_{ce} described in Section 2.1 as the lower limit of the final f_{uh} mask. An image of 339 × 1,000 pixels with associated mask is shown in Figure 3 as an example, where 339 and 1,000 are the number of bins in the y (frequency) and x (time) axes, respectively. Due to the strong auroral kilometric radiations on the top right corner of Figure 3a, the upper hybrid resonance is unidentifiable. Therefore, the human vision intentionally ignored that region when marking the polygon. In order to minimize potential mistakes in image processing, the pixel coordinates of each mask were converted back to frequency and time and plotted on top of the original spectrogram, such as Figure 1, for inspection. This final inspection/verification also prevents bugs in developing complex algorithms. Data and image processing algorithms described in Section 2 were developed in Interactive Data Language (IDL) except for the utilization of LabelMe.

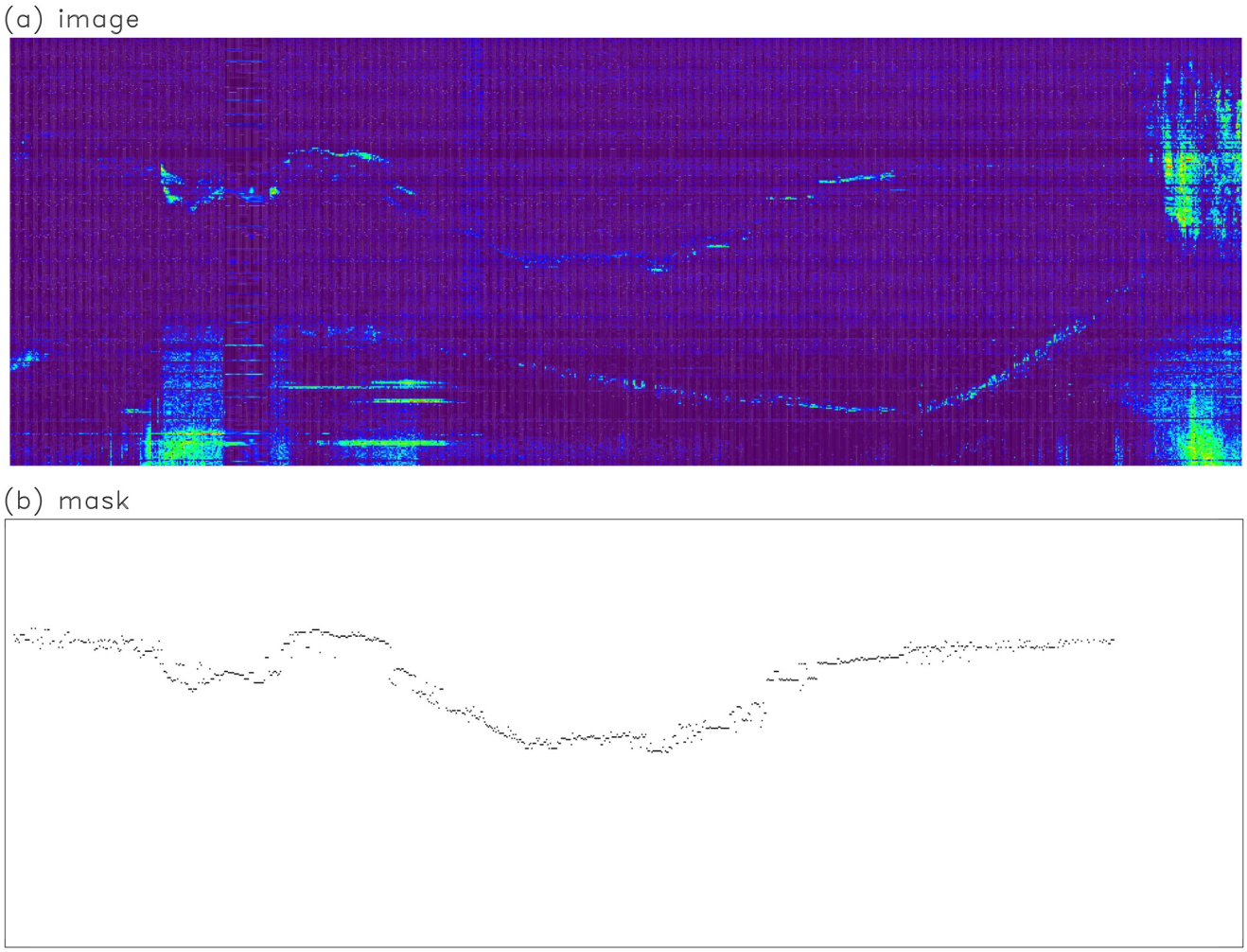

Figure 3. (a) An example after image processing procedures described in Section 2.2 corresponds to Figure 1. (b) The labeled mask associated with panel (a).

2.3. VAP Data Description

The VAP mission, formerly known as the Radiation Belt Storm Probes (RBSP, Stratton et al., 2013), launched on 30 August 2012, was a NASA science mission to study the radiation belts. It consisted of two spacecraft A & B with 618 × 31,414 km orbit, 10.2° inclination, and 8.95 hr orbital period. RBSPB was deactivated on 19 July 2019 before the start of DSX scientific operation in August of the same year, while RBSPA was deactivated on 18 October 2019. The scientific objectives of the VAP mission (Mauk et al., 2013) overlap with DSX. Several conjunction events between DSX and RBSPA were executed, although these are not the focus of this paper. Instead, we took advantage of the long mission duration with ∼14 satellite years from two spacecraft to exercise deep learning models.

The primary VAP data set was obtained from High Frequency Receiver (HFR) spectra level-2 data in survey mode from 10 to 400 kHz. HFR is one component of the Electric and Magnetic Field Instrument Suite and Integrated Science (EMFISIS, Kletzing et al., 2013). An example of a HFR survey spectrogram is shown in Figure 4. Reported HFR spectrograms are not as noisy as those from DSX/NBR data; hence, we did not perform a noise reduction procedure. Six-second density and other parameters inferred by digitizing the traces on the spectrograms are publicly available as EMFISIS level-4 data. The magnetic ephemeris were calculated from various external magnetic field models, and the OP77Q version was adopted in our study. All VAP data described in this paper were downloaded from CDAWeb. From the level-2 data, the HFR spectrograms for every available orbit form the images, and associated f_{uh} masks were extracted from the level-4 data. We note that the accuracy of the reported EMFISIS level-4 data is higher compared to reported IMAGE/RPI densities; however, ∼10% of

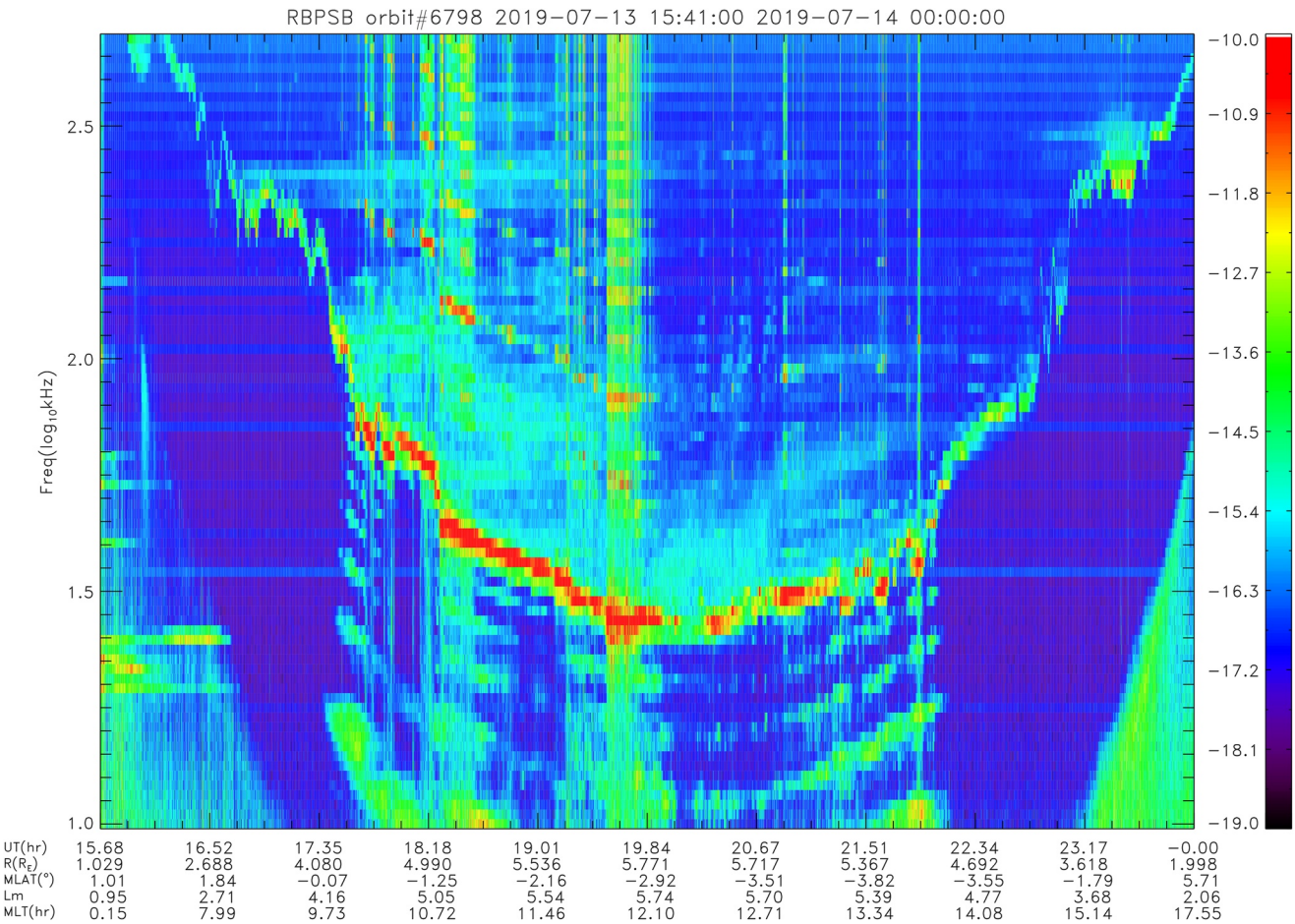

Figure 4. An example of waves survey HFR spectrogram from 10 to 400 kHz.

