Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

DIAS
Institiúid Ard-Léinn Bhaile Átha Cliath | Dublin Institute for Advanced Studies

Modelling of the Solar Wind in Near-Mercury Space

Sarah Kelly

January, 2025

Supervised by Dr Matthew Rutala

Submitted in partial fulfilment of the requirements for the degree of
B.A Physics and Astrophysics

Abstract

This exploratory study investigates the compatibility of the MMESH software with in-situ data from Parker Solar Probe and the Tao+ solar wind propagation model in Mercury's orbital zone. MMESH is designed to temporally shift a propagation model using constant time offsetting and dynamic time warping in order to better align it with given solar wind data. The results of this process are then used to predict future adjustments to the model incorporating multiple linear regression to characterise its errors. No significant improvements were made however the software certainly worked and by making adjustments that ensure optimal results with such variable data as that at Mercury, a very valuable tool is within reach.

Acknowledgements

A massive thank you to Dr. Matthew Rutala for his continued support and encouragement and to the rest of the team at DIAS for their warm welcome. The Parker Solar Probe in-situ data and Tao+ solar wind propagation model data were acquired from AMDA. Context data is from SPICE and the Penticton solar radio telescope in Canada. The MMESH software is the work of Dr. Matthew Rutala.

Contents

List of Figures

List of Tables

1 Introduction

This project is a proof of concept study, investigating the compatibility of the Multi-Model Ensemble System for the outer Heliosphere (MMESH) software with modelling of the solar wind near Mercury. MMESH is a software, developed by Dr. Matthew Rutala in 2024, that uses constant time shifting and dynamic time warping to better align a solar wind model with in-situ data at that planet, in the hopes of improving our understanding and forecasting of the solar wind. The MMESH study includes a Multi-Model Ensemble (MME) incorporating three different solar wind models and from them producing one improved solar wind propagation model. This is not however within the scope nor required in this project.

As in the name, MMESH was designed for the outer heliosphere and a forward propagation of context data collected at Earth. It has been tested on the solar wind in near-Jupiter space, with promising results [16]. Mercury was selected as the planet of interest for this project for a number of reasons. Since it orbits in the inner heliosphere, it tests the backward propagation of Earth-based data. With its close proximity to the Sun, the solar wind conditions at Mercury are unmatched elsewhere in our solar system, providing a unique data base with which to test such solar wind model adjustment programs. Finally, with the upcoming BepiColombo mission due to reach Mercury's orbit in November 2026, accurate prediction of the solar wind in this region is in great interest.

The in-situ data used in this study is from Parker Solar Probe (PSP). It covers multiple date ranges over the years 2019 and 2020 when PSP was at Mercury-like heliolatitude and distance from the Sun. The solar wind propagation model used is the Tao+ model developed in 2005 to better understand the solar wind around planets in our solar system [20].

2 Background and Theory

2.1 The Solar Wind

The solar wind is a continuous outflow of magnetised plasma originating in the Sun's hot corona and travelling along the Sun's magnetic field lines into interplanetary space.

It consists of steady fast streams, variable slow flows, and transient slow and fast coronal mass ejections [12], behind which is the background solar wind, associated with the hydrodynamic expansion of the solar corona [6]. The Sun follows an 11 year cycle, building to a solar maximum which exhibits an increase in factors such as radio flux, solar irradiance, magnetic field strength, solar flares, coronal mass ejections, and geomagnetic activity, which subsequently falls to a solar minimum and a reduction in said properties [5]. It is therefore important to consider the period of time relative to the solar cycle when collecting in-situ data as the solar maximum or minimum can provide misleading inputs.

Accurately predicting solar wind is crucial for many scientific and technological reasons. Space weather events can result in geomagnetic disturbances and an atmospheric drag and a loss of service or satellites in low Earth orbit [13] [8]. Incorrect prediction can jeopardize the stability of observation, communication, navigation and intelligence services on Earth. Correctly forecasting solar wind activity is also paramount to the safety of astronauts, the advancement of space exploration and planetary understanding, and awards the public a greater chance of experiencing the auroras. This project addresses the background solar wind rather than solar events such as CMEs, the reliable understanding of which should not be undermined.

2.2 Solar Wind Propagation Models

The term 'solar wind' was first coined by American physicist Eugene Parker in 1958. Interestingly, in 2017 Parker Solar Probe would be named in his honour. Parker introduced the first supersonic hydrodynamic models, suggesting that the solar corona cannot be static and must expand into space in fast-moving streams of plasma [4] [15]. The first continuous observation of the solar wind was made by the Mariner 2 spacecraft in 1962 which confirmed the existence of a supersonic hydrogen plasma of solar origin [19]. The model was updated to a two-fluid incorporation of both electrons and protons before ultimately including magnetic fields to produce the magnetohydrodynamic (MHD) model [18]. Concurrently, kinetic solar wind models were developed that placed more emphasis on the movement of individual particles rather than the treatment of plasma as a fluid, as more often the case with MHD modelling. However, it was discovered that beyond 0.1 AU the additional accuracies of the kinetic model are not advantageous to the

simulation of quiet solar wind and so MHD modelling is sufficient [3]. There are a number of solar wind propagation models available today, but for this project Tao+ [20] was selected due to the availability of its data at Mercury as well as its use in the Jupiter run of MMESH. The Tao+ solar wind propagation model is a one dimensional MHD model developed by Dr. Chihiro Tao in 2005. It was designed to combat a lack of continuous solar wind monitoring at planets and uses observational data at 1 AU as input. While a 3D solar wind model takes into account helio longitude and latitude as well as distance from the Sun, a 1D model will consider distance from the Sun and assume the wind to be the same over differing longitude and latitudes. It is less computationally expensive and sufficient in the case of studying the solar wind in near-planetary space. To propagate data backwards from Earth to Mercury the model uses the same method as propagating forwards, only the time is reversed incrementally rather than being advanced [20].

The anatomy of MHD modelling consists of the following system of equations; the mass continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{1}$$

where ρ is the mass density and $\mathbf{u}$ is the plasma flow velocity, the equation of state,

$$p = (\gamma - 1)\left(e - \frac{B^2}{2\mu_0}\right) \tag{2}$$

where p is the total plasma pressure, γ is the ratio of specific heats (usually 5/3), e is the internal energy density, μ_0 is the permeability of free space, and B is the magnetic field strength, and several physical laws including Faraday's,

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{3}$$

where $\mathbf{B}$ is the ambient magnetic field and $\mathbf{E}$ is the electric field, Ohm's,

$$\mathbf{j} = \sigma\left(\mathbf{E} + \mathbf{u} \times \mathbf{B}\right) \tag{4}$$

where $\mathbf{j}$ is the plasma current density and σ is the electrical conductivity, and Ampere's laws,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} \tag{5}$$

Finally it includes the momentum equation, the interpretation of which primary differences in the models stem from. Namely for a single-species plasma composed of protons, this is;

$$\frac{\partial\left(m_p n \mathbf{u}\right)}{\partial t}+\rho\left(\mathbf{u}\cdot\nabla\right)\mathbf{u}=-\underbrace{\nabla p}_{\text{pressure}}+\underbrace{\mathbf{j}\times\mathbf{B}}_{\text{Lorentz}}-\underbrace{\frac{GM_{\odot}\rho}{r^2}\hat{\mathbf{r}}}_{\text{gravity}}+\underbrace{\nu\nabla^2\mathbf{u}}_{\text{collision}} \tag{6}$$

where m_p is proton mass, n is plasma number density, G is the gravitational constant, $M_{\odot}$ is the solar mass, r is the radial distance in a heliocentric spherical frame with the $\hat{\mathbf{r}}$ direction pointing radially outward, and ν is collisional frequency. As can be seen on the right hand side of Equation 1, the components represent the (gradient) pressure, Lorentz, gravity and collisional forces, respectively. The Tao+ model does not consider collisional forces, however it does include the remaining three forces on the right hand side of equation 1.

