



# Matching Temporal Signatures of Solar Features to Their Corresponding Solar-Wind Outflows

D. de Pablos<sup>1</sup> · D.M. Long<sup>1</sup> · C.J. Owen<sup>1</sup> · G. Valori<sup>1</sup> · G. Nicolaou<sup>2</sup> · L.K. Harra<sup>3,4</sup>

Received: 3 November 2020 / Accepted: 31 March 2021  
© The Author(s) 2021

## Abstract

The role of small-scale coronal eruptive phenomena in the generation and heating of the solar wind remains an open question. Here, we investigate the role played by coronal jets in forming the solar wind by testing whether temporal variations associated with jetting in EUV intensity can be identified in the outflowing solar-wind plasma. This type of comparison is challenging due to inherent differences between remote-sensing observations of the source and in-situ observations of the outflowing plasma, as well as travel time and evolution of the solar wind throughout the heliosphere. To overcome these, we propose a novel algorithm combining signal filtering, two-step solar-wind ballistic back-mapping, window shifting, and Empirical Mode Decomposition. We first validate the method using synthetic data, before applying it to measurements from the *Solar Dynamics Observatory* and *Wind* spacecraft. The algorithm enables the direct comparison of remote-sensing observations of eruptive phenomena in the corona to in-situ measurements of solar-wind parameters, among other potential uses. After application to these datasets, we find several time windows where signatures of dynamics found in the corona are embedded in the solar-wind stream, at a time significantly earlier than expected from simple ballistic back-mapping, with the best-performing in-situ parameter being the solar-wind mass flux.

**Keywords** Solar-wind, origin · Heliosphere, connection · Coronal hole, observations · Coronal bright point, dynamics

## 1. Introduction

Traditionally, the solar wind is categorized into different types depending on its plasma parameters as measured in situ. As the average solar-wind speed is  $500 \text{ km s}^{-1}$ , fast streams

---

✉ D. de Pablos  
[diego.pablos.18@ucl.ac.uk](mailto:diego.pablos.18@ucl.ac.uk)

<sup>1</sup> Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Surrey, RH5 6NT, UK

<sup>2</sup> Southwest Research Institute, San Antonio, TX 78238, USA

<sup>3</sup> PMOD/WRC, Davos-Dorf, Davos, 7260, Switzerland

<sup>4</sup> ETH-Zürich, Höggerberg campus, HIT building, Zürich, Switzerland

are classified as those moving at  $\geq 600 \text{ km s}^{-1}$ , and slow streams are said to be those moving at  $450 \text{ km s}^{-1}$  or less (Marion, 1980). Fast, low-density solar wind has been shown to be expelled from coronal holes (McComas et al., 2008), which are long-lived, predominantly monopolar regions of the solar corona that are expected to constantly release material into the heliosphere (Krieger, Timothy, and Roelof, 1973). The determination of the specific coronal origin of a given stream, as well as of the particular coronal regions that have important contributions is, however, more complicated. Several authors (see, e.g., Ohmi et al., 2004; Sakao et al., 2007; Stansby et al., 2020b; Rouillard et al., 2020, and the references therein) have suggested a number of methods that intend to aid in the determination of the specific origin of a solar-wind stream, using the range of observations that are currently available. One of the fundamental reasons why the origin of the solar wind is hard to constrain is that a large amount of dynamic processing takes place during propagation, modifying parameters that otherwise might be linked to the specific near-Sun plasma properties of the stream (Abbo et al., 2016). Fortunately, some solar-wind plasma parameters are relatively unaffected by solar-wind dynamical processes, such as the ionic composition of streams (Bochsler, 2007). Ion charge states become fixed in the solar corona, depending on the local electron density and temperature (Wang, 2016), and they can therefore be used as a qualitative tracer for the likely origin of a given stream. In-situ measurements of ion charge-state ratios, along with calculations of the first ionisation potential (FIP) bias of the solar atmosphere through remote measurements, enable a direct comparison of ion populations of different elements that comprise the coronal plasma (Pottasch, 1964; Stansby et al., 2020a). While this can qualitatively help in determining the type of origin of a solar-wind stream, it does not resolve the specific time and location of individual solar-wind parcels. This is problematic for science that requires a verifiable connection between a solar-wind source and its outgoing plasma.

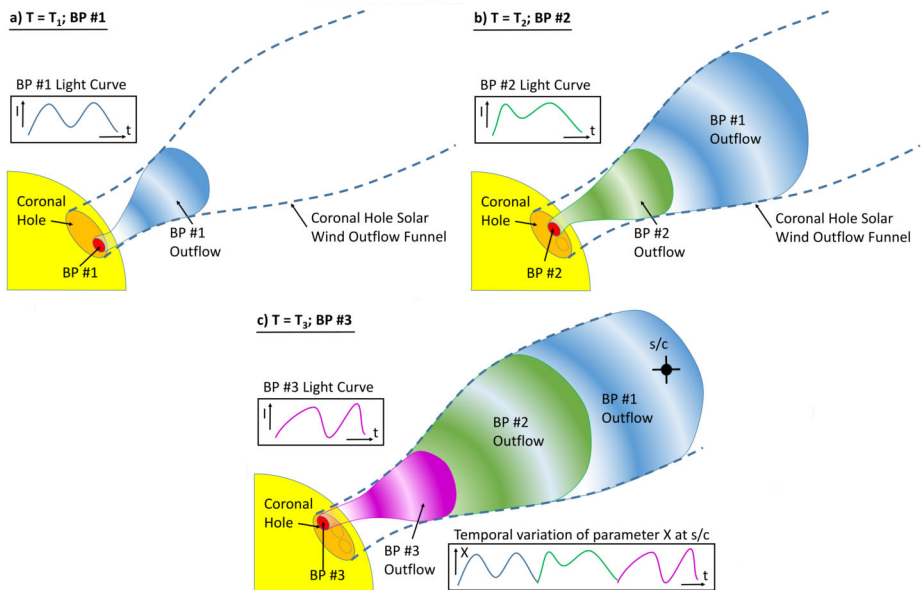
One of the main science objectives of the recently launched *Solar Orbiter* mission (Müller et al., 2020) is to determine the origins of solar-wind streams, shedding light on the origin and modulation of the heliospheric magnetic field and its charged-particle population. For this reason, the spacecraft is equipped with a number of in-situ detectors, as well as a range of imaging instruments. During close approaches at  $\approx 0.3 \text{ AU}$ , the high-resolution imagers will cover up to  $\approx 6\%$  of the visible solar surface, and require a degree of targeting. The ideal pointing location for this science question is the source of the solar wind that is later measured in situ. Due to the delay between plasma ejection from the Sun and detection at the spacecraft location, modelling must be used to forecast likely coronal origins of a solar-wind stream. The current methods estimating likely solar-wind stream footpoints require knowledge of the solar photospheric magnetic field, modelling of solar-wind propagation, and a range of observations of the stream properties. With regards to solar-wind propagation models, some employ magneto-hydrodynamics (MHD) (e.g. Mikić et al., 1999) to resolve solar-wind creation and evolution, and others assume simple ballistic propagation of the solar wind away from a rotating Sun (see, e.g., the solar-wind model by Parker, 1958). While the more rigorous MHD models are useful for generating a time-dependent, 3D representation of solar-wind flows, they take a large amount of computational resources and time, and are unsuitable for quick predictions of stream footpoints for *Solar Orbiter* remote-observing operations. Simpler models that make use of the Parker-spiral-based ballistic back-mapping (e.g. Nolte and Roelof, 1973), paired with a potential coronal-field model (PFSS: Schatten, Wilcox, and Ness, 1969), have been shown to perform almost as well as MHD simulations under simple coronal magnetic-field configurations, such as those corresponding to simpler magnetic-field configurations near the beginning of the solar cycle (see, e.g., Riley et al., 2006; Macneil, 2019).

These simple empirical models are, however, more reliant on output verification to provide realistic information, as they are governed by over-simplified kinematics. To verify a given solution, different techniques have been suggested based on different solar-wind coronal-origin tracers (Cranmer, 2019). These tracers have been taken to be compositional data (ionisation states and ratios) and structural coronal boundaries along with their effect on the outflowing solar wind. A potential solar-wind origin tracer, which has not yet been explored with the aid of the relevant observations, is the potential inclusion of characteristic temporal signatures in the corona, and their propagation as part of the solar wind. As these are numerous and highly variable, they could inject persistent temporal signatures in the outflowing solar wind. When these signatures are sufficiently different from other timescales observed in the corona, they could be used to aid in the determination of a solar-wind source region for a specific stream.

Coronal jets are one example of these ubiquitous coronal dynamics. Solar coronal jets are highly collimated, transient structures observed in coronal plasma that are speculated to be produced due to magnetic reconnection between an underlying closed loop and an overlying open magnetic field (see, e.g., Edmondson et al., 2010; Crooker and Owens, 2012). These jets last between 5 and 20 minutes, and due to their high occurrence rate across the solar atmosphere at any given time (Raouafi et al., 2016), they are thought to contribute a significant amount of mass and energy into the solar wind (Kumar et al., 2019). In remote-sensing observations, coronal jets are linked to sudden localised enhancements of emission intensity in both the X-ray and EUV wavelengths that are coupled with collimated structures. Within in-situ observations, rapid changes in radial magnetic-field direction or *switchbacks* have been observed in the solar wind at distances of up to 1 AU from the Sun (Marsch et al., 1982; Tenerani et al., 2020), with their occurrence rate being larger closer in to the Sun (Kasper et al., 2019). Several authors (see, e.g., Owens et al., 2020; Sterling and Moore, 2020) conclude that a possible driver for these is plasma release from low-lying coronal loops through magnetic reconnection. This would make coronal jets that take place over open field (e.g. a coronal bright-point within a coronal hole) a potential source for these structures.

The *Atmospheric Imaging Assembly* (AIA; Lemen et al., 2012) onboard the *Solar Dynamics Observatory* (SDO; Pesnell, Thompson, and Chamberlin, 2012) provides 12-second cadence, full-disk UV and EUV observations of the solar atmosphere at different temperatures. These observations have been successfully exploited to describe small-scale dynamics occurring in the solar atmosphere (Madjarska, 2019, and the references therein). By making use of the enhanced EUV emission in the SDO/AIA passbands, it is possible to characterise timescales and temporal behaviour of many of these events. We hypothesize that signatures of these dynamics may be captured in outflowing solar-wind plasma parameters, which in turn present unique timescales often linked to solar coronal jets (5- to 20-minute sudden changes, with different timescales of variability found).

Figure 1 shows three panels representing the evolution of a solar-wind stream that is ejected from an equatorial coronal hole and reaches a spacecraft (marked “s/c” on the third panel). Within the coronal hole, a series of bright-point eruptions take place, embedding their characteristic timescales, coloured in the same manner, into the solar wind. Each of the three panels is labelled by a distinct timestamp ( $T_1$ ,  $T_2$ ,  $T_3$ ), for each bright-point eruption and each resulting light curve (BP 1, BP 2, BP 3), such that each of the bright-point eruptions embeds a specific signature on a solar-wind parameter “X”. In this work we argue that because of radial propagation of the solar wind, for a given solar-wind stream with an estimated source point, it is possible to compare dynamics observed at the approximate magnetic foot-points to temporal variations found within solar-wind stream parameters



**Figure 1** Schematic illustrating the hypothesis that the work in this article ultimately seeks to provide a means to address. We argue that physical activity on the Sun, captured as signatures of temporal variations in remotely sensed measurements, may also embed similar temporal variations in some parameters associated with the outflow solar wind, which is then sampled by a spacecraft. Here we illustrate the specific case of an equatorial coronal hole exhibiting three bright points, identified by their light curves (illustrated on the upper left of each panel). Temporal variations either within the light curve of each bright-point, or indeed across the events, may relate to temporal variations in parameters describing the solar wind leaving the Sun. If the coronal hole is ballistically connected to a spacecraft via the associated outflow funnel, and even if the solar wind expands from the localised source into the outflow funnel, as shown in each panel, it is possible that the temporal variations from the source may persist and be identifiable at the spacecraft. There may thus be parameters within the solar-wind flow that have a particular temporal fingerprint leaving the low solar corona and reaching a spacecraft in the form illustrated in the lower right of the third panel. If a match between the temporal variations of the light curves for a given bright-point, or for the coronal hole as a whole, and the observations at a spacecraft can be identified, this would provide evidence supporting the identification of the source region. In this article we address a means to make this identification.

measured at the spacecraft location. Note that Figure 1 depicts magnetic-field-line expansion during solar-wind propagation, which may have a significant effect on the signal morphology by implicitly stretching timescales or lowering amplitudes of variations and thus changing the signal over time, but these effects are not considered in this work. We instead concentrate on extracting a valid light curve containing signatures of coronal dynamics, and compare it against the different solar-wind parameters.

