Dynamic Time Warping as a Means of Assessing Solar Wind Time Series

E. Samara, B. Laperre, R. Kieokaew, M. Temmer, C. Verbeke, L. Rodriguez, J. Magdalenić, S. Poedts

The Astrophysical Journal, 927:187 (19pp), 2022 March 10 https://doi.org/10.3847/1538-4357/ac4af6

Dynamic Time Warping as a Means of Assessing Solar Wind Time Series

E. Samara[1,2], B. Laperre[2], R. Kieokaew[3], M. Temmer[4], C. Verbeke[1,2], L. Rodriguez[1], J. Magdalenić[1,2], and S. Poedts[2,5]
[1] SIDC, Royal Observatory of Belgium, Brussels, Belgium; evangelia.samara@kuleuven.be
[2] CmPA/Dept. of Mathematics, KU Leuven, Leuven, Belgium
[3] IRAP, University of Toulouse, Toulouse, France
[4] University of Graz, Institute of Physics, Graz, Austria
[5] Institute of Physics, University of Maria Curie-Skłodowska, Lublin, Poland
Received 2021 September 9; revised 2021 December 22; accepted 2022 January 12; published 2022 March 15

Abstract

Over the last decades, international attempts have been made to develop realistic space weather prediction tools aiming to forecast the conditions on the Sun and in the interplanetary environment. These efforts have led to the development of appropriate metrics to assess the performance of those tools. Metrics are necessary to validate models, to compare different models, and to monitor the improvements to a certain model over time. In this work, we introduce dynamic time warping (DTW) as an alternative way of evaluating the performance of models and, in particular, of quantifying the differences between observed and modeled solar wind time series. We present the advantages and drawbacks of this method, as well as its application to Wind observations and EUHFORIA predictions at Earth. We show that DTW can warp sequences in time, aiming to align them with the minimum cost by using dynamic programming. It can be applied for the evaluation of modeled solar wind time series in two ways. The first calculates the *sequence similarity factor*, a number that provides a quantification of how good the forecast is compared to an ideal and a nonideal prediction scenario. The second way quantifies the time and amplitude differences between the points that are best matched between the two sequences. As a result, DTW can serve as a hybrid metric between continuous measurements (e.g., the correlation coefficient) and point-by-point comparisons. It is a promising technique for the assessment of solar wind profiles, providing at once the most complete evaluation portrait of a model.

Unified Astronomy Thesaurus concepts: Magnetohydrodynamics (1964); Time series analysis (1916); Solar wind (1534); Space weather (2037)

1. Introduction

Improving the accuracy of space weather forecasting at Earth is directly linked to a better understanding of the capabilities provided by the space weather prediction models currently available to the operational and scientific community. A large number of such models have been developed over the past decades, aiming to reconstruct the solar corona and the heliospheric environment (see, e.g., MacNeice et al. 2018 for an overview of the available models and their capabilities). Some models are fully empirical, such as the PDF model (Bussy-Virat & Ridley 2014), the PROJECTZED model (Riley et al. 2017), and the Analog Ensemble model (Owens et al. 2017). Others require extensive empirical tuning, such as, e.g., the Wang–Sheeley–Arge model (WSA; Arge & Pizzo 2000; Arge et al. 2003, 2004), which is one of the most widely used operational coronal models, or the Empirical Solar Wind Forecast model (ESWF; Reiss et al. 2016), which relates the areas of the coronal holes, as observed in EUV, to the solar wind speed prediction at the Lagrangian point 1 (L1), based on Vršnak et al. (2007). Some models combine empirical approximations with physics (see, e.g., the MULTI-VP model by Pinto & Rouillard 2017), using the PFSS model for the reconstruction of the lower corona and relying on the solution of magnetohydrodynamic (MHD) equations for the calculation of the solar wind properties in the upper corona. Another large category of models are the physics-based models that consist entirely of MHD codes. Such models can either reconstruct the global solar corona alone, such as the Magnetohydrodynamics Around a Sphere (MAS) model (Mikić et al. 1999; Lionello et al. 2003), or the heliospheric domain alone, such as the heliospheric parts of ENLIL (Odstrčil & Pizzo 1999) and EUHFORIA (Pomoell & Poedts 2018). Some models can reconstruct both domains, such as the MAS model for corona and heliosphere (Riley et al. 2011) or the AWSoM model (Meng et al. 2015). Last but not least are models based on tomographic techniques. For example, HelTomo (Jackson et al. 1998; Jackson & Hick 2002; Jackson et al. 2020) is a tool that reconstructs the solar wind, including coronal mass ejections (CMEs), by employing interplanetary scintillation of astronomical radio sources that are viewed through the ambient solar wind plasma.

Space weather prediction models are subject to continuous changes and improvements. Therefore, it is of utmost importance for space weather forecasting applications to record and quantify these changes over time. The differences in performance among different models should also be recorded, e.g., MAS versus ENLIL versus EUHFORIA, or among combinations of different models. For example, an overview of the coupling between the WSA coronal model and the time-dependent MAS solar wind model is given in Linker et al. (2016), and an overview of the coupling between the MULTI-VP coronal model and EUHFORIA-heliosphere is given in Samara et al. (2021). Metrics are the only way of quantitatively

understanding which setup provides the best results, based on the user's needs and goals, every time.

