Leveraging Physics-Informed Neural Networks as Solar Wind Forecasting Models

Nuno Costa[1], Filipa S.Barros[1,2,3], J.J.G.Lima[2,4], Rui F. Pinto[3], André Restivo[1]

1- LIACC / Faculdade de Engenharia da Universidade do Porto, Portugal
2- Instituto de Astrofísica e Ciências do Espaço, CAUP, Porto, Portugal
3- Institut de Recherche and Astrophysique et Planétologie, OMP/CNRS, CNES, University of Toulouse, Toulouse, France
4- Departamento de Física e Astronomia, FCUP, Porto, Portugal

Abstract. Space weather refers to the dynamic conditions in the solar system, particularly the interactions between the solar wind — a stream of charged particles emitted by the Sun — and the Earth's magnetic field and atmosphere. Accurate space weather forecasting is crucial for mitigating potential impacts on satellite operations, communication systems, power grids, and astronaut safety. However, existing models like MULTI-VP require substantial computational resources. This paper proposes a Physics-Informed Neural Network (PiNN) as a faster yet accurate alternative that respects physical laws. PiNNs blend physics and data-driven techniques for rapid and reliable forecasts. Our studies show that PiNNs can reduce computation times and deliver forecasts comparable to MULTI-VP, offering an expedited and dependable solar wind forecasting approach.

1 Introduction

Space weather forecasting aims to predict the effects of solar disturbances on Earth, using models that establish causal relationships across various physical regimes. The solar wind, a flow of charged particles from the Sun's corona, is a primary factor in these disturbances across the solar system. However, accurately predicting the solar wind's journey from the Sun to Earth is challenging due to the absence of direct observations of the numerous physical processes it undergoes as it travels through the upper corona and heliosphere. Solar wind modeling employs specialized models focusing on specific sub-regions or processes and uses magnetic field maps of the Sun's surface as primary input data.

Modeling the coronal portion of the solar wind is challenging due to its complexity. Recent advancements, such as the MULTI-VP model [1], offer promising solutions. MULTI-VP is a computational tool designed to model individual solar wind streams within the solar corona, covering up to 15% of the distance from the Sun to Earth. While the initial computation of these streams is computationally demanding, MULTI-VP uses pre-computed initial conditions to refine solutions through a relaxation process. In a previous paper [2], we have shown that Neural Networks can generate more precise initial conditions, reducing computational time by up to 8%.

While this improvement is promising, MULTI-VP still requires several hours to compute a single solar wind stream. As a result, we considered the feasibility of a surrogate model. This model could be an alternative to MULTI-VP, particularly in scenarios where absolute accuracy is not the primary concern, such as forecasting.

Like MULTI-VP, this model would interact with other heliospheric models like Helio1D [3] and EUHFORIA [4]. These models use MULTI-VP results at around 0.1 AU and extend them to L1 and beyond. However, relying solely on a data-driven surrogate model might produce physically inconsistent results, hindering a seamless connection to these downstream models. In this context, Physics-Informed Neural Networks (PiNNs)[5] combine data-driven techniques with physical laws, ensuring computational efficiency while maintaining physical fidelity. Prior studies have shown the effectiveness of this approach in ensuring adherence to conservation laws and physical constraints by machine learning models[6, 7].

In this paper, we employed a PiNN to develop a surrogate model for MULTI-VP. This model effectively reduces numerical noise from upstream models while improving prediction robustness. By doing so, we ensure that the model captures the complex dynamics from the data and adheres to essential physical principles. Consequently, our approach yields more accurate and physically consistent simulations across various heliospheric modeling tasks.

2 Methodology

2.1 Data Preparation

The MULTI-VP model uses magnetogram data to map the Sun's surface magnetic field and infer a three-dimensional topology of the solar corona's magnetic field [1]. It identifies and traces an ensemble of open magnetic field lines, each representing an elemental solar wind stream. The geometry of each field line is constrained by several physical properties, including the distance from the Sun (R), the position within the flux tube (L), magnetic field amplitude (B), the inclination of the flux tube relative to the Sun (α), and the tube expansion ratio ($A_{exp} = \frac{A}{A_0}$). These geometrical characteristics significantly influence the simulated outputs, which include the plasma density (n), velocity (v), and temperature (T) of the solar wind stream, extending up to 30 solar radii.

