REGRESSION-BASED PHYSICS-INFORMED NEURAL NETWORKS (*Reg-PINNs*) FOR MAGNETOPAUSE TRACKING

Po-Han K. Hou*
Department of Earth Science and Engineering
Imperial College London
London, SW11 2DL
kh124@ic.ac.uk

Sung-Chi C. Hsieh
Department of Electrical Engineering
University of Leicester
Leicester, LE1 7RH
sch45@student.le.ac.uk

ABSTRACT

Previous research in the scientific field has utilized statistical empirical models and machine learning to address fitting challenges. While empirical models have the advantage of numerical generalization, they often sacrifice accuracy. However, conventional machine learning methods can achieve high precision but may lack the desired generalization. The article introduces a Regression-based Physics-Informed Neural Networks (Reg-PINNs), which embeds physics-inspired empirical models into the neural network's loss function, thereby combining the benefits of generalization and high accuracy. The study validates the proposed method using the magnetopause boundary location as the target and explores the feasibility of methods including Shue et al.[1998], a data overfitting model, a fully-connected networks, Reg-PINNs with Shue's model, and Reg-PINNs with the overfitting model. Compared to Shue's model, this technique achieves approximately a 30% reduction in RMSE, presenting a proof-of-concept improved solution for the scientific community.

Keywords physics-informed neural networks · regression · magnetopause tracking

1 Introduction

Over the past 50 years, various theories have been developed to describe the pressure equilibrium boundary between Earth's magnetosphere and the solar wind, known as the magnetopause. These include an investigation of the pressure balance between the magnetosphere and solar wind in 1969 [11] and an empirical model for the magnetopause proposed in 1997 and 1998 [12, 13], which has become the widely used approach over the past 25 years.[1] Additionally, a method using support vector regression (SVR) was introduced in [16] to estimate the magnetopause position.

Traditional empirical models excel in generalization but sacrifice precision, while machine learning models achieve high precision during training but may lack the generalization capability of empirical models. Thus, this study aims to combine the strengths of both approaches by proposing a novel fitting algorithm called Regression-based Physics-Informed Neural Networks (Reg-PINNs).

Reg-PINNs are inspired by Physics-Informed Neural Networks (PINNs), which have been used for high-resolution time evolution or spatial state simulations of physical phenomena governed by ordinary and partial differential equations. This study introduces a new approach that utilizes a regression-based equation form for training PINNs, combining empirical physics-based least squares fitting with neural networks to enhance both precision and generalization. The proposed method was tested against the widely used empirical model in [13], achieving approximately a 30% reduction in root mean square error (RMSE).

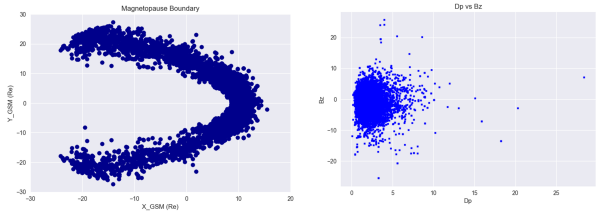
2 Data Collection and Preprocessing

A total of 34,998 magnetopause in-situ crossing data points were derived from NASA OMNIWeb Data Explorer [8], including THEMIS (28,634) [15], Geo-Tail (5,764), ISEE, and IMP 8. As shown in Figure 1a, the inverted C-shaped data pattern represents the in-situ crossing data points, where the satellite detected a change of particle composition, ion density and temperature. Meanwhile, Figure

*Department of Space Science and Engineering at National Central University, Chungli, Taiwan

1b illustrates the relationship between the magnetic field and the solar wind dynamic pressure (D_p). For each data point, the corresponding timestamps were identified in the 5-minute averages and integrated with solar wind ion flux, density, composition, and flow pressure recorded at five-minute intervals. To examine the model's generalizability within a specific region, we integrated the processed data for each bin.

Each bin encompasses north-south interplanetary magnetic field (IMF B_z) values spanning 3 nT, with a shift of 1 nT for each subsequent bin. For D_p, each bin includes two values with a shift of 0.5 nPa for each subsequent bin. The dataset range used in this study is consistent with that in [13], specifically -18 nT < B_z < 15 nT and 0.5 nPa < D_p < 8.5 nPa. The data preprocessing stage ensured the dataset was suitable for training and comparison with empirical models.



