Using Recurrent Neural Networks to improve initial conditions for a solar wind forecasting model

Filipa Barros, Paula Graça, J.J.G. Lima, Rui Pinto, André Restivo, Murillo Villa

Engineering Applications of Artificial Intelligence 133 (2024) 108266

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Research paper

Using Recurrent Neural Networks to improve initial conditions for a solar wind forecasting model

Filipa S. Barros [a,c,e,*], Paula A. Graça [b], J.J.G. Lima [c,d], Rui F. Pinto [e,f], André Restivo [a], Murillo Villa [b]

[a] LIACC/Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, Porto, Portugal
[b] INESC TEC, Rua Dr. Roberto Frias, Porto, Portugal
[c] Instituto de Astrofísica e Ciências do Espaço, CAUP, Rua das Estrelas, Porto, Portugal
[d] Departamento de Física e Astronomia, FCUP, Rua do Campo Alegre 687, Porto, Portugal
[e] Institut de Recherche and Astrophysique et Planétologie, OMP/CNRS, CNES, University of Toulouse, Avenue Colonel Roche 9, Toulouse, France
[f] LDEEE, DAp/AIM, CEA Saclay, Gif-sur-Yvette, Paris, France

ARTICLE INFO

Keywords:
Solar wind
Solar-terrestrial relations
Neural networks
Artificial intelligence
Magnetohydrodynamics
Machine-learning

ABSTRACT

Solar wind forecasting is a core component of Space Weather, a field that has been the target of many novel machine-learning approaches. The continuous monitoring of the Sun has provided an ever-growing ensemble of observations, facilitating the development of forecasting models that predict solar wind properties on Earth and other celestial objects within the solar system. This enables us to prepare for and mitigate the effects of solar wind-related events on Earth and space.

The performance of some simulation-based solar wind models depends heavily on the quality of the initial guesses used as initial conditions. This work focuses on improving the accuracy of these initial conditions by employing a Recurrent Neural Network model. The study's findings confirmed that Recurrent Neural Networks can generate better initial guesses for the simulations, resulting in faster and more stable simulations. In our experiments, when we used predicted initial conditions, simulations ran an average of 1.08 times faster, with a statistically significant improvement and reduced amplitude transients. These results suggest that the improved initial conditions enhance the numerical robustness of the model and enable a more moderate integration time step.

Despite the modest improvement in simulation convergence time, the Recurrent Neural Networks model's reusability without retraining remains valuable. With simulations lasting up to 12 h, an 8% gain equals one hour saved per simulation. Moreover, the generated profiles closely match the simulator's, making them suitable for applications with less demanding physical accuracy.

1. Introduction

Space Weather forecasting is a topic in rapid development that has been the target of numerous new machine-learning (ML) applications (Camporeale, 2019). It aims to understand the physical conditions of Earth's and other solar system bodies' spatial environments and determine and predict the impacts of primarily solar-driven disturbances on infrastructure and biological systems (Lilensten and Belehaki, 2009). Complex chains of events in different regions between the Sun and Earth are frequently responsible for space weather phenomena, requiring the development of models that observe causality ties between a varied collection of events across a wide variety of physical regimes.

In particular, the solar wind – a flow of charged particles that originates in the inner solar corona layers and spreads outward, filling the entire heliosphere – is a time-varying, non-homogeneous flow with the potential to cause a variety of disturbances throughout the whole solar system, some of which have the potential to impact planetary atmospheres significantly. Therefore, forecasting the solar wind accurately and with sufficient lead time is extremely important, but doing it accurately across the entire Sun-to-Earth path and with ample enough lead time is a daunting task.

The development of solar wind forecasting models currently faces a number of obstacles of great importance. To create solar wind models, practitioners depend on spacecraft observations – both remote near-Sun and in-situ near-Earth – having a significant gap in between. Furthermore, the solar wind transits between several physical regimes as it propagates from the former to the latter, being affected by a chain

* Corresponding author at: LIACC/Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, Porto, Portugal.
E-mail address: fbarros@fe.up.pt (F.S. Barros).

https://doi.org/10.1016/j.engappai.2024.108266
Received 12 May 2023; Received in revised form 30 January 2024; Accepted 11 March 2024
Available online 15 March 2024

of different physical processes in its evolution across the interplanetary space (Rouillard et al., 2021). The heliophysics community has been devoting great efforts to addressing this knowledge gap by the launch of space missions such as Solar Orbiter (Müller et al., 2020) and by the development of models that let us establish connections through all these layers of the solar atmosphere (Rouillard et al., 2020). In practical terms, the solar wind modelling problem is usually split into parts, with specialized models focusing on specific sub-regions or sub-processes. Most notably, the magnetically-dominated solar corona (up to about 10% of the Sun–Earth distance) and, just above, the heliosphere (on which thermal pressure and global rotation take over). Maps of the magnetic field at the surface of the Sun (magnetograms) usually constitute the main input data of these models. Modelling the coronal part of the solar wind is usually a much more complex task than that of its heliospheric counterpart. For this reason, practitioners have traditionally used empirical (or semi-empirical) models to circumvent the coronal complexity and provide direct estimations of the solar wind speed at the heliospheric boundary (Arge et al., 2003). The standard alternative has historically been to use complex numerical magnetohydrodynamical (MHD) models of the solar corona that bring invaluable insight into solar wind propagation (Priest, 2014; Solanki et al., 2006; Lionello et al., 2008), but at a very high computational cost. A new approach that bridges these two methods took shape recently by developing and implementing the *MULTI-VP* solar wind model (Pinto and Rouillard, 2017). *MULTI-VP* has been shown to produce good quality predictions of the solar wind at a much smaller computational cost in comparison to standard 3D MHD models (Samara et al., 2021; Poirier et al., 2020). Computational performance is paramount in space weather forecasting, so as to allow for the longest possible lead time and also to make ensemble forecasting possible.

MULTI-VP computes vast collections of individual (or elemental) solar wind streams that flow along open magnetic field lines across the solar corona and up to about 15% of the Sun–Earth distance. These elemental flows sample the whole three-dimensional solar atmosphere or any specific region of interest. Each solar wind solution is constrained by the geometry of the magnetic flux tubes it crosses, namely by the amplitude of the surface field and by its expansion factor and inclination profiles (provided as a function of distance to the Sun measured along its curvilinear coordinate). The code then solves for a one-fluid flow taking into account a coronal heating model, heat conduction, and radiative cooling.