A number of metrics proposed through the years are being used by the scientific community. Owens (2018) presents an overview by categorizing such metrics in two large groups: point-by-point metrics and time window metrics. The former category includes the so-called error functions, such as the mean-square error (MSE), the mean absolute error, the rms error, or general skill scores constructed by the aforementioned. All of them are widely used to compare the amplitudes between observed and predicted time series in a point-by-point manner. Another subcategory of point-by-point metrics is binary metrics. Binary metrics show whether large-scale structures of the solar wind (e.g., high-speed streams, abbreviated as "HSSs") have arrived at the point of interest or not. This technique is based on hit/miss statistics between the observed and predicted time series once specific requirements are fulfilled (e.g., HSS arrives within a specific time window and has an an amplitude and gradient value above a specific threshold). For example, by setting a solar wind velocity threshold at 500 km s^{-1} and by taking into account a specific interval in time (e.g., 2 days), this binary technique indicates whether a solar wind feature was captured by both the observed and predicted time series (hit), whether it was observed but not predicted (miss), whether it was predicted but not observed (false alarm), or, lastly, whether neither observations nor the forecast indicated the arrival of an event (true negative; see, e.g., Reiss et al. 2016; Hinterreiter et al. 2019).

The mentioned metrics are necessary for quantifying the differences between observed and predicted time series. They are not sufficient, though, to present a complete picture of the comparison between observations and predictions. The most significant deficiency is found in their weakness of quantifying time uncertainties. A solution to this problem can be given by using time window metrics. The simplest time window approach is to conduct a case-by-case analysis in which the time differences between the arrival/ending times of observed and modeled events can be recorded. Nevertheless, this is a time-consuming procedure that requires a priori definitions of each event as discussed in Owens (2018). The same author introduces an alternative time window approach, the so-called scale-selective approach, based on which agreements between observations and predictions are taken into account at a range of time scales. As the time scales become increasingly coarse, false alarms and missed events are canceled out. This technique allows an assessment of the time scales at which the forecast provides a specific accuracy level.

In this paper, we introduce the dynamic time warping (DTW) technique as an additional method for assessing solar wind time series. It is a powerful tool that combines the qualities of both point-by-point and time window metrics and, as a result, is very useful for providing a more complete assessment of the relation between predictions and observations. DTW finds matching solar wind signatures (even weak ones of slow speed, or HSSs) in order to quantify their underestimated or overestimated arrival times, their amplitudes, and their durations. Even though this technique was introduced many decades ago, and has been used extensively in other scientific fields, we will show how it can be applied to the evaluation of solar wind time series, for the better understanding of modeled solar wind profiles (long-term and short-term) and their comparisons with measurements. DTW can also be applied to the evaluation of time series during CME arrivals at specific points of interest. Nevertheless, the reasons that led us to focus on solar wind time series are summarized as follows:

1. The solar wind forecast at Earth (or at any other point in the heliosphere) usually deals with long and variable time series (in the order of days, weeks, or even months, depending on the goals). For the correct prediction and assessment of both the fast and the slow solar wind, we need to evaluate the whole range of the available data set. This procedure is much more complicated than assessing CME signatures for which the arrival of the shock/magnetic cloud is well defined during a limited time interval (in the order of hours) and, thus, allows the easy quantification of both the amplitude and the time delay.
2. When we focus on the assessment of fast streams in the solar wind, the majority of the modeled HSSs arrive later or earlier than observed (see Hinterreiter et al. 2019). As a result, there is always a time difference between the observed and predicted large-scale variations in solar wind time series. DTW is an ideal technique for evaluating these variations, since it aligns time series by warping them in time.
3. In some cases, it is not clear how and if the predicted data should be matched with the observed data (see the example in Figure 1(a)). Hence, a technique that quantifies the overall performance of two sequences is required, regardless of the identified structures.

Figure 1 shows a comparison of the observed (by the Wind satellite; Ogilvie et al. 1995) and the predicted (by EUHFORIA) solar wind time series for Carrington rotations (CRs) 2197 and 2198. It covers the time interval from 2017 November 6 to 2017 December 30. Based on the Richardson-Cane list (Cane & Richardson 2003; Richardson & Cane 2010), only one CME was recorded influencing Earth during that period. This was on 2017 December 25, so we will not comment on the time interval after this date, to avoid any uncertainties. Between 2017 November 6 and 2017 December 24, seven HSSs were identified, based on the criteria proposed by Jian et al. (2006). We cross-validated the associated coronal holes on the Sun by checking the Atmospheric Imaging Assembly (Lemen et al. 2012) images from the Solar Dynamics Observatory (Pesnell et al. 2012), as well as NOAA full-Sun drawings.[6] Figure 1(a) shows an example of an unclear HSS reconstructed case. More specifically, it is uncertain whether the first modeled HSS (in red; see the black arrow) corresponds to the first observed HSS (with a maximum velocity of up to $\approx$700 km s^{-1}), which impacted Earth between 2017 November 6 and 2017 November 12, or to the second, slower HSS observed between 2017 November 15 and 2017 November 18 (with a maximum velocity of up to $\approx$500 km s^{-1}). Therefore, instead of trying to determine whether an HSS arrived at Earth, and which one, it is more practical to determine how similar the time series are overall. In Figure 1(b), we notice three HSSs influencing Earth between 2017 December 3 and 2017 December 24. The first two were modeled arriving later than expected by EUHFORIA. For the third HSS, we are uncertain if it was actually predicted by the model. These examples show that assessing the time series as a whole, instead of trying to

[6] https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-imagery/composites/full-sun-drawings/boulder/