For the Multi-Model Ensemble used at Jupiter, three propagation models were chosen; the ENLIL [14], HUXt [11] [10] and Tao+ [20]. They were selected due to their varying input parameters, dimensionality, and approaches to propagating the solar wind beyond Earth. The MME was beyond the scope of this projected. It is instead a pilot study of how well the process used to temporally shift a solar wind propagation model when being compared to in-situ data works in near-Mercury space.

2.3 Input Data

It is essential that the input data used is as accurate a representation as possible of the solar wind that surrounds the planet in question. Achieving such can be done by restricting the helio-latitudinal, helio-longitudinal and radial ranges when collecting data. This project limits heliolatitude to $-3.4^{\circ} < \theta < 3.4^{\circ}$ and spherical distance to 0.2 AU $< r <$ 0.6 AU, with Mercury's perihelion and aphelion sitting at 0.31 and 0.47 AU, respectively, from the Sun.

There are a number of spacecraft that have transited near-Mercury space including the Mariner 10, MESSENGER and Parker Solar Probe missions. For this exploratory project, only the PSP data are used for simplicity. While incorporating more spacecraft data would yield better predictions, understanding the nuances of data from each spacecraft is time consuming.

The fastest object ever built, Parker Solar Probe was launched on the 12th of August

2018 and first reached Mercury's orbit in around October of 2018. Using Venus gravity assists, it has since been orbiting the Sun and on the 24th of December 2024 successfully made its closest approach of 3.8 million miles [7]. The data for this project reflects PSP at Mercury's distance and heliolatitude from the Sun for the years 2019 and 2020. This is at solar minimum and allows for testing on more ambient than usual conditions.

Parker solar probe comprises of four instrument suites, one of which, the Solar Wind Electrons Alphas and Protons investigation, is built to measure various properties of the solar wind. Its key components include the Solar Probe Cup (SPC) designed to collect protons and alpha particles and the Solar Probe Analysers (SPAN-A and SPAN-B) aiming to measure electrons and ions [9]. Since protons are the dominant ion component in the solar wind, the total solar wind density is approximated to the proton density. Other parameters measured and incorporated into this project include the magnitude of the solar wind flow speed u_{mag} and the interplanetary magnetic field B_{IMF}. The solar wind dynamic pressure is independently calculated from $p_{dyn} = m_p n u_{mag}^2$, where m_p is the proton mass and n the proton density.

In addition to in-situ solar wind data at Mercury, this study uses context data in order to characterise model errors and assist future prediction. Context data at Earth includes solar radio flux, smooth solar radio flux and recurrence index. Recurrence index in this case refers to the correlation coefficient of the solar wind speed with that of 28 days prior. This data is taken from the Pendicton solar radio telescope in Canada. The context data for Mercury includes target-Sun-Earth radius, target-Sun-Earth longitude and target-Sun-Earth latitude and is acquired from SPICE.

2.4 MMESH

MMESH is a framework designed to minimise the disagreements in solar wind timing and magnitude between the in-situ data and the propagation models. The system can take as input any number of solar wind propagation models along with simultaneous in-situ data. It then characterises timing uncertainties in two ways; as a value that remains constant over a whole epoch or as a dynamic value that continually changes. The goal of this framework is to identify and remove timing biases from each model.

The optimisation of model and data agreement is characterised by the maximisation of the correlation coefficient r. To improve accuracy, MMESH also identifies uncertainty

by the optimisation metric $r + (1 - \sigma_T/\Delta T)$. r being the correlation coefficient, σ_T represents the timing uncertainty being 34% or one standard deviation of the distribution of timing offsets. This is then divided by ΔT defined as half the largest allowed magnitude of timing uncertainty, a number between 0 and 1 with the latter being most favourable.

MMESH independently quantifies timing uncertainties, as although most models provide some uncertainty in modelled arrival time, the methods used are not equivalent, making cross-model comparisons difficult. For every combination of model and spacecraft the source of model error is investigated in hopes of discerning which components of the solar wind that model is not accurately predicting. This allows for a quantitative comparison of individual models. The model's performance in magnitude is measured by taking the residuals between the aligned model and data time series, however, before any considerations are given to magnitude, the temporal corrections are made.

As previously mentioned, constant time offsetting is the first, most simple, method of correcting for timing uncertainties. It involves shifting the model forwards or backwards by a constant ΔT in order to maximise the correlation coefficient between the propagated time series and the in-situ measurements of a particular component of the solar wind, for example u_{mag}. The correlation coefficient is calculated as a function of ΔT, $2n+1$ times, beginning at $-n$ and continuing $-n + \Delta T$ to $n - \Delta T$ and n. With a maximum offset of about $n = 3$ days and a realistic ΔT of 4 hours, a positive temporal offset illustrates that the model leads the data, a negative indicating it lags [16].

Dynamic time warping is a significantly more complicated process than constant time offsetting. In order to identify shock or event like features in the data and model, they are both binarised, so a steep gradient in the chosen solar wind parameter would be represented by a 1, everything else by a 0. These 'jumps' in the model are subsequently matched to a suitably corresponding jump in the data using the dynamic time warping (DTW) class of algorithms.

DTW is a pattern matching algorithm that succeeds in non-linearly shifting, compressing and stretching one time series to minimise the distance between itself and another time series. It is used extensively in speech recognition and data-mining but has only recently been applied to space weather, with promising results [2]. Although the functions of MMESH can be applied to any property of the solar wind, the binarisation technique should be most effective with solar wind flow speed as this parameter distinctly

displays shocks and shock-like structures after it has been binarised [16]. It is assumed that the input model parameters are similarly aligned with one another and misaligned with the data so that the warping implemented after running u_{mag} can also be applied to these parameters.

DTW works by constructing a two-dimensional cost matrix, where each cell represents the Euclidean distance between points in the series, or the alignment cost of matching a point in the first series with another point in the second. After a number of permutations, a non-linear path is constructed through this matrix that will result in minimising the overall distance between the two time series. Figure 1 displays the effects of DTW. As made clear by the dotted grey lines, matching extrema in the model and data are located after computing the path of best fit. The points between these matched extrema are then linearly interpolated. The result is a warped model time series that aligns more closely with the data.

To enable the analysis of model timing uncertainties in the absence of in-situ data, MMESH supports simple- and multiple-linear regression models to describe these uncertainties. Simple linear regression involves a single predictor variable explaining the target variable while multiple linear regression uses a number of predictor variables to describe the target variable as a linear combination of such predictors.

Correlation can be found between model timing uncertainties and particular physical parameters in the solar wind, a fact that MMESH leverages to apply uncertainty to models where no simultaneous in-situ data is available. Those physical parameters are specifically TSO angle in heliolatitude and heliolongitude, solar cycle phase and modelled u_{mag}. This technique has proven to be successful in the previous application of MMESH to the solar wind near Jupiter; in this case, the multiple linear regression explains 12% of ENLIL's variation in measured timings, 37% in the HUXt model, and 20% in the Tao+ model.

After independently analysing and adapting each model to the data, a multi-model ensemble is developed which combines all propagation models. MMESH equally weights the contribution of each model, proven to be more robust than a more complicated dynamic weighting scheme [16]. MMESH for Jupiter includes the ENLIL, HUXt, and Tao+ solar wind propagation models in its final MME, where multiple linear regression is used to estimate the timing uncertainties to be propagated throughout. Since this