(a) Distribution of crossing data. (b) D_p v.s. B_z.

Figure 1: Magnetopause crossing dataset visualization.

3 Methodologies

There have been numerous studies discussing the construction of empirical models for magnetopause location, including notable works in 1993 [10], 1997 and 1998 [12, 13], and 2002 [3]. In this study, we propose the use of Reg-PINNs, which aim to combine the strengths of both empirical models and machine learning methods, achieving high accuracy and generalization capabilities.

3.1 Physics-Inspired Empirical Model

The fitting function of the magnetopause boundary, as shown in this study adopts the formulation proposed by Shue [12] as expressed in Equation 1 and utilize the parameters from [13] as the baseline, as shown in Equation 2. Shue's model integrated data from satellites such as ISEE, AMPTE/IRM, and IMP 8, and employs traditional statistical analysis for individual parameters. It is known that the magnetopause position is primarily influenced by the B_z and D_p [10, 7, 5], while the distance from the magnetopause to the Earth serves as the dependent variable. Therefore, Shue utilizes the distance relative to the Earth's magnetopause, along with B_z, D_p, and the angle θ between the Earth-Sun connection, in his model.

Figure 2b illustrates a significant variation when B_z transitions from -10 nT to 0 nT. This segment can be explained as the southward IMF reconnecting on the dayside, subsequently compressing and pushing towards the magnetopause in the direction of Earth.[2, 14] That is the reason why Shue imports a hyperbolic tangent function to fit the curve to follow the theorem.

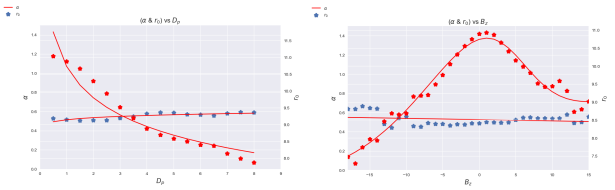
$$r = r_0 \left(\frac{2}{1 + \cos(\theta)} \right)^{\alpha} \tag{1}$$

$$r_0(B_z, D_p) = 10.22 + 1.29 \cdot \tanh\left(0.184 \cdot (B_z + 8.14)\right) D_p{}^{\frac{1}{6.6}}, \tag{2}$$

$$\alpha(B_z, D_p) = (0.58 - 0.007 B_z)[1 + 0.024 \ln(D_p)] \tag{3}$$

3.2 Data-Overfitting Model

The most significant difference in parameter fitting compared to the original formulation in [13] is the decay of r_0 relative to B_z when $B_z > 0$ nT, as shown in Figure 2b. This phenomenon was not observed in Shue's dataset, which led [13] to utilize a single hyperbolic tangent function for curve fitting. However, using the same THEMIS satellite dataset, studies in [9] and [6] have observed this phenomenon. Consequently, we have reevaluated and modified the original parameter formulation to employ two hyperbolic tangent functions, despite the lack of foundational theoretical support. We performed individual fitting of each variable using the least squares method as in [12]. The results are shown in Equation 4 and 2, with a contour plot visualized in Figure 3a:



(a) (α & r_0) v.s. D_p. (b) (α & r_0) v.s. B_z.

Figure 2: Parameters of the empirical model.

$$r_0 = (9.332 + 1.308 \cdot \tanh\left(0.213(B_z + 11.191)\right) - 0.568 \cdot \tanh\left(0.479(B_z - 7.188)\right)) \left(D_p\right)^{\left(\frac{-1}{6.22}\right)}, \tag{4}$$

$$\alpha = \left(0.493 - 3.5 \times 10^{-4} \cdot B_z\right) \left(D_p\right)^{\left(\frac{1}{11.92}\right)} \tag{5}$$

3.3 Regression-based Physics Informed Neural Networks

Neural networks have been used for fitting problems for many years, with the Universal Approximation Theorem [4] being the main theoretical support for using machine learning in fitting tasks. The primary objective of implementing neural networks is to project a high-dimensional manifold of features onto the desired solution. The term "universal" signifies that regardless of the mathematical expression $y = f(\theta)$ that needs to be solved, or the pattern of f, neural networks can approximate a solution model that conforms to the desired spatial features. In other words, it can approximate any continuous function to arbitrary accuracy on a compact input domain.