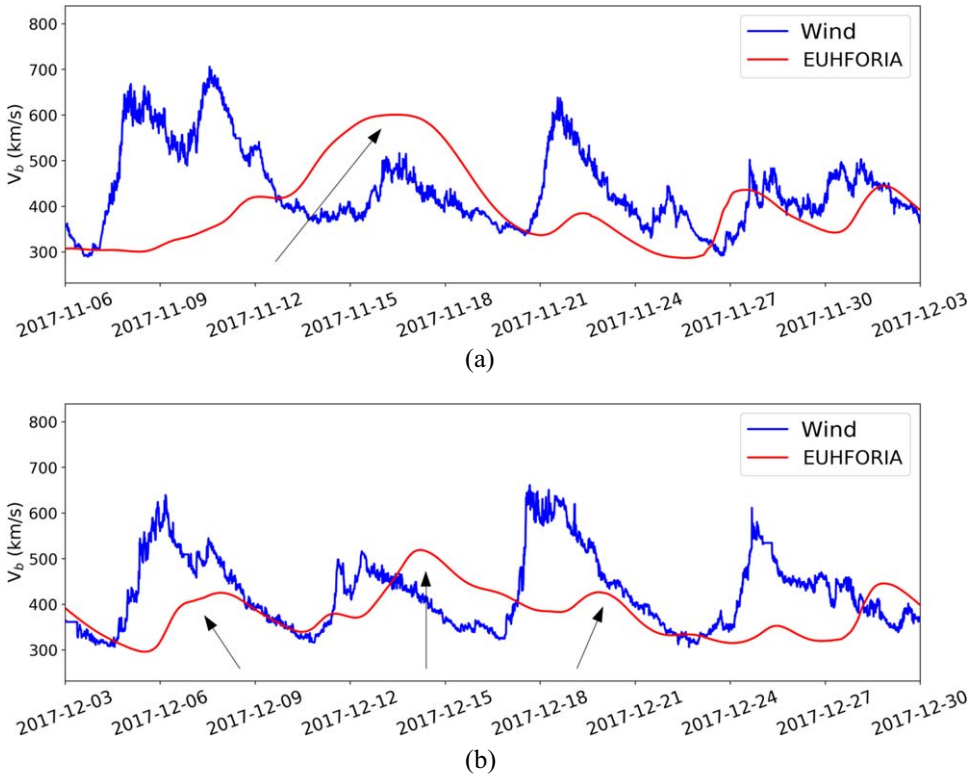


Figure 1. Solar wind speed time series of the observed (blue) and predicted (red) data for CR 2197 (panel (a)) and CR 2198 (panel (b)). The arrow in panel (a) points to the unclear modeled HSS structure, and we are not sure if it corresponds to the observed HSS between 2017 November 6 and 2017 November 12 or to the second, slower HSS observed between 2017 November 15 and 2017 November 18. The first two arrows in panel (b) show the two HSSs that are reproduced by the model but arrive late compared to observations. The third arrow in the same panel corresponds to the uncertain HSS case, and we are not sure if this was predicted by the model.

evaluate uncertain cases on a one-by-one basis, is often the optimal solution.

2. Benefits, Drawbacks, and Restrictions of DTW

DTW is a well-known technique for estimating the similarities between two time series that have similar patterns but differ in time (Keogh & Pazzani 2001; Górecki & Luczak 2013, and references therein). It was initially developed for speech recognition purposes, where specific words are recognized by their audio signal profiles (Itakura 1975; Sakoe & Chiba 1978; Myers et al. 1981; Müller 2007). Over the years, it was met with great interest by other scientific fields, such as meteorology, robotics, medicine, music processing, and manufacturing, and it is widely used in data mining for time series clustering and classification purposes (Keogh & Pazzani 1998, 2001, and references therein). DTW is a distance measure, similar to the Euclidean distance. The main difference from the latter is that DTW can manage time distortions (Zhang et al. 2017) and does not always obey the triangular inequality (Vidal et al. 1988). More specifically, it allows the drifting of the vector components along the time axis when a comparison between two sequences is made. The sequences are eventually nonlinearly warped along the time dimension to match each other (Müller 2007; Górecki & Luczak 2013). A recent study by Laperre et al. (2020) used this technique for evaluating the Dst forecast with machine learning. In this work, for the first time, we will apply the DTW technique as a means of quantifying the differences between observed and modeled time series in solar wind forecasting. For this, we will use DTW in two different ways. The first way is based on the so-called cumulative cost or DTW score, a single number that DTW produces that translates to an estimation of the minimum "effort" that the technique put in to align the observed and predicted time series of the solar wind parameters. After DTW has estimated the best (and less costly) alignment between the data points, we can explicitly quantify the differences in time and amplitude between the two sequences. This is the second way in which the method can be exploited, from which we can derive a number of relevant statistics.

2.1. Properties and Requirements of DTW

DTW determines how similar two time series are by providing a temporal alignment between them, in an optimal way and under certain constraints (Müller 2007). Strictly speaking, DTW is not a metric, as it violates the triangular inequality.[7] Nevertheless, it can be used as one by obeying the following three principles (Müller 2007; Jeong et al. 2011):

1. The first and last point of one sequence should be matched with the first and last point of the other sequence (but it is not necessary for their matches to be unique).
2. The mapping of the elements should be monotonically increasing (it cannot go backwards in time).
3. There should be no data gaps, namely, every point should be matched with at least one other point (the continuity rule).

[7] For the definition of a distance measure as a metric, see https://mathworld.wolfram.com/Metric.html.