Observations extracted from coronal dynamics can, however, be expected to be highly nonlinear, not being fully described through a combination of strictly linear processes, as well as non-stationary, as their statistical properties can be expected to change over time. More specifically, in the remote-sensing observations, the occurrence of particular jetting events or brightenings cannot be estimated through purely linear means, and their statistical properties are expected to broadly differ depending on local activity. Equivalently, many different processes may take place in the solar wind, increasing the complexity of the time-series that is later observed in situ. Examples of such processes are plasma oscillations, turbulence, and solar-wind inter-stream dynamics (Velli and Grappin, 1993). Due to this

complexity, it is necessary to make use of a signal decomposition technique that allows for both nonlinearity and non-stationarity. Here we choose to employ Empirical Mode Decomposition (EMD: Huang et al., 1998; Gao, Hongfei, and Zhu, 2011), which decomposes an original signal into a series of Intrinsic Mode Functions (IMFs), depending on timescales that are captured within them. Applications of EMD in solar and plasma physics are varied; Long et al. (2017) utilise EMD in the characterisation of a trans-equatorial loop oscillation from remote-sensing images. Deng and Feng (2013) and Kolotkov et al. (2017) employ the method to characterise long-term flare characteristics, and Stangalini et al. (2017) use it for the study of MHD waves.

In this article, we present an algorithm that combines the extraction of a relevant signal from a coronal jet found within remote observations, linkage to in-situ plasma measurements and windowing of the data for the application of EMD, allowing the extraction and comparison of temporal characteristics embedded within remote and in-situ datasets linked to a coronal-jet eruption. The algorithm is first tested for robustness by considering potential problems that come as part of the observations, such as noise or additional signal buildup, and it is then applied to SDO/AIA and *Wind* observations of the solar corona and the solar wind, respectively.

While we aim in the future to use this technique on remote-sensing and in-situ observations made closer in to the Sun, the implementation shown here makes use of both remote and in-situ observations taken from the solar corona and at the Lagrange  $L_1$  point, using light curves and in-situ parameters, respectively. Furthermore, the current approach is intended primarily for the verification of a potential connection, rather than the prediction of a best region to be observed.

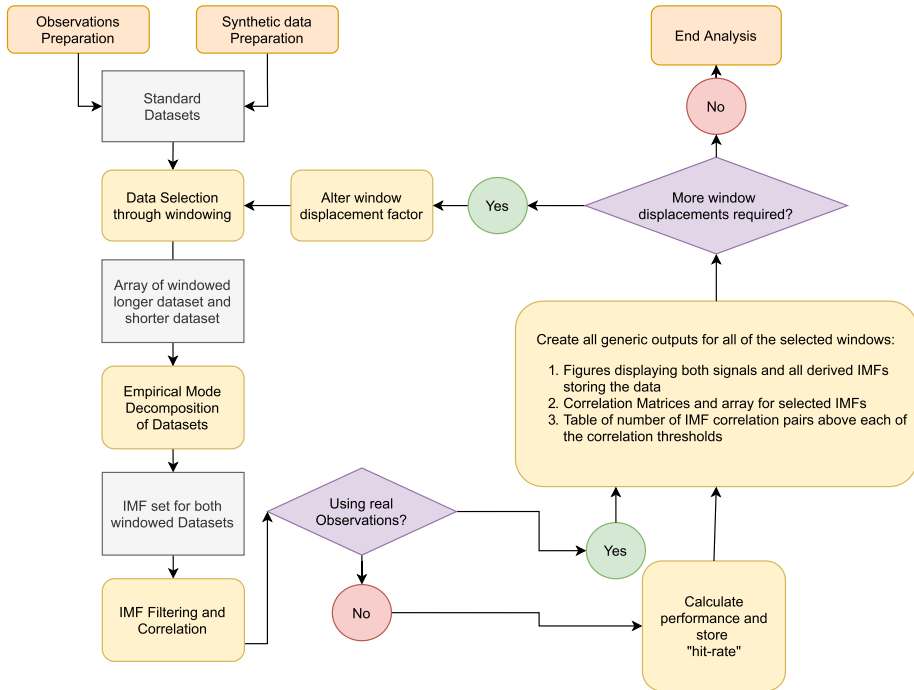
## 2. Algorithm Overview

The main purpose of this algorithm is to find matching temporal signatures in two datasets of different duration. It achieves this through the use of a decomposition method that retains all the information in the signal through selective windowing of the longer of the two datasets. Figure 2 displays the steps taken from the input of a (real or synthetic) dataset to the conclusion of the algorithm analysis. To apply the algorithm, we begin by generating and preparing the datasets that are to be compared. We extract data from the long timeseries using a shifting window with the same duration as the short timeseries. For the short and each of the windows of the long dataset, a signal decomposition method is deployed to extract timescales of variability present in both in the form of functions, ignoring timescales not deemed relevant to the user. We calculate the Pearson R-correlation (Freedman, Pisani, and Purves, 2007), a measure of the linear correlation between two variables, for functions with relevant timescales, and generate a correlation matrix. This procedure is repeated until the information included in the entirety of the long dataset is exhausted. Afterwards, the windows with statistically significant correlation identify the portions of the long timeseries where timescales of the short-duration signal are present.

For a general application of the algorithm, there is an assumption that the two datasets have the potential for a physically meaningful correlation. Similar to extracting the linear correlation between two functions, it is important for users of the algorithm to ensure that the datasets they apply the method to can depend on one another either linearly or non-linearly.

### 2.1. Data Standardisation, Normalisation, and Windowing

To prepare and standardise the data, we begin by ensuring that the two datasets are continuous by linearly interpolating missing values. After this, if the sampling frequency of



**Figure 2** Flowchart describing steps taken in timeseries analysis in this work. The observations require preparation before starting the process, with the following steps describing an object (grey rectangles), and an action that is applied on it (yellow rounded rectangles). Details on each of the steps may be found in the text. IMF is used as “Intrinsic Mode Function”.

the datasets is different, we re-sample the higher-cadence timeseries to the lower cadence through decimation.

Once the datasets are on the same cadence and have been normalised, a window of the length of the short timeseries is used to extract windows from the long timeseries. This window is shifted by one, or any number of, time step at a time, generating several windows of the long timeseries for later analysis. These windows are created until no more new windows of the same temporal duration as the short timeseries are possible.

## 2.2. Empirical Mode Decomposition (EMD)

When the data have been prepared and the relevant windows have been outlined for a specific pair of datasets, EMD can be used to determine the timescales of variability found in them. Appendix A explains the application of EMD for any input signal in more detail. Briefly, the end result of applying EMD to a signal is a set of Intrinsic Mode Functions (IMFs), which are generated depending on characteristic time scales of oscillations present in the data. The first IMF captures the fastest timescales in the signal, and successive IMFs contain increasingly slower components. After all of the IMFs are successfully extracted, only the residual (trend) of the signal is left. Due to the approach taken by EMD, the original signal can be reconstructed exactly by taking the sum of all of its IMFs, and it is possible to make use of IMFs that capture specific timescales exclusively through consideration of

their periodicity. EMD has furthermore been shown to outperform more traditional signal decomposition methods such as Wavelet or Fourier analysis (Huang et al., 1998).

### 2.3. Intrinsic Mode Function Selection and Correlation

Since not all of the derived IMFs provide information that is relevant to the signal of interest, it is useful to consider only those that contain dynamics occurring within a range of periodicities.  $P$ , the implied periodicity of an IMF, is calculated as

$$P = 2T/n, \quad (1)$$

where  $T$  is the temporal duration of the IMF, and  $n$  is the total number of maxima and minima in the IMF.

In our application to observations, we exclude IMFs with implied periodicities smaller than 5 minutes or larger than 100 minutes, as oscillations that are too fast are likely linked to background noise, and oscillations that are too slow and too similar to the trend may give oversimplified timeseries with respect to the dynamics of interest. After selecting the IMFs that are deemed as relevant for a given window, it is possible to correlate these against IMFs derived from the short dataset, effectively generating a correlation matrix similar to the one shown in Figure 3.

## 3. Algorithm Testing with Artificial Data

When using synthetic data, it is possible to define the performance of the technique for typical scenarios. By generating a synthetic short dataset with a characteristic signature and embedding it within a long dataset, we can test whether high degrees of correlation in the IMFs of the short dataset and a window of the long dataset are found where the pulse is embedded, or if the high correlation was instead a false positive.

We employ a correlation threshold  $\text{corr}_{\text{thr}}$  to define a “hit” as a window where temporal signatures of the two datasets match with an absolute correlation value stronger than the given  $\text{corr}_{\text{thr}}$ . Using this definition, a single window can display multiple hits in the case where several IMF pairs are correlated more strongly than  $\text{corr}_{\text{thr}}$ . By deciding how strict we are with IMF pair correlation, such that the number of false positives is minimised with respect to the number of correct identifications, we are able to optimise the performance for different use cases.

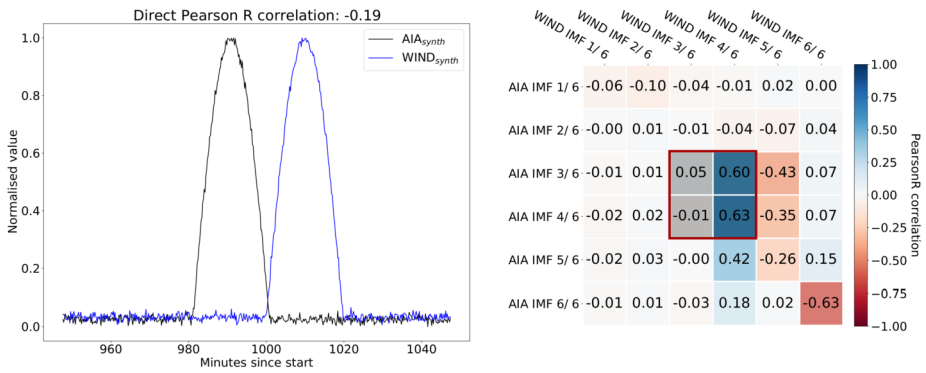
For any given  $\text{corr}_{\text{thr}}$ , we may consider hits within the bounds of the real peak location to be  $n_{\text{correct}}$ , and total hits as  $n_{\text{total}}$ , declaring the hit-rate as;

$$\text{Hit rate} = \frac{n_{\text{correct}}}{n_{\text{total}}} 100. \quad (2)$$

By using the hit-rate, we may quantify how well the synthetic pulse location was predicted for a specific  $\text{corr}_{\text{thr}}$ , taking values from 0–100% depending on how well the real peak location was identified.

### 3.1. Results from Artificial Data Tests

For application and assessment of the performance of the algorithm under realistic constraints, artificial timeseries representative of expected coronal and solar-wind properties



**Figure 3** Left: normalised value over time for a window of  $Wind_{synth}$  (blue) and  $AIA_{synth}$  (black), with white noise up to 1% of the value of the peak, with the calculated linear correlation between the two on top. Right: Combined correlation matrix for the timeseries on the left, for all calculated IMFs, with colour linked with value of the linear correlation shown on each IMF pair. IMFs that are relevant for the analysis are highlighted with a red box.

were generated as follows. A *short* timeseries (referred to as  $AIA_{synth}$  hereafter) that displays properties of a typical solar signal of interest, i.e. an AIA spatially averaged intensity lightcurve of a jet observation, and a *long* timeseries ( $Wind_{synth}$  hereafter), within which the synthetic solar-wind data were embedded, were created. We created  $AIA_{synth}$  as a 12-second cadence timeseries with a duration of 1.5 hours, with a constant value of 40 arbitrary units. We then added a Gaussian pulse with a duration of 20 minutes (1200 seconds) inside  $AIA_{synth}$  to broadly replicate characteristics of a jetting region as observed by SDO/AIA. We created the  $Wind_{synth}$  signal with a cadence of 3 seconds and a duration of 24 hours, taking a constant value of 10 arbitrary units, which are a similar cadence and average value for proton-density measurements at 1 AU. We then decimated  $Wind_{synth}$  to a cadence of 12 seconds, to match that of the  $AIA_{synth}$  signal, and added the Gaussian pulse from within  $AIA_{synth}$  to  $Wind_{synth}$ .

