Regression-based Physics-informed Neural Network (*Reg-PINN*) for Magnetopause Tracking

Po-Han K. Hou[1,2] and Sung-Chi C. Hsieh[3]

[1]Department of Earth Science and Engineering, Imperial College London, London, SW11 2DL
[2]Department of Space Science and Engineering, National Central University, Chungli, Taiwan
[3]Department of Electrical Engineering, University of Leicester, Leicester, LE1 7RH
`pohan.hou24@imperial.ac.uk`

Abstract

Previous research in the scientific field has utilized statistical empirical models and machine learning to address fitting challenges. While empirical models provide numerical generalizations, they often sacrifice precision. Traditional machine learning methods can provide high accuracy but may lack necessary generalization. We present a regression-based physics-informed neural network (Reg-PINN) that integrates physics-inspired empirical models into the neural network's loss function, thus merging the advantages of generalization and precision. 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 network, Reg-PINN with Shue's model, and Reg-PINN with the overfitting model. Compared to Shue's model, this technique achieves an approximately 30% reduction in RMSE, presenting a proof-of-concept improved solution for the scientific community.

Keywords loss function · regression model · magnetopause

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 [14] and an empirical model for the magnetopause proposed in 1997 and 1998 [15, 16], which has been widely acknowledged as the leading approach for the past 25 years [1]. Additionally, a method using support vector regression (SVR) was introduced in [19] 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-PINN).

Reg-PINN 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 differential equations. This study introduces a new approach that utilizes a regression-based equation form for training PINNs, combining physics-inspired fitting with neural networks to enhance both precision and generalization. The proposed method was tested against the widely used empirical model in [16] as a baseline, 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 [11], including THEMIS (28,634) [18], Geo-Tail (5,764), ISEE, and IMP 8. In Figure 1, the inverted C-shaped data pattern represents the in-situ crossing data points under GSM coordinates and illustrates the distribution of the north-south 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 the same time intervals for its consistency. 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 [16], 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.
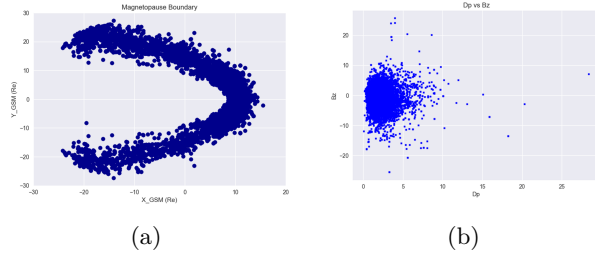

(a)   (b)

Figure 1: Magnetopause crossing data visualization, including (a) distribution of crossing data for origin is the Earth position and (b) D_p v.s. B_z distribution.

3. Methodologies

There have been numerous studies discussing the construction of empirical models for magnetopause loca-

tion, including notable works in 1993 [13], 1997, 1998 [15, 16], and 2002 [3]. In this study, we propose the use of Reg-PINN, which aims to combine the strengths of both empirical models and machine learning methods by achieving high accuracy and generalization capabilities.

3.1. Physics-Inspired Empirical Model

The position of the magnetopause is predominantly determined by B_z and D_p [13, 10, 6]. Consequently, Shue employs data from satellites such as ISEE, AMPTE/IRM, and IMP 8, utilizing statistical analysis of these factors to ascertain the radial distance r from the Earth-Sun line at an angle θ, measured in Earth radii (Re). The fitting function of the magnetopause boundary, as presented in this study, adopts the formulation proposed by Shue [15] under polar coordinates, where r_0 is the standoff distance between Earth and the subsolar point of the magnetopause, and α represents the tail flaring, as expressed in Equation 1, and utilizes the parameters from [16] as the baseline, as shown in Equation 2.

Figure 2b illustrates a significant variation when B_z transitions from -10 nT to 0 nT. This phenomenon is explained by the southward IMF reconnecting on the dayside, subsequently compressing and pushing towards the magnetopause in the direction of Earth [2, 17]. That is the reason why Shue imports a hyperbolic tangent function to fit the curve to follow the theorem.