These *MULTI-VP*'s elemental solutions can, in principle, be computed *ab initio*, but at an unnecessarily high computational cost (one full 3D run consists of several thousands of elemental solutions). Instead, we provide pre-computed initial conditions – initial guesses that consist of full solar wind profiles – that the code then drives towards a final (relaxed) solution. Currently, this relaxation process consumes a significant fraction of the computational load required to obtain a solar wind solution. Consequently, improving the quality of the initial guesses should lead to invaluable gains in total computing time as well as smoother initial relaxation transients.

This manuscript investigates whether and how we can use machine-learning to enhance the initial guesses used in *MULTI-VP* to improve its computational performance and serve as a potential basis for a future surrogate model. In particular, this paper addresses whether "Elemental solutions predicted by a Recurrent Neural Network (RNN) are good enough to be used as initial guesses for MULTI-VP thus decreasing the average time needed to perform a simulation".

2. Previous ML applications

ML has been applied to various space weather phenomena like solar wind forecasting, solar flare detection, and understanding the solar activity cycle.

Numerous works have leveraged ML to enhance the initial data of simulations. Notable examples include the work of Kochkov et al. (2021) where fluid dynamics simulations are accelerated and improved by the use of deep learning, as well as Lattimer et al. (2020)'s and Watson (2019)'s. In the latter two works, artificial neural networks (ANNs) are combined with physically derived models to predict the chaotic Lorenz'96 system.

ML has evidently established itself as a valuable tool in space weather and physics simulations. ML's power lies in its ability to analyse vast datasets, unveiling hidden patterns that elude traditional methods. In recent years, specific ML techniques have gained attention, notably Recurrent Neural Networks (RNNs).

A significant breakthrough in this context is Sexton et al. (2019) which developed an RNN capable of forecasting the planetary Kp-index for up to 24 h, surpassing existing models. Leveraging three decades of historical data from NASA's Omni Virtual Observatory, this RNN incorporated various solar wind parameters, including Bz, n, V, $|B|$, rB, and rBz, along with Kp values. The RNN excelled in predicting the Kp-index, demonstrating strong correlations ranging from 0.8189 for short-term forecasts to 0.8211 for nine-month predictions, with the lowest Root Mean Square Error (RMSE). This highlighted the RNN's reliability and predictive prowess in solar weather analysis.

Several other works in the field emphasize the critical role of initial conditions in physical simulations and models.

For instance, in cosmological *zoom-in* simulations (Hahn and Abel, 2011), multiscale initial conditions are generated with multiple levels of refinement. In this case, a hybrid Poisson solver, which achieves significantly higher accuracy than previous methods, makes it applicable to simulations involving two-component baryons and dark matter. This solver accurately reproduces correlation functions, density profiles, fundamental halo properties, and sub-halo abundances.

Crocce et al. (2006) delve into the impact of using initial conditions in numerical cosmology simulations. Their research highlights the superiority of second-order Lagrangian perturbation theory (2LPT) over Zel'dovich approximation (ZA) initial conditions, as 2LPT substantially reduces nonlinearities and enhances the precision of cosmological simulations.

Prunet et al. (2008) describe a software package used to produce initial conditions in large-scale cosmological simulations. The codes were validated up to resolutions of 4096^3 initial resolution elements and used to generate the initial conditions of large hydro-dynamical and dark matter simulations, proving to be more versatile than previous solutions.

In another work, Brown and Gnedin (2021) propose an approach of customization to the root grid zoom-in initial conditions utilized for simulations of galaxy formation, achieving a twofold CPU runtime improvement due to dimensionality reduction.

Jasche and Wandelt (2013) describe a probabilistic physical model of a nonlinearity evolved density field. The researchers reconstruct the current density and velocity fields from large-scale structure surveys using Bayesian exploration, along with an entire propagation of the observational uncertainties in the initial conditions. The statistical consistency of the reconstructed initial conditions with the inputs of the Gaussian simulation is demonstrated by tests. This demonstrates that statistical approaches based on physical models of large-scale structure distribution are now becoming feasible for realistic current and future surveys.

In summary, incorporating ML, especially RNNs, to improve initial conditions for solar wind simulations holds the potential for a significant advancement in the efficiency of these simulations. There is optimism that, owing to their robust and memory-based architecture, as well as their probable capability to identify intricate patterns and correlations among solar wind parameters, RNNs could empower the scientific community to extract valuable insights from extensive datasets, establishing them as indispensable tools in space weather research.

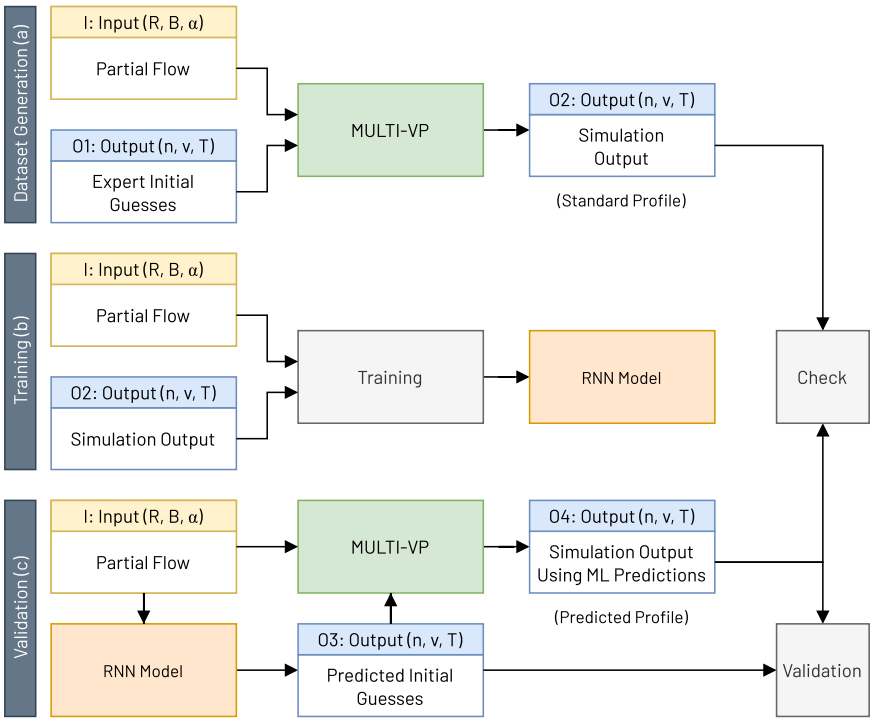


Fig. 1. Three steps of the experimental setup: (a) We generated the dataset by running *MULTI-VP* for a large set of partial flows, (b) we trained a model capable of predicting better initial guesses, and (c) we used that model and checked if the results of the simulation were similar to the original results, and if the predicted guesses were closer to the simulation output.