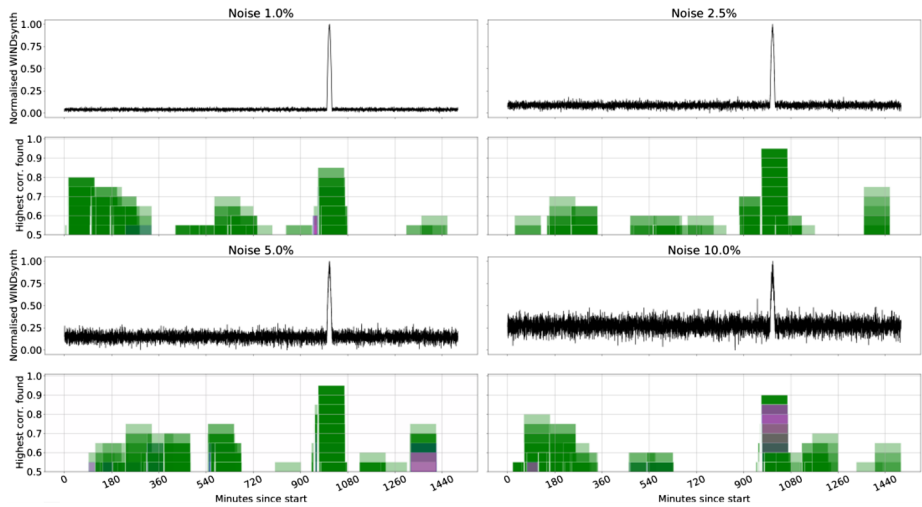
After generating this simple short and long signal pair, for which we show results from the algorithm run on in Appendix B, different tests were used to evaluate the performance of the algorithm under some more realistic constraints. These were the addition of white noise to the datasets, and the addition of a secondary pulse within  $Wind_{synth}$ .

### 3.1.1. Effects of Varying Levels of White Noise on Signal Recognition

By considering a range of white-noise levels, we derive information about how well our technique would perform in a turbulent medium, as is the case for both the solar corona and the solar wind. Gaussian noise was added to both the  $AIA_{synth}$  and  $Wind_{synth}$  datasets, with a standard deviation of 1, 2.5, 5, and 10% the amplitude of the embedded Gaussian pulse.

In the left panel of Figure 3 we show  $AIA_{synth}$  (black), and a window of  $Wind_{synth}$  (blue), with their calculated Pearson correlation of  $\approx -0.19$ , where 1% of the peak value is added as white noise to either signal. In the right panel of Figure 3 we show the corresponding correlation matrix for the derived IMFs for either dataset, with IMFs relevant for analysis highlighted with a red square. Each of the squares is coloured depending on the absolute Pearson correlation value for the specific IMF pair, and the numbers show the value rounded to two decimal places. We observe that, out of the four IMF pairs that are within the red square and thus considered relevant, two show a high degree of correlation ( $> 0.6$ ). We also





**Figure 4** Normalised values of  $Wind_{synth}$  (top), and windows (bottom) where considered IMF pair correlation is above a given  $corr_{thr}$ , for signals with up to 1, 2.5, 5, and 10% of Gaussian noise compared to the amplitude of the embedded Gaussian pulse. For the bottom panels, the colour of the bar segments indicates how many IMFs display a correlation above the given level, with green for one pair and magenta for two. Darker-bar segments indicate an overlap of different windows, due to the small time shift between consecutive windows.

note that, while directly calculating the Pearson correlation between the two datasets leads to a very small linear anti-correlation of  $\approx -0.19$ , the IMF matrix is effectively able to extract a stronger correlation, which may be expected of the two datasets by eye, as timescales included in both datasets are expected to behave similarly. With regards to the significance of the results, we have added the two-tailed p-values correlations of another case study in Appendix C, using a similar number of measurements, such that the significance results are thus also relevant for these test cases.

The correlation matrix shown on the right side of Figure 3 is useful to determine the likeness of the two empirically decomposed datasets for any specific window, but is lacking in terms of determining how well the signal was identified overall. For this reason, we examine the entirety of the  $Wind_{synth}$  dataset, and we identify all the windows where strong correlations were found.

Figure 4 shows the number of IMF pairs (colour of bar in bottom plot for each) that reach a given correlation (y-axis) for a given time window. We derive each of the panels in Figure 4 from an  $AIA_{synth}$ ,  $Wind_{synth}$  signal pair with a given white-noise percentage compared to the peak of 1, 2.5, 5, and 10%. Within each of the panels, the top displays the normalised  $Wind_{synth}$  signal over time, and the bottom displays the IMF correlation achieved at each of the derived windows. The width of the bar shown in the bottom was set to be equal to the temporal duration of  $AIA_{synth}$ , and the colour was set to green for cases where there was one hit ( $AIA-Wind$  IMF pair), and magenta for two. The opacity of the bar increases when hits are found consecutively. As we can see in Figure 4, most of the hits found at correlation thresholds higher than 0.85 are located in windows where the Gaussian pulse was embedded.

For summarised information about the results, Table 1 describes the hit-rate for each of the noise levels, at each of the correlation-threshold values. We see that 100% hit-rate is achieved for all noise profiles when the  $corr_{thr}$  is taken to be at least 0.85, implying that

**Table 1** Hit rate as calculated for a given Gaussian noise percentage at each of the correlation thresholds.

	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
1%	14%	16%	22%	20%	24%	33%	40%	100%	0%	0%
2.5%	24%	35%	51%	76%	84%	96%	100%	100%	100%	100%
5%	56%	60%	67%	74%	84%	93%	100%	100%	100%	100%
10%	61%	69%	80%	90%	90%	94%	97%	100%	100%	0%

selecting the window of any match above this  $\text{corr}_{\text{thr}}$  will always yield a correct identification of the Gaussian pulse. On the other hand, we can see that more-strict correlation thresholds of 0.9 or 0.95 lead to the cases with 1% and 10% noise, respectively, failing to identify any IMF pairs as valid.

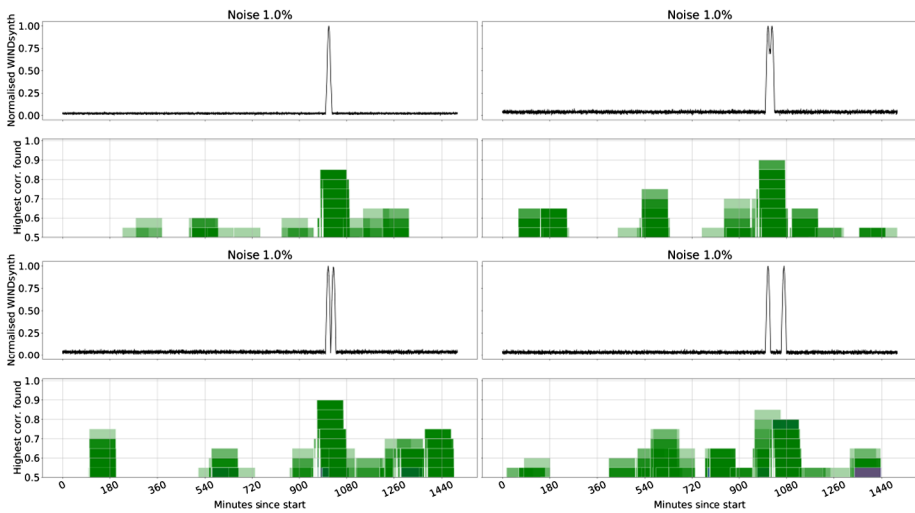
### 3.1.2. Effects of Additional Signals on Signal Recognition

Due to the large number of dynamic events taking place in the solar atmosphere, it is important to consider the effects of multiple pulses being captured in the in-situ dataset. These pulses may add complexity to the original remote-sensing signal, and could change the temporal signatures that are observed as part of the in-situ data, changing the morphology of the derived IMFs. As the separation of signals depends on the timing of their release into the solar wind, it is necessary to explore the effect of an additional pulse on IMF pair correlation and hit-rate with varying distances between the peaks. We choose to use a small amount of white noise of 1%, similarly to the first test case in Section 3.1.1.

To perform this test, the pulse found within  $\text{AIA}_{\text{synth}}$  was isolated and duplicated, with the original and the copy being embedded in the  $\text{Wind}_{\text{synth}}$  timeseries. The time difference between the peaks of the two signals was altered, selecting values from a quarter of the Gaussian width up to three times the signal width, providing a range of  $\text{Wind}_{\text{synth}}$  shapes. Extending the definition of hit-rate from Equation 2 for this test, we considered hits as correct when either of or both of the Gaussian pulse peaks were included in the selected window.

Figure 5 shows the hit-rate for a range of peak distances between the two signals on  $\text{Wind}_{\text{synth}}$ , with background noise taking values of up to 1% of the peak signal value. Within each of the four test cases, the top panel shows the modified  $\text{Wind}_{\text{synth}}$  signal, and the bottom shows windows where valid IMFs show correlations at each of the thresholds with the bars, with width equal to the duration of  $\text{AIA}_{\text{synth}}$  and the colour of the segments being determined by the amount of short–long IMF pairs found to be correlated more strongly that each of the thresholds. The distances between the peaks of the duplicated signals tested here are taken to be, starting from the top left and following clockwise 0.25, 0.75, 1, and 3 times the width of the original signal. All of the generated signals in this test can be directly compared to the performance of the first case on the previous test (top-left panel in Figure 4), as the white-noise level is taken to be 1% in both. We can see that most of the hits are found at the correct time if  $\text{corr}_{\text{thr}}$  is set to 0.8.

Table 2 displays the original results from 1% noise on top of the simple signals on the first row, copied from Table 1, followed by the hit-rate results for the different peak-separation distances shown in Figure 5. Table 2 shows that, unlike in the previous 1% noise test signal, a correlation threshold of 0.8 is enough for the hit-rate of all cases to be 100%, compared to the  $\text{corr}_{\text{thr}}$  of 0.85 that is required on the case with 1% noise. In Table 2 we can also see that every one of the duplicated-peak cases performs better than



**Figure 5** Normalised values of  $Wind_{synth}$  (top), and windows (bottom) where considered IMF pair correlation is above a given  $corr_{thr}$ , for signals with up to 1% of Gaussian noise compared to the amplitude of the embedded Gaussian pulse, and a duplicated pulse in  $Wind_{synth}$  positioned at 0.25, 0.75, 1, and 3 times the width of the original pulse from the centre of the original pulse. For the bottom panels, the colour of the bar segments indicates how many IMFs display a correlation above the given level, with green for one pair and magenta for two. Darker bar segments indicate an overlap of different windows, due to the small time shift between consecutive windows.

**Table 2** Hit rate as calculated for the simple case with 1% noise, and for each of the given duplicated signal peak separation, at each of the correlation thresholds.

	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
1%	14%	16%	22%	20%	24%	33%	40%	100%	0%	0%
0.25x	57%	78%	86%	94%	100%	100%	100%	100%	0%	0%
0.75x	38%	47%	56%	64%	88%	91%	100%	100%	100%	0%
1x	50%	53%	57%	61%	72%	80%	100%	100%	100%	0%
3x	52%	64%	73%	80%	92%	92%	100%	100%	0%	0%

the simple 1% noise test, with higher calculated hit-rates across all correlation thresholds.

### 3.2. Discussion of Tests

After assessing the performance of the technique under different scenarios, we have established the conditions under which the method performs poorly. Generally, a correlation threshold above 0.7 shows hit-rates above 80% in all cases, except the one with 1% Gaussian noise. This is likely because too many IMFs are ignored in this case, as the signal is over-simplified with such a small amount of noise. Based on this assessment, we expect IMF Pearson correlations above 0.7 to be relevant, and correct in three out of every four cases, therefore implying strong temporal signature correlation between the two datasets.

## 4. Case Study of SDO and Wind Datasets

After the testing of the algorithm with synthetic datasets, the same methodology was employed to match temporal signatures of an observational case which used remote-sensing and in-situ measurements of solar features. In the case of coronal-jet signals and the injection of their temporal signatures into the solar wind, the assumptions we make for application of the algorithm are:

- i) That dynamics observed as light-emission-intensity changes in the 193 Ångstrom pass-band of SDO/AIA impose particular timescales that then propagate with the outflowing stream.
- ii) That these timescales do not suffer significant changes during propagation out to 1 AU.
- iii) That the Parker-spiral model of the solar wind can be utilised to predict solar-wind origins within a 12-hour error bar.