available EMFISIS density data were removed in our study due to an obvious mismatch between the spectrogram and reported f_{uh} by human inspection. The mismatched orbits with start and stop times are provided in Data Set S1. We greatly encourage readers who use the public data to double-check the data, especially for scientific researches and publications, because unintentional errors can occur to any manual process. Readers can refer to Kurth et al. (2015) regarding potential f_{uh} identification errors. Combining both A and B spacecraft, 7532 image and mask sets remain for our exercise corresponding to 7.7 satellite years. In order to reduce computational training time, each image is reduced to 82 × 100 pixels, where 82 is the number of reported HFR frequency steps and 100 is the number of bins in the x (time) axis. The vertical resolution is ∼0.0208 in $\log_{10}$ kHz, while the horizontal resolution is ∼320 s. The reported HFR survey mode has a time resolution of ∼6.5 s. The CONGRID function in IDL was used to shrink the size of a time array. An example is shown in Figure 5.

3. Model Architecture and Hyperparameters

Our initial exercise was performed on a simple neural network with only three convolution layers trained on IMAGE/RPI images (Su & Carilli, 2023). Each image was decomposed into subimages using a sliding window setup. The results were encouraging, prompting for a more sophisticated network. Next, 50 convolution layers of residual networks, ResNet50, developed for deep residual learning in image recognition (He et al., 2015) was trained on VAP/HFR images. We finally settled on U-Net, a CNN developed for biomedical image segmentation, because it outperforms a sliding window setup with a shorter runtime. Since the development of U-Net in 2015, various versions have been shared publicly. A PyTorch version of U-Net without architecture modification is adopted for our investigation. Interested readers can refer to the original paper by Ronneberger et al. (2015) regarding the model architecture. Although these networks are suitable for multi-class labels, only one class, the f_{uh} mask, is necessary for our problem.

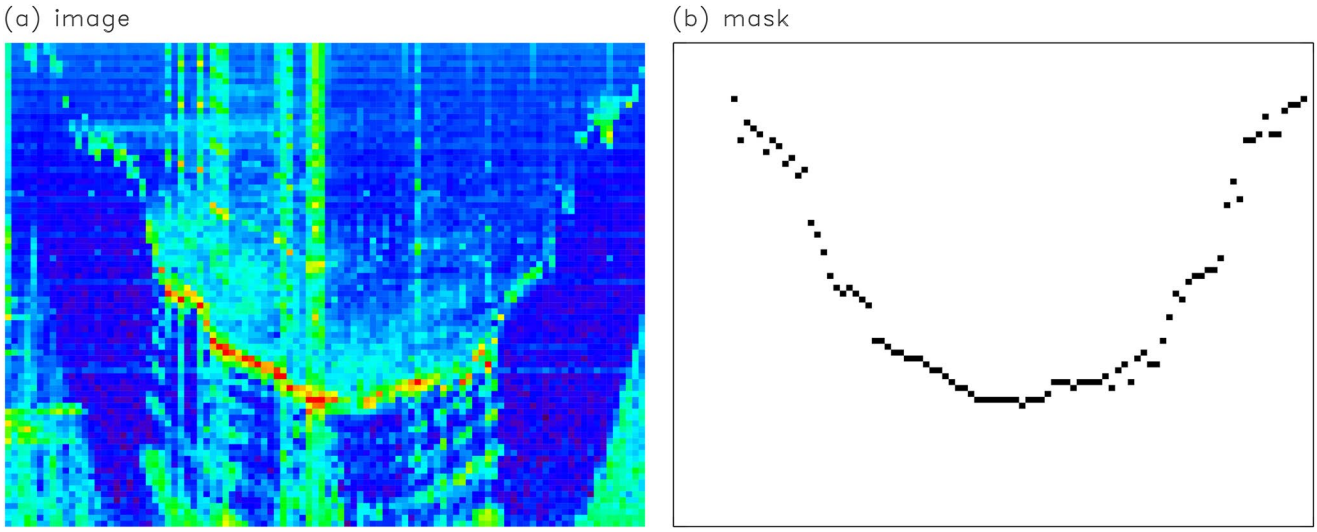

Figure 5. An image (a) and mask (b) set for image segmentation exercise corresponding to the survey spectrogram shown in Figure 4.

Our problem, similar to many biomedical images, is training the model to minimize the error between predicted and labeled masks. Stochastic gradient descent (SGD) has been proven to be an efficient and effective optimization method in recent deep learning advances. Adam (Kingma & Ba, 2015) is adopted as the optimization algorithm in our study. The name Adam came from Adaptive moment estimation in combination of two popular methods: adaptive gradient algorithm (AdaGrad) and root mean square propagation (RMSProp adopted in the original PyTroch U-Net package). Adam requires only first-order gradients with little memory requirement. The PyTorch version is straightforward to implement. A smaller learning rate for the convolution layers is often used in practice when applying SGD. Although a thorough parameter search was not performed, the magic learning rate, 3×10^{-4}, is selected throughout our study based on our initial analysis by varying the rate slightly up and down for comparison.

For a single-class problem, a combination of the binary cross-entropy and dice losses (also known as a combo loss) has been adopted in the original PyTorch U-Net package. We learned that the selection of the loss functions is crucial in order to increase the prediction accuracy, because the weighting parameters are adjusted/updated to reduce the loss in each evaluation. Loss functions are an active research area, and many functions are shared publicly by the community. Ma (2020) presented a systematic taxonomy to sort existing loss functions for image segmentation into four categories: (a) distribution-based loss, such as the cross-entropy loss; (b) region-based loss, such as the dice and Tversky losses; (c) boundary-based loss, such as the Hausdorff distance loss; and (d) compound loss, that is, a combination of different types of loss functions. A comprehensive comparison of different loss functions and various combinations is not the focus of this manuscript, however, we took the suggestion from Ma that using compound loss functions is a better choice overall. After trying a few combinations and parameters, we settled on the weighted sum of the binary focal cross-entropy loss, the focal Tversky loss, and the Hausdorff morphological erosion loss.

We adopted the Kornia version of the focal loss function (Binary Focal LossWithLogits), which is a variant of the binary cross entropy loss by down-weighting the contribution of easy examples enabling learning of harder examples. The focal loss $(-\boldsymbol{\alpha}(1-\boldsymbol{p})^{\boldsymbol{\gamma}}\mathbf{log}(\boldsymbol{p})$, where p is the model's estimated probability) is parameterized by α and γ, which control the class weights and degree of down-weighting of easy-to-classify pixels, respectively (Yeung et al., 2021). The focal loss simplifies to the binary cross-entropy loss when $\gamma = 0$. $\alpha = 0.6$ and $\gamma = 2$ are used in our study.

The Tversky index, $\mathrm{TI} = \mathrm{TP}/(\mathrm{TP} + \alpha\mathrm{FP} + \beta\mathrm{FN})$, is closely related to the dice score, where TP, FP, and FN represent true positives, false positives and false negatives. TI enables optimization for output imbalance by assigning weights α and β to FP and FN, respectively. TI simplifies to the dice score when $\alpha = \beta = 0.5$. When the dice loss function is applied to class imbalanced problems, the resulting segmentation often exhibits high precision, $\mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$, but low recall scores, $\mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$. Assigning a greater weight to FN usually results in a better

balance of precision. Hence, $\alpha = 0.3$ and $\beta = 0.7$ are set for our study, as suggested by Yeung et al. (2021). The focal Tversky loss is defined as $(1 - \mathrm{TI})^{\gamma}$, where $\gamma > 1$ increases the degree of focusing on harder examples. The focal Tversky loss simplifies to the Tversky loss when $\gamma = 1$. Because the governing equation is easy to apply, the algorithm of the focal Tversky loss function with $\gamma = 2$ used in our study was simply modified from the dice loss function included in the original PyTorch U-Net package. The predicted mask during each iteration was converted to a sigmoid function between 0 and 1 before calculating the loss.

In mathematics, the Hausdorff distance (HD) measures how far two subsets of a metric space are from each other. It is widely used in evaluating medical image segmentation methods, and is selected as our prediction metrics. However, existing segmentation methods do not attempt to reduce HD directly. Karimi and Salcudean (2019) proposed three methods to estimate HD loss from the segmentation probability maps produced by a convolution neural network. We apply morphological erosion on the difference between the true and predicted segmentation maps described in Section II.C of Karimi and Salcudean. The parameter α determines how strongly larger segmentation errors are penalized, while the parameter K denotes the total number of erosions. We adopted the class HausdorffERLoss as part of the HausdorffLoss package shared by PatRyg99 on GitHub. As reported by Karimi and Salcudean, $\alpha = 2$ and $K = 10$ are used for our study.