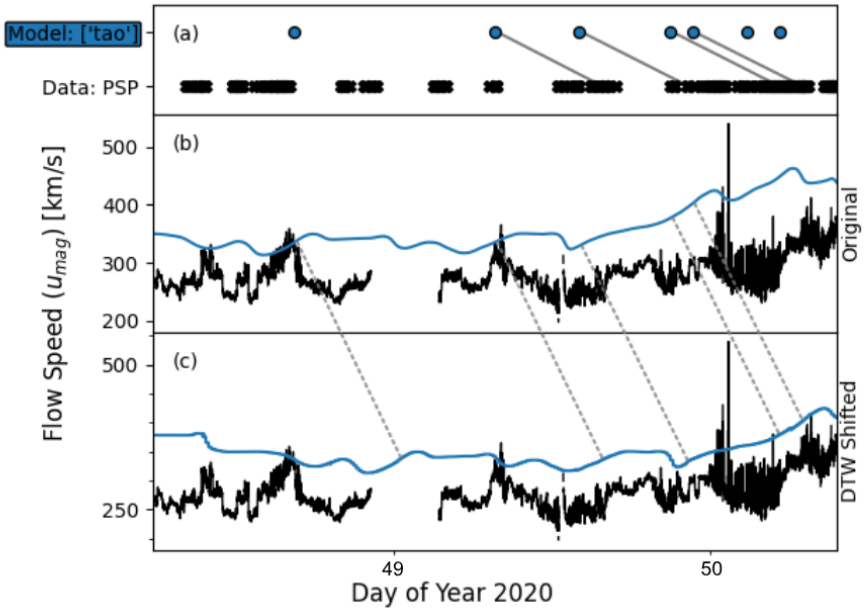

Figure 1: The effects of dynamic time warping for flow velocity and temporal span 17/02/2020 - 20/02/2020. The points in (a) mark the selected 'jumps' after binarisation in both the model and data, the lines mark where an apex in the model has been matched with that in the data. (b) displays the data (in black) and initial Tao model (in blue). (c) shows how the model has been shifted after the DTW process, the data unchanged. The dotted lines link corresponding points in the original and warped models.

study aims to demonstrate MMESH's compatibility with Mercury, a MME was neither within the scope nor required.

The MME at Jupiter shows promising results. The predicted flow speed u_{mag} of the MME (r=0.49) outperforms ENLIL by 110% (r=0.23), HUXt by 7% (r=0.46) and Tao+ by 51% (r=0.32) in correlation coefficient and achieves a centered root-mean-square difference (RMSD=32.8) 28% lower than ENLIL (RMSD=45.9), 14% lower than HUXt (RMSD=38.1) and 9.1% lower than Tao+ (RMSD=36.1) [16].

2.5 Solar Wind Modelling in Near-Mercury Space

This project centered on Mercury for a number of reasons. Due to its proximity to the Sun, the solar wind conditions in near-Mercury space are very dynamic, the combination of which with its relatively strong global magnetic field is unmatched in our solar system. Figures 2 and 3 compare data for four parameters of the solar wind at both Jupiter and Mercury. Although the data sets here originate from different epochs, their stages in the solar cycle were taken into consideration. Since 2016 is just after a solar maximum and

2024 is just before the next, the years exhibit very similar magnitudes of solar radio flux. In addition to the vastly different scales of magnitude in the solar wind parameters at Jupiter and Mercury, the variability of the solar wind at Mercury is clearly much more severe. Testing a solar wind model in a wide variety of space weather environments is very beneficial in highlighting the aspects of a model that work well and those that need improvement. Since at Mercury the space weather conditions are extreme, a relatively successful prediction of solar wind in this region is assuring for a solar wind model.

Unlike Jupiter, Mercury is inside the Earth's orbital zone which means the context data collected at Earth must be hindcasted to Mercury rather than forecasted in the case of Jupiter. Although the same processed is used, only reversed, it is worthwhile exploring where problems may arise.

Additionally, in November 2026, BepiColombo, a joint European Space Agency (ESA) and Japan Aerospace Exploration Agency (JAXA) mission, is due to arrive at the little-studied planet. This mission's primary goals include investigating Mercury's interior, surface, exosphere, and magnetosphere to understand the planet's origin as well as testing Einstein's theory of general relativity [1] [17]. Better understanding the environment a spacecraft is in is paramount to the success of such missions, not only in knowing how to build the instruments but also in directing the spacecraft on the most advantageous trajectory. Should a spacecraft like BepiColombo find itself in hazardous space weather environments, damage or destruction is likely as well as biased measurements.

3 Computational Method

3.1 Reading in the Data

The data for this project was found on Amda, a website which makes available a wind range of space physics data for download. It provided both the Parker Solar Probe in-situ data and the Tao+ solar wind propagation model for Mercury. Separate functions for both data and model were then created to read in the downloaded data files and save them as workable data frames, including a start and stop time to be determined at the users discretion. The PSP data frame is resampled and the model interpolated so as to make both data frames a similar size. Since some of the longer date ranges were too computationally expensive for my device, adjusting the resampling and interpolating of

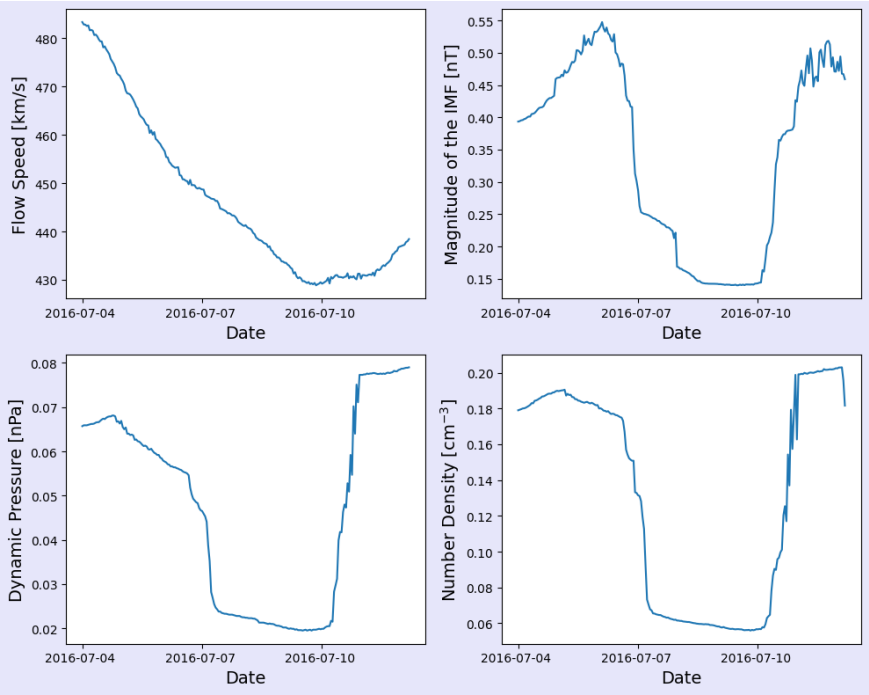

Figure 2: Properties of the solar wind at Jupiter for the temporal span 04/07/2016 - 12/07/2016. Data is from *Juno*.

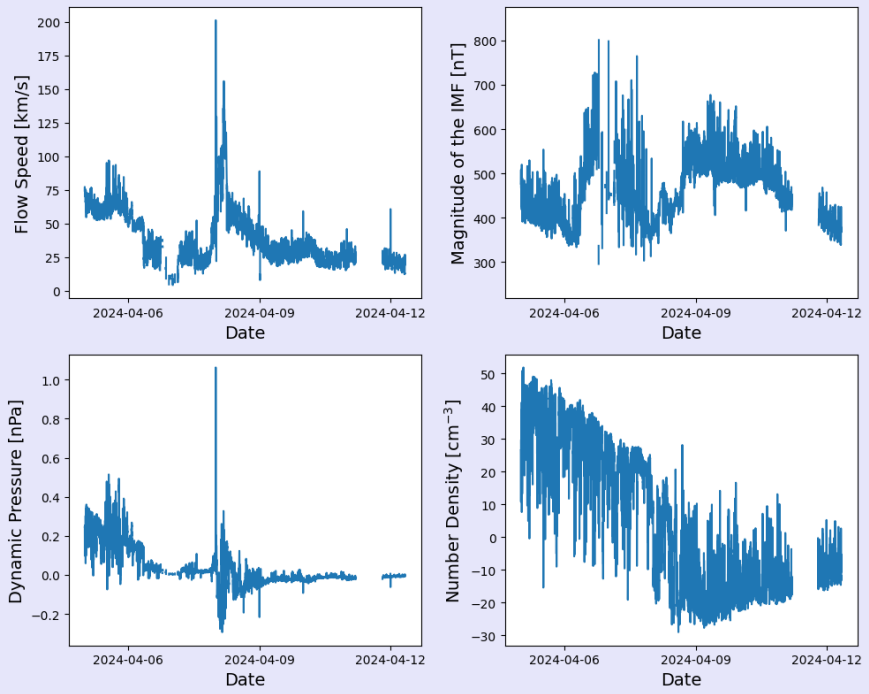

Figure 3: Properties of the solar wind at Mercury for the temporal span 05/04/2024 - 12/04/2024. Data is from *PSP*.

the data sets was a convenient way of reducing the number of data points and allowing longer date ranges to run. In the case of PSP data, the dynamic pressure is calculated, a parameter that is readily available for the model. A padding of one day is added to both

the beginning and end of the model sample as to deduce whether or not the start and end points are apexes. A configuration file defines variables such as distance to the Sun, heliolatitude, start and stop time for each separate trajectory and the final cast and the paths containing the downloaded data.