3. Data and methodology

MULTI-VP takes in magnetogram data (maps of the magnetic field at the surface of the Sun) from which the global three-dimensional topology of the solar coronal magnetic field is deduced. It then selects and traces an ensemble of open magnetic field lines (those rooted at the Sun's surface but extending outwards into the interplanetary medium). Each line will host an elemental solar wind solution (or wind stream). Each elemental wind stream is fully constrained by the geometry of its driving field line (*i.e.*, the profiles of magnetic field amplitude (B) and inclination (α) in respect to the vertical direction as a function of the distance to the Sun (R)) and by a chosen thermal model derived from previous theoretical work that takes into account coronal heating, heat conduction, and radiative cooling. The outputs of the simulator consist of a profile of plasma density (n), velocity (v), and temperature (T) given as a function of the distance to the Sun along each single one of the field lines traced (up to about 30 solar radii).

Table 1 summarizes these input and output parameters. The first column is the line index, and the following three represent the geometry of open magnetic flux tubes derived from reconstructions of the coronal magnetic field based on magnetograms: R represents the radial coordinate radius, B stands for the magnetic field intensity, and α denotes the inclination angle of the flux-tube in respect to the vertical direction. The last three correspond to the physical quantities computed by the *MULTI-VP* simulation (or equivalently to their corresponding initial conditions): n, the number of protons per unit volume (number of ionized H protons), v, the speed-oriented along the line, and T, the temperature at that point in space. For simplicity's sake, we will refer to the first three of these values (R, B, and α) as the input and the last three (n, v, and T) as the output.

Although theoretically not needed, for performance reasons, *MULTI-VP* receives not only the simulation input but also initial guesses for the output values. The closer these initial guesses are to the final output, the faster the simulation will relax into its final state. Currently, these initial guesses are provided by a solar wind expert.

Table 1
Magnetic flux-tubes geometry and physical quantities calculated by *MULTI-VP* for a single solar wind profile over its 640 abscissas (and their respective units).

index	$R\ [R_{sun}]$	$B\ [G]$	$\alpha\ [\mathrm{deg}]$	$n\ [\mathrm{cm}^{-3}]$	$v\ [\mathrm{km/s}]$	$T\ [MK]$
1	R_1	B_1	α_1	n_1	v_1	T_1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
640	R_{640}	B_{640}	α_{640}	n_{640}	v_{640}	T_{640}

Since the used data is a many-to-many sequence of multiple spatiotemporal steps as input mapped to a sequence with multiple steps as output and there are intricate patterns between the provided inputs, an RNN model, in detriment to a simple ANN model, has been chosen to deal with such connections. Even-though RNNs are known to have vanishing and exploding gradient limitations, Weerakody et al. (2021), these are mostly due to long-term dependencies, Lin et al. (1996), which our data does not possess.

As such, we hypothesize that a RNN model, trained with several *MULTI-VP* simulation results, should be able to infer good initial guesses for n, v, and T from the values of R, B, and α.

To test this hypothesis, we randomly selected five different dates from a large pool of simulations, each date containing several distinct instances of solar wind profiles, with each of these corresponding to a single independent simulation of an elementary solar wind stream (driven along a given magnetic flux-tube) in the format depicted in Table 1.

The whole process can be seen in Fig. 1. The top section (a) depicts the dataset generation. *MULTI-VP* receives a partial flow as input (I), as well as an initial guess hand-crafted by an expert (O_1). By running *MULTI-VP* for each one of these pairs ($\langle I, O_1 \rangle$), we will generate a new dataset of paired input and output files ($\langle I, O_2 \rangle$). The middle section (b) shows how this dataset ($\langle I, O_2 \rangle$) can be used to train an RNN model capable of predicting initial guesses for each partial data flow. The bottom section (c) depicts the testing and validation process of the work, where the output of the model, coupled with the partial

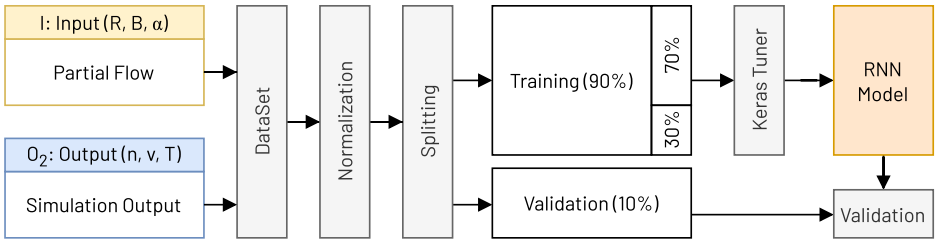


Fig. 2. Training Process and model validation: Creating the dataset from simulation runs, normalizing that dataset, and splitting into training, testing and validation. Keras Tuner was used to generate the RNN model.

flow ($\langle I, O_3 \rangle$), is used by the *MULTI-VP* simulator to generate the final outputs (O_4).

In the end, we compared the outputs of the initial simulation (O_2) and the outputs of the new simulation (O_4) to check if the simulation results were congruent; the final results of the simulation (O_4) were also compared with the predicted initial guesses (O_3) to check if the predictions were closer than the expert initial guesses. The speed-up between running the simulation using expert or predicted initial guesses was also measured.

The training process is detailed in Fig. 2. We started by generating a dataset from $\langle I, O_2 \rangle$ and splitting it (90%/10%) into training and validation datasets. We normalized the data using SciKit's QuantileTransformer. We then used Keras's RandomSearch Tuner (with a 70%/30% split between training and testing data) to find the Recurrent Neural Network that minimized our loss, in this case, the minimum squared error (MSE).