We developed the HD metrics based on the Euclidean distance transform function included in SciPy. This function calculates the distance transform of the input by replacing each non-zero element with its shortest distance to any zero-valued element. The HD maps with non-zero elements from predicted and true masks are combined to obtain the mean value with 95% confidence interval $(\bar{s} + \mathbf{1.96}\sigma/\sqrt{\sum_i s_i})$ representing the symmetric point-wise distance between two masks, where s_i, $\bar{s}$, and σ represent each non-zero distance, the averaged distance, and its standard deviation, respectively. The unit of s_i is the image pixel; hence, the resulting values are associated with the image resolution.

4. Results and Comparison

As described in Section 2.1, there are a total of 1467 image and mask sets including partial spectrograms with data gaps from the DSX mission. Thirty-two images from the end-of-mission (EOM) period between 22 and 31 May 2021 are reserved as the test data set. The remaining 1435 sets from 9 August 2019 to 21 May 2021 are randomly separated into two parts for training and validation, respectively. The batch size is a hyperparameter of gradient descent that defines the number of training samples to work through before the model's internal parameters are updated. It can be one of the crucial steps to reach model's peak performance. Within the training set, 10 images per batch were selected for our exercise. Due to limited GPU memory, higher numbers (>10) of batch have not been tested. However, results based on 10 images/batch are better than those from 1 or 5 image(s)/batch.

The number of epochs is a hyperparameter which defines the number times that the learning algorithm works through the entire training/validation set. The training runtime over 20 epochs took ∼130 min on a 24 GB Nvidia GPU. The training results from each epoch were saved and tested on EOM images. While the probability values are between 0 and 1 in the model, the minimum probability value to consider a predicted mask pixel is set at 0.5. The minimum averaged HD value defines the epoch of the best-case scenario in our study. The mean HDs with 95% confidence interval as a function of epoch from two exercises are shown in Figure 6. These two exercises have the same training/validation samples and hyperparameters.

The model generally reaches the best-case scenario faster with increasing numbers of training samples. Due to limited DSX data, all available data were included in our exercise, for example, 1425 training and 10 validation samples. We note that the training results are different each time when the model was executed with the same hyperparameters, because the weighting parameters are updated internally based on the combination of weighted losses during each evaluation. However, the minimum averaged HD typically reaches ∼5 with the same set of training/validation samples. It is noted that the number of samples for these exercises is very small compared to other CNN applications, ideally millions of image samples with 5% reserved for validation and 10% for testing. By increasing the number of validation samples with decreasing training samples, the model performance with limited test samples is not noticeably improved in our exercises.

Based on the best-case scenario marked as the red circle in Figure 6, we present one test case with the median averaged HD value, 2.056 ± 0.318, from 20 completed spectrograms of the 32 test samples in Figure 7. The

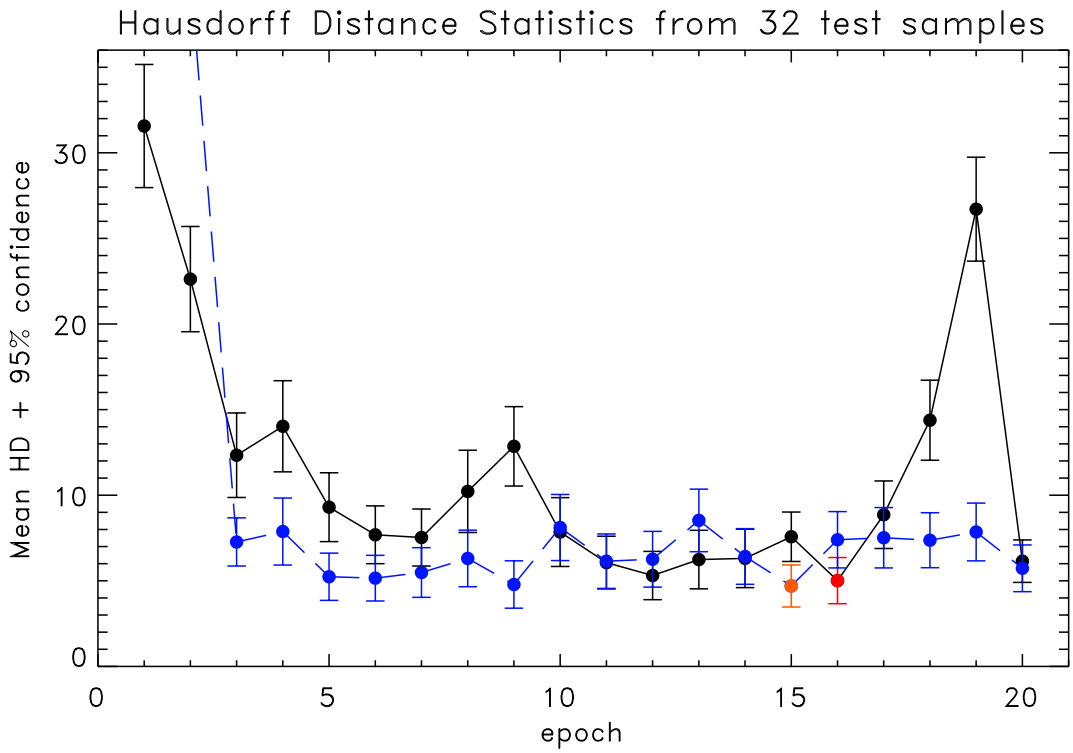

Figure 6. The statistics of Hausdorff distance with 95% confidence interval as a function of training epoch from two different exercises, where the circles represent the mean value, while the vertical bars indicate the 95% confidence interval. The red and orange circles represent the minimum mean HD values, that is, the best training scenario, of each exercise.

digitized images on Panels (a)–(c) are the same, but the black dots in Figures 7b and 7c represent the labeled true mask (also shown as the black dots in Figure 7d) and predicted mask (also shown as the blue dots in Figure 7d), respectively. We note that the predicted mask can report multiple frequency pixels per time step. The final predicted values shown as the red dots in Figure 7d are extracted from the pixel with the maximum intensity or the higher frequency pixel if the intensities are the same. The final predicted values are then converted to predicted n_e as a function of time to compare with the "true" n_e obtained from the semi-automatic feature extraction described in Section 2.2 (i.e., Figure 8c.)

The comparisons between different model outputs to the true logarithmic n_e are shown in Figure 8. The x-axis of all three panels represents the truth data, while the y-axes in Figures 8a–8c represent predicted values from the empirical model (Ozhogin et al., 2012) based on IMAGE/RPI data, the updated neural network-based plasmaspheric model (Chu et al., 2017b) trained from VAP/EMFISIS data, and the U-Net trained from DSX/NBR data, respectively. The diagonal dashed blue line represents the ideal scenario when the predicted values match the truth values. The red lines present the 17th and 83rd percentiles of the data points, approximately the standard deviation of distributions. The prediction efficiency for each comparison is labeled on top of each panel.

$$\mathrm{PE} = 1 - \frac{\mathrm{MSE}}{\sigma_o^2} = 1 - \left[\frac{1}{N}\sum_{i=1}^{N}(p_i - t_i)^2\right] \Big/ \left[\frac{1}{N}\sum_{i=1}^{N}\left(t_i - \bar{t}\right)^2\right]$$

where the mean square error (MSE) is calculated between the predicted and truth n_e, while the standard deviation (σ_o) is obtained from the distribution of truth data. We note that PE calculated in Figure 8c includes only points with coexisting true and predicted values. For this reason, the HD associated with the morphology is selected as our model metric rather than MSE.

The predictions from the U-Net shown in Figure 8c are the closest to the truth, because the model is trained from a data set of the same instrument in a similar space environment of test images. A conclusion can be inferred from our experience that image segmentation would fail if the model were trained from images based

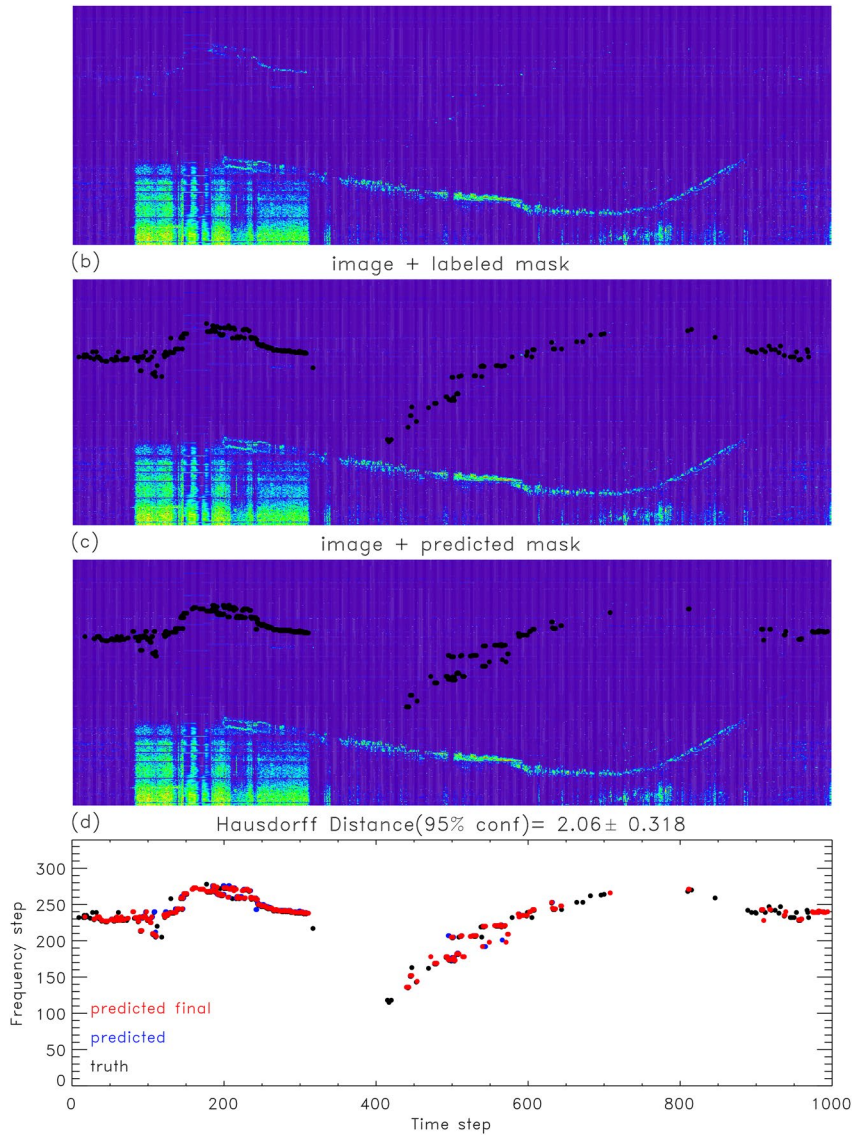

Figure 7. (a) The digitized test image for the time period between 01:17:00 and 06:23:20 UTC on 25 May 2021. The image is repeatedly shown in panels (b) and (c). (b) The labeled mask (black dots) plotted on top of the test image. (c) The predicted mask plotted on top of the test image. (d) Comparison between predicted (blue) and true mask (black), while the final predicted pixel per time step is represented by the red dots.