The first step to select data for a case study was to explore coronal-hole solar-wind streams near solar minimum, at a time where SDO observations were available. The chosen period was the second week of November 2016, when a coronal-hole solar-wind stream was identified through low charge-state ratios, high bulk flow velocity, and low proton density in solar-wind measurements at  $L_1$ . This time period was checked using the [Cactus Coronal Mass Ejection \(CACTus\) catalogue](#) to ensure that there was no large-scale, Earth-facing activity that may have affected the persistence of small-scale coronal activity in the solar-wind stream. The second step was to identify the likely origin of this solar-wind stream. The likely origin was identified to be a large southern coronal-hole extension, observed in SDO/AIA EUV images of the solar atmosphere. This coronal-hole extension was large, spanned the solar Equator, and showed no active regions surrounding it.

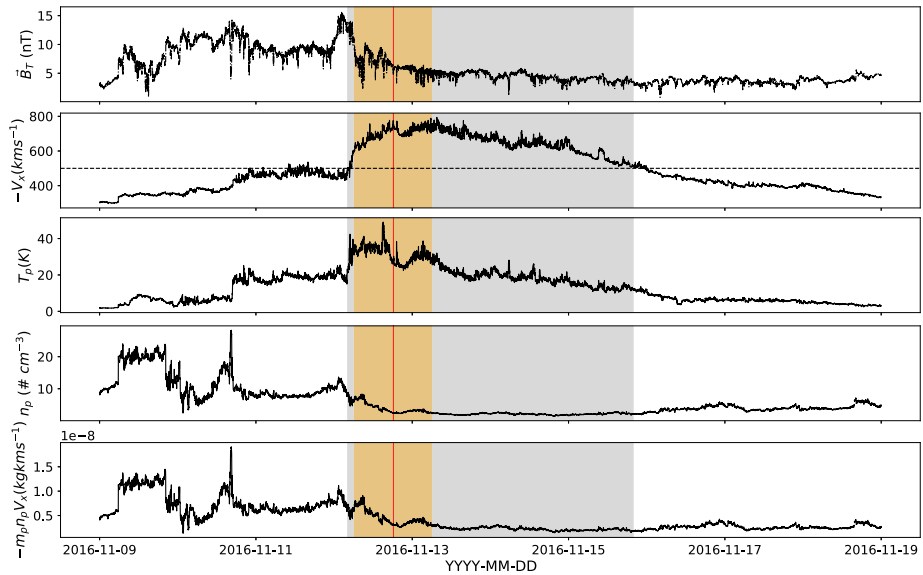
We first required a characterisation of observational signatures of relevant temporal signatures in the solar corona. Through investigation of these measurements, we inferred events that could potentially affect solar-wind temporal characteristics. The measurements that we utilised for this purpose were light-emission-intensity measurements of the solar corona on the remote-sensing side, and proton density, temperature, mass flux, solar-wind velocity, and magnetic-field strength on the in-situ side.

Due to the uncertainty in time of ejection and duration of travel, for the 1.5 hours of remote-sensing observations, we collected 24 hours of in-situ observations to apply the algorithm to.

### 4.1. Observations

We use in-situ solar-wind measurements of plasma bulk-flow velocity from the 3DP instrument (3DP: Lin et al., 1995) onboard the *Wind* spacecraft (Harten and Clark, 1995), along with the spacecraft position at the time to perform ballistic back-mapping of the solar-wind stream to the solar corona. We highlight measurements used in the back-mapping with the red bar in Figure 6 (18:05 to 18:15 UT on the 12 November), and the 24 hours of measurements used for analysis in orange. The back-mapped stream shown in red was selected to be at the peak of outflowing velocity of a coronal-hole solar-wind stream, which lasts several days and is illustrated in grey.

For the application of our algorithm, we used data from 06:00 UT on 12 November to 06:00 UT on 13 November 2016, and considered several different measurements from the *Wind* spacecraft. From the 3DP instrument, we considered the solar-wind velocity  $V_x$ , proton number density  $N_p$ , proton temperature  $T_p$ , and calculated mass flux radially away from the

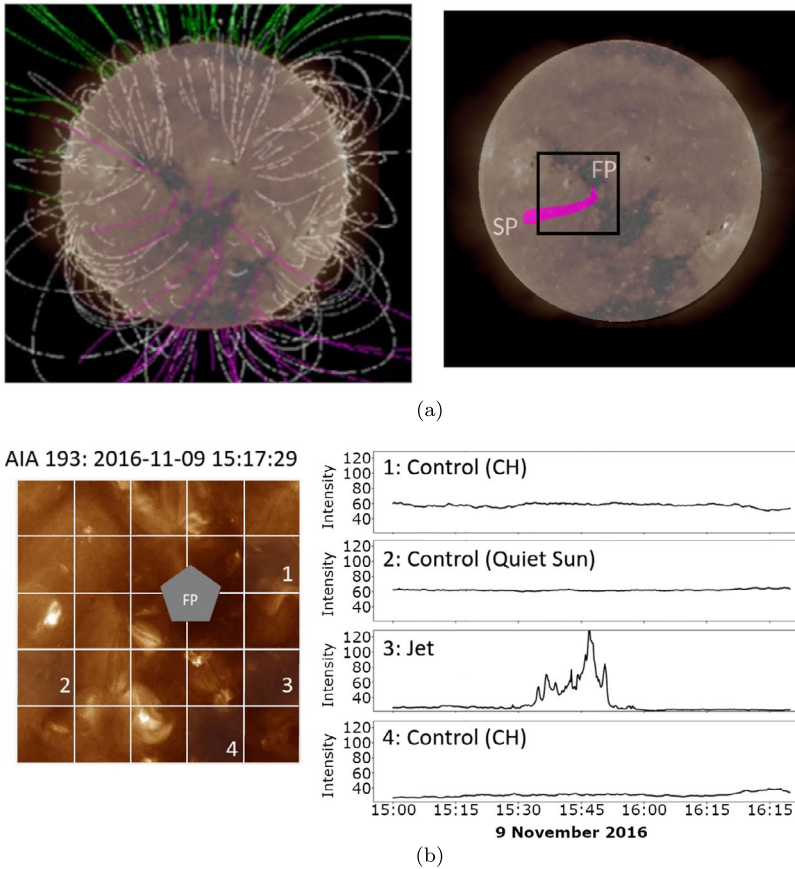


**Figure 6** Solar-wind stream from 9 to 19 November 2016. From the top to the bottom panel, we show total magnetic-field strength, solar-wind radial velocity, proton temperature, number density, and mass flux. Highlighted in grey the portion of solar wind with coronal-hole plasma-like properties: high bulk flow velocity (above  $500 \text{ km s}^{-1}$ ) and low proton density. The red vertical line highlights data which are used for ballistic back-mapping (18:05–18:15 12 November 2016). The orange column highlights data that is used in the analysis, selected up to 12 hours before and 12 hours after data used for back-mapping.

Sun,  $M_f = -N_p m_p V_x$ . From the *Magnetic Field Investigation* instrument (MFI: Lepping et al., 1995) we calculated the total magnetic-field magnitude as  $B_T = \sqrt{B_x^2 + B_y^2 + B_z^2}$ .

To construct the remote-sensing signal, we employed EUV measurements of the 193 Å passband from SDO/AIA. These provided us both with context information for the magnetic-footpoint location and with observations needed to characterise dynamics in the upper solar atmosphere. The coronal region of interest was selected using results from a combination of ballistic back-mapping and coronal-field-line tracing, leading us to employ SDO/AIA 193 measurements from 15:00 to 16:35 UT on 9 November 2016.

We obtained remote-sensing measurements of the photospheric magnetic field as synoptic magnetograms. This is a data product that is published by the *Global Oscillation Network Group* network (GONG: Harvey et al., 1996). The time of the selected GONG synoptic map was 18:00 UT on 9 November 2016, as this was the closest synoptic map to the calculated time of solar-wind ejection of  $\approx 16:00$ . These magnetograms contain measurements of the magnetic-field strength and polarity and were used as the boundary condition for a Potential Field Source Surface model of the coronal magnetic field. The PFSS model solver that we employed can be found in the “pfss” package as part of SolarSoftWare (de Rosa, 2010). This package allows for the derivation of potential-field solutions to the coronal magnetic-field, with the free parameter being the “Source Surface”: defined as a height where the solar-wind kinetic pressure is expected to dominate, such that any coronal loop that extends to the arbitrary height is considered “open”. The source-surface height is usually set at around 2–2.5 solar radii above the solar surface. In this analysis we selected a source-surface height of 2.5 solar radii to minimise the error associated with selecting a lower source surface and potentially back-mapping to a region of the photosphere that has closed magnetic field.



**Figure 7** **a** Left: PFSS extrapolation overlaid on 193 Å observation of the solar atmosphere, showing closed field lines in white, and open field lines in green (positive polarity) and pink (negative polarity). Right: Coronal-field-line tracing from source-surface height of 2.5 solar radii, at the sourcepoint “SP”, to the field footpoints in the solar photosphere, “FP”, with the black bounding box representing the field of view used in the case study. **b** Left: Cutout of an SDO/AIA 193 image of the area utilised in the case study, with numbered regions, and an indicated footpoint location for *Wind* observations (FP) with a grey pentagon. Right: Average emission intensity in SDO/AIA 193 for numbered regions outlined to the left, considering only values at least a standard deviation above the regional mean within each respective square for one hour and a half of observations.

## 4.2. Data-Driven Modelling and Image Processing

To link remote and in-situ observations of the stream, we applied a two-step back-mapping technique, commonly used to produce mapped solar-wind sourcepoint locations at the photosphere (see, e.g., Fazakerley, Harra, and van Driel-Gesztelyi, 2016, and the references within). This technique makes use of a PFSS model, which combined with ballistic back-mapping (see Nolte and Roelof, 1973), allows for the establishment of a connection between a given solar-wind stream, and a time and location of ejection from the solar atmosphere. We employed an empirical correction to travel time [ $t_{\text{real}} = \frac{4}{3}t_{\text{ballistic}}$ ], which was first discussed by Neugebauer et al. (1998), and later tested on coronal-hole solar-wind streams by Macneil

(2019), showing promising results with regards to footpoint location when compared to the original calculated source region.

Figure 7a shows the PFSS extrapolation of the solar photospheric magnetic field at the relevant time, and overlays SDO/AIA 193 measurements to guide the reader in the relative position of open (closed) field lines (white), on top of the equatorial coronal hole. On the right of Figure 7a, a black square is overlaid on the figure, representing the location of the bounding box identified and used by the study, as well as calculated footpoints for the ten minutes of *Wind* observations used in the ballistic back-mapping, after following the relevant open magnetic-field lines to the photosphere.

We took these magnetic footpoints to be within a  $250 \text{ arcsec}^2$  bounding box, from which a cutout of the selected pixels was selected to extract the right panel of Figure 7b. The size of this bounding box was chosen to be similar to the field of view of the *High Resolution Imager* that is part of EU1, onboard *Solar Orbiter*, when the spacecraft is at 0.3 AU (Rochus et al., 2020). These extreme-ultraviolet light-emission-intensity images were processed using the `aia_prep.pro` routine, available in the SolarSoftWare library, before extraction of the lightcurves.

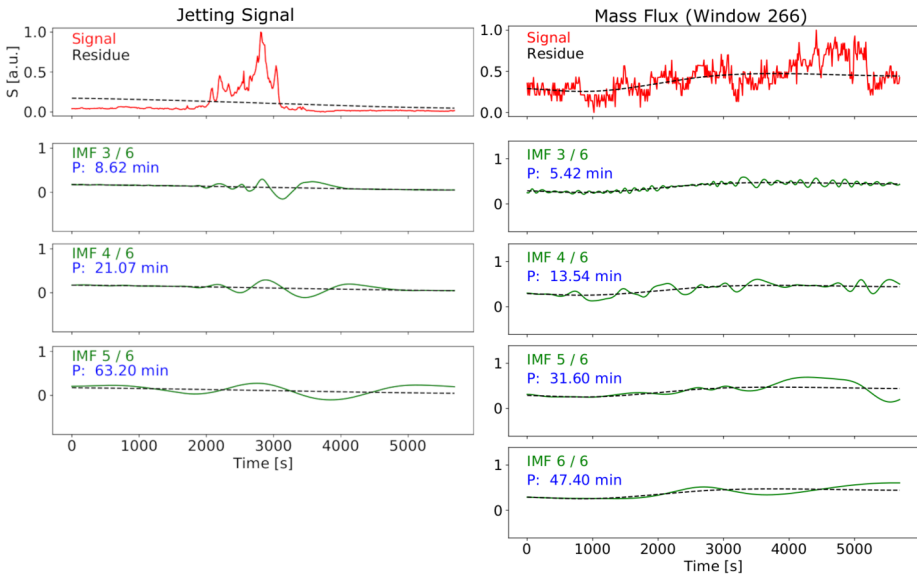
From this  $250 \text{ arcsec}^2$  cutout we create a grid with a set of five rows and columns, separating each of the image into 25 squares, such that the small-scale dynamics present within each of them could be well resolved by SDO/AIA, while some context was retained. Within each of these squares, we extracted a proxy for activity within the waveband as a lightcurve by considering the average of pixels that had a brightness at least one standard deviation above the regional mean. The right panel of Figure 7b shows four EUV emission-intensity curves (lightcurves) for each of the numbered squares. The third outlined region “Jet” shows data from a part of the coronal hole where a coronal bright point emerged and produced four brightening events that showed jet-like features (collimated intensity enhancements, fast flows), with durations over 10–20 minutes, lasting about 20 minutes overall. The other numbered regions were extracted to be used as control datasets for both the signal extraction and analysis methods, and chosen to be different from the “Jet” region in background intensity and expected timescales.