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

$$\begin{cases} r_0(B_z, D_p) &= 10.22 + 1.29 \cdot \tanh\left(0.184 \cdot (B_z + 8.14)\right) (D_p)^{\frac{1}{6.6}}, \\ \alpha(B_z, D_p) &= (0.58 - 0.007 B_z)[1 + 0.024 \ln(D_p)] \end{cases} \tag{2}$$

3.2. Data-Overfitting Model

The most significant difference in parameter fitting compared to the original formulation in [16] 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 [16] to utilize a single hyperbolic tangent function for curve fitting. However, using the same THEMIS satellite dataset, studies in [12] and [9] 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 [15] with the result shown in Equation 3. It is worth noting that the power law relationship between D_p and r_0, as mentioned by Schield in 1996 [14], refers to the case of an ideal dipole magnetic field in a vacuum, where the relationship between the magnetopause boundary position r_0 and the solar wind dynamic pressure D_p should follow a negative exponential power law with an exponent of 1/6. This value serves as a reference for comparing the actual relationship between r_0 and D_p under realistic conditions. Shue, on the other hand, calculated an exponent of 1/6.6 in 1998 [16], while in this study, the calculated exponent is 1/6.22.
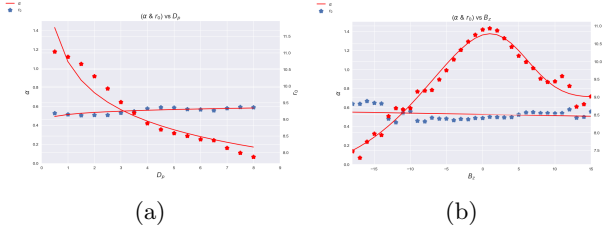

Figure 2: Parameters for fitting the empirical model. (a) (α & r_0) v.s. D_p, and (b) (α & r_0) v.s. B_z.

$$
\begin{cases}
r_0 &= \left(9.332 + 1.308 \cdot \tanh\left(0.213(B_z + 11.191)\right)\right. \\
&\quad \left. -0.568 \cdot \tanh\left(0.479(B_z - 7.188)\right)\right)(D_p)^{\left(\frac{-1}{6.22}\right)}, \\
\alpha &= \left(0.493 - 3.5 \times 10^{-4} \cdot B_z\right)(D_p)^{\left(\frac{1}{11.92}\right)}
\end{cases}
$$
$$(3)$$

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

$$\sup_{\theta \in K} |f(\theta) - NN(\theta)| < \epsilon \qquad (4)$$

The fitting for the $NN(\theta)$ consists of multiple non-linear and linear combinations with respect to a controllable cost function to optimize the model parameters or grant a discipline to follow the desired pattern. We used a regression pattern based on data distribution that was inspired by physics and added an external term for the loss function called L_{reg}. This makes the total loss function L_{total} used in Reg-PINN as follows:

$$L_{total} = L_{data} + \lambda L_{reg} \qquad (5)$$

where L_{data} represents the mean square error (MSE) between the predicted values and the known data points, L_{reg} stands for 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 for the desired importance of the regression model. In the experiment, $\lambda = 1$ was used.

Algorithm 1 Reg-PINN
1: **Input:** Dataset $D = \{[x_1, x_2, x_3, \ldots, x_n] \rightarrow [y_1, y_2, y_3, \ldots, y_m]\}$
2: **Output:** Regularized model obtained via the empirical formula (regression form)
3: Set $\mathbf{Y}_{reg} = f_{reg}(\mathbf{X})$.
4: Set maximum iteration (K), threshold for model $(\epsilon_{threshold})$, learning rate (η), weighted parameters (λ).
5: Set $NN(\mathbf{X})$.
6: **Initialize:** iteration = 0, $\epsilon = \infty$.
7: **while** $\epsilon > \epsilon_{threshold}$ and iteration $< K$ **do**
8: $\quad \mathbf{Y}_{NN} = NN(\mathbf{X})$
9: $\quad L_{NN} : L_{NN} = \frac{1}{m}\sum_{i=1}^{m}\left(y_{NN_i} - y_i\right)^2$
10: $\quad L_{reg} : L_{reg} = \frac{1}{m}\sum_{i=1}^{m}\left(y_{reg_i} - y_i\right)^2$
11: $\quad L_{total} = L_{NN} + \lambda L_{reg}$
12: $\quad W_i \leftarrow W_i - \eta \cdot \frac{\partial L_{total}}{\partial W_i}$
13: $\quad \epsilon = L_{total}$
14: $\quad$ iteration = iteration + 1
15: **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-PINN (Shue), which inherits Shue's physics-inspired constraints, stabilises when $B_z > 0$ nT (see).
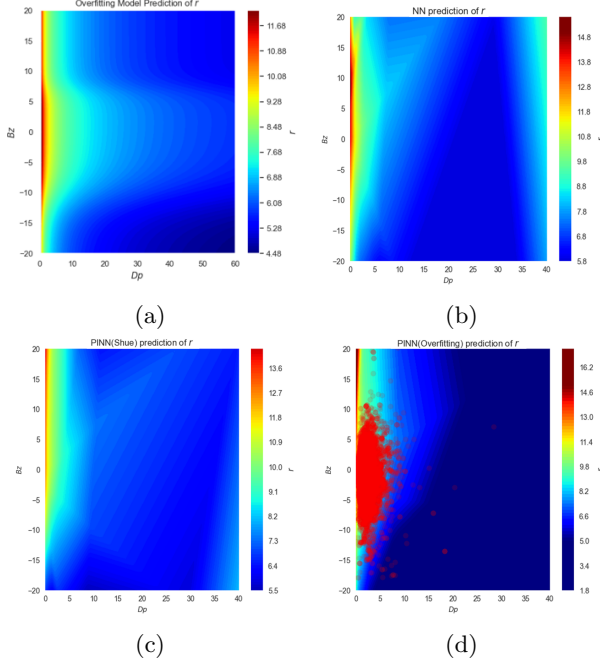