on different instruments and environments of test images. The predicted n_e from the Chu model saturates at ~1,700 cm^{-3}, because the highest frequency of the VAP/HFR only reached 400 kHz, but DSX/NBR measured up to 1 MHz. Moreover, the magnetic latitude coverage of VAP was smaller than that of DSX due to the difference in inclination of the satellites. We note that Hartley et al. (2018) obtained higher densities above the HFR limitation based on plasmaspheric hiss observations of VAP. The Arase data are planned to be included in the training exercise of the Chu model in the near future.

Although NBR has its heritage from IMAGE/RPI, the IMAGE's orbits (1,000 × 46,000 km with 90° inclination and 14.27 hr orbital period) dramatically differed from DSX's. Unsurprisingly, the Ozhogin empirical model based on IMAGE/RPI data statistics did not perform as well as the U-Net. We demonstrated that U-Net is a powerful tool in deep learning of image segmentation. By deciphering space wave instrument data based on supervised learning, satisfactory results can be obtained even with limited training samples.

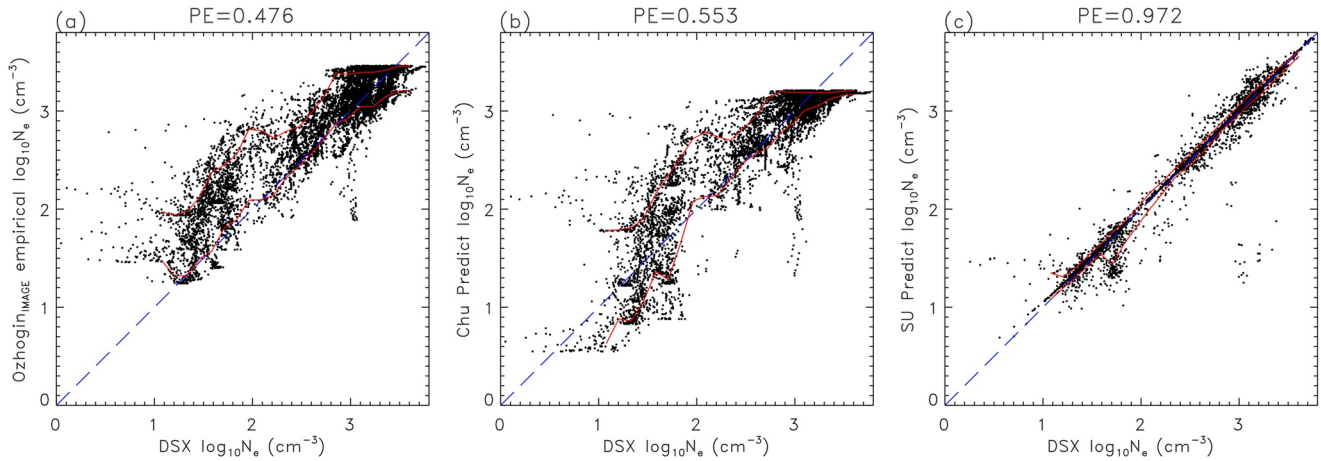

Figure 8. (a) Ozhogin empirical model prediction versus true logarithmic densities. (b) Chu neural network prediction versus true densities. (c) U-Net prediction versus true density. The red lines represent the 17th and 83rd percentiles of distributions. The time period of data is from May 21 to 31 2021, covering 32 satellite orbits.

5. Synthetic Image Generator

Feature-learning image recognizers such as CNN have been criticized for their lower accuracy in identifying previously unseen or excessively noisy feature examples, highlighting the importance of preparing the training data set. From our experience, three primary bottlenecks in building a sufficiently large collection of training spectrogram images were identified: (a) data availability; (b) exhaustive human labor in identifying features from training images; and (c) the noises, that is, the environmental background or instrument noise which is labeled as "0" on the mask.

We propose to utilize existing empirical models combined with statistical distributions derived from large data sets to help resolve the bottlenecks. An algorithm capable of generating realistic synthetic images with labeled f_{uh} would provide an unlimited data supply for advanced training. Although each empirical model has inherent limitations, we are able to generate various versions of synthetic f_{uh} traces using a combination of density models to acquire f_{pe} and the IGRF + OP77Q model to obtain f_{ce}.

Model 1: Ozhogin et al. (2012) represents a smooth field-aligned plasmaspheric density profile without a plasmapause. This model is driven by L-shell and MLAT.

Model 2: Densities in the plasmasphere, plasmapause and plasma trough are prescribed by Carpenter and Anderson (1992). This model is driven by 5 parameters: L-shell, MLT, day of the year, K_P index in the preceding 24 hr, and 13-month averaged sunspot number. Densities in the dusk sector and at $L < 2.25$, which are not specified, can be filled in by Model 1.

Model 3: Sheeley et al. (2001) estimate densities in the plasmasphere and plasma trough; however, plasmapause locations are not specified. This model is driven by L-shell and MLT. Densities at $L < 3$ or $L > 7$ can be filled in by Model 2. The inner and the outer plasmapause boundaries can be determined from Moldwin et al. (2002) and O'Brien and Moldwin (2003). Moldwin et al.'s model is driven by MLT and the maximum K_P from the previous 12 hr. We adopted the simplest format of O'Brien and Moldwin, which is only driven by MLT, although more sophisticated inputs can be specified by max(K_P) between −36 and −2 hr, $\log_{10}$|min(Dst)| from the previous 24 hr, or $\log_{10}$[max(AE)] from the previous 36 hr.

The VAP mission (2012–2019) carrying two identical wave receiver suites was selected as our choice of statistical analysis because of its large data volume as described in Section 2.3. Data set 1 is based on the f_{uh} signal intensity and neighboring pixels of relative intensities as a function of L-shell. Intensities of the remaining pixels from all available spectrograms are binned as a function of frequency, L_M, MLT, and MLAT. Because the wave behaviors are different above and below f_{uh}, two sets of statistics are generated: one above the f_{uh} curves (Data set 2); another below (Data set 3). We note that the histogram of each bin cannot be represented by any analytical function; however, the percentiles of distributions can be interpreted to represent geomagnetic activity levels without including other indices.

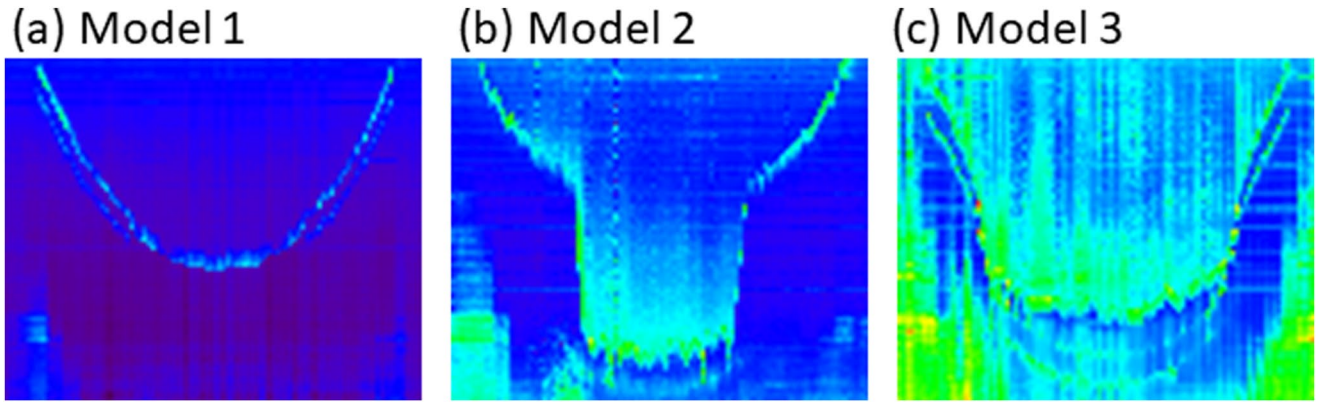

Figure 9. Based on VAP/HFR data statistics, synthetic images generated from (a) Model 1, (b) Model 2, and (c) Model 3.