3.2 Selecting Date Ranges

In order to ensure the solar wind data being used is reflective of that at Mercury, limits are placed on Parker Solar Probe's distance from the Sun and helio latitude, $0.307 \text{ AU} < r < 0.467 \text{ AU}$ and $-3.4° < \theta < 3.4°$ respectively. The corresponding dates for each epoch are found and recorded in the configuration file. Figure 4 demonstrates Mercury and Parker Solar Probe's orbits and figure 5 the relevant epoch's of PSP's orbit.

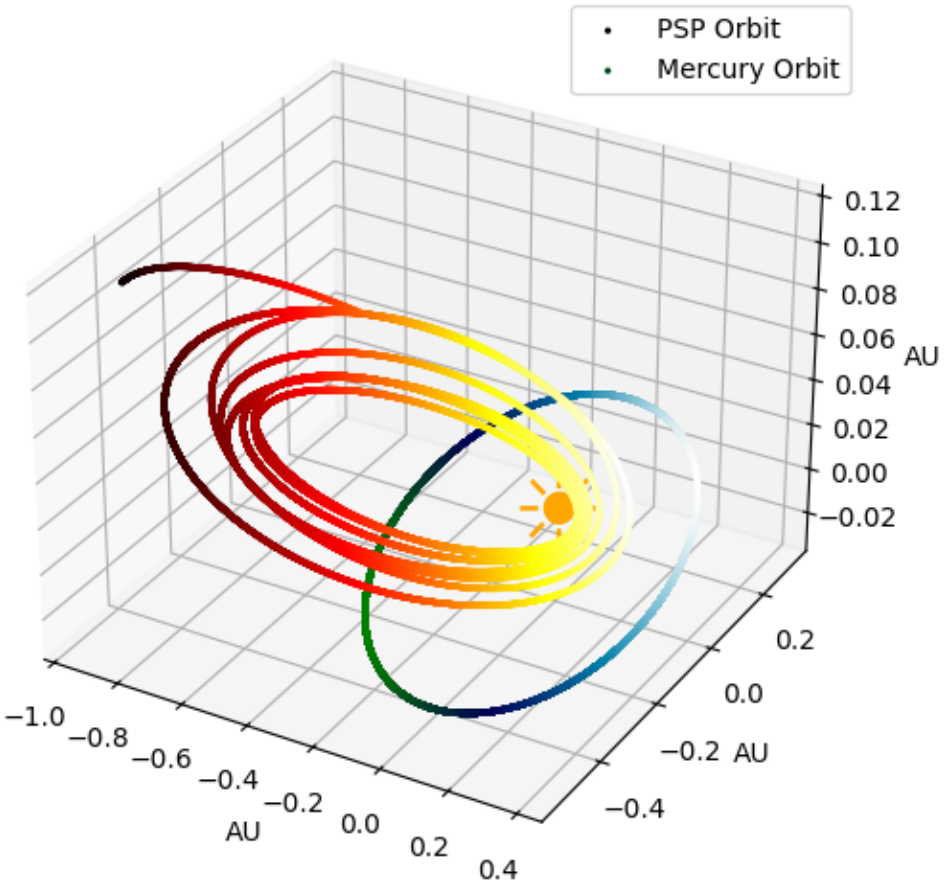

Figure 4: Mercury and PSP orbits in heliocentric inertial coordinates.

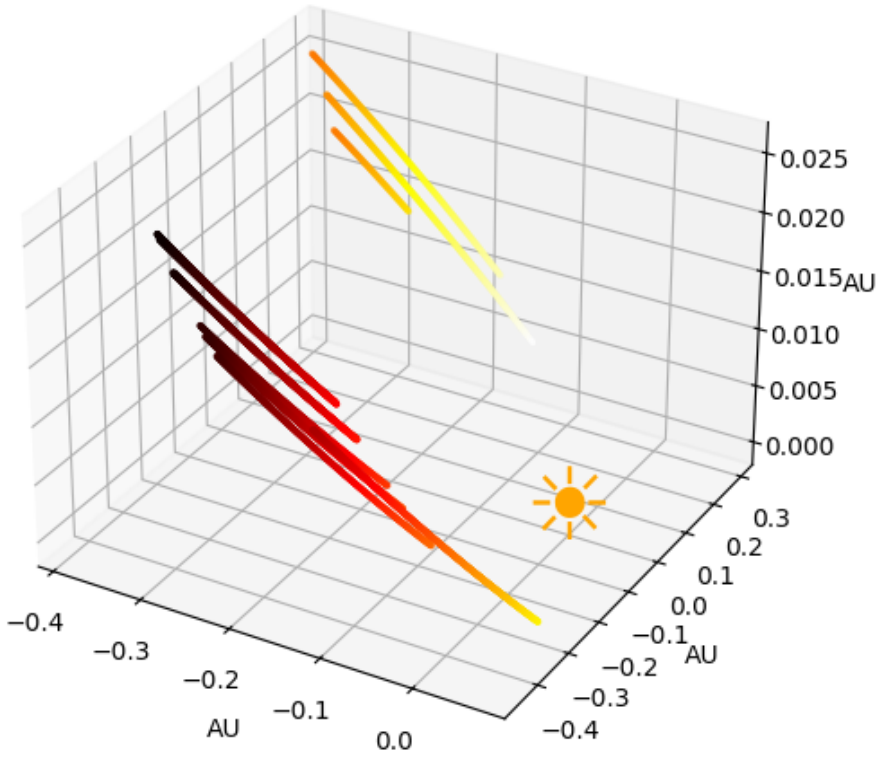

Figure 5: PSP at roughly Mercury's distance and heliolatitude from the Sun in heliocentric inertial coordinates.

3.3 Adding Context

In order to be able to replicate space weather at Mercury, the software needs some context, the information from which it can use to propagate the solar wind backwards. This information includes the parameters target Sun earth distance, longitude and latitude which represent the location of Mercury. The data collected at 1 AU includes solar radio flux, smooth solar radio flux and the recurrence index. A unique function reads this data into a data frame before it is again suitably saved.

3.4 Connecting the Data and Model

In preparation for the dynamic time warping process, the data and model must first be smoothed and subsequently binarised. This is done by computing the rolling mean of a normalised solar wind parameter, for example flow speed, before calculating its derivative. The standard deviation of the derivative is found, and the smoothing window increased until the standard deviation drops below an initially specified value. In the case of Jupiter, as previously mentioned, there is much less variability in the solar

wind parameters so the threshold under which the standard deviation must drop is set to 1. On Mercury however, the threshold was altered until the function ran at a value of 1.5.

The binarisation is done by calculating the z-score of the derivative in order to characterise the variability of the data. When this z-score exceeds an initially specified significance level, it is denoted by a 1. Every other point is valued at 0. Again, due to the extreme variability in the Mercury in-situ data compared to that at Jupiter and the fact that the data at Jupiter much more clearly exhibits extrema, the constant significance level, measured across the full duration of each time series, was changed from 3σ to 1σ.

3.5 Dynamic Time Warping

As previously described in section 2.4, dynamic time warping involves shifting individual data points in a time series, respecting a maximum time deviation, so that it will better align tie points in the model with the data. In this case the maximum time deviation has been set to 3 days. The distances between tie points are calculated and a confusion matrix returns accuracy, precision, recall and F1 score for each shift the model makes. The correlation coefficient, standard deviation and root mean squared deviation are also computed along with time delta widths at 1σ, 2σ and 3σ levels.