To employ the RandomSearch tuner, we supplied it with a hyper model and specified various hyperparameters for the search. We set 500 training epochs and a maximum of 100 trials, representing the combinations of hyperparameters to be tested. Each trial involved two executions to assess the performance of a particular combination on the objective. The range for the number of layers was set between 5 and 100, the number of hidden units ranged from 16 to 128, and the activation functions included relu/tanh (rectified linear unit/hyperbolic tangent) and sigmoid. For model training, we utilized the ADAM optimizer, as outlined in Kingma and Ba (2014). The choice of this optimizer was driven by its computational efficiency, minimal memory requirements, resilience in the face of diagonal gradient rescaling, and suitability for handling large datasets with intricate parameter configurations.

4. Results

A dataset of 12,980 simulated wind profiles over five non-consecutive days was collected, with each day containing 2,596 unique profiles that had already been processed by *MULTI-VP* (see Fig. 1a). After removing 1,760 profiles due to corruption or missing data, there were 11,220 profiles left for further analysis. To validate the performance of our model, we set aside 10% of the profiles (1,122) and used the remaining data to generate and train the RNN model (see Fig. 1b).[1]

The final trained RNN model found by Keras RandomSearch tuner had 92 layers, with the hidden units ranging between 16 and 124 and the activation functions varying between the three mentioned before. The training MSE of the best model was 8.52×10^{-2}. For each of the 1,122 wind profiles set aside for validation, we employed this RNN model to generate initial guesses ($O3$), which were then used to rerun the simulations (see Fig. 1c). The final validation MSE on these results was the same as the training one. RMSE was also used to measure the loss of the predicted profiles. The validation RMSE was of approximately 0.29. According to Moriasi et al. (2007), these MSE and RMSE values are considered low.

To ensure the reliability of our results, we performed a sanity check by comparing the simulation outputs using expert initial guesses ($O2$) with those generated by the ML model ($O3$). Our analysis revealed that the differences between the simulation outcomes were well within the simulator's precision margin, thus providing further evidence of the accuracy of the results.

Fig. 3 displays the results of running the simulation for the 1,122 validation profiles using the expert initial guesses ($O2$) and the predicted initial guesses ($O3$). Although the model is less discerning than the simulation, the predicted values for n, v, and T are overall consistent with the simulation outputs. Notice that the figure shows at least one instance – in orange – where the simulator has misbehaved. Albeit rare, this situation can occur whenever the initial guess provided to the system is inappropriate and leads to transients that are to strong for the numerical scheme to handle. The model, however, seems to not have learnt that erroneous behaviour and, in consequence, to have predicted a regular solar wind profile.

4.1. Simulator experiments

To gain insight into the simulation process, we analysed the results of running simulations using both expert initial guesses and predicted initial guesses. Fig. 4 illustrates the difference between the two runs for a specific profile. The lines represent the values of the variables (n, v, and T) at each of the 640 data points, with the red line indicating the initial guesses and the green line indicating the final output of the simulator.

The analysis revealed that the simulation converged correctly in both cases, consistent with our earlier findings. Furthermore, we observed that the difference between the red and green lines was generally smaller using the predicted values for most profiles, indicating that the predicted initial guesses were closer to the final solution. Additionally, the convergence process seemed smoother and more straightforward for most cases, with a smoother initial transient. Notably, the density profile tended to be closer to the final solution than the standard profiles. We arrived at these observations by analysing multiple graphs of other profiles as well.

4.2. Temporal evolution

We then analysed the temporal evolution of the relative variations of each quantity between two consecutive data outputs for each run. Specifically, we calculated the ratio $\left(X_i - X_{i-1}\right)/X_{i-1}$ for a given quantity X at a fixed position in the upper part of the numerical

[1] There were around 2 GB of training data, and around 0.2 GB of validation data. The model was trained on an Intel® Core™ i5-10300H CPU @ 2.50 GHz processor with 16 GB of RAM. It took around 36 h to train the RNN for 500 epochs.