With Data sets 2 and 3 and numerous random processes, synthetic images can be generated to represent background broadband waves without f_{uh} resonance based on VAP magnetic ephemeris. Independently from the background spectrogram, synthetic f_{uh} as a function of time for each orbit is generated by randomly perturbing outputs from Models 1–3. Similar to the magic eraser concept of recent Pixel smartphones, the locations of synthetic f_{uh} with neighboring pixels are erased from the background image, and replaced by randomly perturbed statistics of Data set 1 based on the perturbed model frequencies. Other discrete or narrow band signals, such as multiple gyro-harmonics or a faint signal below f_{uh}, can be randomly added according to modeled frequencies. Currently, instrument noise has not been included in our synthetic generator, except the consistent noise in the statistics. Instrument noise is typically uniform and easier to remove from real images or add to the synthetic generator. Three synthetic images representing different percentile levels are shown in Figure 9 as examples. Although synthetic images might still look artificial by human vision from experienced scientists, our intention is to explore whether the computer vision trained by an U-Net is able to distinguish between fake and real images.

The U-Net was trained from two sets of images: (Exercise I) 7382 real images based on actual VAP satellite measurements and (Exercise II) 8955 synthetic images with VAP orbital information from the generator. 10% of the images were randomly selected for internal validation. The model hyperparameters are the same as those described in Sections 3 and 4. Trained models from both exercises were tested on the same 150 real images not included in the first training set of images. Although there are more training images in these exercises, the training runtime took ∼60 min for 8–10 epochs on the same video card because the image resolution 82 × 100 is much lower than those presented in Section 4. Two cases of 150 test samples from the best scenarios are demonstrated in Figure 10.

The average and median HDs from 150 test images of Exercise I are 0.853 and 0.675, respectively, while 2.244 and 1.682 resulted from Exercise II. Quantitatively, results from Exercise II epoch 3 are ∼2.6 times worse than those from Exercise I epoch 7. The metrics of mean HD with an associated 95% confidence interval as a function of epoch for these two exercises are displayed in Figure 11.

The synthetic image generator is also modified to utilize the statistics of DSX/NBR data with DSX orbital information, however, the results trained from fake images are ∼8 times worse than those trained from real NBR images (not shown). One of the reasons may be due to a much smaller volume of the NBR data set. Our exercise implies that synthetic images are less sophisticated than real images from nature. Although we have not completely resolved the bottlenecks in machine learning, we demonstrate that a generator based on probabilities and randomization may be a good starting point for a generative adversarial network (GAN).

6. Summary and Discussion

The investigation demonstrated in this paper reveals a potential application of computer vision in deciphering wave instrument data to derive f_{pe}, one of the most important space plasma parameters, in real time. To date, we have performed tests using various CNNs, and U-Net has been proven to outperform methods with a sliding box setup with a quicker runtime. Moreover, the choice of loss functions is critical to improve the prediction

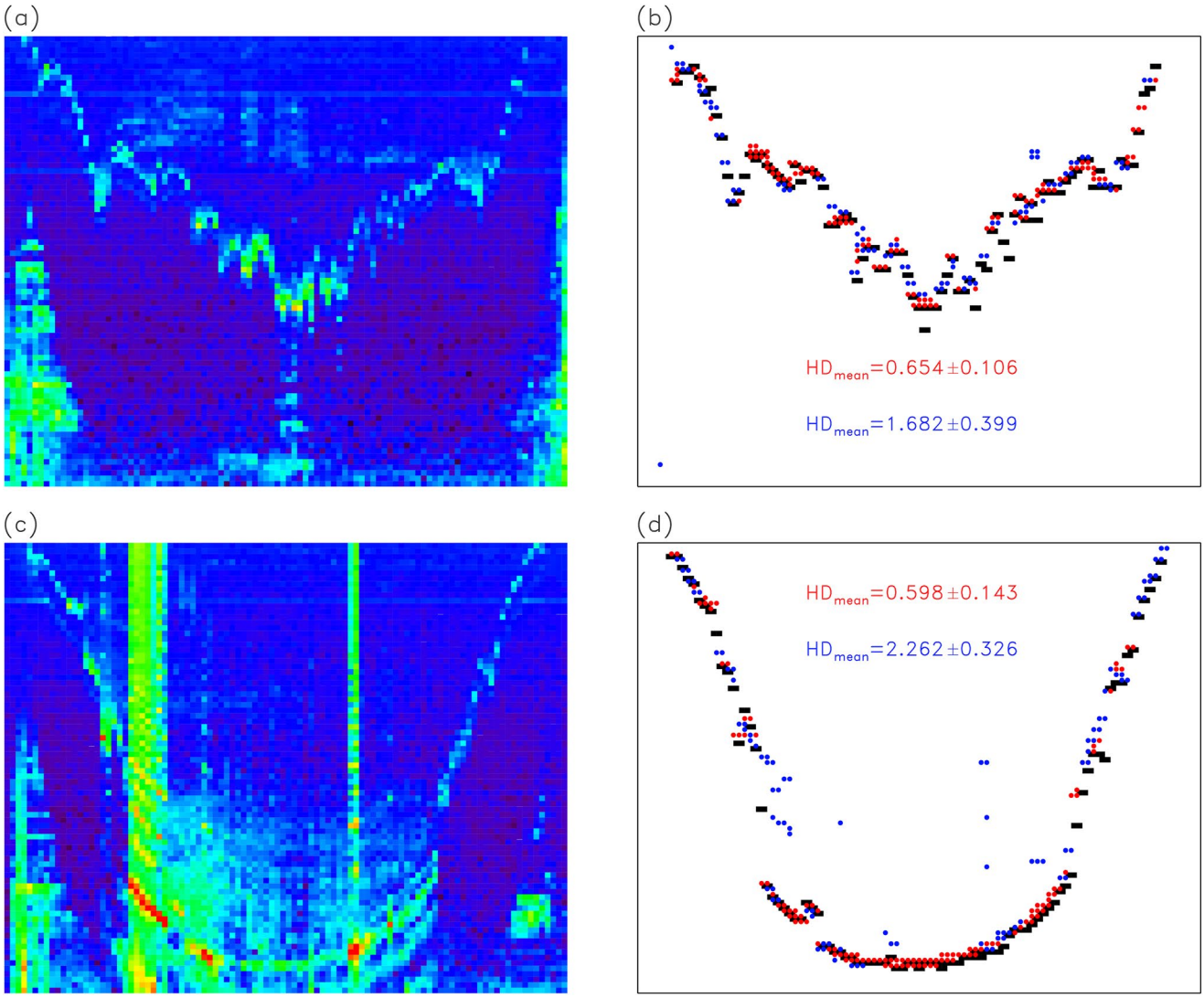

Figure 10. Two test images from actual measurements are shown in panels (a) and (c). The corresponding labeled masks are represented by black squares in panels (b) and (d), while the predicted masks are represented by the red and blue dots of Exercises I and II, respectively.

accuracy. Satisfactory results are obtained from a weighted sum of the binary focal cross-entropy loss, the focal Tversky loss, and the Hausdorff morphological erosion loss. New and improved networks and numerical codes are continuously being developed and shared by the deep learning community. Interested readers are encouraged to perform further evaluations by incorporating off-the-shelf packages or developing their own algorithms.

Although less than 1 year of NBR data are available due to data gaps and human bias may not be completely avoided, the f_{pe} and f_{ce} distributions for the entire DSX mission are acquired as the byproduct of our image processing algorithms, including semi-automated f_{uh} extraction routines. Approximately 100 human hours were required to extract the f_{uh} feature based on a more traditional semi-automated algorithm. By comparison, it would take 13 min on a 24 GB Nvidia GPU to completely analyze 1467 orbits of 339 × 1,000 spectrograms based on a trained deep-learning model. Figures 12a–12c summarize the density distribution as a function of L_M, MLAT, and MLT, respectively, while Figures 12d–12f display f_{ce} distributions. Data forming Figure 12, more accurate than any model output, are included as the Data Set S2 to benefit future machine-learning training and testing, empirical, physics-based, and assimilative model development and validation, or any scientific research requiring fundamental plasma parameters.

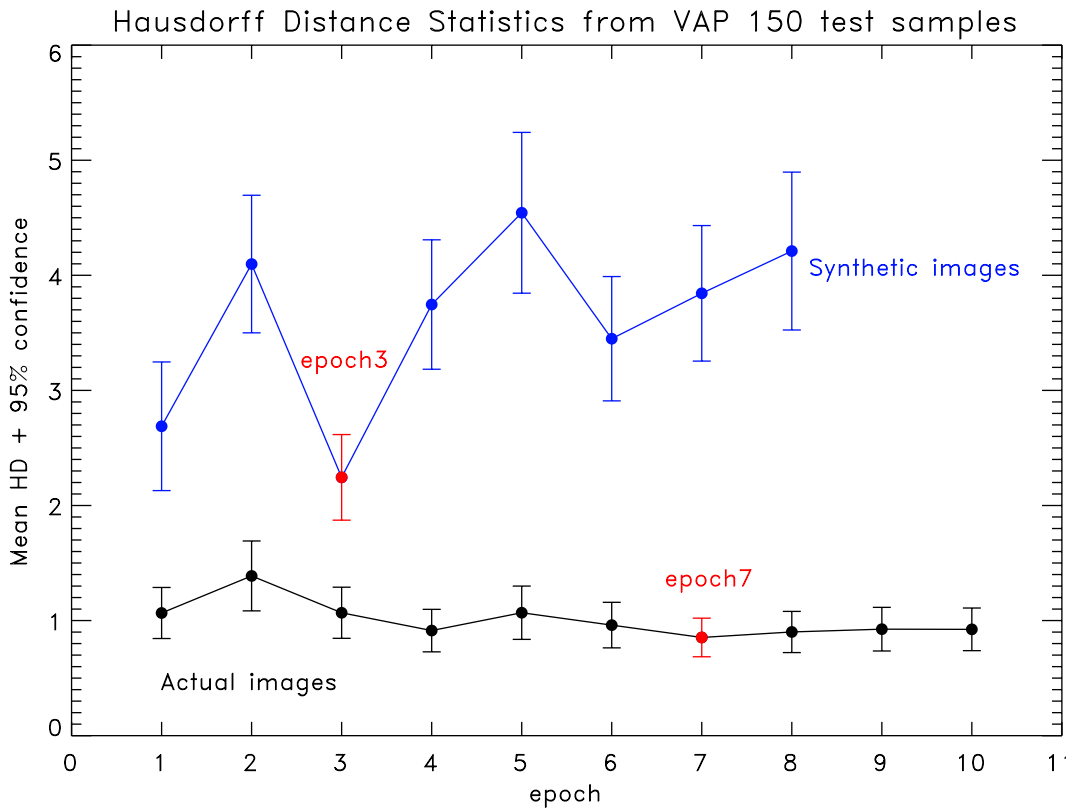

Figure 11. The mean HD with 95 confidence interval as a function of epoch for Exercises I (black) and II (blue). The minimum HD, that is, the best scenario, is marked by a red dot for each exercise.