Let $f(\theta)$ be a continuous function defined on a compact subset $K \subseteq \mathbb{R}^n$. Given any $\epsilon > 0$, there exists a neural network function $NN(\theta)$ such that

$$\sup_{\theta \in K} |f(\theta) - NN(\theta)| < \epsilon \tag{6}$$

The function $NN(\theta)$ can be expressed as a composition of multiple layers, each of which is a linear transformation followed by a non-linear activation function. The general form of a $NN(\theta)$ is shown in 7:

$$NN(\theta) = \sigma(W_L \sigma(W_{L-1} \cdots \sigma(W_1 x + b_1) \cdots) + b_{L-1}) + b_L \tag{7}$$

where W_i and b_i are the weights and biases of the i-th layer, respectively, and σ is the activation function.

The key difference between Reg-PINNs and standard neural networks lies in the loss function. The total loss function L_{total} used in Reg-PINNs is defined as:

$$L_{total} = L_{data} + \lambda L_{reg} \tag{8}$$

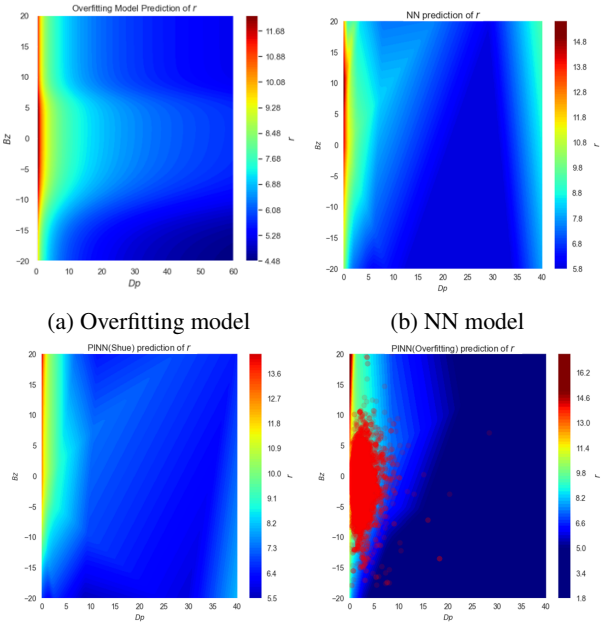
where L_{data} represents the mean square error (MSE) between the predicted values and the known data points, L_{reg} represents the MSE between the regression model output and the network output (serving as a regularization term with a desired pattern), and λ is a hyperparameter that acts as a weighting parameter to balance the contributions of the two loss components. In the experiment, $\lambda = 1$ was used.

Algorithm 1 Reg-PINNs

Input: Dataset $D = \{[x_1, x_2, x_3, \ldots, x_n] \rightarrow [y_1, y_2, y_3, \ldots, y_m]\}$
Output: Regularized model obtained via the empirical formula (regression form)
Set $\mathbf{Y}_{\text{reg}} = f_{\text{reg}}(\mathbf{X})$.
Set maximum iteration (K), threshold for model $(\epsilon_{\text{threshold}})$, learning rate (η), weighted parameters (λ).
Set $NN(\mathbf{X})$.
Initialize: iteration = 0, $\epsilon = \infty$.
while $\epsilon > \epsilon_{\text{threshold}}$ and iteration $< K$ **do**
    $\mathbf{Y}_{\text{NN}} = NN(\mathbf{X})$
    $L_{\text{NN}} : L_{\text{NN}} = \frac{1}{m} \sum_{i=1}^{m} (y_{\text{NN}_i} - y_i)^2$
    $L_{\text{reg}} : L_{\text{reg}} = \frac{1}{m} \sum_{i=1}^{m} (y_{\text{reg}_i} - y_i)^2$
    $L_{\text{total}} = L_{\text{NN}} + \lambda L_{\text{reg}}$
    $W_i \leftarrow W_i - \eta \cdot \frac{\partial L_{\text{total}}}{\partial W_i}$
    $\epsilon = L_{\text{total}}$
    iteration = iteration + 1
end while

4 Results and Discussions

The experimental results are shown in Figure 3. The contour plot of the overfitting model in Figure 3a and the pure machine learning model in Figure 3b exhibit a bulged pattern under $D_p < 10$ nPa in the contour plot. In contrast, the Reg-PINNs (Shue), which inherit Shue's physics-inspired constraints, stabilize when $B_z > 0$ nT (see Appendix: Contour plot of Shue's model).