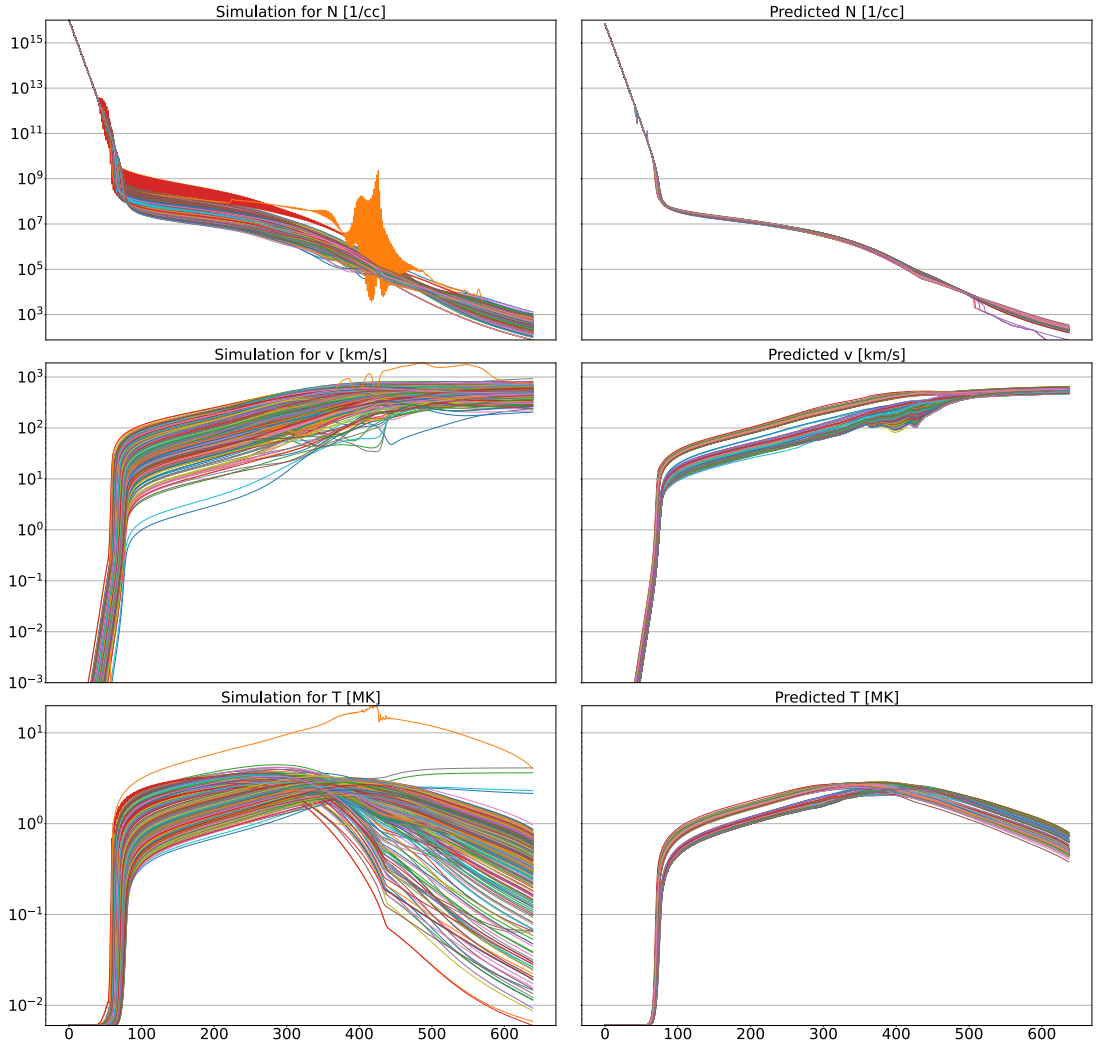


Fig. 3. Comparison of the n [cm^{-3}], v [km/s], and T [MK] (top to bottom) outputs from the original simulation dataset (on the left, ($O2$)) and the predicted ones from the RNN model (on the right, ($O3$)) using a logarithmic scale. The x-axis corresponds to the grid-point number and y-axis to the variables' values for each of those points.

domain. Fig. 5 shows the evolution of this ratio for a specific profile. By examining the evolution for this and other profiles, we observed that using the predicted values resulted in a faster convergence of the simulation compared to using standard profiles.

The simulations using predicted initial guesses generally seem to generate lower amplitude transients, which is twofold beneficial from a numerical point of view because it makes the calculation more robust (less likely to produce transients excessively stiff for the solver), and because it allows for the code to keep a more moderate integration time step.

We also noted a series of oscillations at well-defined frequencies that follow the initial impulse. These do not represent numerical noise: it is a well-defined oscillatory mode – which corresponds to an acoustic-type compressible mode or, more precisely, to the slow MHD mode – that is excited as a secondary response to perturbations in the structure of the transition region between the chromosphere and the solar corona (*i.e.*, between the colder and denser, and the hotter and rarefied parts of the atmosphere, see Pinto et al. (2009), Griton et al. (2020)).

4.3. Performance

The simulations exhibit different initial transients, with smoother transients observed when using predicted input flows. Additionally, the integration time step automatically adapts to the complexity of the simulated flow, causing it to vary between the two realizations. Consequently, the number of total iterations required for the code to converge also differs. As such, to measure performance, we used the number of iterations rather than the elapsed time. The speed-up factor (sf) was calculated using the formula $sf = nis/nip$, where nis represents the number of iterations using standard profiles and nip represents the number of iterations using predicted profiles. The mean speed-up was obtained and found to be 1.0796.

A box plot of the natural logarithm of the speed-up per file is shown in Fig. 6. The median value is slightly above 0, and most of the Q1–Q3 values are positive, indicating a small positive improvement in the time needed to run the simulations. However, further analysis was required to understand whether there were statistical differences between the speed-up using standard and predicted profiles. A significant number of outliers were observed on both sides of the result.

The *Shapiro–Wilk* test was used to check for normality of the measured effect, and a *p-value* of 0.0 was obtained, indicating that the distribution was not normal. A quantile–quantile plot confirmed that the distribution was heavy-tailed. Further analysis revealed that the number of iterations from both groups had the same shape and scale, enabling a *Wilcoxon* signed-rank test to be performed. The *p-value* score