As previously mentioned in section 2.4, the optimisation metric used is $r + (1 - \sigma_T/\Delta T)$. By passing the previously calculated shift statistics through the equation, the optimal alignment is determined. A constant time offset is found from the best shift and applied to the whole model while dynamic offsets, specific to the model data, are also extracted. A plot is generated displaying the data together with the original model and separately with the improved model, highlighting the tie points between the two.

3.6 Forecasting

In order to predict the solar wind at Mercury with relevant error during periods without in-situ solar wind data, context data collected at Earth as well as information about Mercury's location are required. Multiple linear regression is used to compare the model time series errors to the context data. The result is a formula containing each element of the context data along with its unique coefficient. This coefficient signifies

how strongly that variable affects the solar wind model error. A negative coefficient results in reducing error while a positive will increase it. Ordinary least squares is then used to predict the errors at another time. The model itself is shifted temporally with the sample that was compared to in-situ data.

4 Results and Discussion

Overall, the results of this project, for a proof of concept study, are optimistic. The program definitely works and, I believe, with a few Mercury-specific tweaks to the software, a very useful tool is within reach.

Prefacing with a brief description of the Taylor diagram; it works by relating the standard deviation of the modelled time series, the correlation coefficient of the modelled time series relative to the measured time series, and the centered root-mean-square difference between the modelled and measured time series to one another by analogy with the law of cosines allowing all three quantities to be displayed as a single point. A model's accuracy is reflected by how closely its graphical representation aligns with the semi-circular pattern of the observed data [16].

4.1 Varying Parameters of the Solar Wind

I will begin by analysing how the model performed under analysis of four solar wind parameters, namely IMF magnitude, number density, dynamic pressure and flow velocity. The temporal span covered ranges from the 20^{th} to the 27^{th} of March 2019, early in Parker Solar Probe's journey. As is evident from figure 6, the models have a general idea of rises and falls in the parameter values. It's worth noting that although the main focus of the MMESH software is flow velocity, the initial Tao+ model performs the most poorly of all four parameters in this case, a trend that is consistent throughout other epochs. Although it does produce the highest correlation coefficient, flow velocity also results in a very high standard deviation and from the plot cannot be seen to predict any major peaks in the data in comparison to the other parameters.

Despite being one of the main forms of statistical measure for this experiment at Jupiter, I believe in Mercury's case, due to the extreme variability in the solar wind data, correlation coefficient does not reflect well the real scenario. The negligible changes in all

cases can be seen in table 1 and in figure 7. This is despite the fact that the plots suggest improvement. If the data is constantly increasing and decreasing, measuring how a continuous model compares is arduous and a more holistic approach must be considered. In the case of Jupiter, simply smoothing the data performed effectively but after an attempt to smooth the data enough for these statistics to work at Mercury, the result is a data set that cannot be binarised. It would perhaps be worth considering replacing the smoothening step with a fitted function or similar. The standard deviation for this epoch however performed very well and is a sign of a strong Tao+ model prediction of the magnitude of solar wind parameter peaks and troughs. The success of the Tao+ model can also be seen from the Taylor diagrams in figure 7 and the proximity of the markers to the black line representing the data.

Another noteworthy result is the inaccuracy of the IMF magnitude at Mercury, it being roughly half the value of the data, however this is not a consistent trend. IMF magnitude is the only parameter whereby a dynamic time warping shift is utilised. I don't believe this is due to the parameter itself but instead that the trend of constant time shifting over DTW in the Mercury run of MMESH may be on account of the restriction to shorter epochs. This restriction is due to the computer's inability to run longer date ranges but the fact is worth noting and might possibly be due to physical reasons. Since the distance between the orbits of Mercury and Earth is roughly 3.5 times shorter than the distance between that of Earth and Jupiter, it is plausible that the solar wind better maintains uniformity as it travels the shorter distance. Should constant time shifting prove to work well in Mercury's case, making temporal improvements to the Tao+ model would be a relatively simple task.

Finally, despite all being of the exact same date range, none of the parameters have undergone the same temporal adjustment. A worthwhile investigation would be to consider a MME-like multi-parameter ensemble which takes the alignment of least errors across a number of solar wind parameters. As will be evident in the following section, this was not an issue at Jupiter as the models showed distinct maxima, however, with more ambiguous solar wind models at Mercury, it would certainly be beneficial.

Overall, the MMESH software is clearly almost equally compatible with all four parameters of the solar wind.

Parameter	r before	r after	σ before	σ after	σ of data
IMF magnitude	-0.07620	-0.07666	5.56097	5.56034	6.39722
Number density	-0.07778	-0.08176	19.53530	19.54994	20.92490
Dynamic pressure	0.09132	0.08823	4.02294	4.02439	3.86351
Flow velocity	0.36068	0.36080	30.09804	30.10544	29.52385

Table 1: Correlation coefficient (r) of the modelled time series relative to the measured time series and the standard deviation (σ) of the modelled time series before and after temporal adjustment as well as of the data for the temporal span 20/03/2019-27/03/2019. A conservative error of 20% is estimated.

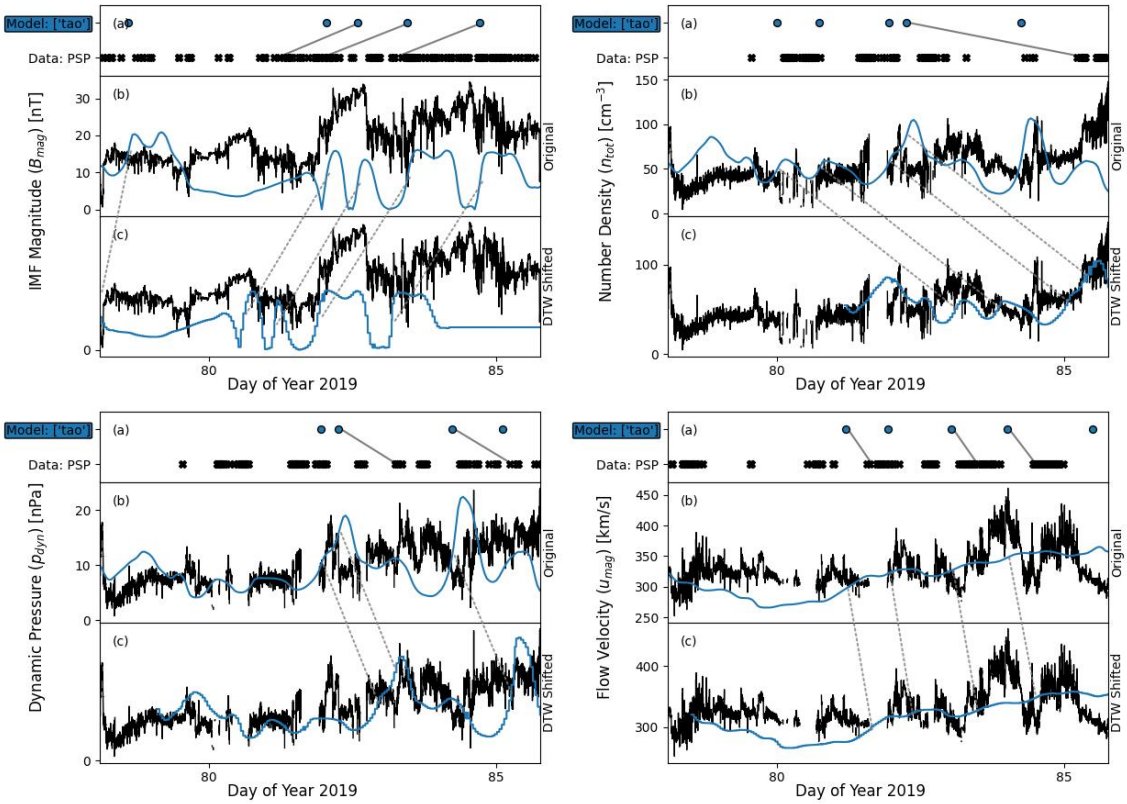

Figure 6: The effects of constant time offsetting and dynamic time warping for the solar wind parameters IMF magnitude, number density, dynamic pressure and flow velocity. The temporal span is 20/03/2019-27/03/2019.