(a) Overfitting model (b) NN model

(c) Reg-PINNs (Shue) (d) Reg-PINNs(Overfitting)

Figure 3: Contour plot of different models.
P.S. (D_p v.s. B_z) scatter points has been added on Figure 3d

The performance of the proposed Reg-PINNs was evaluated using root mean square error (RMSE) as the primary metric. Table 1 presents the RMSE values for the fitting based on empirical analysis, using Shue's model as the baseline, an overfitted model, and an overfitted model with parameters sampled using Markov Chain Monte Carlo (MCMC). These models were tested with both the full set of 34,998 data points and the dataset used in [13]. The results indicate that the statistical model shows a small amount of improvement, even with the modified overfitting approach.

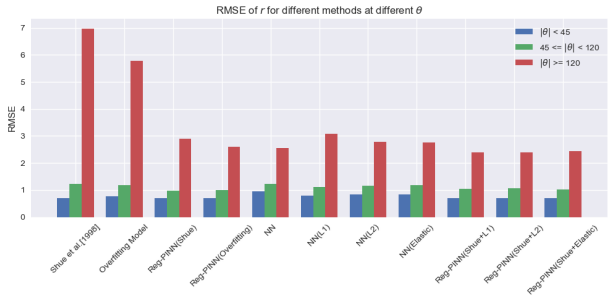
Table 1: RMSE comparison of different models

Model	RMSE (All)	RMSE (Shue)
Shue: Baseline	1.316 Re	1.347 Re
Overfitting (O.F.)	**1.192** Re	1.297 Re
MCMC	1.195 Re	**1.283** Re

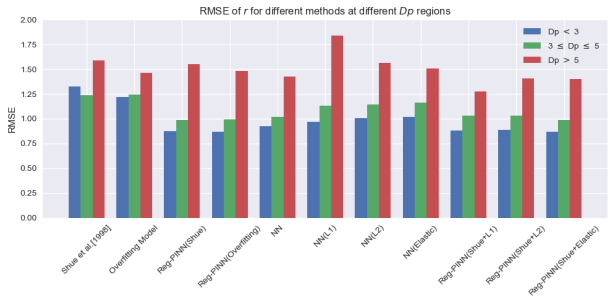
Table 2 compares the ML-based method to the baseline, divided into two groups: the first column shows results with 80% of data used for training and 20% masked for testing, while the second column shows results with 20% of data used for training and 80% masked for testing. The vanilla NN achieved the best performance with 80% of the data for training, but it lost precision when trained on a smaller dataset. In contrast, the Reg-PINNs demonstrated consistent improvement in capturing the magnetopause position across varying data amounts, indicating its generalizability and precision.

Table 2: RMSE comparison of different models

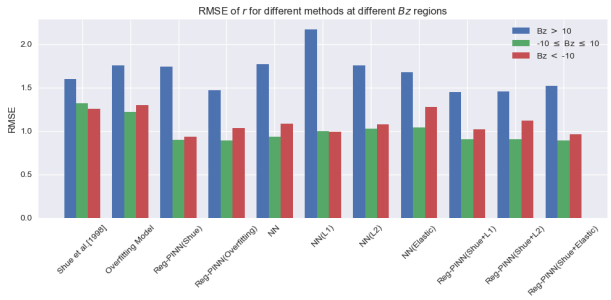
Model	RMSE (20%)	RMSE (80%)
Baseline	1.243 Re	1.266 Re
Vanilla NN	**0.866** Re	1.073 Re
Reg-PINNs (Shue)	0.889 Re	0.932 Re
Reg-PINNs (O.F.)	0.873 Re	**0.918** Re



(a) RMSE for comparing θ



(b) RMSE for comparing D_p



(c) RMSE for comparing B_z

Figure 4: RMSE for comparing different physical parameters

Figure 4 presents the performance of each test model, Shue's 1998 model, overfitting model, Reg-PINNs(Shue, overfitting, Shue - L1, Shue - L2, Shue - Elastic), vanilla neural networks(L1, L2, and Elastic), and under varying θ, D_p, and B_z parameters, with 50% of the data used for training and 50% reserved as unseen data for testing. In most parameter regions, the ML-based method performs competitively with the conventional approach, as shown in Figure 4b for $D_p \leq 5$ nPa and Figure 4c for $B_z \leq 10$ nT. However, the proposed method emphasize the benefits of integrating neural networks with empirical models, as this enables the inclusion of low-probability events such as $|B_z| > 10$ nT, $D_p > 5$ nPa, and $|\theta| > 120^o$, with probabilities of 0.62%, 1.55%, and 1.56%, respectively. The most notable improvement for the ML-based approach is observed in Figure 4a for the ML-based approach handling $|\theta| > 120^o$.