We applied the algorithm to the four selected remote-sensing datasets, as well as to each of the five parameters in the in-situ dataset shown in Figure 6 for the period shown as shaded orange. The remote-sensing data were taken to be the short dataset (AIA), and the in-situ parameters were taken to be the long dataset (*Wind*). Similarly to the synthetic cases, each dataset was normalised before analysis.

### 4.3. Results

Both of the columns in Figure 8 display a signal decomposed using EMD, as well as their calculated IMFs with periods [ $P$ ] found to be between 5 and 100 minutes, as calculated using Equation 1, and in line with timescales of coronal bright-point jets (Madjarska, 2019), along with the IMF number on the top left of each IMF. On the top panel of either column, we show the signal in red, and the residual in a dashed black line.

On the left panel of Figure 8, we show the “jetting” signal. On the top panel we find a clear signal compared to the background, with a localised peak around the centre, and several features of varying magnitude within it, which are captured in the IMFs. IMF 3/6 captures dynamics that take place on approximately ten-minute timescales, capturing four sudden brightening events of different intensities. IMF 4/6 is constituted by dynamics that have a timescale of 20 minutes, namely the two main features of the jetting signal. IMF 5 captures the entire coronal bright point event, which shows an estimated period of 63 minutes, similar to the implied duration of the event.



**Figure 8** Left: Resulting IMF decomposition for normalised jetting signal of a region of the solar atmosphere that displays jetting activity on top of a coronal bright point. Right: Resulting IMF decomposition for a window of the proton-density measurements, which shows a high degree of correlation against the jetting signal. Top panel of each column shows original signal (red), along with residual (black), with panels below displaying all derived Intrinsic Mode Functions with the residual values added, from smallest to largest implied periodicities  $P$ , accompanied by IMF number.

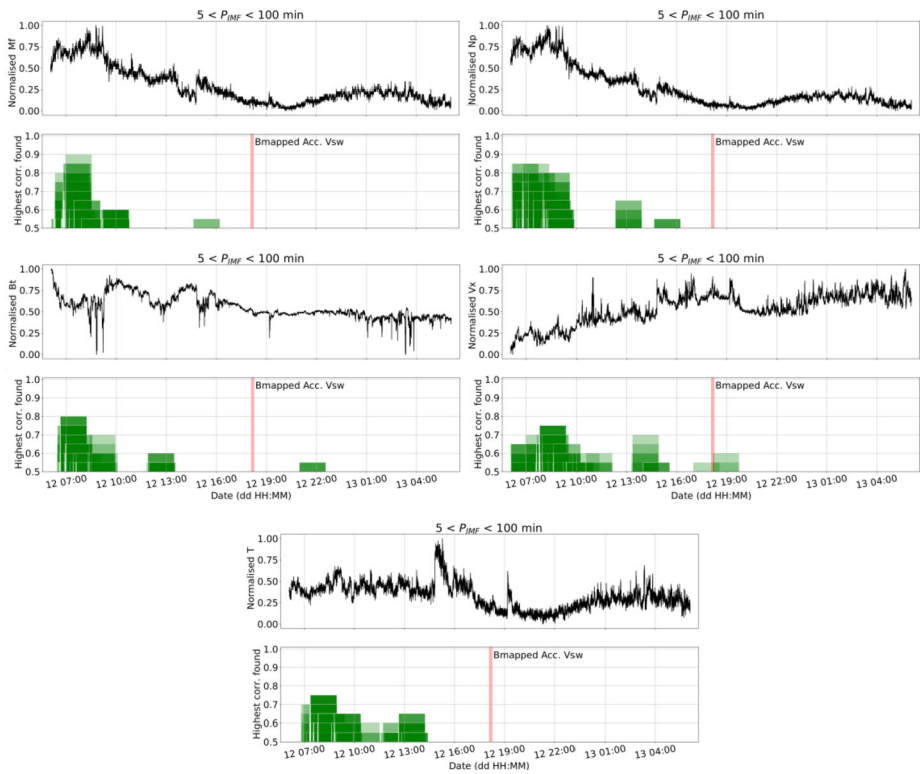
On the right column of Figure 8, we show a window of the mass flux signal that shows a high degree of correlation to the remote-sensing dataset when applying the algorithm. Of all of the tested in-situ parameters, the mass flux provides the highest correlation between the IMFs that are considered. On the top panel of Figure 8, we do not find a structure as clear as on the jetting signal. IMFs 3, 4, 5 and 6 show implied periodicities of 5, 12, 32, and 63 minutes, respectively, and are therefore all considered for correlation.

For each of the *Wind* parameters, the algorithm described in Section 2 was applied, correlating them against the remote-sensing jet observations, and thus yielding time shifts for which temporal signatures found within relevant IMFs match up with those also found in the remote data.

Figure 9 shows the results from the algorithm being applied to  $M_f$  (top left),  $N_p$  (top right),  $B_T$  (bottom left),  $V_x$  (bottom right), and  $T_p$  (centre bottom), ordered depending on the highest correlation achieved between valid IMFs. Within each of the five plots, the top part shows the normalised in-situ dataset over time, and the bottom part shows the highest Pearson correlation achieved for valid IMF pairs over time. The width of the bars is equal to the temporal duration of the remote-sensing signal, and the colour of each of the segments corresponds to the number of IMF pairs which are correlated more significantly than each of the specific thresholds. The red vertical bar shown in all of the panels is co-temporal with the radial solar-wind velocity used for pinpointing the origin of the solar-wind stream.

Broadly, the highest correlations for all variables in Figure 9 are reached near the start of the timeseries, approximately 10 to 11 hours before the back-mapped time (07:00 UT on the



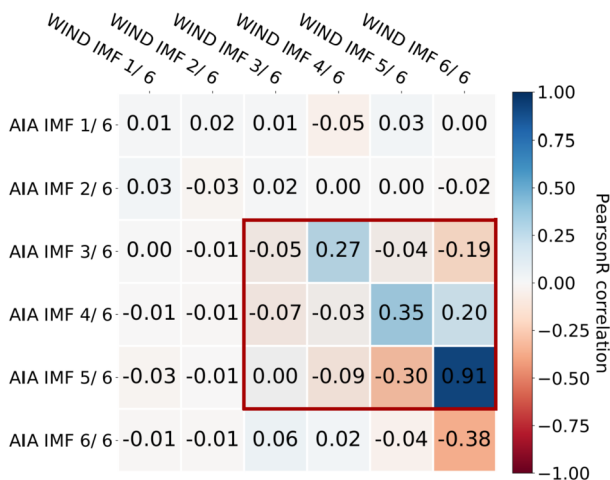


**Figure 9** Normalised values of in-situ parameters (top), and windows (bottom) where considered IMF pair correlation to the jetting signal is above a given  $corr_{thr}$ , for the set of *Wind* measurements used in the algorithm, decimated to 12 seconds and then normalised. The width of the green bars is equal to the temporal duration of the jetting dataset (1.5 hours), and the red bar shows the time used for ballistic back-mapping. Variables shown are as follows: top left, solar-wind mass flux [ $M_f$ ]; top right, proton number density [ $N_p$ ]; bottom left, Total magnetic-field strength [ $B_T$ ]; bottom right, solar-wind radial velocity [ $-V_x$ ]; bottom centre: proton temperature [ $T_p$ ].

12 November). The variable that shows the strongest correlation is the mass flux [ $M_f$ ], with a window displaying a Pearson correlation between one IMF pair higher than 0.9. The rest of the parameters shown here are correlated less significantly, displaying similar timescales to the jetting observations between 07:00 UT and 10:00 UT on 12 November.

Figure 10 contains the IMF correlation matrix for a window of the normalised mass flux dataset that corresponds to the peak correlation reached on the top-left panel of Figure 9. The intensity of the blue/red colour in Figure 10 is linked to the strength of the calculated Pearson linear correlation between IMFs. The fifth AIA IMF is positively correlated with the sixth *Wind* IMF with a high value of 0.91. The remainder of the correlations highlighted in the red rectangle are below 0.4, and therefore are not strong enough to be relevant for the purpose of finding similar timescales within the two datasets. Appendix C shows the corresponding p-values for the Pearson R-values that we show in Figure 10, which are indicative of the significance of these results.

**Figure 10** Pearson linear correlation matrix for IMFs derived from a highly correlated window of the mass flux. Colour of the square indicates positive or negative correlation, with the intensity of the shade related to the absolute value. The highlighted rectangle in red, spanning *Wind* IMFs 3–6 and AIA IMFs 3–5, shows IMF pairs that were considered for correlation.



## 5. Discussion

### 5.1. Technique Performance on Artificial Datasets

We have described an algorithm that is able to correctly identify the location of a simple Gaussian-like pulse under idealised circumstances in synthetic observations. The performance appears to be poor when the utilised correlation threshold is below 0.7. We therefore see that the technique requires a strong, clear pulse, that needs to stand out well above background oscillations, and benefits from a degree of complexity on the background noise up to the case with 10% noise.

When including an additional pulse in the dataset, the algorithm performance is better for larger separation between the pulses. This is likely due to a large overlap between the signals significantly deforming the original shape of the relevant window, affecting the derived IMF shape and by extension timescales that are detected.

### 5.2. Technique Performance on Observational Dataset

We have shown that our method can distinguish between coherent enhancements and random fluctuations of a signal by accounting for expected periodicity of events that occur in a remote-sensing dataset (see Figure 8), while still considering the amplitude of the decomposed variability thanks to EMD. For our case study, the correlation matrix for the window showing the highest correlation for the two datasets overall, the solar-wind mass flux, is shown in Figure 10. We discuss the scientific implication for this result in the next subsection. We also note that the only IMF pair that is highly correlated is that with the slowest variability for either dataset, and see similar behaviour on all of the other matching windows, where the highest-order IMFs of either dataset are the ones found to be correlated more strongly. This is a different behaviour from that observed from the synthetic case, where a cluster containing several different IMF pairs shows strong correlation between them (see Figure 3).

### 5.3. Science Results from Observational Case Study

After applying our algorithm to observations, we observe that the best match for the timescales of the jetting eruption occur in the solar-wind dataset some 10–12 hours be-

fore the time identified using solar-wind back-mapping. This high correlation is seen most notably in the mass flux but is identifiable also in the proton number density, total magnetic field, and across other analyzed parameters with varying degrees of linear correlation values. Overall, these results imply that the solar wind from the coronal ejection could have reached 1 AU rather faster than would be expected from the simple ballistic back-mapping. The latter technique was applied in the case presented here with the “empirical correction factor” of  $4/3$  originally suggested by Neugebauer et al. (1998). These authors employed this factor to force a closer match between mapping of coronal-hole boundaries and their corresponding solar-wind streams, and this factor has been subsequently used in other studies of solar-wind propagation such as Macneil (2019). However, given that is a rather simplistic approach, it may be that one of the primary uses for our technique would be fine-tuning solar-wind travel times for use in studies of the link between the solar corona and the heliosphere. Indeed, application of the technique for different events and over different radial-distance ranges (either between the corona and spacecraft such as *Solar Orbiter* and PSP or between pairs of such spacecraft in the solar-wind) offers the possibility of refining the value of the average empirical correction factor, and determining its dependence on distance range. In this sense, the technique developed here could offer the opportunity to determine the degree of acceleration that the solar wind may experience at different distances. Nevertheless, in the case study presented here, our results imply a solar-wind travel time of 64 hours instead of the 75 hours predicted by the simple back-mapping with a  $4/3$  correction factor. Application of this factor implies that the solar wind has an average speed over its journey to 1 AU that is only 75% of that observed. Performing the back-mapping without the  $4/3$  factor, the calculated travel time (for solar-wind traveling at constant speed – the observed speed at 1 AU) would be 56 hours. Our result clearly falls between these two values, and so in this case we can conclude that the result would be consistent with the solar wind having an average speed which is 87.5% of the observed speed at 1 AU. Since this speed is 17% faster than use of the  $4/3$  empirical factor would imply, it appears that the solar-wind stream had experienced less acceleration over its journey than the case studied by Neugebauer et al. (1998). This means that a smaller empirical correction factor of  $64/56 = 8/7$  would have performed better than the  $4/3$  acceleration factor for this solar-wind stream.