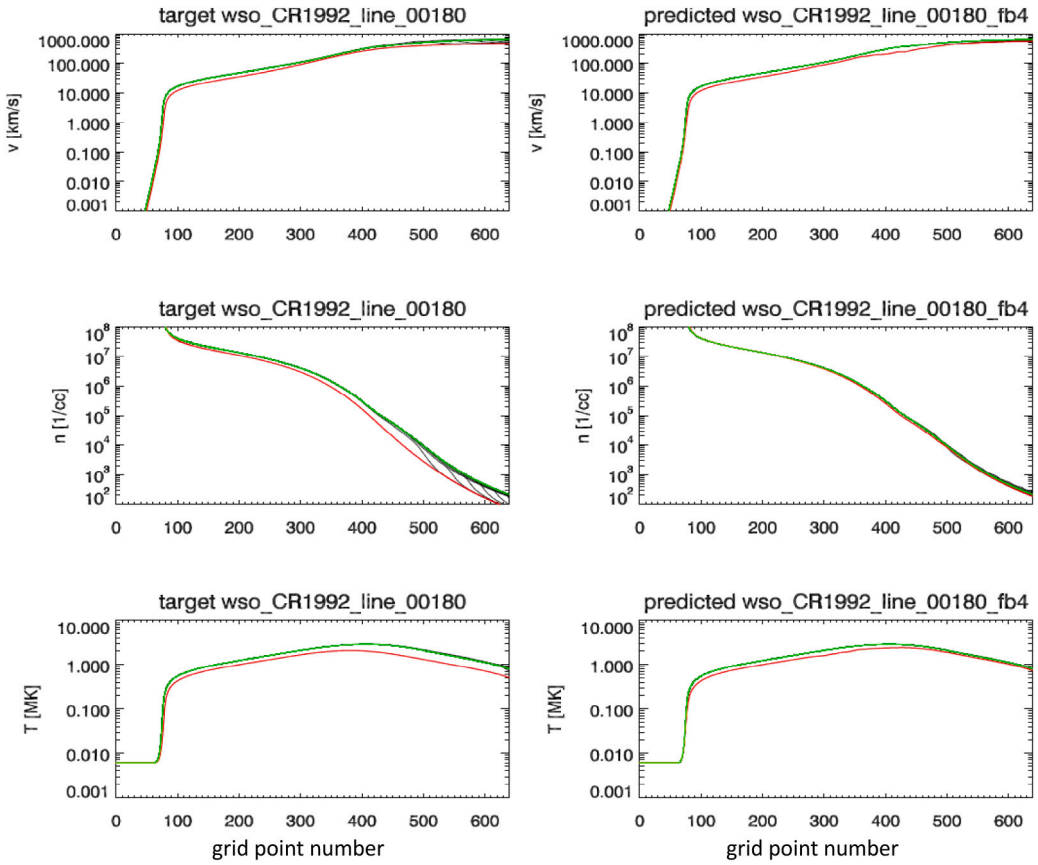


Fig. 4. Illustrative example of comparison of standard and predicted profiles performance using one of the tested predicted profiles. The standard profile can be seen on the left panels, while the new profile can be seen on the right. Red lines represent the initial condition, green lines represent the final solutions, and black lines represent some intermediate instants, set with the same time-cadence (Δt) on both sides. The abscissa represents the grid-point number of the simulation, like in Fig. 3.

obtained was 6.7335^{-29}, allowing us to reject the null hypothesis, as the result is significant at $p < 0.05$. This supports the hypothesis that using the predicted profiles has a statistically significant positive effect on the performance of the simulation.

4.4. Error analysis

To validate our hypothesis that the improved performance of the simulation was due to the better initial guesses generated by our RNN model, we conducted an additional analysis to compare the accuracy of the expert and predicted initial guesses. Specifically, we calculated the average difference between the simulation output and each initial guess for every abscissa. Fig. 7 displays the results, which indicate that the average error was consistently smaller for the predicted initial guesses compared to the expert guesses. For clarity, the plots exclude the lowest layers of the solar atmosphere (below gridpoint 100) that consists of very low temperature, very dense plasma, and with very low velocities, that is, with no wind flow yet developed (cf. Fig. 4). Furthermore, this dense layer spontaneously develops small amplitude oscillations that have no impact on the state of the computed solar wind flow, but could be misidentified as an error in the predicted profiles. This finding provides further evidence that the RNN model can generate more accurate initial guesses, supporting our hypothesis.

4.5. Validation threats

We identify the following threats to the validity of our results (Campbell and Cook, 1979; Campbell and Stanley, 2015):

Data integrity, representativeness and bias (external). because *MULTI-VP* uses massive volumes of data, the number of the dimensions of the ML modelling features makes it difficult to ensure data's integrity and representativeness. Because our data is manually selected by humans, there is also a tendency for it to be biased. Another important matter is that the data comes pre-processed, which might imply some noise or over processing.

Explainability challenges (internal). ML models (primarily neural network-based ones) are not easy to comprehend and are frequently perceived as black boxes. The complexity and architecture of neural networks make it challenging to evaluate the varying selection process and explainability of driving factors. Even if these models outperform traditional ones, the lack of explainability may cause ML models to be restricted in use by specialized data scientists. Since the model would be used by astronomers and astrophysicists, a GUI of some sort could be presented as a solution to this threat.

Parameter and method selection (internal). ML models involve scaling, normalization, parameter optimization, randomization, activation functions. How these parameters and hyperparameters are selected when developing models can impact test error estimation – in our case, the MSE – and the absolute error.

5. Discussion and conclusion

Machine-learning has become a popular approach for addressing space weather problems, and in this work, our aim was to investigate

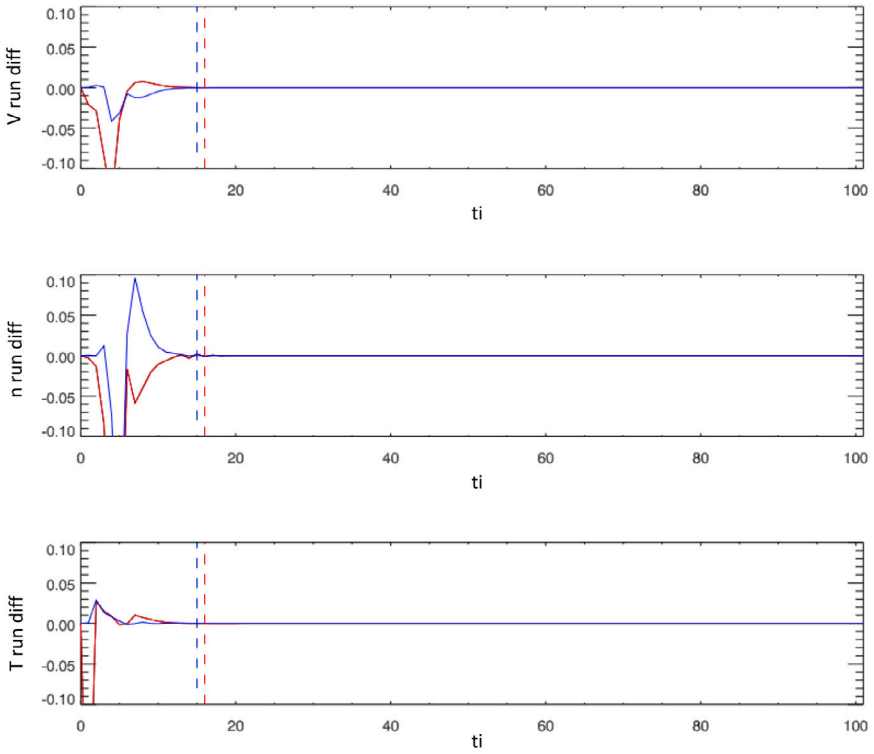
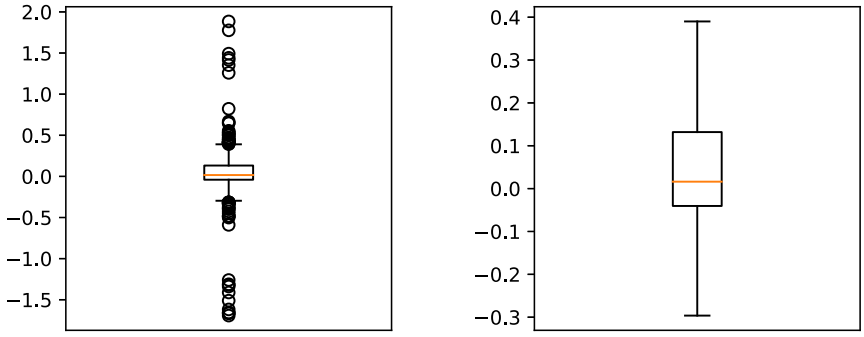


Fig. 5. Illustrative example of convergence on one of the tested profiles. Each graph displays the relative variations of n, v, and T, measured at a reference altitude between successive code outputs ("running differences"). The profiles generated using expert initial guesses are shown in red, while the ones generated using predicted initial guesses are shown in blue. The abscissa (x-axis) shows the elapsed time in code units. The vertical dashes indicate the instant when convergence was detected, which is defined as the point when the relative variations of all three quantities oscillate less than a predefined threshold value.