Although the ML-based approach yields lower error, this article focuses on physics-inspired techniques supported by real physical governing principles for modeling. Under these circumstances, we present the overfitting model, Reg-PINNs (Shue), Reg-PINNs (Shue-L2), and a vanilla neural network to evaluate performance, as shown in Figure 5. The decision to use Reg-PINNs with Shue's model is based on its alignment with the underlying theoretical framework, offering a more "physics-informed" perspective, with an error reduction rate of 32.5% compared to Shue's model. While numerous new machine learning algorithms can achieve high accuracy in fitting tasks, few are grounded in real physics. The experiment further utilize an overfitting model within Reg-PINNs to provide evidence of Reg-PINNs' robustness and precision over purely empirical models or data-driven approaches when applied to new datasets, providing a simple, proof-of-concept for enhancing applied scientific models with empirical foundations.

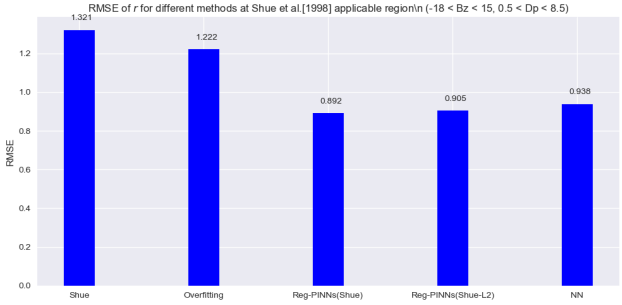


Figure 5: Overall RMSE for evaluation under baseline applicable region

5 Conclusion

The proposed algorithm, Reg-PINNs, mitigates the precision constraints of empirical methods and the generalization difficulties in machine learning. Experimental evaluations focus on the feasibility of Reg-PINNs with both a physics-inspired model and an overfitted model, demonstrating that Reg-PINNs extend the capabilities of standard PINNs by incorporating algebraic equations into model training, extending the usual scope of ODEs/PDEs. Furthermore, the Reg-PINNs forecast the magnetopause position on unseen data with an error reduction of about 30%. By providing the freedom to integrate domain-specific empirical models with neural networks, this approach demonstrates potential for wide-ranging scientific applications, including finance, securities, materials science, mechanical engineering, chemistry, and other applied sciences. Across fields, this flexibility makes it possible to model and forecast complex phenomena across disciplines, which opens new avenues for scientific advancements and practical applications, contributing to progress and innovation across diverse fields.

Acknowledgments

We would like to express our heartfelt appreciation to Dr. Chun-Yu Lin, associate researcher at National Center for High-Performance Computing - National Applied Research Laboratories, for exchanging ideas, insights into the concept of PINNs and for engaging in fruitful discussions. In addition, we would like to extend our gratitude to Prof. Jih-Hong Shue, Professor at National Central University (NCU), for providing the data used in Shue et al. [1998], and to Mr. Yu-Wei Chen and Dr. Pai-Sheng Wang, researchers at the Space Environment Laboratory (NCU), for their contributions to reinforcing the understanding of the magnetopause.

Appendix: Contour plot of Shue's model

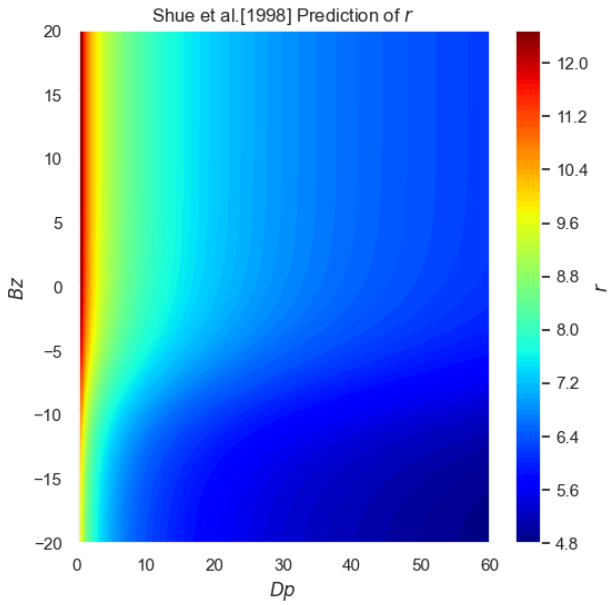


Figure 6: Shue's empirical model [13]