We performed the EMD analysis using a number of parameters categorizing the solar wind. As we noted above, of these tested parameters, the solar-wind mass flux is the parameter that shows the greatest similarity in EMD timescales with the coronal-hole jetting signal. This may indicate that variation of the mass-loss rate from the coronal hole, which manifests as variation in the mass flux detected at a spacecraft in the solar wind, has a broad relationship with the intensity variations in the bright-point light curve (cf. Figure 1). The good correlation for the mass flux may then be due to this being a conserved quantity, decreasing at a rate directly proportional to radial distance.

#### 5.4. Challenges in Data Collection

Several assumptions made when extracting the timeseries must be considered. These are with regards to the kinematics of plasma near and away from the solar origin, signal persistence and propagation, impact of solar activity on performance, or are imposed due to constraints of the modeling.

In terms of assumptions related to the kinematics of the plasma, there could be problems with the calculation of the back-mapped time and location. The first potential problem is that the time taken for jet plasma to reach from the low corona to the point where the solar wind is generated, the “source surface”, is not considered here. Additionally, although we have

argued above that the technique could be used to refine the empirical correction factor and more accurately identify a travel time, problems may result if in reality that correction factor, reflecting the degree of acceleration of the wind, changes rapidly from stream to stream.

These issues can be minimised by making use of in-situ measurements taken closer in to the corona, such as those provided by PSP and *Solar Orbiter*, while the use of a different model for validation of the estimated emission time may help in terms of propagation within the low-lying solar atmosphere.

Regarding assumptions related to the persistence of the signal during propagation, there is an important assumption that the morphology of the Intrinsic Mode Functions, and therefore of the maxima and minima present in the original signal and respective frequencies that are extracted, would be maintained from emission in the corona to the measurement at *Wind*. Furthermore, the relationship between light-emission-intensity changes in the corona and plasma density and temperature has been shown not to be one-to-one. In spite of the idea of modifying signal shape slightly being discussed in the case where the duplicated signal overlaps with the original signal in Section 3.1.2, it could be further tackled by testing the effects of modifications of signal morphology through stretching the signal before insertion in the longer dataset, as well as through use of one-dimensional solar-wind models (e.g. Pinto and Rouillard, 2017) to understand likely evolution and expansion of the solar-wind and magnetic field, and thus of the signal.

While we have tested the effects of one additional Gaussian pulse on synthetic observations, we have not considered the effects of a more active underlying corona. The current test case is purposely made on as quiet conditions as possible for available observations, and further work would need to be done in order to constrain and differentiate potential source regions of solar-wind streams for more complex coronal configurations.

Regarding assumptions made in the modelling of the coronal magnetic field, the selection of a representative source-surface height for PFSS models has been studied extensively in the past. More recently, it was demonstrated that in order for the polarity of a coronal-hole source region to match that of the solar wind captured at PSP, different modifications to the height of the source surface were required (see Badman et al., 2020, and the references therein), with current work agreeing with a non-spherical source surface based on near-Sun observations from PSP (Panasenco et al., 2020). Similarly to the plasma-kinematics problem, making use of PSP or *Solar Orbiter* data to verify the connection through, e.g., polarity on the two domains can help in deciding what the best-performing source-surface height is for the specific stream.

## 5.5. Implementation of the Technique in *Solar Orbiter* Operations

As it stands, the technique we propose has been tested to work on simple synthetic data and on a single observational dataset. We show an approach for comparing timescales at a predicted solar-wind stream footpoint to those present within the outgoing solar-wind stream. For use in the *Solar Orbiter* data-assimilation pipeline in the future, it will be necessary to investigate parameter space in more detail, and create a model for solar-wind emission around the predicted footpoints, identifying likely areas and contributions. Relevant parameters for remote measurements that can be refined are the selected fields of view, the size of sub-regions, and the magnetic-field-line expansion in the low corona leading to possibly different footpoints. With regards to in-situ observations, the two free parameters are the variables being considered and whether there is a “stretching” factor enforced by field-line expansion to fill the heliosphere.

## 6. Conclusions

We have created an algorithm that, through the use of windowing and EMD, is able to compare the timescales of variability of two different timeseries. We have tested the algorithm with synthetic datasets over a range of noise levels and found that if using a correlation threshold of at least 0.7 we ensure temporal signatures are found where expected over 75% of the time. We have investigated the effects of additional signals on the performance of the algorithm and have found that it was still able to extract relevant signatures with an even lower correlation threshold required.

When we have applied the algorithm to SDO/AIA and *Wind* observations, we find similar variability timescales at a time 10–11 hours earlier than expected on the basis of the solar-wind stream used for back-mapping, and that the best-performing parameter is the mass flux. The former finding implies that the ballistic back-mapping underestimates solar-wind speed by about 15% at 1 AU when using an empirical correction factor of 4/3 to simulate the effects of solar-wind acceleration, suggesting that a smaller empirical correction should be instead applied. We have argued that the latter finding is supported by the conservation of mass flux when assuming the magnetic-field lines to be frozen into the plasma for high plasma- $\beta$  environments.

In the future, this technique is to be used with PSP and *Solar Orbiter* data. To date, there have been no suitable remote-observation configurations for PSP perihelia such that the footpoints of a stream are on-disk at the relevant time. The application of this technique to the closer-in PSP observations is therefore postponed for future work.

Apart from making use of data from spacecraft found closer to the Sun, future work will explore signal-shape modification through compressive effects in the solar wind by compressing and stretching the relevant remote-sensing signal before applying EMD and comparing to the in-situ observations. This approach would effectively increase the possible dynamical features that can be correlated, enhancing the applicability of this algorithm, as well as allowing for more realistic constraints.

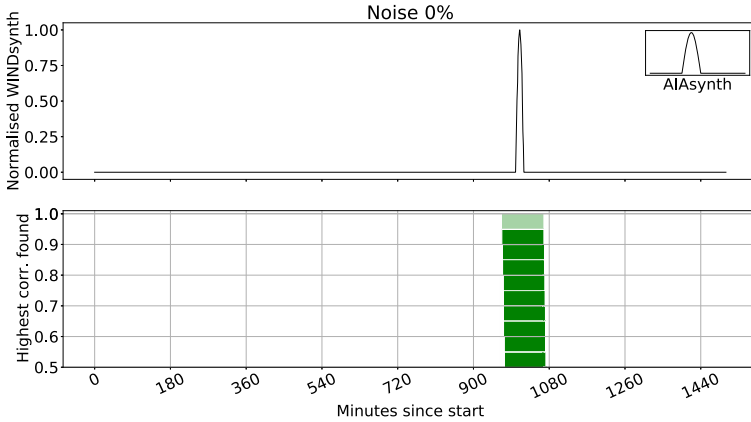
Additionally, and as a final planned application of the algorithm, we plan to apply it to in-situ datasets found within the same Parker spiral. With this application we will be able to verify its performance with equivalent properties for different types of solar-wind stream, at varying distances to the Sun. As we have bulk-flow-velocity measurements of the solar wind at the spacecraft that is closer to the Sun of the two, ballistic back-mapping will have a useful solar-wind speed value. We will explore whether the times of statistically significant temporal signature correlation are when expected, or at a different time altogether. This will inform us on how under/overestimated the Parker-spiral constant solar-wind speed model velocity is, and by extension aid in the study of solar-wind origins and propagation.

## Appendix A: Empirical Mode Decomposition (EMD) Definition

The most evident advantage to using EMD over more traditional techniques is that, for any data, the sum of all IMFs plus residual is exactly equal to the input data, independently of data resolution and number of IMFs that are derived.

EMD analysis considers a given time series to be composed of a combination of oscillations, which may not display purely sinusoidal behaviours. These oscillation packets can be isolated from the timeseries depending on their characteristic variability time scales. For an input signal  $x(t)$ ,  $EMD(x(t))$  is calculated as follows:

- i) Identify all local maxima and minima in the given timeseries (now taken as  $x_m(t)$ ).



**Figure 11** Locations where the correlation between the IMFs of  $AIA_{synth}$  and  $Wind_{synth}$  display high correlation for signals with no additional noise. On the top right, the  $AIA_{synth}$  signal is shown. The x-axis is time since start in minutes, and the y-axis is the correlation threshold reached. Colour of the bars indicates how many IMFs display a correlation above the given level, with green for one pair.

- ii) Interpolate between all identified local maxima, to estimate an upper envelope  $x_{high_m}(t)$ , and between all local minima to estimate a lower envelope  $x_{low_m}(t)$ .
- iii) Compute the mean of the two envelopes:  $m(t) = \frac{x_{high}(t)+x_{low}(t)}{2}$  and subtract it from the given timeseries:  $d(t) = x_m(t) - m(t)$ , then set  $x_{m+1}(t) = d(t)$ .

Steps i–iii are then repeated on the newly defined  $x_{m+1}(t)$  until the resulting signal  $x_0$  satisfies the following criteria:

- i) That the number of extremes and the number of zero crossings are equal, or differ at most by one.
- ii) That the mean value  $m(t)$  of the upper and lower envelope is zero at all times.
- iii) That the standard-deviation difference between two consecutive repetitions is not smaller than a predetermined value.

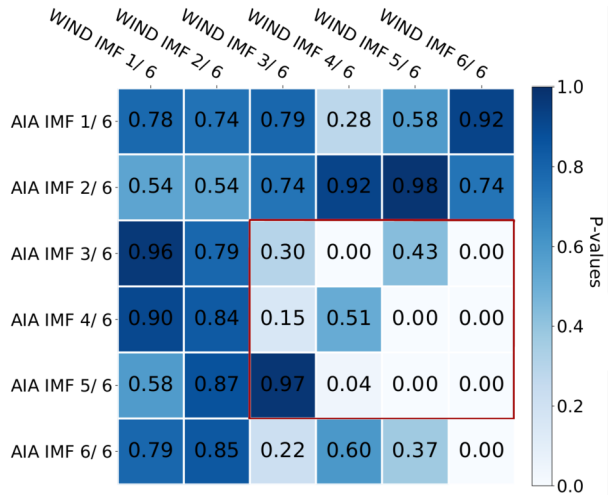
When all of these criteria are met, the resulting signal  $x_0(t)$  is said to be the first IMF for the input signal. When this resulting signal is subtracted from the original data, the residual can be taken as the new data, enabling the process to start again. If the criteria are met after following the steps and iterating upon them, a second-order IMF [ $IMF_1(t)$ ] may be defined. This process is then repeated, extracting new IMFs defined as  $IMF_n(t)$ , up to the point where the residual becomes a monotonic function (the trend of  $x(t)$ ).

### Appendix B: Algorithm Accuracy Under Ideal Conditions

In order to determine reliability of the algorithm, we perform a simple test where we embed a single Gaussian on a long and a short dataset. We apply the smallest possible window displacement, and we expect there to be several windows around the location of the embedded signal where the match is optimal, generating 100% hit-rate as the rest of the signal does not correlate strongly.

The top panel of Figure 11 shows the  $Wind_{synth}$  dataset, without any additional noise. On the top right, we show  $AIA_{synth}$ , a simple Gaussian pulse, whose IMFs we correlate against

**Figure 12** Matrix containing two-tailed p-values for IMF pairs calculated with the mass flux observational test case and the jetting signal shown in Section 4. Intensity of the blue colouring is linked to p-value. Red highlight denotes IMF pairs that are used in analysis.



the normalised  $Wind_{synth}$  signal. The bottom panel of Figure 11 displays correlation reached over time, specified in minutes since start, with the width of the bars being equal to the temporal duration of  $AIA_{synth}$ , and the green colour declaring that one IMF pair reaches each of the relevant correlation thresholds.

