Physics informed Multi-Modal Networks for Solar Wind modeling

Jorge D. Enciso

November 11, 2024

Abstract

Contents

1 Introduction

The solar wind phenomena is a vastly known event that arises from ionized solar outbursts into the stellar medium, commonly stimulating planets' magnetosphere. [4] The first record of a related event dates from 1859, popularly named "The Carrington event" after the English astronomer Richard Carrington (1826, 1875), who settled the intuition upon solar flares and geomagnetic fluctuations. Individuals from all around the glove sighted the Northern Lights, subtle variations in atmosphere's hues' tones, caused by ionized particles flowing through the ionosphere.

Several attempts to model Solar Wind's dynamics can be empirically pinpointed : analytical methods [], numerical modelling [3,5,10], and machine learning approaches [1, 7, 9]. All of them with different purposes.

Traditional models and methodologies fail to account for the anysotropic nature of the Solar Wind. The Magnetohydrodynamics Model assumes a Boltzmann-Maxwellian distribution, doted with standard thermodynamical behavior. On the other hand, the Solar Wind is a collisionless plasma, precluding its inclusion solely within those boundaries. Thus, the present work emphasizes the usage of learnable neural operators engineered to analytically and empirically fit kinetic descriptions.

Moreover, kinetic frames based on Maxwellian electromagnetism tend to be computationally expensive for the microscopic and macroscopic frame:

$$n_\alpha(\mathbf{r}, t) := \int f_\alpha(\mathbf{r}, \mathbf{v}, t) d^3 v \tag{1}$$

$$\tau(r, t) := \sum q_\alpha n_\alpha \tag{2}$$

$$\mathbf{u}_\alpha(\mathbf{r}, t) := \frac{1}{n_\alpha(\mathbf{r}, t)} \int v f_\alpha(\mathbf{r}, v, t) d^3 v \tag{3}$$

$$j(\mathbf{r}, t) := \sum q_\alpha n_\alpha \mathbf{u}_\alpha \tag{4}$$

where n_α is the particle density, u_α the particle's drift velocity, $\tau(r, t)$ is the charge density, $j(r, t)$ is the current density, and $f_\alpha(r, v, t)$ is the Boltzmann's density function.

Yet modern physics informed machine learning research schemes enable faster inference time and less computational load: embedding systems' equations within the loss function and constraining network's backbone based on empirical approaches.

2 Related Work

2.1 Solar Wind modelling

2.2 Physics informed Vlasov modeling

3 Kinetic Models

In order to fully assume the magnetohydrodynamical model, interactions must converge to the Maxwell-Boltzmann thermodynamical distribution, implying a high colisionality degree, contrary to empirical observations of the Solar Wind. Therefore, the utilization of such a constrictive framework must be carefully evaluated before setting it up as a modeling constraints.

Moreover, trying to answer quasy-isotropic behavior on near-earth L1-Lagrange point, different research papers assume the bi-Maxwellian temperature distributions, carrying out the Maxwell-Boltzmann thermodynamical distribution. Nevertheless, this assumption seems overly ideal. As a possible option, the inclusion of numerical distributions could bring insights into the interactions between plasma ionized particles.

That's why, this work departs from the traditional approach, seeking the description of Solar Wind kinetic model without magnetohydrodynamical systems nor temperature distributions, fully describing these interactions with a physics informed data-driven structure as a substitution of empirical assumptions.

3.1 Boltzmann's equation

Within this frame, different kinetic statistical descriptors are suitable for the task as long as they make the correct assumptions. Therefore, the usage of Boltzmann's equation works perfectly within certain boundaries:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \mathbf{a} \cdot \nabla_{\mathbf{v}} f_\alpha = \left(\frac{\delta f}{\delta t} \right)_{col} \tag{5}$$

This result is widely known for describing the Maxwell-Boltzmann distribution under defined conditions, so an alternative interpretation that accounts for colisionless plasma should be integrated:

3.2 Vlasov's equation

The Vlasov's equation is the colisionless variant of the Boltzmann's equation. Therefore, it is a suitable kinetic model for the Solar Wind, accounting for key properties of it.