We encourage the space science community to continue seeking and developing generative adversarial networks which may be trained to distinguish between fake and real images, and eventually be utilized to mass produce realistic synthetic images. The probability approach described in Section 5 may serve as an initial guess solution to assist the network in converging on a final solution more rapidly in order to reduce computational time. Self-supervised or unsupervised learning without human-annotated labels, an active area of artificial intelligence research, may be able to resolve the machine learning bottlenecks, although the methodology is more complex and results may not be as accurate as supervised learning in the current state-of-the-art.

Although many empirical, physics-based, assimilative, and machine learning-based space environment models have been developed, in situ measurements are the most accurate way to obtain the fundamental plasma parameters. Multi-satellite missions require enormous resources in development, in operation, and for analysis. With additional development and refinement, we propose that our proof-of-concept study may be matured to a level suitable for incorporation into onboard signal processing units of future wave receivers. With a flexible onboard software design, the latest trained model could be updated easily during flight as the instruments continuously collect new data to improve training results. In addition, the automatically derived plasma frequencies could be directly downlinked and fed into any assimilative model, significantly improving nowcast/forecast accuracy. While expecting 100% accuracy may not be realistic, the goal is to make incremental progress toward reducing the requirement for human labor and the potential for human-in-the-loop induced operational errors during future missions.

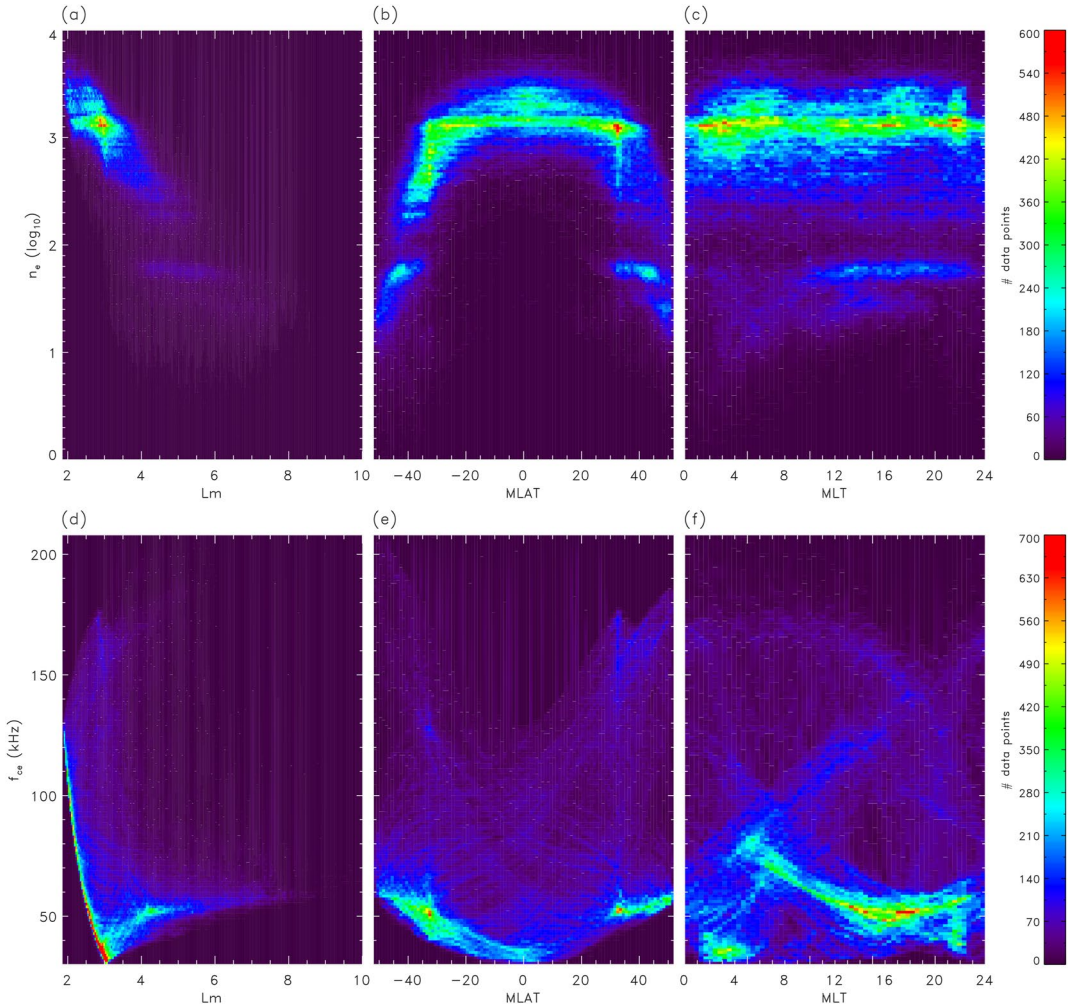

Figure 12. The distribution of electron density $\log_{10}(n_e)$ with respect to (a) L-shell, (b) MLAT, and (c) MLT. The bin sizes are 0.025 for $\log_{10}(n_e)$, 0.02 for L_M, 1° for MLAT, and 0.375 hr for MLT. The f_{ce} distribution with respect to (d) L-shell, (e) MLAT, and (f) MLT. The bin sizes are 1.1125 kHz for f_{ce} and 0.04 for L_M. The MLAT and MLT bins are the same as those in (b) and (c), respectively. The colorbar represents the number of data points in each bin.

Data Availability Statement

The VAP and IMAGE wave data with associated electron densities, cyclotron frequencies, and satellite ephemeris were downloaded from the Coordinated Data Analysis Web (CDAWeb) hosted by the NASA Space Physics Data Facility (SPDF) (n.d.), https://cdaweb.gsfc.nasa.gov.The EMFISIS level-4 data were downloaded with filename "rbsp-x-density_emfisis-l2_yyyymmdd_v1.5.xx.cdf," while the HFR survey level-2 data were downloaded with filename "rbsp-x_hfr-spectra-emfisis-l2_yyyymmdd_v1.x.x.cdf." The 13-month averaged sunspot numbers were obtained from the Royal Observatory of Belgium (n.d.), https://www.sidc.be/silso/newdataset and the K_P indices were obtained from the World Data Center for Geomagnetism, Kyoto University (n.d.), https://wdc.kugi.kyoto-u.ac.jp/kp. TLEs of DSX were downloaded from Space-Track.org. (n.d.), https://space-track.org, and the predicted magnetic ephemeris from the orbit propagator was generated from the software included in AE9/AP9/SPM (n.d.), https://www.vdl.afrl.af.mil/programs/ae9ap9/. The image processing and all figures shown in this manuscript were generated using IDL, while CNNs were adapted and developed in Python. The third-party software in Python used in this study are described in the text and their sources are included in references. Milesial. U-Net: Semantic segmentation with PyTorch (n.d.), https://github.com/milesial/Pytorch-UNet; PatRyg99. HausdorffLoss (n.d.), https://github.com/PatRyg99/HausdorffLoss. Wkentaro. Labelme: Image polygonal annotation with Python (n.d.), https://github.com/wkentaro/labelme.

Acknowledgments
This work is supported by AFRL Space Vehicle Directories. The authors thank the community for making the processed data from NASA's science missions publicly available to benefit this research. The first author thanks colleagues Jay Albert for supplying recent references on neural network-based plasmaspheric density model, Ron Caton for proof-reading the manuscript, and Richard Selesnick for providing useful comments. Additionally, we would like to thank the reviewers, including Dr. David Hartley, for their positive feedback resulting in the improved manuscript. The views expressed are those of the authors and do not reflect the official guidance or position of the United States Government, the Department of Defense or of the United States Air Force. The appearance of external hyperlinks does not constitute endorsement by the United States Department of Defense (DoD) of the linked websites, or the information, products, or services contained therein. The DoD does not exercise any editorial, security, or other control over the information you may find at these locations.