As we can see in Figure 11, the only time where matching signatures are found within the derived IMFs is when the two signals overlap exactly, illustrated by the centre of the green bar coinciding with the peak of the Gaussian. Note that the colour of the match is slightly lighter as there is exactly one timestep where the two timeseries overlap exactly, and thus a correlation of exactly unity is achieved. This test shows that our algorithm successfully finds the similarity between the two datasets, with 100% hit-rate.

### Appendix C: Significance of Correlation Values for Wind Case Study

In order to verify the statistical significance of the calculated correlations, we extracted the two-tailed p-values for the correlation matrix shown in Figure 10. The two-tailed p-value gives the probability that the null hypothesis cannot be rejected. In the case of the Pearson R-correlation, the null hypothesis is that the absolute value of a correlation value could be achieved by a different pair of datasets, therefore the p-value is a measure of the statistical significance of any calculated correlation. In this work, we take a p-value below 0.05 to be showing that the found correlation is statistically significant.

We show p-values for the *Wind* case study in Section in Figure 12, with lighter colours referring to lower p-values, and darker colours referring to higher p-values. We note that in pairs of strong or moderate correlation within Figure 9 (e.g. AIA IMF 5: *Wind* IMF 6, with a Pearson R-correlation of 0.9, or AIA IMF 4: *Wind* IMF 5, with a Pearson R-correlation of 0.35), the p-value is low, and therefore the calculated Pearson R-correlation is statistically significant. As the number of measurements is equal for all other parameters and correlations shown in Figure 9, as well as throughout the synthetic datasets from Section 3, we expect moderate and strong correlation values (above 0.3) to be statistically significant throughout this work.

**Acknowledgements** The authors thank the reviewer whose comments helped to significantly improve the article. This work utilises data from the National Solar Observatory Integrated Synoptic Program, which is operated by the Association of Universities for Research in Astronomy, under a cooperative agreement with the National Science Foundation and with additional financial support from the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration, and the United States Air Force. The GONG network of instruments is hosted by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory. D. de Pablos acknowledges use of NASA/GSFC's Space Physics Data Facility's CDAWeb service to access data ([cdaweb.sci.gsfc.nasa.gov](http://cdaweb.sci.gsfc.nasa.gov)). The authors are grateful to the ACE, GONG, and AIA instrument teams for producing and making the data used in this study publicly available. D. de Pablos thanks the Science Technology and Facilities Council (STFC) for support via funding given in his PhD studentship. D.M. Long is grateful to the STFC for the award of an Ernest Rutherford Fellowship (ST/R003246/1). C.J. Owen and G. Nicolau are grateful to the STFC for support via a consolidated grant to UCL/MSSL (ST/S000240/1). G. Valori acknowledges the support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135 and of the STFC grant number ST/T000317/1. This research made use of Astropy, a community-developed Python package for general astronomy (Astropy Collaboration et al., 2018), HelioPy, a Python package for space plasma physics data downloading and processing (Stansby et al., 2020c), and of Sunpy, a community-developed Python package for heliophysics (The SunPy Community et al., 2020).

## Declarations

**Disclosure of Potential Conflicts of Interest** The authors declare that they have no conflicts of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abbo, L., Ofman, L., Antiochos, S.K., Hansteen, V.H., Harra, L., Ko, Y.-K., Lapenta, G., Li, B., Riley, P., Strachan, L., von Steiger, R., Wang, Y.-M.: 2016, Slow solar wind: observations and modeling. *Space Sci. Rev.* **201**, 55. DOI.
- Badman, S.T., Bale, S.D., Martínez Oliveros, J.C., Panasenco, O., Velli, M., Stansby, D., Buitrago-Casas, J.C., Réville, V., Bonnell, J.W., Case, A.W., Dudok de Wit, T., Goetz, K., Harvey, P.R., Kasper, J.C., Korreck, K.E., Larson, D.E., Livi, R., MacDowall, R.J., Malaspina, D.M., Pulupa, M., Stevens, M.L., Whittlesley, P.L.: 2020, Magnetic connectivity of the ecliptic plane within 0.5 AU: potential field source surface modeling of the first Parker Solar Probe encounter. *Astrophys. J. Suppl.* **246**, 23. DOI.
- Bochsler, P.: 2007, Minor ions in the solar wind. *Astron. Astrophys. Rev.* **14**, 1. DOI.
- Cranmer, S.R.: 2019, Solar-wind origin. In: *Oxford Research Encyclopedia of Physics*. DOI.
- Crooker, N.U., Owens, M.J.: 2012, Interchange reconnection: remote sensing of solar signature and role in heliospheric magnetic flux budget. *Space Sci. Rev.* **172**, 201. DOI.
- de Rosa, M.: 2010, PFSS extrapolation software: FAQ and USERS' GUIDE. [www.lmsal.com/~derosa/pfsspack/#usersguide](http://www.lmsal.com/~derosa/pfsspack/#usersguide).
- Deng, L.-H., Feng, S.: 2013, Empirical mode decomposition of long-term solar flare activity. In: Lu, W., Cai, G., Liu, W., Xing, W. (eds.) *Proc. 2012 Internat. Conf. Inform. Tech. Software Eng.*, Springer, Berlin, 263. 978-3-642-34522-7.
- Edmondson, J.K., Antiochos, S.K., DeVore, C.R., Lynch, B.J., Zurbuchen, T.H.: 2010, Interchange reconnection and coronal hole dynamics. *Astrophys. J.* **714**, 517. DOI.
- Fazakerley, A.N., Harra, L.K., van Driel-Gesztelyi, L.: 2016, An investigation of the sources of Earth-directed solar wind during Carrington rotation 2053. *Astrophys. J.* **823**, 145. DOI.
- Freedman, D., Pisani, R., Purves, R.: 2007, *Statistics (International Student Edition)*, 4th edn. WW Norton, New York.



- Gao, P.-X., Hongfei, L., Zhu, W.W.: 2011, Periodicity of flare index revisited using the Hilbert–Huang transform method. *New Astron.* **16**, 147. DOI.
- Harten, R., Clark, K.: 1995, The design features of the GGS wind and polar spacecraft. *Space Sci. Rev.* **71**, 23.
- Harvey, J.W., Hill, F., Hubbard, R.P., Kennedy, J.R., Leibacher, J.W., Pintar, J.A., Gilman, P.A., Noyes, R.W., Title, A.M., Toomre, J., Ulrich, R.K., Bhatnagar, A., Kennewell, J.A., Marquette, W., Patrón, J., Saá, O., Yasukawa, E.: 1996, The global oscillation network group (GONG) project. *Science* **272**, 1284. DOI.
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Sniin, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: 1998, The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc., Math. Phys. Eng. Sci.* **454**, 903. DOI.
- Kasper, J.C., Bale, S.D., Belcher, J.W., Berthomier, M., Case, A.W., Chandran, B.D.G., Curtis, D.W., Gallagher, D., Gary, S.P., Golub, L., Halekas, J.S., Ho, G.C., Horbury, T.S., Hu, Q., Huang, J., Klein, K.G., Korreck, K.E., Larson, D.E., Livi, R., Maruca, B., Lavraud, B., Louarn, P., Maksimovic, M., Martinovic, M., McGinnis, D., Pogorelov, N.V., Richardson, J.D., Skoug, R.M., Steinberg, J.T., Stevens, M.L., Szabo, A., Velli, M., Whittlesey, P.L., Wright, K.H., Zank, G.P., MacDowall, R.J., McComas, D.J., McNutt, R.L., Pulupa, M., Raouafi, N.E., Schwadron, N.A.: 2019, Alfvénic velocity spikes and rotational flows in the near-Sun solar wind. *Nature* **576**, 228. DOI.
- Kolotkov, D.Y., Smirnova, V.V., Strelakova, P.V., Riekhokainen, A., Nakariakov, V.M.: 2017, Long-period quasi-periodic oscillations of a small-scale magnetic structure on the Sun. *Astron. Astrophys.* **598**, 2. DOI.
- Krieger, A.S., Timothy, A.F., Roelof, E.C.: 1973, A coronal hole and its identification as the source of a high velocity solar wind stream. *Solar Phys.* **29**, 505. DOI.
- Kumar, P., Karpen, J.T., Antiochos, S.K., Wyper, P.F., DeVore, C.R., DeForest, C.E.: 2019, Multiwavelength study of equatorial coronal-hole jets. *Astrophys. J.* **873**, 93. DOI.
- Lemen, J.R., Title, A.M., Akin, D.J., et al.: 2012, The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 17. DOI.
- Lepping, R.P., Acuña, M.H., Burlaga, L.F., Farrell, W.M., Slavin, J.A., Schatten, K.H., Mariani, F., Ness, N.F., Neubauer, F.M., Whang, Y.C., Byrnes, J.B., Kennon, R.S., Panetta, P.V., Scheifele, J., Worley, E.M.: 1995, The WIND magnetic field investigation. *Space Sci. Rev.* **71**, 207. DOI.
- Lin, R., Anderson, K., Ashford, S., Carlson, C., Curtis, D., Ergun, R., Larson, D., McFadden, J., McCarthy, M., Parks, G., et al.: 1995, A three-dimensional plasma and energetic particle investigation for the wind spacecraft. *Space Sci. Rev.* **71**, 125.
- Long, D.M., Valori, G., Pérez-Suárez, D., Morton, R.J., Vásquez, A.M.: 2017, Measuring the magnetic field of a trans-equatorial loop system using coronal seismology. *Astron. Astrophys.* **603**, 1. DOI.
- Macneil, A.R.: 2019, Solar wind particle populations at 1 AU: examining their origins in advance of the Solar Orbiter mission. PhD thesis, University College London. [discovery.ucl.ac.uk/id/eprint/10066843](https://discovery.ucl.ac.uk/id/eprint/10066843).
- Madjarska, M.S.: 2019, Coronal bright points. *Liv. Rev. Solar Phys.* **16**, 2. DOI.
- Marion, J.B.: 1980, *Physics and the Physical Universe*. ISBN 9780471034308. [books.google.co.uk/books?id=QKjvAAAAMAAJ](https://books.google.co.uk/books?id=QKjvAAAAMAAJ).
- Marsch, E., Mühlhäuser, K.-H., Rosenbauer, H., Schwenn, R., Neubauer, F.M.: 1982, Solar wind helium ions: observations of the Helios solar probes between 0.3 and 1 AU. *J. Geophys. Res.* **87**, 35. DOI.
- McComas, D.J., Ebert, R.W., Elliott, H.A., Goldstein, B.E., Gosling, J.T., Schwadron, N.A., Skoug, R.M.: 2008, Weaker solar wind from the polar coronal holes and the whole Sun. *Geophys. Res. Lett.* **35**, 1. DOI.
- Mikić, Z., Linker, J.A., Schnack, D.D., Lionello, R., Tarditi, A.: 1999, Magnetohydrodynamic modeling of the global solar corona. *Phys. Plasmas* **6**, 2217. DOI.
- Müller, D., St. Cyr, O.C., Zouganelis, I., Gilbert, H.R., Marsden, R., Nieves-Chinchilla, T., Antonucci, E., Auchère, F., Berghmans, D., Horbury, T.S., Howard, R.A., Krucker, S., Maksimovic, M., Owen, C.J., Rochus, P., Rodríguez-Pacheco, J., Romoli, M., Solanki, S.K., Bruno, R., Carlsson, M., Fludra, A., Harra, L., Hassler, D.M., Livi, S., Louarn, P., Peter, H., Schühle, U., Teriaca, L., del Toro Iniesta, J.C., Wimmer-Schweingruber, R.F., Marsch, E., Velli, M., De Groof, A., Walsh, A., Williams, D.: 2020, The solar orbiter mission - science overview. *Astron. Astrophys.* **642**, A1. DOI.
- Neugebauer, M., Forsyth, R.J., Galvin, A.B., Harvey, K.L., Hoeksema, J.T., Lazarus, A.J., Lepping, R.P., Linker, J.A., Mikić, Z., Steinberg, J.T., von Steiger, R., Wang, Y.-M., Wimmer-Schweingruber, R.F.: 1998, Spatial structure of the solar wind and comparisons with solar data and models. *J. Geophys. Res.* **103**, 14587. DOI.
- Nolte, J.T., Roelof, E.C.: 1973, Large-scale structure of the interplanetary medium. *Solar Phys.* **33**, 241. DOI.
- Ohmi, T., Kojima, M., Tokumaru, M., Fujiki, K., Hakamada, K.: 2004, Origin of the slow solar wind. *Adv. Space Res.* **33**, 689. DOI.
- Owens, M., Lockwood, M., Macneil, A., Stansby, D.: 2020, Signatures of coronal loop opening via interchange reconnection in the slow solar wind at 1 AU. *Solar Phys.* **295**, 37. DOI.