The method finds the optimal alignment between two sequences by finding the path through the DTW cost matrix that minimizes the total cumulative cost among all the other possible paths (Keogh & Pazzani 2001; Müller 2007; Ratanamahatana & Keogh 2004). This path is called the warping path, and it characterizes the mapping between the two time series of interest. The DTW cost matrix is filled based on the following equation:

$$D(i, j) = \delta(s_i, q_i) + \min\{D(i-1, j-1), D(i-1, j), D(i, j-1)\}, \quad (1)$$

where $D(i, j)$ is the cumulative DTW cost or distance, and $\delta(s_i, q_i) = |s_i - q_i|$ corresponds to the Euclidean distance between the point s_i from one time series and the point q_i from the other time series. The first element of the array $D(0, 0)$ is equal to $\delta(s_0, q_0)$. After the DTW cost matrix has been filled, the warping path can be efficiently found. For that, dynamic programming is used to evaluate the recurrence that defines the cumulative DTW distance $D(i, j)$ as the Euclidean distance, δ(i,j), found between two elements and the minimum of the cumulative distances of the adjacent elements (Ratanamahatana & Keogh 2004; Górecki & Luczak 2013).

2.2. Drawbacks of DTW and Ways of Eliminating Them

The weak point of DTW is the so-called "pathological alignment problem," namely, the fact that a data point in one time series can be linked to a large subsection of points in the other time series (Keogh & Pazzani 2001; Górecki & Luczak 2013, and references therein). These pathological alignments are called singularities. Many techniques have been proposed to alleviate this problem. Three of the most widely used techniques can be summarized as follows (Keogh & Pazzani 2001):

1. *Windowing* (Berndt & Clifford 1994): the allowable elements of the cumulative matrix are restricted to those that fall into a warping window. In other words, a data point in one time series cannot be matched with all of the data points in the second time series, but can only be matched with the points found in a specific time window. Although this approach constrains the maximum size of a singularity, it does not prevent their occurrence.
2. *Slope weighting* (Kruskall & Liberman 1983; Sakoe & Chiba 1978): if the equation that calculates the accumulated cost is replaced with $D(i, j) = \delta(i, j) + \min\{D(i-1, j-1), XD(i-1, j), XD(i, j-1)\}$, where X is a positive real number, we can constrain the warping by changing the value of X. As X gets larger, the warping path is increasingly biased toward the diagonal.
3. *Step patterns or slope constraints* (Itakura 1975; Myers et al. 1981): we can replace the cumulative cost equation with $D(i, j) = \delta(i, j) + \min\{D(i-1, j-1), D(i-1, j-2), D(i-2, j-1)\}$, which corresponds to the step pattern. Using this equation, the warping path is forced to move one diagonal step for each step parallel to an axis.

Besides these three methods, a number of DTW variants have been proposed with the aim of reducing singularities. For example, the derivative DTW (DDTW; Keogh & Pazzani 2001), the weighted DTW (WDTW; Benedikt et al. 2008), and the value-derivative DTW (VDDTW; Kulbacki & Bak 2002) are some of these techniques. The outputs from these three methods were also tested for the cases investigated in this work, but they yielded ambiguous results. Therefore, for the purposes of this study, we only focus on classic DTW, as it leads to the most transparent output.

It is important to note that even though singularities are considered a drawback of DTW by many authors, we find that for solar wind time series, the best alignment between two sequences cannot be achieved without them. Identical time series have no singularities. As we deal with nonidentical (but similar) time series, we expect singularities to occur as a way of optimally matching the points between them.

3. Data Preprocessing

In order to apply DTW efficiently to the evaluation of modeled solar wind time series, we need to adopt a number of constraints that will best serve our needs. We focus on the evaluation of the solar wind bulk speed, as it is the best-modeled solar wind characteristic in EUHFORIA (v1.0.4; see, e.g., Pomoell & Poedts 2018). The same procedure can be applied to other solar wind signatures, such as, e.g., density, magnetic field, temperature, etc.

3.1. Sensitivity on the Initial and Final Points of the Sequences

We are interested in approximately two years of continuous solar wind data (2017 November–2019 September). The first important task is to separate this interval into smaller periods, for two reasons: first, because it will be easier for future users to compare an upgraded version of EUHFORIA (or any other model) with the current version; and second, because it is faster to evaluate shorter periods, rather than a single, extended one, since DTW has a quadratic ($O(nw)$) complexity (Keogh & Pazzani 2001; Ratanamahatana & Keogh 2004, and references therein). Nevertheless, recent studies have shown that for many applications, DTW complexity is reduced to linear ($O(n)$; Dau et al. 2019), similar to the complexity of the Euclidean distance.

Special attention is needed during the division of the large time interval into smaller periods, since DTW is highly affected by the initial and final points of the sequences. One way to do this is to split the whole range into time segments of CRs. However, this is not always the best solution, as an HSS may be cut by the artificial start and end time of the CR (see Figure 2(a)). For an optimal evaluation with DTW, we require that solar wind features (such as HSSs) are fully covered, preferably with quiet background solar wind times before and after the time period under study (see the comparison between Figures 2(a) and (b)).