4.2 Flow Velocity

As mentioned previously, the flow velocity solar wind parameter is the primary focus of MMESH at Jupiter. Consequently, flow velocity is the parameter by which comparisons are here made between four epochs across the years 2019 and 2020. Figure 8 portrays how the Tao+ solar wind model was adjusted after constant time offsetting

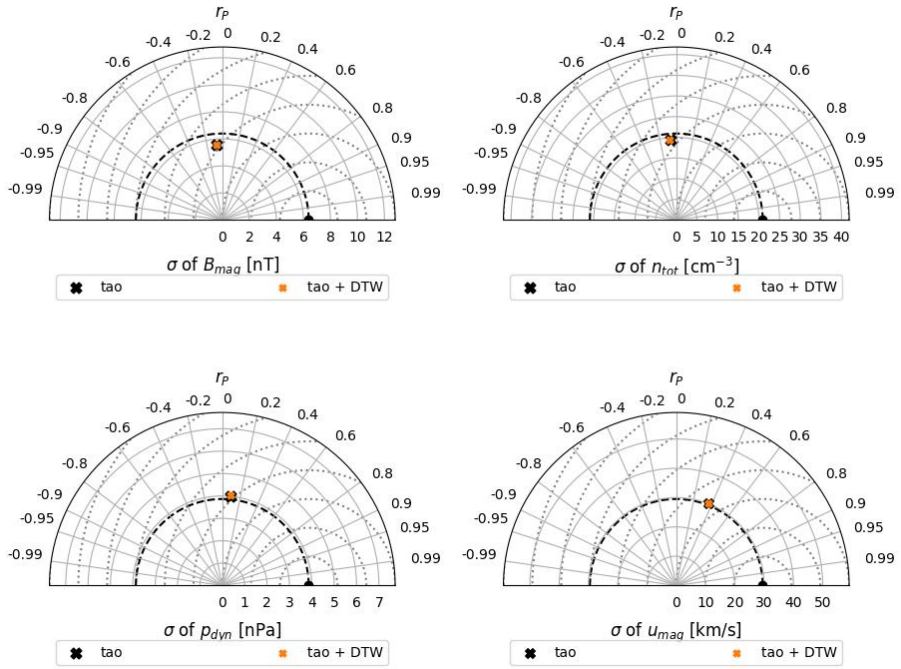

Figure 7: Taylor diagrams for all four parameters of the solar wind for the temporal span 20/03/2019-27/03/2019.

and dynamic time warping. As can be observed, the dotted lines connecting tie points between the original and shifted models are very much parallel, signifying that optimal alignment was achieved with constant time offsetting. It is also clear that some of the tie points are not in fact maxima in the model. This is sometimes a result of the smoothing function which can deviate maxima slightly and is indicative of an incompatibility between the binarisation process and the Tao+ model and data at Mercury.

It can be seen from the first and fourth epoch in figure 8 that the magnitude of the Tao+ solar wind model is wildly inaccurate, exhibiting a root mean square deviation of about 500. This is a trend that follows on from the fourth epoch into 2021 and 2022, restricting the focus of this project to the two years prior. This inaccuracy is also apparent throughout other solar wind parameters and is thought to be due to defects in the Tao+ model.

The statistics of these adjustments can be seen in figure 9 and table 2. For the second two epochs it is clear that little difference was made after DTW. However, there is a major disimprovement in the standard deviation after realignment in the first epoch. A worthwhile modification to the dynamic time warping software would be to refrain from any temporal changes should none be found to be an enhancement of the original

model. This is however not of great concern due to the initial discrepancy between model and data magnitude for this epoch.

The second epoch, despite not having shifted the model by very much, displays the most promising of results with a 19% decrease in standard deviation bringing it closer to that of the data and a shift from a negative correlation coefficient to a positive one.

Figure 10 represents the effects of dynamic time warping on the HUXt solar wind propagation model at Jupiter for a period of time in 2016. The first remarkable distinction between this plot and that at Mercury is the variability of the data, as previously discussed. The second is the form of the HUXt model in comparison to Tao+. HUXt demonstrates very distinct peaks allowing for the 'jumps' to be much more easily identified. As can be seen, every tie point is indeed a maximum, and the program has succeeded in locating each one. Interestingly, the Tao+ model shows a similar pattern for its prediction at Jupiter. Since noticeable peaks are identifiable in the Tao+ model for the remaining three parameters in this study, doubt lies in the competence of the Tao+ model for flow velocity at Mercury. This implies an interest and need within the scientific community for a working version of the multi-model ensemble for flow velocity at Mercury. Finally, it is apparent that dynamic time warping has been implemented to realign the model. This is possibly due to a longer temporal span but may also simply work better with the solar wind at Jupiter, for reasons previously discussed.

Epoch	**r before**	**r after**	**σ before**	**σ after**	**σ of data**
16-24 08 19	-0.38571	-0.32776	19.50252	46.22884	14.23150
14-21 01 20	-0.52742	0.41148	15.19748	12.30698	9.00657
23-30 05 20	0.22860	0.22888	8.37122	8.37068	15.77778
17-19 09 19	0.67428	0.67596	42.19127	42.18734	29.42426

Table 2: Correlation coefficient (r) of the modelled time series relative to the measured time series and the standard deviation (σ) of the modelled time series as well as for the data before and after temporal adjustment for the solar wind parameter flow velocity. A conservative error of 20% is estimated.

4.3 Prediction

After running a date range through MMESH, the model is now shifted temporally by some amount ΔT. Future prediction will be shifted by this same amount. In order to estimate the errors in a future model's performance, multiple linear regression is used on the parameters given in the context data set.

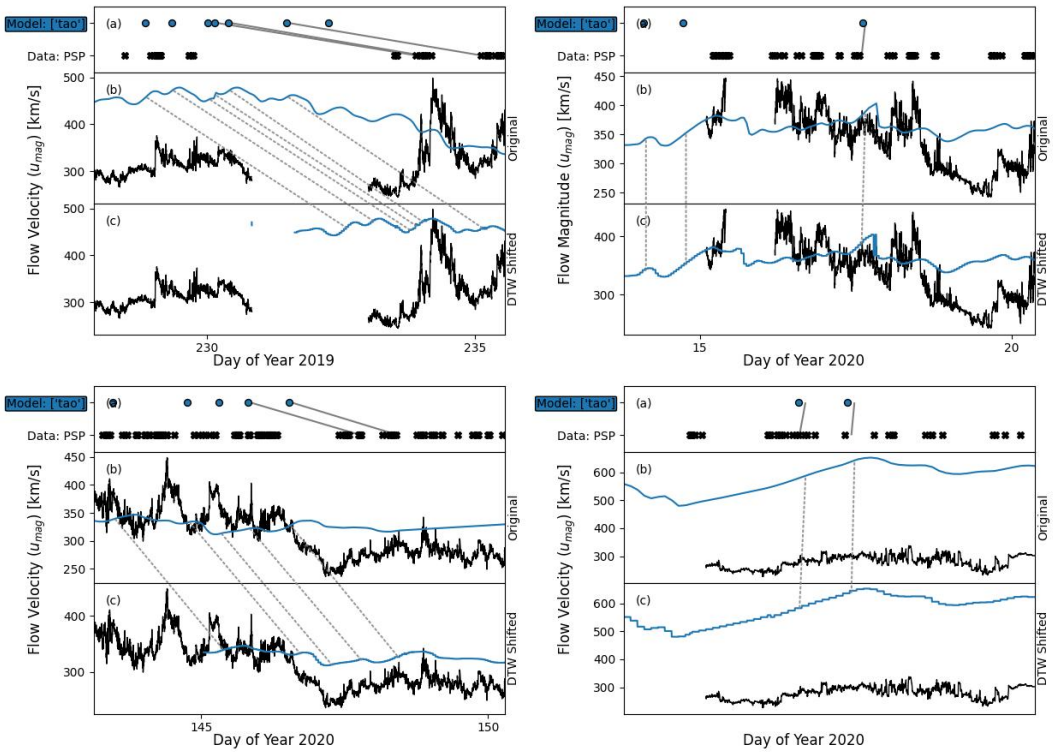

Figure 8: Shifted Tao+ solar wind propagation model for the parameter flow velocity covering the following temporal spans respectively; 16/08/2019 - 24/08/2019, 14/01/2020 - 21/01/2020, 23/05/2020 - 30/05/2020, 17/09/2020 - 19/09/2020.