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_\alpha = 0 \tag{6}$$

In this case, governed by electromagnetic interactions. Therefore, replacing the force $\mathbf{F}$ with the Lorentz Force:

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_\alpha + \frac{q(\mathbf{E} + \mathbf{v} \times \mathbf{B})}{m} \cdot \nabla_{\mathbf{v}} f_\alpha = 0 \tag{7}$$

4 Physics Informed Machine Learning

Usually, solely data-driven methods require huge volumes of high-quality data with black box algorithms that prevent interpretability, a key factor for physical modeling. The most recent research [**?**] advocate for physical informed virtual loss function terms, leading off to the vastly known field of Phyics Informed Neural Networks.

This models where created as numerical methods for solving Partial Differential Equations that embed physical modeling. It further extended to all fields that required PDE solving as a computationally efficient numerical alternative to other solvers.

The most recent research, using PINNs, develop a way to not just learn functions, but to learn functional operators: Lagrangian, Hamiltonian, etc. This research [] creates a neural network that takes the general coordinates q and $\dot{q}$ as input, to provide the lagrangian as an output, and leverage the auto-differentiation capabilities of modern machine learning frameworks to enforce the euler-lagrange equation:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0 \tag{8}$$

As usage of q and $\dot{q}$ as inputs is very convenient for the auto-differentiation engines, making it further computationally feasible.

4.1 Learnable Splines for physical quantities

Implicit PINN research usually enriches datasets with high resolution simulation data. High resolution data plays an important role within the machine learning PDE solvers. Numerical integration methods' accuracy decays as the data resolution does so. Given the unfulfilled knowledge about solar wind dynamics and low resolution nature of the data, this work utilizes polynomial interpolation methods with empirical backup [5] to convert the discrete problem into continuous.

4.2 Approximating the density function with PINNs

4.3 The Multi Modal aspect

Empirical studies demonstrated the strong correlation between imagery features and magnetic disturbances, all tied to Coronal Holes, propitious to be the source of ionic outburst. [8] Moreover, one of the most recent studies extracts coronal features with segmentation algorithms and use them as a basis for geomagnetic forecasting. [2] It is noticeable the incremental amount of research that utilizes this data source as a cornerstone of their work.

Following this path, the present work exploits the representational power of this dynamical system offered by corongraph imagers, broadening prior researchs to the extend of kinetic models aided by deep neural networks.

5 Data

5.1 DSCOVR: Deep Space Climate Observatory

DSCOVR, a joint mission between NASA and the National Oceanic and Atmospheric Administration, is a crucial observational platform for monitoring space weather [6]. Launched in 2015, DSCOVR's primary mission is to monitor and provide advanced warning of potentially hazardous space weather events such as solar flares and coronal mass ejections that could impact Earth.

It is equipped with two key instruments for measuring both energetic particle incidence and magnetic field parameters: the Faraday cup and the magnetometer from the PlasMag instrument [6]. The readings from these two sensors are crucial for virtually analyzing plasma dynamics near the L1 Lagrange point. These readings will be used as part of the core model data due to their real-time availability.

5.2 ACE: Advanced Composition Explorer

ACE, launched in 1997, provides continuous measurements of the solar wind and interstellar particles. It is equipped with several instruments designed to study the composition of solar and galactic particles, which are crucial for understanding the space weather environment. ACE's data helps in predicting geomagnetic storms and contributes to our understanding of the heliosphere.

5.3 WIND

The WIND spacecraft, launched in 1994, is part of the Global Geospace Science initiative. It provides comprehensive measurements of the solar wind, magnetic fields, and energetic particles. WIND's data is essential for understanding the fundamental processes of the solar wind and its interaction with the Earth's magnetosphere.

5.4 SOHO: Solar and Heliospheric Observatory

SOHO, a joint project of ESA and NASA, was launched in 1995. It is designed to study the Sun from its core to the outer corona and the solar wind.

5.4.1 LASCO: Large Angle and Spectrometric Coronagraph Observatory

LASCO, one of the instruments on SOHO, observes the solar corona by creating an artificial eclipse. It is instrumental in detecting coronal mass ejections, which are significant drivers of space weather.

6 Methods

6.1 Neural Architecture Search (NAS)

As an intermediate step before training, a hyperparameter grid search algorithm based in Bayesian multi-objective optimization is employed to find the neural architecture. The Pareto front is plotted, selecting minimization of neural complexity and loss function.