We now consider a simple example with random time series. Figure 3 shows how DTW behaves in four different cases. In the first case, we compare two identical time series. In the second case, we shift one of the two sequences one value along the x-axis, keeping the same initial and final points. In the third case, we shift one of the two sequences one value along the y-axis, maintaining the exact same pattern. In the last case, we keep the two time series as in the initial example, but we only shift the first element of the red time series up one value along the y-axis. The DTW cost matrix for each case is shown as a green heat map. The x-axis and y-axis of the heat maps correspond to the index numbers of the elements in each time

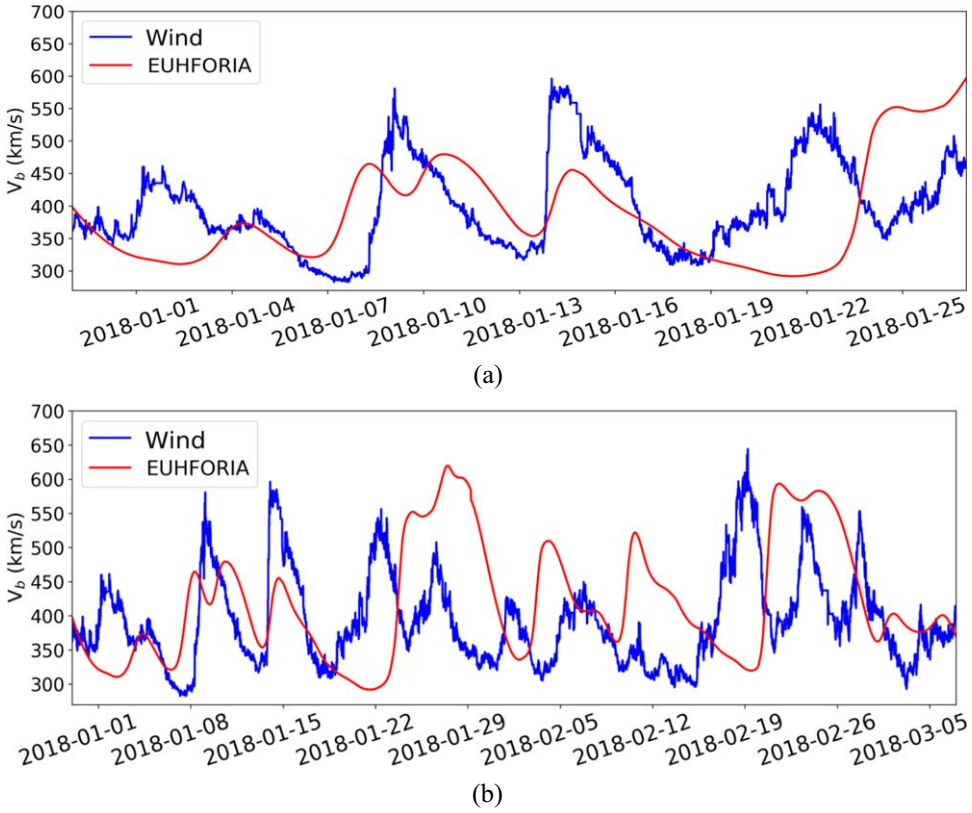


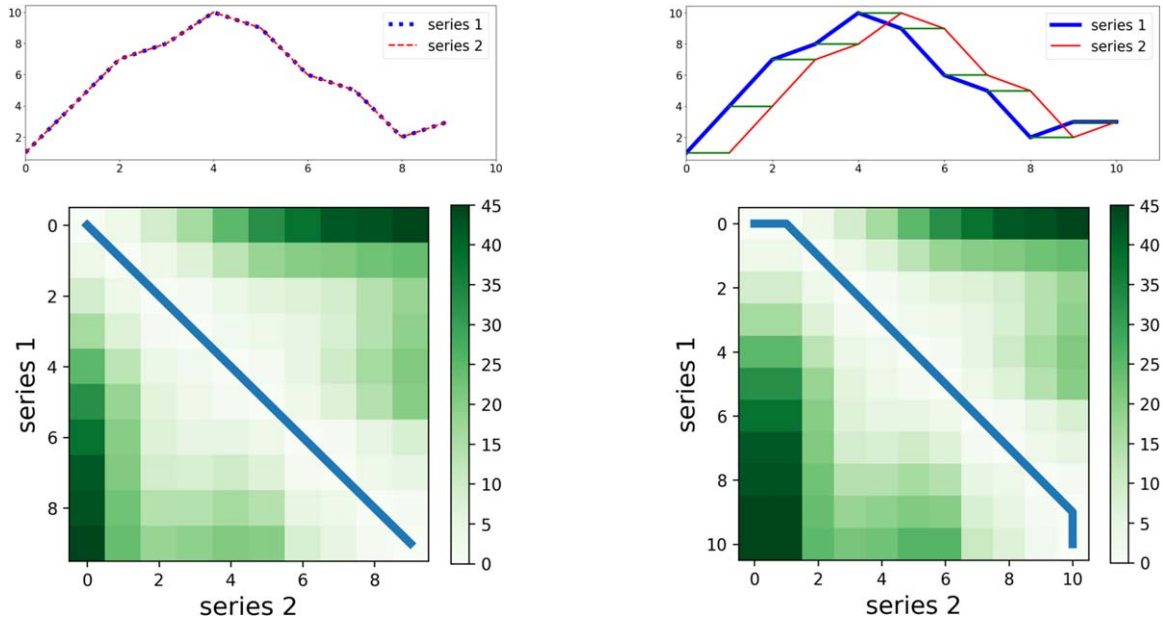
Figure 2. An example of the extension of a not-well-selected time interval, where the influence of an HSS is still ongoing at the end of that interval. (a) CR 2197, where the effect of an HSS is still ongoing. (b) The extension of CR 2197 until a point at which the observed and predicted data are very close to each other during quiet solar wind levels.

series. The darker shades of green correspond to higher estimated costs, based on Equation (1). Each matrix cell corresponds to a link between two points of the two series. The last (bottom right) cell in the DTW cost matrix indicates the DTW score, which is a representation of the minimum effort the technique put in to align the two time series.