(a) With outliers **(b)** Without outliers

Fig. 6. Natural logarithm of the ratio between the number of iterations needed to simulate each profile using expert and predicted guesses as a box plot. Values above zero represent profiles where using predicted guesses had a positive outcome on the performance of the simulation. The box extends from the Q1 to the Q3 of the data, with a line at the median. The whiskers extend from the box by 1.5x the interquartile range (IQR), and any data points beyond the whiskers are considered outliers and represented as fliers in (a).

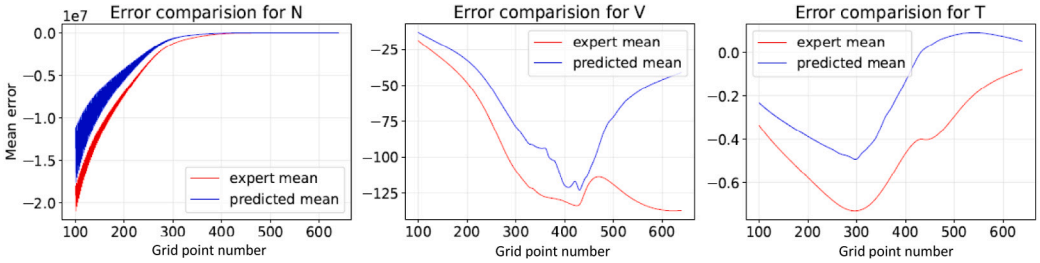


Fig. 7. Comparison of the average difference between the simulator output and the expert/predicted initial guesses. The average error is consistently smaller for the predicted initial guesses compared to the expert guesses.

the potential of ML techniques to enhance the performance of *MULTI-VP*'s simulator. To this end, we utilized a neural network approach to formulate our central hypothesis.

Specifically, we designed and tested a novel neural network model comprising 92 hidden layers, which uses flow data as inputs to predict n, v, and T values for each flow. We obtained results that closely matched the expected ones, thereby confirming the applicability of these predicted values in *MULTI-VP* simulations.

Next, we used our RNN model to predict 1,122 randomly chosen profiles, and used them as initial guesses in the *MULTI-VP* simulator, comparing their performance against simulations using expert initial guesses. Our analysis indicated that in most cases, there was a small but statistically significant improvement in the convergence time of simulations. To validate our results, we employed the *Wilcoxon* signed-rank test, achieving a high level of statistical significance with a *p-value* of 6.7335^{-29}.

Since the RNN can be trained with years of simulations that encompass several solar cycles, the evolving solar wind patterns can be captured in a single training. Even though the simulation convergence time improvement was modest, since the RNN model does not need to be retrained to be used, it still has valuable applications. With a range of up to 12 h per simulation, an improvement of 8% is a gain of 1 h per simulation. Furthermore, since the generated profiles are very close to the ones produced by the simulator, these can be used as a first instantaneous result for applications that can tolerate lower physically accuracy. Moreover, the same model typology and approach can be used for different applications where initial conditions are required by physical simulators.

In addition to improving simulation time, the study revealed that the predicted initial guesses were closer to the actual final solutions on average. This closer proximity to the simulated solution resulted in a more stable and efficient convergence process. Overall, using our ML technique for generating initial conditions resulted in a twofold benefit: faster simulation times and improved simulation accuracy.

5.1. Future work

The RNN model developed in this study has shown promising results in predicting initial guesses for *MULTI-VP* simulations. While our statistical analysis supports our main hypothesis that the RNN model provides better guesses than current expert guesses on average, we also observed that in some cases, the model performed worse than expected. We have identified several potential reasons for this:

Corrupted profiles: We identified several corrupted profiles during our work, which may have influenced the training process and resulted in a suboptimal RNN model. To address this issue, we suggest using Generative Adversarial Networks (GAN) or Variational Autoencoders (VAE) to identify outliers and corrupted profiles and remove them from the training process.

Different types of solar wind: There are different types of solar winds, and it may be challenging for an RNN model to distinguish between them. Fig. 3 showed that the model struggled to predict values for different types of profiles. To address this issue, we suggest using clustering techniques to identify different types of profiles and train separate models for each type of solar wind.

Error estimation: The mean squared error (MSE) metric may not be the best metric for training these types of models, since a low error does not always guarantee a physically realistic result. We suggest exploring the use of Physics-informed Neural Networks (PINN) in a future work.

In conclusion, our work has room for improvement not only in validation but also in optimization and enrichment. We are confident that future research will yield even better results, potentially leading to the development of a physically-informed surrogate ML model capable of approximating the output of *MULTI-VP* without running the simulation. This would be especially useful in cases where a quick approximation of the simulation output is needed.

CRediT authorship contribution statement

Filipa S. Barros: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Paula A. Graça:** Software, Writing – review & editing. **J.J.G. Lima:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition, Methodology, Project administration. **Rui F. Pinto:** Conceptualization, Data curation, Project administration, Resources, Supervision, Visualization, Writing – review & editing, Funding acquisition. **André Restivo:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing, Methodology. **Murillo Villa:** Software, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was partially funded by the Portuguese Foundation for Science and Technology (FCT, Portugal), under the research grant number UI/BD/153350/2022 as well as through the research grants UIDB/04434/2020 and UIDP/04434/2020.