(a)   (b)

(c)   (d)

Figure 3: Contour plot of different models, including (a) overfitting, (b) Vanilla NN, (c) Reg-PINN (Shue), and (d) Reg-PINN (Overfitting) along with (D_p v.s. B_z) distribution for comprehensively visualization.

The performance of the proposed Reg-PINN was evaluated using root mean square error (RMSE) as the primary metric. Table 1 shows the RMSE 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 [16]. The results show that the statistical model provides only a marginal improvement, even with the modified overfitting approach.

Table 1: RMSE comparison of different empirical models

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

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. Each experiment was structured with [3 x 27 x 81 x 27 x 9 x 1] layers, an RMSProp optimizer, and 500 epochs. Figure 4d illustrates that Reg-PINN's loss remains consistently higher than NN's loss. 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-PINN demonstrated consistent improvement in capturing the magnetopause position across varying data amounts, indicating its generalizability and precision.

Table 2: RMSE comparison of different ML models

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

(a)   (b)

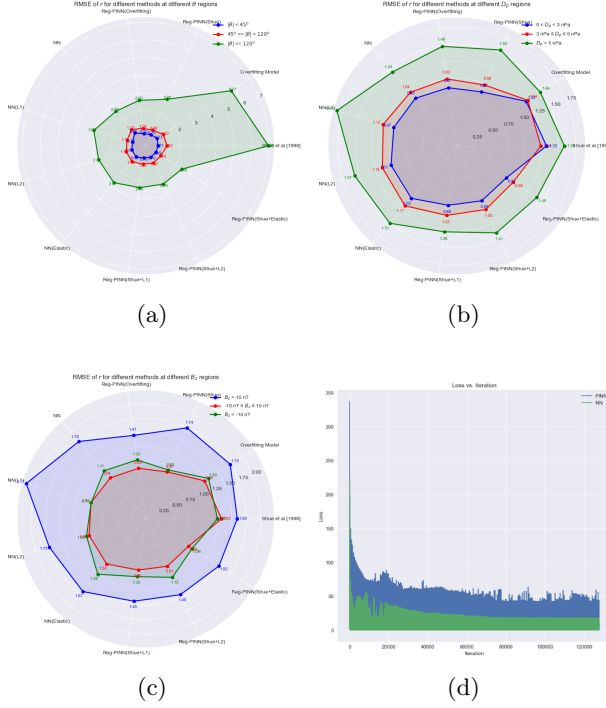(c)   (d)

Figure 4: RMSE for evaluating different physical parameters, including (a) θ, (b) D_p, and (c) B_z, along with (d) Loss vs. Iteration for NN and Reg-PINN on this case.

Figure 4 presents the performance of each test model: Shue's 1998 model, the overfitting model, Reg-PINN (Shue, overfitting, Shue-L1, Shue-L2, and 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 illustrated in Figure 4b for $D_p \leq 5$ nPa and Figure 4c for $B_z \leq 10$ nT. However, the proposed method emphasizes 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 significant improvement for the ML-based approach is observed in Figure 4a for the ML-based approach handling $|\theta| > 120^o$. In figure 4b, the result for $D_p > 5$ nPa shows that the NN method performs the worst

among the models, indicating potential data insufficiency under extreme conditions. With the help of [16], which has considered extreme solar wind conditions, the Reg-PINN (Shue) can follow the underlying constraint, resulting in better predictions under machine learning optimization. This also indicates that the model can learn from the pattern of a physics-inspired model, providing a robust approach for acquiring the magnetopause position.