In the first case (Figure 3(a)), where identical time series are considered, the DTW score is zero and the warping path is the diagonal of the cost matrix. This means that every element in sequence 1 was only matched with its exact corresponding (and identical) element in sequence 2. In the second example (Figure 3(b)), the DTW score is again equal to zero, but the warping path deviates from the diagonal, due to the time series being shifted along the *x*-axis. This means that the first and last points from one sequence were matched twice with points from the other sequence, thus creating two singularity points. More specifically, the first point from series 1 was matched with two points from series 2 (horizontal shifting in the warping path), while the last point from series 2 was matched with two points from series 1 (vertical shifting in the warping path). The question is how we can distinguish between the first and second cases, when no heat maps/warping paths are provided, e.g., in large data sets, when we want to quickly evaluate the sequences of interest. The answer is given by calculating the sum of the diagonal elements of the cost matrix. For example, when we compare identical time series, the sum of the elements along the diagonal should be zero, opposite to the case where the time series are shifted along the *x*-axis.

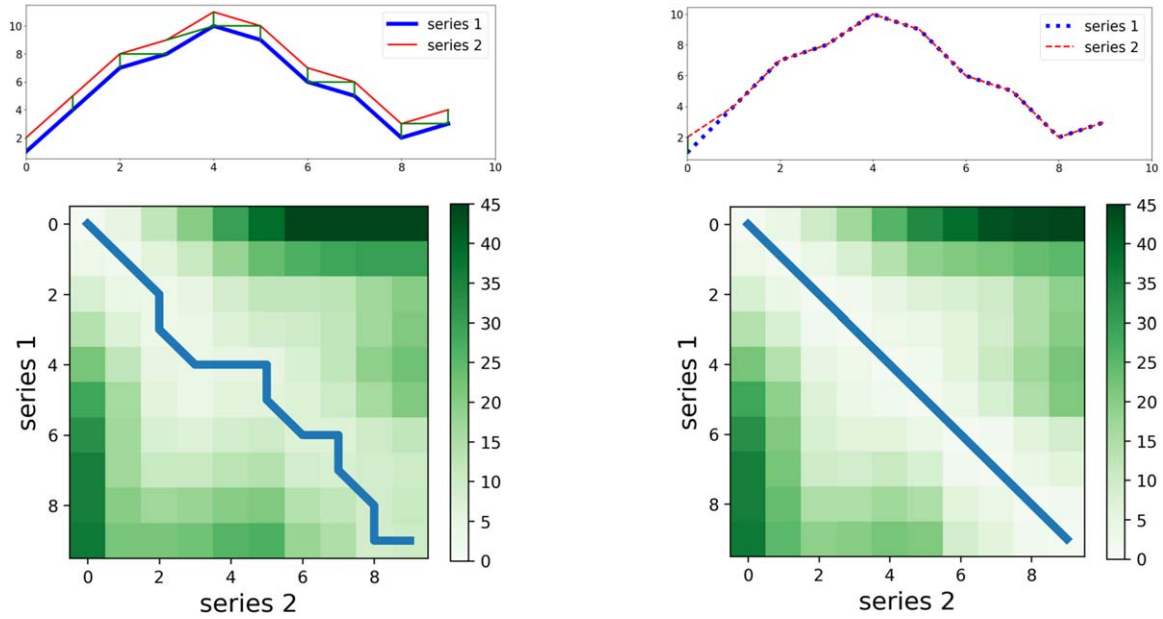
In the third case (Figure 3(c)), in which we have shifted the red sequence along the *y*-axis, we see that the DTW score and the sum of the diagonal are different to zero. Moreover, the warping path deviates a lot from the diagonal. To understand better how DTW behaves during the vertical shifting of a sequence, we compare in Figure 3(d) the same time series as in Figure 3(a), having shifted only the first element of the red sequence along the *y*-axis. The DTW score and the sum of the diagonal in this case are again different to zero, but the warping path is the diagonal itself.

In Figure 4, we follow the same procedure as in Figure 3, but for sequences from our data set. We employ Wind data from the time period 2017 December 3 to 2017 December 30 (CR 2198). In Figure 4(a), we show how DTW behaves when comparing the observed solar wind bulk speed with the same data set, assuming the ideal scenario in which the observed and predicted time series are exactly the same. The *x*-axis in the time series plots corresponds to the index number of the elements that actually describe time, thus the "time elements" label. In Figure 4(b), the observed data set has been shifted forward by one day. The gap that is created between elements 0 and 144 (1 day) is then filled with the value at time 0 of the nonshifted (blue) time series. The same happens for the gap that is created between elements 3744 and 3888 (the last day) of the nonshifted (blue) time series. It is filled with the value at time 3888 of the shifted (red) time series. The DTW score in this case is zero, and the warping path follows the same pattern as described in Figure 3(b). Nevertheless, when comparing the performance between the observed and predicted time series, we rarely have the initial and final points of the sequences