6.2 Loss function

The neural network is trained using the standard gradient-descent based optimization algorithm: Adam. One Cycle learning optimizer is used to facilitate the learning rate plateau. Weight decay is implemented as a regularizer.

The loss function consists on a virtual term, three integral boundary condition terms, and a l2 regularizer term:

$$\mathcal{L}_{virtual}\left(\mathbf{E},\mathbf{B},\mathbf{v},\mathbf{r};\theta\right) = \left(\frac{\partial f_\alpha}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{r}} f_\alpha + \frac{q(\mathbf{E}+\mathbf{v}\times\mathbf{B})}{m}\cdot\nabla_{\mathbf{v}} f_\alpha\right)^2 \tag{9}$$

$$\mathcal{L}_{BC_1}\left(x;\theta\right) = \left(\int_{x\sim X} f_\alpha(\mathbf{v},\mathbf{r},t) d^3\mathbf{v} d^3\mathbf{r} - 1\right)^2 \tag{10}$$

$$\mathcal{L}_{BC_2}\left(x;\theta\right) = \left(\int_{x\sim X} f_\alpha(\mathbf{v},\mathbf{r},t) d^3\mathbf{v} - n_\alpha(\mathbf{r},t)\right)^2 \tag{11}$$

$$\mathcal{L}_{BC_3}\left(x;\theta\right) = \left(\sum_\alpha q_\alpha n_\alpha \mathbf{u}_\alpha - j(\mathbf{r},t)\right)^2 \tag{12}$$

$$\mathcal{L}_{regularizer}\left(\theta\right) = \mathbb{E}_{w\in\theta}\left[w^2\right] \tag{13}$$

Getting our final objective function:

$$\mathcal{L}\left[\mathcal{L}_{virtual},\mathcal{L}_{BC},\mathcal{L}_{regularizer}\right] = \alpha_1\mathcal{L}_{virtual} + \alpha_2\sum_i \mathcal{L}_{BC_i} + \lambda\mathcal{L}_{regularizer} \tag{14}$$

7 Results

8 Discussion

9 Conclusion

This work demonstrated the effectiveness of Physics-Informed Neural Networks to model Solar Wind's behavior tied to diverse empirical and analytical formulations. By embedding the governing physical equations directly into the loss function, we can efficiently train a neural network to approximate the solution of this complex, non-linear system. Future work will focus on improving the kernel design and extending the approach to more sophisticated MHD scenarios, such as the interactions between Solar Wind and Sun's magnetosphere.

References

[1] Yasser Abduallah, Jason TL Wang, Prianka Bose, Genwei Zhang, Firas Gerges, and Haimin Wang. A deep learning approach to dst index prediction. *arXiv preprint arXiv:2205.02447*, 2022.

[2] Daniel Collin, Yuri Shprits, Stefan J. Hofmeister, Stefano Bianco, and Guillermo Gallego. Forecasting high-speed solar wind streams from solar images. 2024.

[3] A. Ekenbäck and M. Holmström. MHD modeling of the interaction between the solar wind and solar system objects. 3732:554–562, 2006.

[4] J. T. Gosling. *The Solar Wind*. Elsevier, 2007.

[5] Takuma Matsumoto. Full compressible 3D MHD simulation of solar wind. *Monthly Notices of the Royal Astronomical Society*, 500(4):4779–4787, 11 2020.

[6] National Aeronautics and Space Administration. Dscovr mission overview, 2023.

[7] AS Parnowski. Regression modeling method of space weather prediction. *Astrophysics and Space Science*, 323(2):169–180, 2009.

[8] Matthew Shelby, Scott Scharlach, Petar Matejic, RJ Everett, and Colton Morgan. An optical analysis of sunspots as predictors of geomagnetic storms. 2023.

[9] Iris Yan. Early prediction of geomagnetic storms by machine learning algorithms. 2024.

[10] Man Zhang, Xueshang Feng, Huichao Li, Ming Xiong, Fang Shen, Liping Yang, Xinhua Zhao, Yufen Zhou, and Xiaojing Liu. Numerical modeling of solar wind and coronal mass ejection in the inner heliosphere: A review. *Frontiers in Astronomy and Space Sciences*, 10, 2023.