References

Ace, H. (2021). Plasmaspheric dynamics studied using a three-dimensional machine learning-based plasma density model in the inner magnetosphere and the ionosphere (UVM Honors College Senior Theses, 380). Retrieved from https://scholarworks.uvm.edu/hcoltheses/380

Albert, J. M. (2005). Evaluation of quasi-linear diffusion coefficients for whistler mode waves in a plasma with arbitrary density ratio. *Journal of Geophysical Research*, *110*(A3), A03218. https://doi.org/10.1029/2004JA010844

Angerami, J. J., & Carpenter, D. L. (1966). Whistler studies of the plasmapause in the magnetosphere: 2. Electron density and total tube content near the knee in magnetospheric ionization. *Journal of Geophysical Research*, *71*(3), 711–725. https://doi.org/10.1029/JZ071i003p00711

Benson, R. F., Webb, P. A., Green, J. L., Garcis, L., & Reinisch, B. W. (2004). Magnetospheric electron densities inferred from upper-hybrid band emissions. *Geophysical Research Letters*, *31*(20), L20803. https://doi.org/10.1029/2004GL020847

Berchem, J., & Etcheto, J. (1981). Experimental study of magnetospheric convection. *Advances in Space Research*, *1*(1), 179–184. https://doi.org/10.1016/0273-1177(81)90104-6

Bianco, S., Haas, B., & Shprits, Y. Y. (2023). PINE-RT: An operational real-time plasmasphere model. *Frontiers in Astronomy and Space Sciences*, *10*, 1116396. https://doi.org/10.3389/fspas.2023.1116396

Bust, G. S., Garner, T. W., & Gaussiran, H. (2004). Ionospheric data assimilation three-dimensional (IDA3D): A global, multisensory, electron density specification algorithm. *Journal of Geophysical Research*, *109*(A11), A11312. https://doi.org/10.1029/2003JA010234

Carpenter, D. L. (1966). Whistler studies of the plasmapause in the magnetosphere: 1. Temporal variations in the position of the knee and some evidence on plasma motions near the knee. *Journal of Geophysical Research*, *71*(3), 693–709. https://doi.org/10.1029/JZ071i003p000693

Carpenter, D. L., & Anderson, R. R. (1992). An ISEE/whistler model of equatorial electron density in the magnetosphere. *Journal of Geophysical Research*, *97*(A2), 1097–1108. https://doi.org/10.1029/91JA01548

Chu, X., Bortnik, J., Li, W., Ma, Q., Denton, R., Yue, C., et al. (2017a). A neural network model of three-dimensional dynamic electron density in the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, *122*(9), 9183–9197. https://doi.org/10.1002/2017JA024464

Chu, X. N., Bortnik, J., Li, W., Ma, Q., Angelopoulos, V., & Thorne, R. M. (2017b). Erosion and refilling of the plasmasphere during a geomagnetic storm modeled by a neural network. *Journal of Geophysical Research: Space Physics*, *122*(7), 7118–7129. https://doi.org/10.1002/2017JA023948

Coordinated Data Analysis Web (CDAWeb), Space Physics Data Facility. (n.d.). Coordinated data analysis web (CDAWeb), space physics data facility. [IMAGE/RPI and VAP Dataset]. Goddard Space Flight Center, NASA. Retrieved from https://cdaweb.gsfc.nasa.gov

Denton, R. E., Menietti, J. D., Goldstein, J., Young, S. L., & Anderson, R. R. (2004). Electron density in the magnetosphere. *Journal of Geophysical Research*, *109*(A9), A09215. https://doi.org/10.1029/2003JA010245

Denton, R. E., Wang, Y., Webb, P. A., Tengdin, P. M., Goldstein, J., Redfern, J. A., & Reinisch, B. W. (2012). Magnetospheric electron density long-term (>1 day) refilling rates inferred from passive radio emissions measured by IMAGE RPI during geomagnetically quiet times. *Journal of Geophysical Research*, *117*(A3), A03221. https://doi.org/10.1029/2011JA017274

Fung, S. F., Benson, R. F., Galkin, I. A., Green, J. L., Reinisch, B. W., Song, P., & Sonwalkar, V. (2022). Radio-frequency imaging techniques for ionospheric, magnetospheric, and planetary studies. In Y. Colado-Vega, D. Gallagher, H. Frey, & S. Wing (Eds.), *Understanding the space environment through global measurements* (pp. 101–216). Elsevier. https://doi.org/10.1016/B978-0-12-820630-0.00006-4

Galkin, I. A., Khmyrov, G. M., Kozlov, A. P., Tilton, J. C., Fung, S. F., & Plaza, A. (2005). A pre-attentive vision model for data prospecting. In *2nd international conference on cybernetics and information technologies, systems, and applications* (pp. 24–29). CITSA.

Gallagher, D. L., Craven, P. D., & Comfort, R. H. (2000). Global core plasma model. *Journal of Geophysical Research*, *105*(A8), 18819–18833. https://doi.org/10.1029/1999JA000241

Goldstein, J., De Pascuale, S., Kletzing, C., Kurth, W., Genestreti, K. J., Skoug, R. M., et al. (2014). Simulation of Van Allen Probes plasmapause encounter. *Journal of Geophysical Research: Space Physics*, *119*(9), 7464–7484. https://doi.org/10.1002/2014JA020252

Hartley, D. P., Cunningham, G. S., Ripoll, J.-F., Malaspina, D. M., Kasahara, Y., Miyoshi, Y., et al. (2023). Using Van Allen Probes and Arase observations to develop an empirical plasma density model in the inner zone. *Journal of Geophysical Research: Space Physics*, *128*(3), e2022JA031012. https://doi.org/10.1029/2022JA031012

Hartley, D. P., Kletzing, C. A., De Pascuale, S., Kurth, W. S., & Santolik, O. (2018). Determining plasmaspheric densities from observations of plasmaspheric hiss. *Journal of Geophysical Research: Space Physics*, *123*(8), 6679–6691. https://doi.org/10.1029/2018JA025658

Hasegawa, T., Matsuda, S., Kumamoto, A., Tsuchiya, F., Kasahara, Y., Miyoshi, Y., et al. (2019). Automatic electron density determination by using a convolutional neural network. *IEEE Access*, *7*, 163384–163394. https://doi.org/10.1109/ACCESS.2019.2951916

He, K., Zhang, Z., Ren, S., & Sun, J. (2015). Deep residual learning from image recognition. *Computer Vision and Pattern Recognition (cs.CV)*. https://doi.org/10.48550/arXiv.1512.03385

Huang, S., Li, W., Shen, X.-C., Ma, Q., Chu, X., Ma, D., et al. (2022). Application of recurrent neural network to modeling Earth's global electron density. *Journal of Geophysical Research: Space Physics*, *127*(9), e2022JA030695. https://doi.org/10.1029/2022JA030695

Huba, J. D., & Krall, J. (2013). Modeling the plasmasphere with SAMI3. *Geophysical Research Letters*, *40*(1), 6–10. https://doi.org/10.1029/2012GL054300

IRENE-AE9/AP9/SPM. (n.d.). IRENE-AE9/AP9/SPM: Radiation belt and space plasma specification models [Software]. https://www.vdl.afrl.af.mil/programs/ae9ap9/

Jorgensen, A. M., Ober, D., Koller, J., & Friedel, R. H. W. (2011). Specification of the Earth's plasmasphere with data assimilation. *Advances in Space Research*, *47*(12), 2152–2161. https://doi.org/10.1016/j.asr.2010.06.013

Karimi, D., & Salcudean, S. E. (2019). Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. In *Image and Video Processing (EESS.IV)*. https://doi.org/10.48550/arXiv.1904.10030

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *Machine Learning*. https://doi.org/10.48550/arXiv.1412.6980

Kletzing, C. A., Kirth, W. S., Acuna, M., MacDowall, R. J., Torbert, R. B., Averkamp, T., et al. (2013). The electric and magnetic field instrument suite and integrated science (EMFISIS) on RBSP. *Space Science Reviews*, *179*(1–4), 127–181. https://doi.org/10.1007/s11214-013-9993-6

Kroll, A. (2000). Grey-box models: Concepts and application. In *New frontiers in computational intelligence and its applications* (Vol. 57, pp. 42–51). IOS Press.

Kurth, W. S., De Pascuale, S., Faden, J. B., Kletzing, C. A., Hospodarsky, G. B., Thaller, S., et al. (2015). Electron densities inferred from plasma wave spectra obtained by the wave instrument on Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *120*, 904–915. https://doi.org/10.1002/2014JA020857

J. W. LaBelle, & R. A. Treumann (Eds.) (2006). *Geospace electromagnetic waves and radiation*, *Lecture notes in physics* (Vol. 687). Springer. https://doi.org/10.1007/b11580119

LeDocq, M. J., Gurnett, D. A., & Anderson, R. R. (1994). Electron number density fluctuations near the plasmapause observed by the CRRES spacecraft. *Journal of Geophysical Research*, *99*(A12), 23661–23671. https://doi.org/10.1029/94JA02294