(a) DTW score = 0, Sum(diag) = 0

(b) DTW score = 0, Sum(diag) = 18



(c) DTW score = 10, Sum(diag) = 55

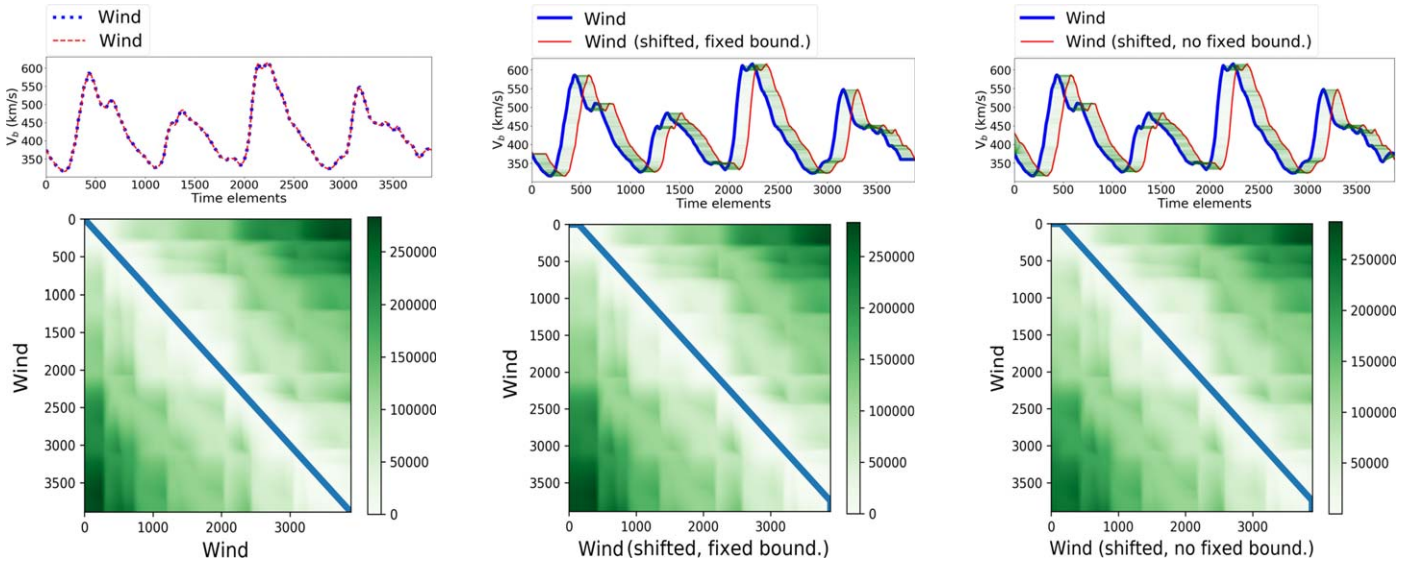
(d) DTW score = 1, Sum(diag) = 10

Figure 3. Alignments, warping paths, and heat maps of random time series. The DTW score and the sum of the cost-matrix diagonal are also calculated for each case. The DTW alignments are shown in green when one series is shifted compared to the other. (a) The two series are identical. (b) Series 2 from panel (a) has been shifted along the *x*-axis with respect to series 1. The initial and final points stay fixed. (c) Series 2 from panel (a) has been shifted along the *y*-axis with respect to series 1. (d) Series 2 is identical to series 1, except for its initial point, which has been shifted one value along the *y*-axis.

matched in the way shown in Figure 4(b). The same sequences as in Figures 4(a) and (b) are shown in Figure 4(c), but now the shifting occurs by filling in the 1 day shifted interval with the data recorded by Wind one day earlier. As a result, the initial and final points of the sequences are not the same anymore. The DTW score in this case is 5242.86. This number does not have a meaning yet; it only reflects the fact that the cost of aligning the two time series is much larger than in the cases of Figures 4(a) and (b), and it arises due to the alignments of the first 144 and final 144 points.

3.2. The Importance of the Applied Smoothing

The second important task is to apply an optimal smoothing to the observed time series. This is a subjective procedure that depends on the goals of the study, the data set, and the user. Usually the real data, as recorded by Wind at L1, contain high-frequency fluctuations, opposite to the modeled time series, which are described by a smooth trend (Figure 1). For a proper comparison between the observations and predictions, it is optimal to smooth the observed time series in a similar way to the modeled ones, as local minima and maxima influence the



(a) DTW score = 0, Sum(diag) = 0

(b) DTW score = 0, Sum(diag) $\approx 18.3 \cdot 10^6$

(c) DTW score = 5242.86, Sum(diag) $\approx 33.7 \cdot 10^6$

Figure 4. The same as Figure 3, but for time series observed by Wind during CR 2198. The DTW alignments are shown in green when one series is shifted compared to the other. (a) The two time series are identical. (b) The red series was shifted by 1 day compared to the blue one. The initial and final points of the two time series overlap. (c) The red series was shifted by 1 day compared to the blue one, while their initial and final points do not overlap. On the contrary, the 1 day data gap has been filled in with data observed by Wind 1 day earlier.

DTW result and generate more singularities. This happens because the method aims to match local fluctuations to parts of the modeled time series where no fluctuations have been detected. Figure 5(a) shows such an example. The alignment of the points is not optimal, and the cumulative DTW cost is larger than for the case where the fast fluctuations have been smoothed. However, we should also be careful not to smooth out important features. An analytic example of how different smoothing influences the results is presented in Figures 5(b), (c), and (d). In this figure, we apply a time-centered smoothing, based on moving averages of 6 hr, 12 hr, and 24 hr, to the Wind time series shown in Figure 4(c). In the event that the smoothing window exceeded the size of the array, we applied data padding, by reflecting about the edges of the first/last elements.