References

[1] Australian Bureau of Meteorology, Space Weather Services. Satellite observations. `https://www.sws.bom.gov.au/Satellite/3/2`. Accessed: 2024-10-31.

[2] Paul A. Cassak and Stephen A. Fuselier. Reconnection at earth's dayside magnetopause. In Walter Gonzalez and Eugene Parker, editors, *Magnetic Reconnection*, volume 427 of *Astrophysics and Space Science Library*, pages 213–276. Springer, Cham, 2016.

[3] J.K. Chao, J.-H. Shue, H.C. Fu, K.K. Khurana, C.T. Russell, and P. Song. Models for the size and shape of the earth's magnetopause and bow shock. In L.-H. Lyu, editor, *Space Weather Study Using Multipoint Techniques*, volume 12 of *COSPAR Colloquia Series*, pages 127–135. Pergamon, 2002.

[4] Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks*, 6(4):911–917, 1995.

[5] D. H. Fairfield. Structure of the magnetopause: Observations and implications for reconnection. *Space Science Reviews*, 23(3):427–448, 1979.

[6] Sheng Li, Yue-Ying Sun, and Chia-Hsien Chen. An interpretable machine learning procedure which unravels hidden interplanetary drivers of the low latitude dayside magnetopause. *Space Weather*, 21(3):e2022SW003391, 2023.

[7] G. Lu, L. R. Lyons, and A. D. Richmond. Simulations of the polar ionosphere for varying interplanetary magnetic field conditions. *Journal of Atmospheric and Terrestrial Physics*, 56(8):949–958, 1994.

[8] NASA Goddard Space Flight Center. OMNIWeb Data Explorer. `https://omniweb.gsfc.nasa.gov/`. Accessed: 2024-11-15.

[9] Z. Nemecek, J. Safrankova, L. Prech, E. Grigorenko, J. D. Richardson, and J. Šafránková. Magnetopause position under different conditions. In *AGU Fall Meeting Abstracts*, pages SM13B–1812, 2010.

[10] E. C. Roelof and D. G. Sibeck. Magnetopause shape as a bivariate function of interplanetary magnetic field bz and solar wind dynamic pressure. *Journal of Geophysical Research: Space Physics*, 98(A12):21421–21450, 1993.

[11] M. A. Schield. Pressure balance between solar wind and magnetosphere. *Journal of Geophysical Research (1896-1977)*, 74(5):1275–1286, 1969.

[12] J.-H. Shue, J. K. Chao, H. C. Fu, C. T. Russell, P. Song, K. K. Khurana, and H. J. Singer. A new functional form to study the solar wind control of the magnetopause size and shape. *Journal of Geophysical Research: Space Physics*, 102(A5):9497–9511, 1997.

[13] J.-H. Shue, J. K. Chao, H. C. Fu, C. T. Russell, P. Song, K. K. Khurana, and H. J. Singer. Magnetopause location under extreme solar wind conditions. *Journal of Geophysical Research: Space Physics*, 103(A8):17691–17700, 1998.

[14] B. U. Ö. Sonnerup. Magnetic field reconnection at the magnetopause: An overview. In Edward W. Jr. Hones, editor, *Magnetic Reconnection at the Magnetopause*, volume 30 of *Geophysical Monograph Series*, pages 92–99. American Geophysical Union, 1984.

[15] University of California, Los Angeles, Institute of Geophysics and Planetary Physics. THEMIS Overview Data. `https://themis.igpp.ucla.edu/overview_data.shtml`. Accessed: 2024-10-31.

[16] Y. Wang, Y. Zhou, Q.-G. Zong, H. Zhang, P. Song, L. Xie, and S. Fu. A new three-dimensional magnetopause model with a support vector regression machine and a large database of multiple spacecraft observations. *Journal of Geophysical Research: Space Physics*, 118(5):2173–2184, 2013.