Although the ML-based approach yields lower error, the research focuses on physics-inspired techniques supported by real physical governing principles for modeling. Under these circumstances, we present the overfitting model, Reg-PINN (Shue), Reg-PINN (Shue-L2), and a vanilla neural network to evaluate performance, as shown in Figure 5. The decision to use Reg-PINN 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-PINN to provide evidence of Reg-PINN' robustness and precision over empirical or data-centric approaches when applied to new datasets, providing a simple proof of concept for enhancing applied scientific models with empirical foundations.
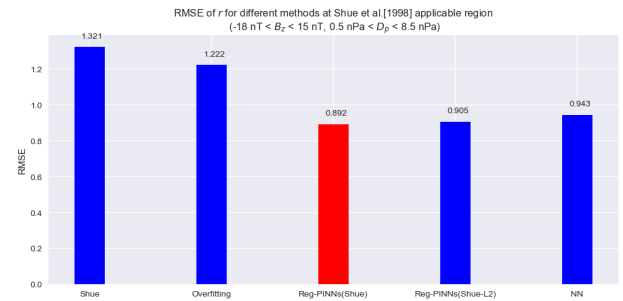

Figure 5: Selected model RMSE evaluated under the baseline applicable region

Table 3: RMSE comparison of different λ with Reg-PINN(Shue)

Model	RMSE (20%)	RMSE (80%)
Reg-PINN ($\lambda = 1$)	**0.866** R_e	0.932 R_e
Reg-PINN ($\lambda = 2$)	0.946 R_e	1.063 R_e
Reg-PINN ($\lambda = 5$)	1.092 R_e	1.118 R_e
Reg-PINN ($\lambda = 0.1$)	0.884 R_e	**0.927** R_e
Reg-PINN ($\lambda = 0.5$)	0.878 R_e	0.939 R_e
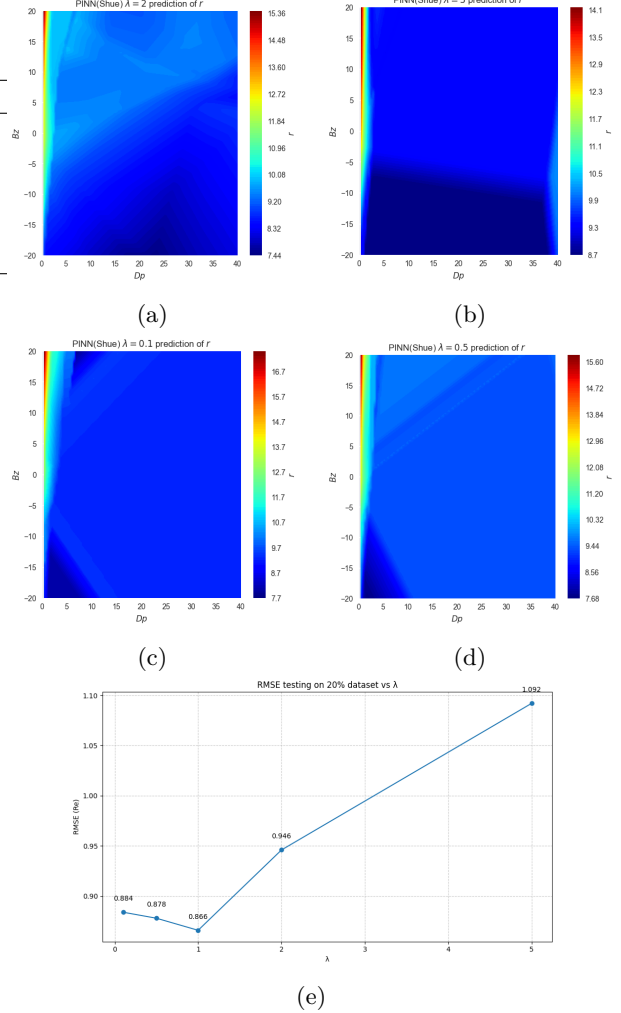
Table 3 and Figure 6 illustrate the different configurations of the weight (λ) for regression loss on Reg-PINN (Shue) alongside their corresponding contour plot, revealing distinct levels of influence from Shue's model as shown in Figure 7. It is worthy of mention that for the case of $\lambda = 2$, 5, the results indicate a significant increase in RMSE compared to the cases of $\lambda = 0.1$, 0.5. This pattern indicates that disproportionately prioritising the regression loss may result in an undue focus on data fitting, compromising the model's fundamental physical restrictions and perhaps causing numerical instability or overfitting to noise within the dataset. Consequently, the model exhibits poor generalisation, resulting in elevated RMSE values. This behaviour can be seen as an irreducible RMSE of the model, which seems to converge to about 0.86 Re across different parameter configurations. This irreducible error presumably arises from constraints in the model's construction, the lack of robust theoretical physics support, and the inherent uncertainties linked to measurement from the satellite. In the scenario of evaluating 20% masked data, the default value of $\lambda = 1$ consistently yields the lowest RMSE. Conversely, when testing 80% of the data with 20% allocated for training, the minimum RMSE is attained with $\lambda = 0.1$. However, in this instance, $\lambda \leq 1$ yields comparable performance, whereas the performance declines as λ increases.

