ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

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 solar wind coronal 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 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 behavior through space 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 to effectively emulate the solar wind profiles' behavior from the Sun to the Earth.

Modeling the coronal portion of the solar wind is challenging due to its complexity. Recent advancements, such as the MULTI-VP model [1, 2], 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. Leveraging magnetogram-based data but also pre-computed initial conditions of the dynamic system, this model calculates the behavior of the solar wind stream properties along open magnetic field lines from the Sun's surface up until around 30 solar radii, integrating a one-fluid MHD approach and accounting for critical physical processes such as coronal heating, heat conduction, and radiative cooling. To fully model the behavior of solar wind, MULTI-VP interfaces with other heliospheric models like Helio1D [3] and

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

EUHFORIA [4]. These models use MULTI-VP results at around 0.1 Astronomical Units (AU) and extend them to L_1 and beyond [1].

Despite its accuracy, MULTI-VP still produces noisy outputs for some solar wind profiles. Moreover, it is computationally intensive, typically requiring several hours to process inputs and generate forecasts. To mitigate the latter challenge, our recent studies have shown that Neural Networks (NNs) can generate more precise initial conditions, reducing computational time by up to 8% [6]. However, the ongoing need for efficient computation highlights the relevance of exploring data-driven approaches that can offer faster inference times, particularly in scenarios where achieving high accuracy is not the primary concern, such as forecasting, leading us to consider the feasibility of a surrogate model.

However, relying solely on a data-driven surrogate model may lead to physically inconsistent results, complicating its integration with downstream models. Physics-Informed Neural Networks (PiNNs) [7] address this issue by combining data-driven techniques with physical laws. This approach ensures both computational efficiency and physical accuracy, shown in previous research where PiNNs effectively adhere to conservation laws and physical constraints [8, 9].

In this paper, we employed PiNNs to develop surrogate models for MULTI-VP. These models effectively reduced computational cost and numerical noise compared to the simulator while still maintaining prediction robustness. As such, we ensured that the surrogates captured the complex dynamics of the data and adhered to essential physical principles. Consequently, our approach yields more physically consistent simulations across various heliospheric test scenarios.

2 Methodology

2.1 Data Preparation

The MULTI-VP model uses magnetogram-based 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 defined 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}). 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 [6]. For each stream, both the geometric properties (R, B, α) and the MULTI-VP's simulated outputs (n, v, T) were defined 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

[1]The Sun-Earth L_1 Lagrange Point, situated about 1.5 million kilometers from Earth, represents one of the many equilibrium points for small mass objects under the gravitational influence of the Sun and the Earth [5].

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

wind streams from five distinct solar events, but using solely the latter 540 grid points, removing the initial turbulent part of the data domain. We refined our previous approach by incorporating input information for L and A_{exp}. These properties play a key role in defining 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 Centimetre–gram–second (CGS) system of units to ensure consistency with the physical properties that need to be verified. Afterward, we analyzed the quantiles, density distribution, and other relevant statistical measures of our data to choose the most appropriate normalization techniques for each variable. Given the skewness of our data, we applied a logarithmic transformation to all variables but α and T, which did not have that behavior and were, thus, standardly normalized. Additionally, we applied an absolute value transformation to both B and α. For the magnetic field B, the direction was irrelevant in this context, so we focused only on the positive values. For α, any negative values indicated errors in the data that were corrected to maintain consistency with expected positive values.

We also noticed unphysical sinusoidal patterns in the MULTI-VP outputs, likely associated with numerical errors. As such, we smoothed out the outputs n, v, and T with a Butter Low-Pass Filter, which effectively removes the high-frequency components of the outputs associated with the unphysical patterns.

2.2 Model Definition

We used a wide network architecture[2], 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 [10], as commonly done in the literature, to achieve greater stability in training.