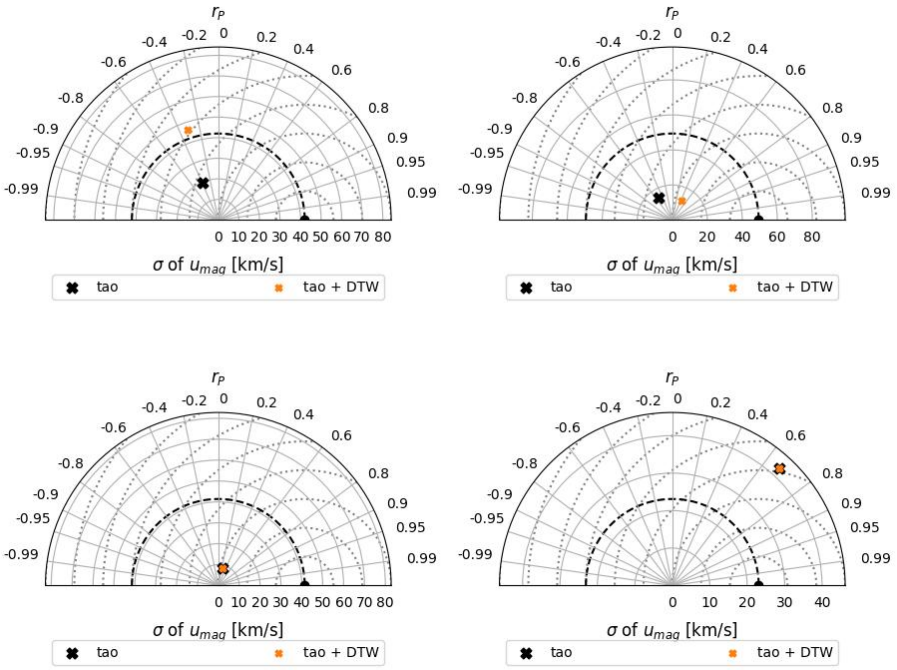

Figure 9: Taylor diagrams for flow velocity covering the following temporal spans respectively; 16/08/2019 - 24/08/2019, 14/01/2020 - 21/01/2020, 23/05/2020 - 30/05/2020, 17/09/2020 - 10/09/2020.

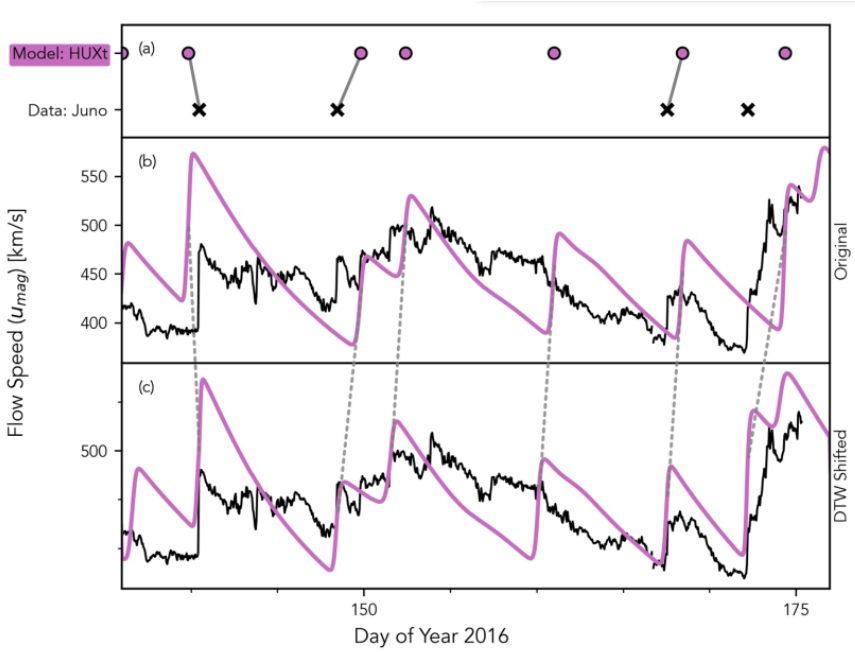

Figure 10: The effects of dynamic time warping on the HUXt solar wind propagation model in near-Jupiter space. The in-situ data is from the *Juno* spacecraft and the temporal span is 05/15/2016 - 29/06/2016.

After obtaining the coefficient values previously described over four epochs, their averages and associated errors, displayed in table 3, were acquired. The larger the coefficient, the more impact this parameter will have on the error. A negative coefficient indicates that a greater value in this parameter reduces errors, a positive meaning a greater value increases errors. So target Sun Earth radius has the greatest role to play and as this quantity increases, so do the errors. Recurrence index also plays a significant role as expected as it conveys how rapidly the solar wind is changing over time, determined by position in the solar cycle. Over the date ranges studied, this combination of parameters accounts for 43.7% of the variation in measured timings over the first time period (i.e. $R^2 = 0.437$) and 63%, 92.3% and 87.3% for subsequent epochs demonstrating significant confidence in the error values attained.

Figure 11 shows the data, the original Tao+ model and how the model has been shifted after MMESH, with error, for the date range 02-04-2019 to 10-04-2019. It is immediately evident that the magnitude of IMF magnitude, dynamic pressure and number density is very inaccurate, a consequence of the fact that PSP was slightly outside of Mercury's orbital zone at this time. This is on account of the project being restricted to very short time periods when PSP was in Mercury's orbital zone, usually between

Parameter	Average Coefficient	Standard Error
target Sun Earth latitude	3252.75	± 679.35
target Sun Earth longitude	264.81	± 55.33
target Sun Earth distance	15094.49	± 3156.75
solar radio flux	-0.015	± 0.026
smooth solar radio flux	-125.36	± 18.95
recurrence index	9181.34	± 1167.61

Table 3: Average coefficients and standard errors of context data parameters after multiple linear regression.

two and ten days. It is also necessary that the tested and predicted temporal span data be in the same file causing a restriction induced by device limitation. This could easily be rectified on a device with a higher capacity, however the present example can still provide insight into possible improvements.

It is noticeable from the plots in figure 11 that some of the peaks and troughs align better with the data. From table 4 it is evident that the standard deviation of the adjusted model more closely aligns with that of the data in only one case. The correlation coefficient of dynamic pressure and number density was modified from a negative value in the original model to a positive one in its adjusted counterpart, suggesting improvement. The errors generated from the MLR are evident in this plot, however they do not cover the range of data due to reasons previously discussed.

Measurement	Data	Original	Adjusted
Flow Speed σ	57.3247900	50.1406930	52.3669769
IMF Magnitude σ	19.5337322	4.6819823	4.1673982
Dynamic Pressure σ	9.3632651	2.1756614	2.0736311
Number Density σ	72.8201241	8.3912381	7.1487983
Correlation Values			
Parameter		**Original**	**Adjusted**
Flow Speed r		-0.0831349	-0.2858669
IMF Magnitude r		0.1123027	0.0016141
Dynamic Pressure r		-0.0789822	0.0626859
Number Density r		-0.1680102	0.0665382

Table 4: Standard deviation of the data, the original Tao+ model and the predicted adjustment to the Tao+ model. Correlation coefficient of the original model with the data and the adjusted model with the data. Temporal span is 02/04/2019 - 10/04/2019. A conservative error of 20 % is estimated.