Figure 6: Contour plot for different λ on Reg-PINN(Shue), including (a) $\lambda = 2$, (b) $\lambda = 5$, (c) $\lambda = 0.1$, (d) $\lambda = 0.5$ along with (D_p v.s. B_z) distribution for visualization and (e) line chart for varying λ for weighted regularization (20% masked data point)

5. Conclusion

The proposed algorithm, Reg-PINN, mitigates the precision constraints of empirical methods and the generalization difficulties in machine learning. Experimental evaluations focus on the feasibility of Reg-PINN with both a physics-inspired model and an overfitted model, demonstrating that Reg-PINN extends the capabilities of standard PINNs by incorporating

algebraic equations into model training, extending the usual scope of ODEs/PDEs. Furthermore, the Reg-PINN forecasts the magnetopause position on unseen data with an error reduction of about 30%. By enabling the integration of domain-specific empirical models with neural networks, this approach demonstrates potential for a wide range of scientific applications, such as remodeling the Fama-French three-factor model for stock returns [7] and understanding Hall-Petch strengthening in materials science [8, 5]. This flexibility allows for modeling and forecasting complex phenomena across disciplines, opening up new opportunities for scientific advancements and practical applications.

Acknowledgments

Appendix: Contour plot of Shue's model
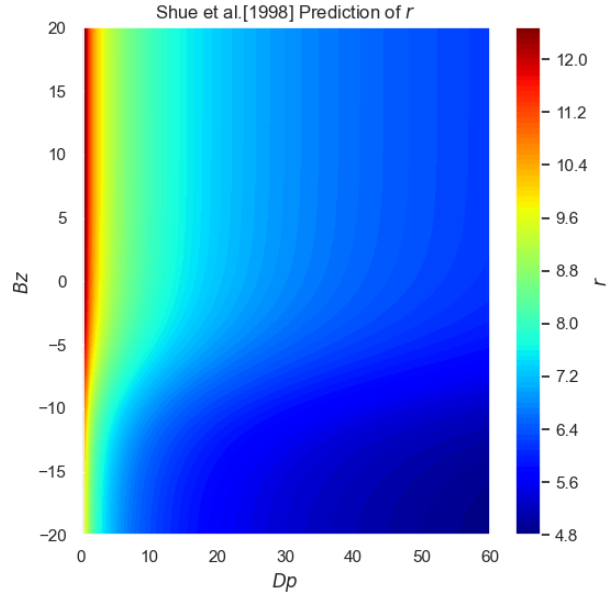
Figure 7: Shue's empirical model [16]

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] P. A. Cassak and S. A. Fuselier. Reconnection at earth's dayside magnetopause. In W. Gonzalez and E. 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] T. Chen and H. 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] Z. C. Cordero, B. E. Knight, and C. A. Schuh. Six decades of the hall–petch effect – a survey of grain-size strengthening studies on pure metals. *International Materials Reviews*, 61(8):495–512, 2016.

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

[7] E. F. Fama and K. R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

[8] N. Hansen. Hall–petch relation and boundary strengthening. *Scripta Materialia*, 51(8):801–806, 2004.

[9] S. Li, Y.-Y. Sun, and C.-H. Chen. An interpretable machine learning procedure which unravels hidden interplanetary drivers of the low latitude dayside magnetopause. *Space Weather*, 21(3):e2022SW003391, 2023.

[10] 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.

[11] NASA Goddard Space Flight Center. Omniweb data explorer. https://omniweb.gsfc.nasa.gov/. Accessed: 2024-11-15.

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

[13] 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.

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

[15] 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.

[16] 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.

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

[18] 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.

[19] 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.