In previous work, we trained a neural network to predict the solar wind's physical properties at 1 AU [2]. For each stream, both the geometric properties (R, B, and α) and the MULTI-VP's simulated outputs (n, v, T) were defined at over 640 data points spanning from the Sun's corona to 30 solar radii. In this work, all models were trained and validated using the same dataset comprising solar wind streams from five distinct solar events. We refined our previous approach by incorporating additional input information for L and A_{exp}. These properties play a key role in determining the physical properties used in the PiNN.

We have also implemented a comprehensive normalization process to enhance the model's performance further. All data was converted to the CGS system of units to ensure consistency with the physical properties that need to be verified.

Afterward, we employed a thorough analysis of the behavior of our data to choose the most appropriate normalization techniques for each variable. We also smoothed out the outputs n, v, and T with a Butter Low-Pass Filter to address the unphysical sinusoidal patterns observed in the MULTI-VP outputs.

2.2 Model Definition

We used a wide network architecture[1], capable of handling 640 data points across five variables. It consists of an initial hidden layer with 2056 neurons, followed by two layers of 1024 neurons each, and another hidden layer of 2056 neurons. The final output comprises 640 data points for the three output variables. Batch normalization is employed to manage gradient flow, and ReLU activation functions are used. Weights were initialized using a Xavier uniform distribution.

The model uses the AdamW optimizer with an initial learning rate of 10^{-2}. During training, the learning rate was reduced via a learning rate scheduler on plateau, and weight decay was applied with a very small value of 10^{-8}, considering the high complexity of the system we are modeling. For the supervised loss $\mathcal{L}_s$, we used a Smoothed L1-Loss with a quadratic term under a threshold, making the resulting model less sensitive to outliers.

2.3 Physical Properties as Physics-Informed Losses

Due to the inherent complexity and nonlinearity of solar and space weather phenomena, deriving closed-form solutions that accurately represent the entire system's behavior is challenging. Therefore, numerical methods and machine learning models are often used for predictions and simulations. The MULTI-VP model involves numerous complex calculations and simplifications to produce valid results. While data-driven machine learning models are faster, they do not consider physical laws. PiNNs bridge the gap between these approaches by incorporating physical constraints. In this work, we used two physical expressions to help our model achieve more accurate results. The first physical property we incorporated pertains to the principle of mass conservation within the flux tube:

$$m_c = \sigma\left(\frac{n \cdot v}{T}\right) \approx 0 \tag{1}$$

The second property mirrors the momentum conservation characteristics of a simplified Magneto-hydrodynamics System of differential equations, where G represents the gravitational constant and ν stands for a predefined viscosity constant: $p_c = P_{grad} + g_{term} + v_{grad} + \nu_{damp} \approx 0$ (2), where:

$$P_{grad} = \frac{\mathrm{d}(n \cdot T)}{\mathrm{d}L} \cdot n \qquad v_{grad} = \frac{\mathrm{d}v^2}{\mathrm{d}L} - v\frac{\mathrm{d}v}{\mathrm{d}L}$$

$$g_{term} = G\frac{cos(\alpha)}{R^2} \qquad \nu_{damp} = -\nu\left(\frac{\partial^2 v}{\partial L^2} + A_{exp}\frac{\mathrm{d}v}{\mathrm{d}L}\right)$$

[1] Implemented using Pytorch: `https://github.com/biromiro/pi-multivp`

These properties can be seamlessly integrated as new optimization objectives during neural network training as a suitable loss function.

While implementing the mass conservation property as a loss function is straightforward, incorporating momentum conservation proves more complex. The necessity for first and second-order derivatives poses a challenge that cannot be addressed using *autodiff*, as commonly done in PiNN implementations. This limitation arises due to the 1D-grid data format, which differs from the point-wise data typically required. Although converting to a point-wise data format is feasible, it was not pursued because it could significantly increase model complexity, disrupt spatial correlations among proximal inputs, and potentially degrade the performance or complicate the training. Consequently, our approach employs Finite Difference methods, vectorizing the implementation and adapting it to utilize tensor operations [8].