RFP acknowledges support by the ERC synergy grant Whole Sun (#810218), from the French space agency (Centre National des Études Spatiales; CNES; https://cnes.fr/fr) that funds the activity of the space weather team in Toulouse (Solar-Terrestrial Observations and Modelling Service; STORMS; http://storms-service.irap.omp.eu/), and of the European Union's Horizon 2020 research and innovation programme under grant agreement No 87043 (SafeSpace project). The numerical simulations used HPC resources from CALMIP (Grant 2020-P1504). *MULTI-VP* has been implemented on the SafeSpace Space Weather forecasting demonstrator[2] and on the European Virtual Space Weather Modelling Centre[3] (Poedts et al., 2020).

References

Arge, C.N., Odstrcil, D., Pizzo, V.J., Mayer, L.R., 2003. Improved method for specifying solar wind speed near the sun. In: AIP Conference Proceedings, vol. 679, (1), American Institute of Physics, pp. 190–193.

Brown, G., Gnedin, O.Y., 2021. Improving performance of zoom-in cosmological simulations using initial conditions with customized grids. New Astron. 84, 101501.

Campbell, D.T., Cook, T.D., 1979. Quasi-Experimentation. Rand Mc-Nally, Chicago, IL.

Campbell, D.T., Stanley, J.C., 2015. Experimental and Quasi-Experimental Designs for Research. Ravenio Books.

Camporeale, E., 2019. The challenge of machine learning in space weather: Nowcasting and forecasting. Space Weather 17 (8), 1166–1207.

Crocce, M., Pueblas, S., Scoccimarro, R., 2006. Transients from initial conditions in cosmological simulations. Mon. Not. R. Astron. Soc. 373 (1), 369–381.

[2] https://www.safespace-h2020.eu/

[3] accessible via ESA's Space Weather Service Network

Griton, L., Pinto, R.F., Poirier, N., Kouloumvakos, A., Lavarra, M., Rouillard, A.P., 2020. Coronal bright points as possible sources of density variations in the solar corona. Astrophys. J. 893 (1), 64.

Hahn, O., Abel, T., 2011. Multi-scale initial conditions for cosmological simulations. Mon. Not. R. Astron. Soc. 415 (3), 2101–2121.

Jasche, J., Wandelt, B.D., 2013. Bayesian physical reconstruction of initial conditions from large-scale structure surveys. Mon. Not. R. Astron. Soc. 432 (2), 894–913.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kochkov, D., Smith, J.A., Alieva, A., Wang, Q., Brenner, M.P., Hoyer, S., 2021. Machine learning–accelerated computational fluid dynamics. Proc. Natl. Acad. Sci. 118 (21).

Lattimer, B., Hodges, J., Lattimer, A., 2020. Using machine learning in physics-based simulation of fire. Fire Saf. J. 114, 102991.

Lilensten, J., Belehaki, A., 2009. Developing the scientific basis for monitoring, modelling and predicting space weather. Acta Geophys. 57 (1), 1.

Lin, T., Horne, B.G., Tino, P., Giles, C.L., 1996. Learning long-term dependencies in NARX recurrent neural networks. IEEE Trans. Neural Netw. 7 (6), 1329–1338.

Lionello, R., Linker, J.A., Mikić, Z., 2008. Multispectral emission of the Sun during the first whole Sun month: Magnetohydrodynamic simulations. Astrophys. J. 690 (1), 902.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Müller, D., Cyr, O.S., Zouganelis, I., Gilbert, H.R., Marsden, R., Nieves-Chinchilla, T., Antonucci, E., Auchère, F., Berghmans, D., Horbury, T., et al., 2020. The solar orbiter mission-science overview. Astron. Astrophys. 642, A1.

Pinto, R., Grappin, R., Wang, Y.-M., Léorat, J., 2009. Time-dependent hydrodynamical simulations of slow solar wind, coronal inflows, and polar plumes. Astron. Astrophys. 497 (2), 537–543.

Pinto, R.F., Rouillard, A.P., 2017. A multiple flux-tube solar wind model. Astrophys. J. 838 (2), 89.

Poedts, S., Kochanov, A., Lani, A., Scolini, C., Verbeke, C., Hosteaux, S., Chané, E., Deconinck, H., Mihalache, N., Diet, F., et al., 2020. The virtual space weather modelling centre. J. Space Weather Space Clim. 10, 14.

Poirier, N., Kouloumvakos, A., Rouillard, A.P., Pinto, R.F., Vourlidas, A., Stenborg, G., Valette, E., Howard, R.A., Hess, P., Thernisien, A., et al., 2020. Detailed imaging of coronal rays with the parker solar probe. Astrophys. J. Suppl. Ser. 246 (2), 60.

Priest, E., 2014. Magnetohydrodynamics of the Sun. Cambridge University Press.

Prunet, S., Pichon, C., Aubert, D., Pogosyan, D., Teyssier, R., Gottloeber, S., 2008. Initial conditions for large cosmological simulations. Astrophys. J. Suppl. Ser. 178 (2), 179.

Rouillard, A.P., Pinto, R., Vourlidas, A., De Groof, A., Thompson, W., Bemporad, A., Dolei, S., Indurain, M., Buchlin, E., Sasso, C., et al., 2020. Models and data analysis tools for the Solar Orbiter mission. Astron. Astrophys. 642, A2.

Rouillard, A.P., Viall, N., Pierrard, V., Vocks, C., Matteini, L., Alexandrova, O., Higginson, A.K., Lavraud, B., Lavarra, M., Wu, Y., et al., 2021. The solar wind. Solar Phys. Solar Wind 1–33.

Samara, E., Pinto, R.F., Magdalenić, J., Wijsen, N., Jerčić, V., Scolini, C., Jebaraj, I.C., Rodriguez, L., Poedts, S., 2021. Implementing the MULTI-VP coronal model in EUHFORIA: Test case results and comparisons with the WSA coronal model. Astron. Astrophys. 648, A35.

Sexton, E.S., Nykyri, K., Ma, X., 2019. Kp forecasting with a recurrent neural network. J. Space Weather Space Clim. 9, A19.

Solanki, S.K., Inhester, B., Schüssler, M., 2006. The solar magnetic field. Rep. Progr. Phys. 69 (3), 563.

Watson, P.A., 2019. Applying machine learning to improve simulations of a chaotic dynamical system using empirical error correction. J. Adv. Modelling Earth Syst. 11 (5), 1402–1417.

Weerakody, P.B., Wong, K.W., Wang, G., Ela, W., 2021. A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing 441, 161–178.