- Panasenco, O., Velli, M., D'Amicis, R., Shi, C., Réville, V., Bale, S.D., Badnan, S.T., Kasper, J., Korreck, K., Bonnell, J.W., de Wit, T.D., Goetz, K., Harvey, P.R., MacDowall, R.J., Malaspina, D.M., Pulupa, M., Case, A.W., Larson, D., Livi, R., Stevens, M., Whittlesey, P.: 2020, Exploring solar wind origins and connecting plasma flows from the Parker solar probe to 1 AU: nonspherical source surface and Alfvénic fluctuations. *Astrophys. J. Suppl.* **246**, 54. DOI.
- Parker, E.N.: 1958, Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.* **128**, 664. DOI.
- Pesnell, W.D., Thompson, B.J., Chamberlin, P.C.: 2012, The Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 3. DOI.
- Pinto, R.F., Rouillard, A.P.: 2017, A multiple flux-tube solar wind model. *Astrophys. J.* **838**, 89. DOI.
- Pottasch, S.R.: 1964, On the chemical composition of the solar corona. *Mon. Not. Roy. Astron. Soc.* **128**, 73. DOI.
- Price-Whelan, A.M., Sipőcz, B.M., Günther, H.M., Lim, P.L., Crawford, S.M., Conseil, S., Shupe, D.L., Craig, M.W., Dencheva, N., Ginsburg, A., VanderPlas, J.T., Bradley, L.D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T.L., Cruz, K.L., Robitaille, T.P., Tollerud, E.J., Ardelean, C., Babej, T., Bach, Y.P., Bachetti, M., Bakanov, A.V., Bamford, S.P., Barentsen, G., Barmby, P., Baumbach, A., Berry, K.L., Biscani, F., Boquien, M., Bostroem, K.A., Bouma, L.G., Brammer, G.B., Bray, E.M., Breytjenbach, H., Buddelmeijer, H., Burke, D.J., Calderone, G., Cano Rodríguez, J.L., Cara, M., Cardoso, J.V.M., Cheedella, S., Copin, Y., Corrales, L., Crichton, D., D'Avella, D., Deil, C., Depagne, É., Dietrich, J.P., Donath, A., Droettboom, M., Earl, N., Erben, T., Fabbro, S., Ferreira, L.A., Finethy, T., Fox, R.T., Garrison, L.H., Gibbons, S.L.J., Goldstein, D.A., Gommers, R., Greco, J.P., Greenfield, P., Groener, A.M., Grollier, F., Hagen, A., Hirst, P., Homeier, D., Horton, A.J., Hosseinzadeh, G., Hu, L., Hunkeler, J.S., Ivezić, Ž., Jain, A., Jenness, T., Kanarek, G., Kendrew, S., Kern, N.S., Kerzendorf, W.E., Khvalko, A., King, J., Kirkby, D., Kulkarni, A.M., Kumar, A., Lee, A., Lenz, D., Littlefair, S.P., Ma, Z., Macleod, D.M., Mastropietro, M., McCully, C., Montagnac, S., Morris, B.M., Mueller, M., Mumford, S.J., Muna, D., Murphy, N.A., Nelson, S., Nguyen, G.H., Ninan, J.P., Nöthe, M., Ogaz, S., Oh, S., Parejko, J.K., Parley, N., Pascual, S., Patil, R., Patil, A.A., Plunkett, A.L., Prochaska, J.X., Rastogi, T., Reddy Janga, V., Sabater, J., Sakurikar, P., Seifert, M., Sherbot, L.E., Sherwood-Taylor, H., Shih, A.Y., Sick, J., Silbiger, M.T., Singanamalla, S., Singer, L.P., Sladen, P.H., Sooley, K.A., Sornarajah, S., Streicher, O., Teuben, P., Thomas, S.W., Tremblay, G.R., Turner, J.E.H., Terrón, V., van Kerkwijk, M.H., de la Vega, A., Watkins, L.L., Weaver, B.A., Whitmore, J.B., Woillez, J., Zabalza, V., Astropy Contributors (Astropy Collaboration): 2018, The astropy project: building an open-science project and status of the v2.0 core package. *Astron. J.* **156**, 123. DOI.
- Raouafi, N.E., Patsourakos, S., Pariat, E., Young, P.R., Sterling, A.C., Savcheva, A., Shimojo, M., Moreno-Insertis, F., DeVore, C.R., Archontis, V., Török, T., Mason, H., Curdt, W., Meyer, K., Dalmasse, K., Matsui, Y.: 2016, Solar coronal jets: observations, theory, and modeling. *Space Sci. Rev.* **201**, 1. DOI.
- Riley, P., Linker, J.A., Mikić, Z., Lionello, R., Ledvina, S.A., Luhmann, J.G.: 2006, A comparison between global solar magnetohydrodynamic and potential field source surface model results. *Astrophys. J.* **653**, 1510. DOI.
- Rochus, P., Auchère, F., Berghmans, D., Harra, L., Schmutz, W., Schühle, U., Addison, P., Appourchaux, T., Aznar Cuadrado, R., Baker, D., Barbay, J., Bates, D., BenMoussa, A., Bergmann, M., Beurthe, C., Borgo, B., Bonte, K., Bouzit, M., Bradley, L., Büchel, V., Buchlin, E., Büchner, J., Cabé, F., Cadiergues, L., Chaigneau, M., Chares, B., Choque Cortez, C., Coker, P., Condamine, M., Coumar, S., Curdt, W., Cutler, J., Davies, D., Davison, G., Defise, J.-M., Del Zanna, G., Delmotte, F., Delouille, V., Dolla, L., Dumesnil, C., Dürig, F., Enge, R., François, S., Fourmond, J.-J., Gillis, J.-M., Giordanengo, B., Gissot, S., Green, L.M., Guerreiro, N., Guilbaud, A., Gyo, M., Haberreiter, M., Hafiz, A., Hailey, M., Halain, J.-P., Hansotte, J., Hecquet, C., Heerlein, K., Hellin, M.-L., Hemsley, S., Hermans, A., Hervier, V., Hochedez, J.-F., Houbrechts, Y., Ihsan, K., Jacques, L., Jérôme, A., Jones, J., Kahle, M., Kennedy, T., Klaproth, M., Kolleck, M., Koller, S., Kotsialos, E., Kraaikamp, E., Langer, P., Lawrenson, A., Le Clech, J.-C., Lenaerts, C., Liebecq, S., Linder, D., Long, D.M., Mampaey, B., Markiewicz-Innes, D., Marquet, B., Marsch, E., Matthews, S., Mazy, E., Mazzoli, A., Meining, S., Meltchakov, E., Mercier, R., Meyer, S., Moncke, M., Monfort, F., Morinaud, G., Moron, F., Mountney, L., Müller, R., Nicula, B., Parenti, S., Peter, H., Pffiffer, D., Philippon, A., Phillips, I., Plessier, J.-Y., Pylser, E., Rabecki, F., Ravet-Krill, M.-F., Rebellato, J., Renotte, E., Rodriguez, L., Roose, S., Rosin, J., Rossi, L., Roth, P., Rouesnel, F., Roulliy, M., Rousseau, A., Ruane, K., Scanlan, J., Schlatter, P., Seaton, D.B., Silliman, K., Smit, S., Smith, P.J., Solanki, S.K., Spescha, M., Spencer, A., Stegen, K., Stockman, Y., Swzec, N., Tamiatto, C., Tandy, J., Teriaca, L., Theobald, C., Tychon, I., van Driel-Gesztelyi, L., Verbeek, C., Vial, J.-C., Werner, S., West, M.J., Westwood, D., Wiegmann, T., Willis, G., Winter, B., Zerr, A., Zhang, X., Zhukov, A.N.: 2020, The solar orbiter EUVI instrument: The extreme ultraviolet imager. *Astron. Astrophys.* **642**, A8. DOI. ADS.
- Rouillard, A.P., Kouloumvakos, A., Vourlidas, A., Kasper, J., Bale, S., Raouafi, N.-E., Lavraud, B., Howard, R.A., Stenborg, G., Stevens, M., Poirier, N., Davies, J.A., Hess, P., Higginson, A.K., Lavarra, M., Viall,

- N.M., Korreck, K., Pinto, R.F., Griton, L., Réville, V., Louarn, P., Wu, Y., Dalmasse, K., Génot, V., Case, A.W., Whittlesey, P., Larson, D., Halekas, J.S., Livi, R., Goetz, K., Harvey, P.R., MacDowall, R.J., Malaspina, D., Pulupa, M., Bonnell, J., de Witt, T.D., Penou, E.: 2020, Relating streamer flows to density and magnetic structures at the parker solar probe. *Astrophys. J.* **246**, 37. DOI.
- Sakao, T., Kano, R., Narukage, N., Kotoku, J., Bando, T., DeLuca, E.E., Lundquist, L.L., Tsuneta, S., Harra, L.K., Katsukawa, Y., Kubo, M., Hara, H., Matsuzaki, K., Shimojo, M., Bookbinder, J.A., Golub, L., Korreck, K.E., Su, Y., Shibasaki, K., Shimizu, T., Nakatani, I.: 2007, Continuous plasma outflows from the edge of a solar active region as a possible source of solar wind. *Science* **318**, 1585. DOI.
- Schatten, K.H., Wilcox, J.M., Ness, N.F.: 1969, A model of interplanetary and coronal magnetic fields. *Solar Phys.* **6**, 442. DOI.
- Stangalini, M., Giannattasio, F., Erdélyi, R., Jafarzadeh, S., Consolini, G., Crisculoli, S., Ermolli, I., Guglielmino, S.L., Zuccarello, F.: 2017, Polarized kink waves in magnetic elements: evidence for chromospheric helical waves. *Astrophys. J.* **840**, 19. DOI.
- Stansby, D., Baker, D., Brooks, D.H., Owen, C.J.: 2020a, Directly comparing coronal and solar wind elemental fractionation. *Astron. Astrophys.* **640** A28. DOI.
- Stansby, D., Matteini, L., Horbury, T.S., Perrone, D., D'Amicis, R., Berčić, L.: 2020b, The origin of slow Alfvénic solar wind at solar minimum. *Mon. Not. Royal Astron. Soc.* **492**, 39. DOI.
- Stansby, D., Rai, Y., Argall, M., JeffreyBroll, Haythornthwaite, R., Erwin, N., Shaw, S., Aditya, Saha, R., Ireland, J., Lim, P.L., Badman, S., Mishra, S., Badger, T.G., dupuisIRT, tlm: 2020c, heliopython/heliopy: Heliopy 0.15.0. DOI.
- Sterling, A.C., Moore, R.L.: 2020, Coronal-jet-producing minifilament eruptions as a possible source of parker solar probe switchbacks. *Astrophys. J.* **896**, L18. DOI.
- Tenerani, A., Velli, M., Matteini, L., Réville, V., Shi, C., Bale, S.D., Kasper, J.C., Bonnell, J.W., Case, A.W., de Wit, T.D., Goetz, K., Harvey, P.R., Klein, K.G., Korreck, K., Larson, D., Livi, R., MacDowall, R.J., Malaspina, D.M., Pulupa, M., Stevens, M., Whittlesey, P.: 2020, Magnetic field kinks and folds in the solar wind. *Astrophys. J. Suppl.* **246**, 32. DOI.
- The SunPy Community, Barnes, W.T., Bobra, M.G., Christe, S.D., Freij, N., Hayes, L.A., Ireland, J., Mumford, S., Perez-Suarez, D., Ryan, D.F., Shih, A.Y., Chanda, P., Glogowski, K., Hewett, R., Hughitt, V.K., Hill, A., Hiware, K., Inglis, A., Kirk, M.S.F., Konge, S., Mason, J.P., Maloney, S.A., Murray, S.A., Panda, A., Park, J., Pereira, T.M.D., Reardon, K., Savage, S., Sipőcz, B.M., Stansby, D., Jain, Y., Taylor, G., Yadav, T., Rajul, Dang, T.K.: 2020, The SunPy project: open source development and status of the version 1.0 core package. *Astrophys. J.* **890**, 68 DOI.
- Velli, M., Grappin, R.: 1993, Properties of the solar wind. *Adv. Space Res.* **13**, 49. DOI.
- Wang, Y.-M.: 2016, The oxygen charge-state ratio as an indicator of footpoint field strength in the source regions of the solar wind. *Astrophys. J.* **833**, 121. DOI.