It is important to note that these physical properties, especially momentum conservation, do not show high accuracy in the initial turbulent section of the domain due to simplifications. The effectiveness of the surrogate model heavily relies on its accuracy, particularly in the more stable latter part of the domain. Thus, we impose physical constraints from index k onwards in the domain, relying on empirical solutions for earlier sections. The index, where the solutions demonstrated the highest physical compliance, was identified as $k = 348$. With these considerations in mind, and considering val_{phys} as m_c for Equation 1 and as p_c^2 for Equation 2, respectively, the loss functions were defined as follows:

$$\mathcal{L} = \lambda_s \mathcal{L}_s + \lambda_{phys} \sqrt{\frac{1}{640-k} \sum_{i=idx}^{640-k} val_{phys}} \tag{3}$$

We noticed that even when using curriculum training — by slowly increasing the λ_{phys} value — to ease the impact of the physical properties on the optimization process, the momentum conservation equations would completely take over the loss landscape as soon as they began to be optimized. This led to unwanted and oversimplified solutions, given our lack of defined initial and boundary conditions; we needed to make sure that the supervised loss $\mathcal{L}_s$ was enforced when the physical loss $\mathcal{L}_{phys}$ took over. As such, we used $\lambda_s = 10^{log(\mathcal{L}_{phys})-log(\mathcal{L}_s)}$.

3 Results

We developed three distinct PiNNs, along with a new data-driven baseline, to examine the impact of each physical rule on the resulting surrogate model: one for each physical rule (see equations 1 and 2), and another combining both. To assess and compare our models, we used four metrics: the mean coefficient of variation (MCV), indicating variability relative to the mean of the solution; the mean squared error (MSE) compared to MULTI-VP's test set; and the mass ($\mathcal{L}_{mass}$) and momentum ($\mathcal{L}_{mom}$) conservation losses, as previously defined. Results are presented in Table 1.

	MCV	MSE	$\mathcal{L}_{mass}$	$\mathcal{L}_{mom}$
MULTI-VP Prediction	0.392	–	1.47×10^{6}	2.84×10^{4}
Baseline	0.336	$\mathbf{3.60 \times 10^{-2}}$	2.08×10^{6}	2.11×10^{4}
Mass conservation PiNN	0.311	3.73×10^{-2}	1.00×10^{6}	1.89×10^{4}
Mom. conservation PiNN	0.305	5.85×10^{-2}	7.22×10^{6}	$\mathbf{1.71 \times 10^{2}}$
Mass+Mom. cons. PiNN	0.319	3.77×10^{-2}	$\mathbf{9.83 \times 10^{5}}$	9.45×10^{3}

Table 1: Metrics for each of the evaluated models.

Figure 1 provides a view of the models' outputs. Since these models were designed for integration with other frameworks, our focus lies on their behavior in the latter, more stable stages of the domain. Specifically, we present results from data point 400, which corresponds to approximately 5.5 solar radii. We also observe that the MULTI-VP prediction exhibits a high physical loss that can be attributed to existing outliers and numerical noise. However, this does not necessarily indicate that the simulator is physically inaccurate; instead, it suggests that the metrics used may not fully capture all physical properties.