J. F. Lemaire, & K. I. Gringauz (Eds.) (1998). *The Earth's plasmasphere*. Cambridge University Press.
Ma, J. (2020). Segmentation Loss Odyssey. In *Image and video processing (EESS.IV)*. https://doi.org/10.48550/arXiv.2005.13449
Matsuda, S., Hasegawa, T., Kumamoto, A., Tsuchiya, F., Kasahara, Y., Miyoshi, Y., et al. (2020). Detection of UHR frequencies by a convolutional neural network from Arase/PWD data. *Journal of Geophysical Research: Space Physics*, *125*(10), e2020JA028075. https://doi.org/10.1029/2020JA028075
Mauk, B. H., Fox, N. J., Kanekal, S. G., Kessel, R. L., Sibeck, D. G., & Ukhorskly, A. (2013). Science objectives and rationale for the radiation belt storm probes mission. *Space Science Reviews*, *179*(1–4), 3–27. https://doi.org/10.1007/s11214-012-9908-y
McCollough, J. P., Miyoshi, Y., Ginet, G. P., Johnston, W. R., Su, Y.-J., Starks, M. J., et al. (2022). Space-to-space very low frequency radio transmission in the magnetosphere using the DSX and Arase satellites. *Earth Planets and Space*, *74*(1), 64. https://doi.org/10.1186/s40623-022-01605-6
Milesial. U-Net. (n.d.). Milesial. U-Net: Semantic segmentation with PyTorch [Software]. https://github.com/milesial/Pytorch-UNet
Moldwin, M. B., Downward, L., Rassoul, H. K., Amin, R., & Anderson, R. R. (2002). A new model of the location of the plasmapause: CRRES results. *Journal of Geophysical Research*, *107*(A11), 1339. https://doi.org/10.1029/2001JA009211
Mosier, S. R., Kaiser, M. L., & Brown, L. W. (1973). Observations of noise bands associated with the upper hybrid resonance by the IMP 6 radio astronomy experiment. *Journal of Geophysical Research*, *78*(10), 1673–1679. https://doi.org/10.1029/JA078i010p01673
Muldrew, D. B. (1972). Electron resonances observed with topside sounders. *Radio Science*, *7*(8–9), 779–789. https://doi.org/10.1029/RS007i008p00779
Nikoukar, R., Bust, G., & Murr, D. (2015). A novel data assimilation technique for the plasmasphere. *Journal of Geophysical Research: Space Physics*, *120*(10), 8470–8485. https://doi.org/10.1002/2015JA021455
Obana, Y., Miyashita, Y., Maruyama, N., Shinbori, A., Nosé, M., Shoji, M., et al. (2021). Field-aligned electron density distribution of the inner magnetosphere inferred from coordinated observations of Arase and Van Allen Probes. *Journal of Geophysical Research: Space Physics*, *126*(10), e2020JA029073. https://doi.org/10.1029/2020JA029073
Ober, D. M., Horwitz, J. L., & Gallagher, D. L. (1997). Formation of density troughs embedded in the outer plasmasphere by subauroral ion drift events. *Journal of Geophysical Research*, *102*(A7), 14595–14602. https://doi.org/10.1029/97JA01046
O'Brien, T. P., & Moldwin, M. B. (2003). Empirical plasmapause models from magnetic indices. *Geophysical Research Letters*, *30*(4), 1152. https://doi.org/10.1029/2002GL016007
Olson, Q. P., & Pfitzer, K. A. (1974). A quantitative model of the magnetospheric magnetic field. *Journal of Geophysical Research*, *79*(25), 3739–3748. https://doi.org/10.1029/JA079i025p03739
Ozhogin, P., Tu, J., Song, P., & Reinisch, B. W. (2012). Field-aligned distribution of the plasmaspheric electron density: An empirical model derived from the IMAGE RPI measurements. *Journal of Geophysical Research*, *117*(A6), A06225. https://doi.org/10.1029/2011JA017330
PatRyg99. HausdorffLoss. (n.d.). PatRyg99. HausdorffLoss [Software]. Retrieved from https://github.com/PatRyg99/HausdorffLoss
R. F. Pfaff, J. E. Borovsky, & D. T. Young (Eds.) (1998). *Measurement techniques in space plasmas: Particles*, *Geophysical monograph* (Vol. 102). American Geophysical Union. https://doi.org/10.1029/GM102
Pierrard, V., Goldstein, J., André, N., Jordanova, V. K., Kotova, G. A., Lemaire, J. F., et al. (2009). Recent progress in physics-based models of the plasmasphere. *Space Science Reviews*, *145*(1–2), 193–229. https://doi.org/10.1007/s11214-008-9480-7
Pierrard, V., & Stegen, K. (2008). A three-dimensional dynamic kinetic model of the plasmasphere. *Journal of Geophysical Research*, *113*(A10), A10209. https://doi.org/10.1029/2008JA013060
Reid, R. A., Marshall, R. A., Starks, M. J., Usanova, M. E., Wilson, G. R., Johnston, W. R., et al. (2022). Active VLF transmission experiments between DSX and VPM spacecraft. *Journal of Geophysical Research: Space Physics*, *127*(4), e2021JA030087. https://doi.org/10.1029/2021JA030087
Reinisch, B. W., Haines, D. M., Bible, K., Cheney, G., Galkin, I. A., Huang, X., et al. (2000). The radio plasma imager investigation on the IMAGE spacecraft. *Space Science Review, IMAGE Special Issue*, *91*(1/2), 319–359. https://doi.org/10.1023/A:1005252602159
Reinisch, B. W., Song, P., Galkin, I. A., Stelmash, S. S., Roche, K., Khmyrov, G. M., et al. (2023). The VLF transmitter, narrowband receiver, and tuner investigation on the DSX spacecraft. *Journal of Geophysical Research: Space Physics*, *128*(7), e2022JA030435. https://doi.org/10.1029/2022JA030435
Richards, R. G., & Torr, D. G. (1988). Hydrodynamic models of the plasmasphere. In T. E. Moore, J. H. WaiteJr., T. W. Moorehead, & W. B. Hanson (Eds.), *Geophysical monograph series* (Vol. 44, pp. 67–77). https://doi.org/10.1029/GM044p0067
Ripoll, J.-G., Thaller, S. A., Hartley, D. P., Cunningham, G. S., Pierrard, V., Kurth, W. S., et al. (2022). Statistics and empirical models of the plasmasphere boundaries from the Van Allen Probes for radiation belt physics. *Geophysical Research Letters*, *49*(21), e2022GL101402. https://doi.org/10.1029/2022GL101402
Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Computer vision and pattern recognition (cs.CV)*. https://doi.org/10.48550/arXiv.1505.04597
Roth, C. (2022). IRENE: AE9/AP9/SPM radiation environment model version 1.57.004: User's guide. Retrieved from https://vdl.afrl.af.mil/programs/ae9ap9/files/package/IRENE_Users_Guide.pdf
Sheeley, B. W., Moldwin, M. B., Rassoul, H. K., & Andersson, R. R. (2001). An empirical plasmasphere and trough density model: CRRES observations. *Journal of Geophysical Research*, *106*(A11), 25631–25641. https://doi.org/10.1029/2000JA000286
Solar Influences Data Analysis Center, Royal Observatory of Beigiumm. (n.d.). Solar influences data analysis center, royal observatory of Beigiumm [Sunspot Number Version 2.0 Dataset]. Retrieved from https://www.sidc.be/silso/newdataset
Space-Track.org. (n.d.). Space-Track.org. [Two-Line Elements Dataset]. Retrieved from https://space-track.org
Stratton, J. M., Harvey, R. J., & Heyler, G. A. (2013). Mission overview for the radiation belt storm probes mission. *Space Science Reviews*, *179*(1–4), 29–57. https://doi.org/10.1007/s11214-012-9933-x
Su, Y.-J., & Carilli, J. A. (2023). Plasma frequency determination from in-situ wave receiver. Extended abstract. In *The XXXVth general assembly and scientific symposium of the international union of radio science (URSI GASS), IEEE Xplore*. https://doi.org/10.23919/URSIGASS57860.2023.10265524
Thébault, E., Finlay, C. C., Beggan, C. D., Alken, P., Aubert, J., Barrois, O., et al. (2015). International geomagnetic reference field: The 12th generation. *Earth Planets and Space*, *67*(1), 79. https://doi.org/10.1186/s40623-015-0228-9
Trotignon, J. G., Décréau, P. M. E., Rauch, J. L., Vallières, X., Rochel, A., Kougblénou, S., et al. (2010). The WHISPER relaxation sounder and the CLUSTER active archive. In H. Lakkso, M. Taylor, & C. P. Escouber (Eds.), *Studying the Earth's space plasma environment*, *Astrophysics and space science proceeding*. (pp. 185–208). Springer. https://doi.org/10.1007/978-90-481-3499-1_12
Webb, P., Benson, R., Denton, R., Goldstein, J., Garcia, L., & Reinisch, B. (2007). An inner magnetospheric electron density database determined from IMAGE/RPI passive dynamic spectra. *AGU fall meeting abstracts* (Vol. 88, p. SM12A-04).

Wkentaro. Labelme: Image polygonal annotation with Python. (n.d.). Wkentaro. Labelme: Image polygonal annotation with Python [Software]. Retrieved from https://github.com/wkentaro/labelme

World Data Center for Geomagnetism. (n.d.). World data center for geomagnetism, Kyoto university. [Kp index Dataset]. Retrieved from https://wdc.kugi.kyoto-u.ac.jp/kp

Yeung, M., Sala, E., Schönlieb, C.-B., & Rundo, L. (2021). Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. In *Image and video processing (EESS.IV)*. https://doi.org/10.48550/arXiv.2102.04525

Zhelavskaya, I. S., Aseev, N. A., & Shprits, Y. Y. (2021). A combined neural network- and physics-based approach for modeling plasmasphere dynamics. *Journal of Geophysical Research: Space Physics*, *126*(3), e2020JA028077. https://doi.org/10.1029/2020JA028077

Zhelavskaya, I. S., Shprits, Y. Y., & Spasojevic, M. (2017). Empirical modeling of the plasmasphere dynamics using neural networks. *Journal of Geophysical Research: Space Physics*, *122*(11), 11227–11244. https://doi.org/10.1002/2017JA024406

Zhelavskaya, I. S., Shprits, Y. Y., & Spasojevic, M. (2018). Reconstruction of plasma electron density from satellite measurements via artificial neural networks. In E. Camporeale, S. Wing, & J. R. Johnson (Eds.), *Machine learning techniques for space weather* (pp. 301–327). Elsevier Inc. https://doi.org/10.1016/B978-0-12-811788-0.00012-3