The model uses the AdamW [11] 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 0.5 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

[2]Implemented using Pytorch: `https://github.com/biromiro/pinn-multivp`

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

not consider physical laws. PiNNs bridge the gap between these approaches by incorporating physical constraints. In this work, we used two expressions to help our model achieve more physically sound results. The first relates to the principle of mass conservation within the flux tube, meaning mass should remain constant. Mass can be inferred using the equation $\frac{n \cdot v}{B}$, so its standard deviation (σ) should be 0. Therefore, we define the physical property m_c as follows:

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

The second property mirrors the momentum conservation characteristics of a simplified MHD system of differential equations, where G represents the gravitational constant and ν stands for a predefined viscosity constant. As such, the physical property p_c is defined as follows:

$$p_c = P_{grad} + g_{term} + v_{grad} + \nu_{damp} \approx 0 \tag{2}$$

$$P_{grad} = \frac{\partial(n \cdot T)}{\partial L} \cdot n \tag{3}$$

$$g_{term} = G\frac{cos(\alpha)}{R^2} \tag{4}$$

$$v_{grad} = \frac{\partial v^2}{\partial L} - v\frac{\partial v}{\partial L} \tag{5}$$

$$\nu_{damp} = -\nu\left(\frac{\partial^2 v}{\partial L^2} + A_{exp}\frac{\partial v}{\partial L}\right) \tag{6}$$

These properties can be seamlessly integrated as new optimization objectives during neural network training by evaluating them as loss functions. While implementing the mass conservation property that way 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 disrupt spatial correlations among proximal inputs and possibly degrade the performance or complicate the training. Consequently, our approach employs Finite Difference methods by following Eckner's algorithm and adapting it to utilize tensor operations [12].

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 of the modeling equations. Additionally, the effectiveness of the surrogate model heavily relies on its accuracy, particularly in the more stable latter part of the domain. Thus, we only 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 in the training data was identified as $k = 248$. With these considerations in mind, and considering $\mathcal{R}_i$ as either the mass conservation residual (m_c) for Equation 1 or as the momentum conservation residual (p_c) for Equation 2 at grid point i, the loss functions are defined following Equation 2.3, with the $\mathcal{L}_{phys}$ term varying according to what conservation properties were being used (or two terms in the general loss $\mathcal{L}$ for the PiNN tackling both conservations).

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

$$\mathcal{L} = \lambda_s \mathcal{L}_s + \lambda_{phys} \mathcal{L}_{phys} \quad (7) \qquad \mathcal{L}_{phys} = \sqrt{\frac{1}{540-k} \sum_{i=k}^{540} \mathcal{R}_i} \quad (8)$$

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 solely data-driven neural network baseline, to examine the impact of each physical rule on the resulting surrogate: 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 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.

	MCV	MSE	$\mathcal{L}_{mass}$	$\mathcal{L}_{mom}$
MULTI-VP Prediction	0.392	–	1.47×10^6	2.84×10^4
Classical NN 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.

Analyzing the results in Table 1 shows that the models in this paper closely replicate MULTI-VP's outputs under different physical constraints, making them reliable for use in downstream pipelines. They also show less variability by effectively excluding significant outliers. Additionally, these models can act 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.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

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] Evangelia Samara, Rui F. Pinto, Jasmina Magdalenic, Nicolas Wijsen, Veronika Jercic, Camilla Scolini, Immanuel C. Jebaraj, Luciano Rodriguez, and Stefaan Poedts. Implementing the MULTI-VP coronal model in EUHFORIA: Test case results and comparisons with the WSA coronal model. *Astronomy & Astrophysics*, 648:A35, April 2021.

[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] European Space Agency. L1, the first lagrangian point, 2024. Accessed: 2024-08-05.

[6] 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.

[7] 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.

[8] 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.

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

[10] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and Mike Titterington, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 249–256, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.

[11] Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization, January 2019.

[12] Andreas Eckner. Algorithms for unevenly spaced time series: Moving averages and other rolling operators, 2015.