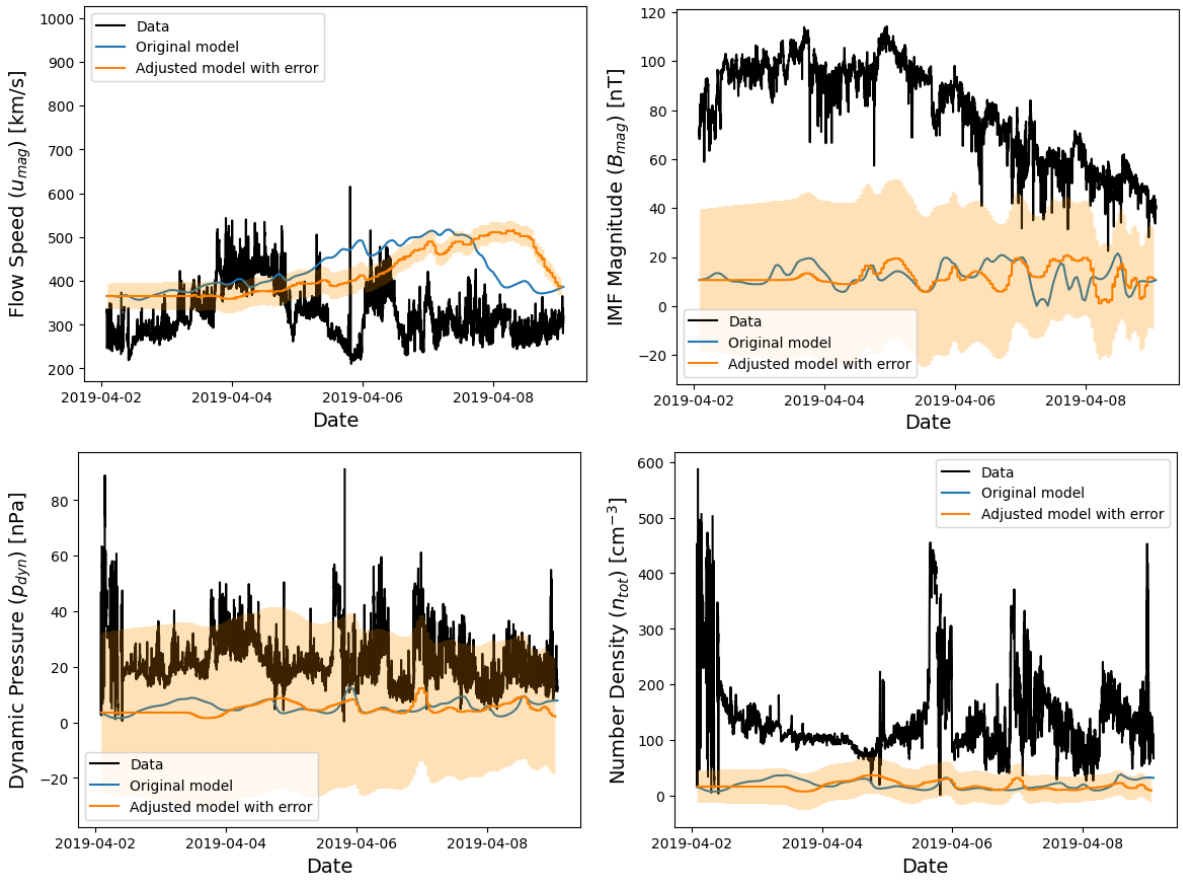

Figure 11: The data, original Tao+ model and predicted Tao+ model with predicted errors for the temporal span 02/04/2019 - 10/04/2019.

5 Conclusion

The results of this exploratory investigation demonstrated significant potential. One of the key obstacles in adapting an outer heliosphere solar wind analysis program to near-Mercury data is the considerable difference in data variability. In the data smoothing process, the derivative of the rolling mean is found, the standard deviation of which must drop below a certain value. At Jupiter this value was set to 1, however, at Mercury the binarisation process failed until this value was increased to 1.5. In the binarisation process, the z-score of the derivative is calculated. Since the extrema in the solar wind model are much more pronounced at Jupiter than at Mercury, the significance level under which this z-score must fall had to be lowered from 3 to 1. The program was subsequently able to identify maxima in the model, however, the adjustments posed problems in later procedures.

It was apparent, once dynamic time warping was introduced, that a number of the selected 'jumps' in both the model and data were not indeed maxima due to this lowering of significance level. This suggests that the introduction of specifying a separate threshold and significance level for model and data would be advantageous. Another adjustment that would likely be beneficial is the fitting of a function to the in-situ data, generating a more holistic impression of this variable.

This minimal version of the MMESH software worked equally well with the four parameters of the solar wind tested; flow velocity, number density, dynamic pressure and IMF magnitude. Notably, despite being of primary focus throughout the MMESH run at Jupiter, the program was least effective with the flow velocity parameter and is thought to be due to inaccuracy in the Tao+ model's approach to predicting this variable. Despite being run for the same temporal span, the four parameters each underwent unique temporal adjustment. This indicates that a multi-model ensemble like multi-parameter ensemble would be a constructive undertaking in modelling the solar wind at Mercury.

Other noteworthy results are the inadequacy of using correlation coefficient and standard deviation in measuring the statistical success of this procedure. These values rarely changed in any significance but is a problem that can be resolved by addressing previously mentioned issues.

The favouring of the use of constant temporal shifting over dynamic time warping

can be seen throughout varying parameters and epochs and is likely due to restrictions to shorter temporal spans, but would be a very noteworthy result should it be due to physical reasons.

Although this exploratory project did not yield major breakthroughs, applying MMESH to near-Mercury space shows promise. With some strategic modifications to the software a powerful tool is within reach, one that would address critical knowledge gaps and equip the scientific community with a valuable resource they currently lack.

References

[1] J. Benkhoff. Bepicolombo - mission overview and science goals. *Space Science Reviews*, 2021.

[2] Eamonn J. Keogh C. Ratanamahatana. Three myths about dynamic time warping data mining. *SDM*, 2005.

[3] T. Holzer E. Leer. Collisionless solar wind protons: A comparison of kinetic and hydrodynamic descriptions. *Journal of Geophysical Research*, 1972.

[4] M. Echim. A review on solar wind modeling: Kinetic and fluid aspects. *Physics, Environmental Science*, 2013.

[5] D. Hathaway. The solar cycle. *Living Reviews in Solar Physics*, 2010.

[6] A. Hundhausen. Coronal expansion and solar wind. *Springer-Verlag*, 1972.

[7] Mara Johnson-Groh. Nasa's parker solar probe makes history with closest pass to sun. *science.nasa.gov*, 2024.

[8] Bodemann & Kalden. Threats awaiting earth observation satellites. *Institute of Electrical and Electronics Engineers*, 2013.

[9] J. Kinnison. Parker solar probe: A mission to touch the sun. *Institute of Electrical and Electronics Engineers*, 2020.

[10] M. Lang L. Barnard, M. Owens. Sir-huxt—a particle filter data assimilation scheme for cme time-elongation profiles. *Space Weather; The International Journal of Research and Applications*, 2023.

[11] M. Owens L. Barnard. Huxt—an open source, computationally efficient reduced-physics solar wind model, written in python. *Frontiers in Physics*, 2022.

[12] E. Marsch. Origin and evolution of the solar wind. *Proceedings of the International Astronomical Union*, 2006.

[13] R. Miteva. Space weather effects on satellites. *Astronomy*, 2023.

[14] D. Odstrcil. Modeling 3-d solar wind structure. *Advances in Space Research*, 2003.

[15] E. Parker. Dynamics of the interplanetary gas and magnetic fields. *The Astrophysical Journal*, 1958.

[16] M. Rutala. A multi-model ensemble system for the outer heliosphere (mmesh): Solar wind conditions near jupiter. *Journal of Geophysical Research: Space Physics*, 2024.

[17] R. Schulz. The bepicolombo mission. *Proceedings of the International Astronomical Union*, 2005.

[18] S. Seuss. A two-fluid, mhd coronal model. *Journal of Geophysical Research: Space Physics*, 1999.

[19] C. Snyder. Particles and fields in interplanetary space. *Eos, Transactions American Geophysical Union*, 1963.

[20] C. Tao. Magnetic field variations in the jovian magnetotail induced by solar wind dynamic pressure enhancements. *Journal of Geophysical Research: Space Physics*, 2005.