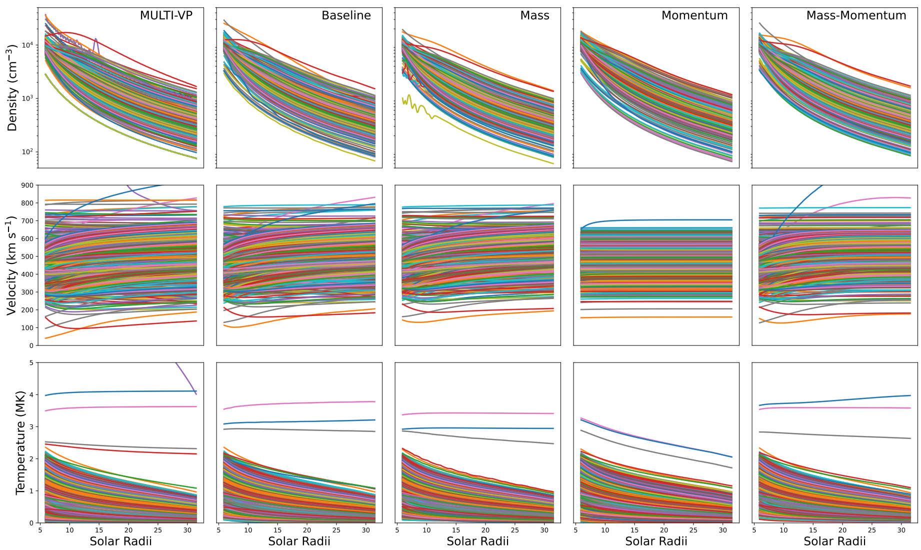

Fig. 1: Density (n), velocity (v), and temperature (T) values, displayed from top to bottom, are shown starting from approximately 5 solar radii. The graph compares test solar wind profiles from MULTI-VP's original samples against the new baseline, as well as three physical models: mass conservation only, momentum conservation only, and both, arranged from left to right.

The models in this paper closely replicate MULTI-VP's outputs under various physical constraints, which would allow them to be used reliably in downstream pipelines. They also exhibit less variability by effectively managing and excluding significant outliers. Moreover, these models can serve as surrogates, reducing MULTI-VP's computation times from hours to seconds.

Increased regularization has improved the model's physical properties while closely resembling MULTI-VP's outputs. This suggests that it is possible to significantly reduce physical loss while maintaining a similar MSE to the baseline, highlighting the delicate balance between physical and unphysical results.

These findings also underscore the importance of accurately representing all physical constraints. Focusing solely on momentum conservation led to significant issues with mass conservation, demonstrating that ignoring any physical constraints or boundary conditions can easily result in unphysical outcomes.

4 Conclusions

This study, explored the feasibility of using physics-informed models as surrogates for the computationally intensive solar wind simulator, MULTI-VP. We demonstrated that these surrogate models are highly effective, providing an alternative to MULTI-VP by significantly reducing the computational time required for solar wind predictions. This reduction in time could broaden the application possibilities. We also found that the physics-informed variants of our data-driven models not only closely mimic MULTI-VP's outputs but also adhere more strictly to the laws of conservation, which are crucial for accurate solar wind prediction. Future work involves further validating these findings by testing the surrogate models with heliospheric models like Helio1D to see if their performance aligns with or surpasses that of MULTI-VP's solutions.

References

[1] Rui F Pinto and Alexis P Rouillard. A multiple flux-tube solar wind model. *The Astrophysical Journal*, 838(2):89, 2017.

[2] Filipa S. Barros, J. Lima, Andre Restivo, Rui Pinto, Paula Graca, and Murillo Villa. Using recurrent neural networks to improve initial conditions for a solar wind forecasting model, 03 2024.

[3] R. Kieokaew, R. F. Pinto, E. Samara, C. Tao, M. Indurain, B. Lavraud, A. Brunet, V. Génot, A. Rouillard, N. André, S. Bourdarie, C. Katsavrias, F. Darrouzet, B. Grison, and I. Daglis. Physics-based model of solar wind stream interaction regions: Interfacing between multi-vp and 1d mhd for operational forecasting at l1, 2023.

[4] Pomoell, Jens and Poedts, S. Euhforia: European heliospheric forecasting information asset. *J. Space Weather Space Clim.*, 8:A35, 2018.

[5] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, February 2019.

[6] Salvatore Cuomo, Vincenzo Schiano di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next, June 2022.

[7] Rui Wang and Rose Yu. Physics-Guided Deep Learning for Dynamical Systems: A Survey, February 2023.

[8] Andreas Eckner. Algorithms for unevenly-spaced time series: Moving averages and other rolling operators. In *Working Paper*, 2012.