Besides the different smoothing, we keep the same temporal resolution for both time series under comparison. This means that both sequences have the same number of elements, even though DTW can also be applied to unequal time series (see, e.g., Wong & Wong 2003; Zhang et al. 2021, and references therein). In our analysis, we adopted a 10 minute resolution and a 12 hr smoothing window as the ideal setup for Wind and EUHFORIA time series (see Section 5).

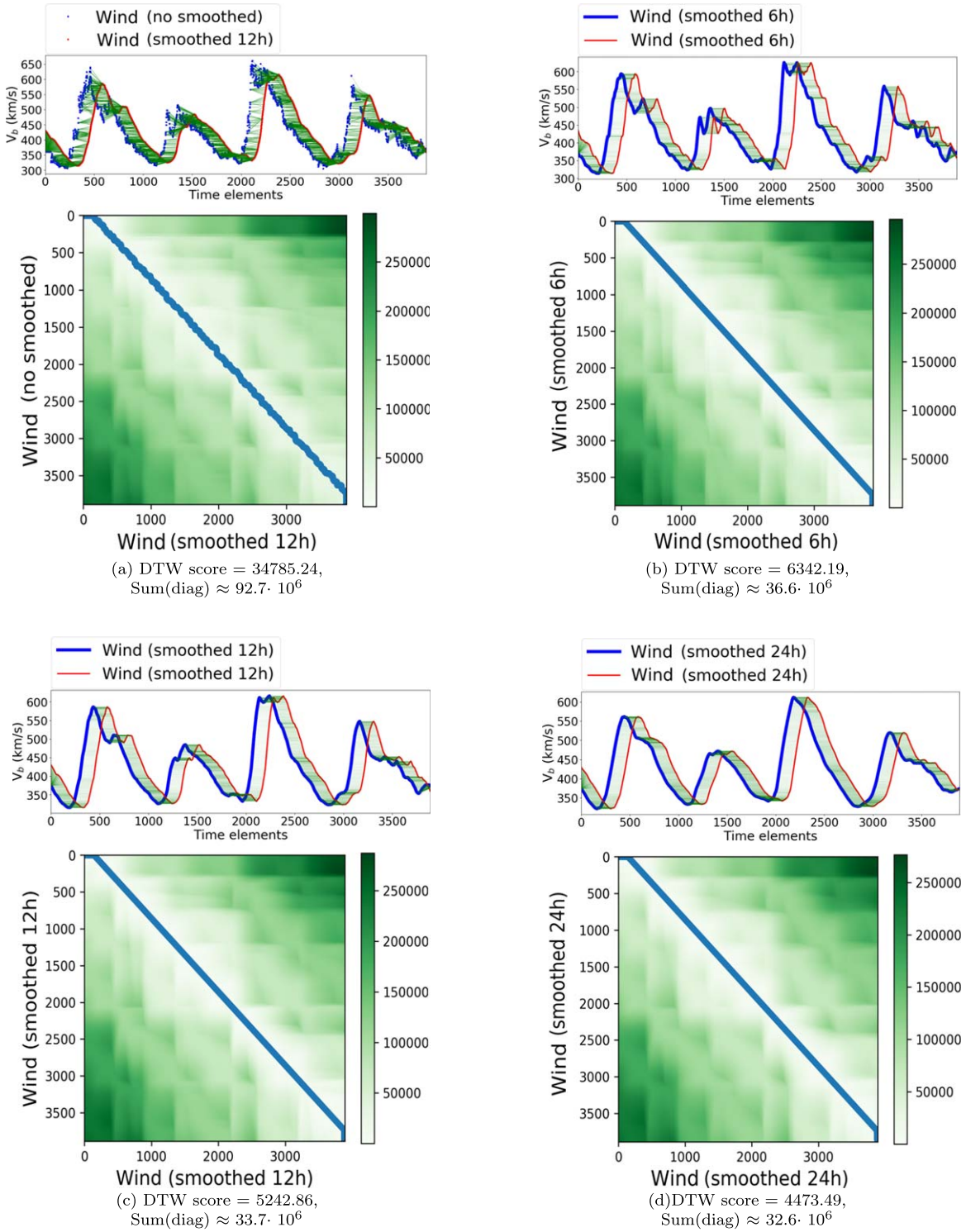
3.3. Window Constraint

After the time series preprocessing, we discuss the employed DTW constraints. The first restriction comes from the fact that the time warping of the sequences needs to be done within a specific time interval. We have to restrict all of the possible matches of the points within a specific time window, otherwise the temporal alignment between them could be indefinite. For solar wind forecasting purposes, it is undesirable for a point at day one to be matched with a point at day five, six, seven, or more if the temporal uncertainty of the predictions lies within a smaller time window. The approximate maximum Δt in EUHFORIA between the arrival of the observed and modeled HSSs for the time interval of interest (2017 November–2019 September) is ≈2 days, so a time window of ±2 days is applied for our purposes. Some other studies have shown that there is a mismatch by at least 1 day and up to 3 days between the arrival of modeled and observed solar wind structures (see, e.g., Owens et al. 2008; MacNeice 2009; Gressl et al. 2014; Jian et al. 2015; Reiss et al. 2016; Temmer et al. 2018). Setting a window like this also reduces the computational time in calculating the cost matrix. No constraints regarding slope weighting or step pattern are further imposed because we do not want to bias the alignment of the points toward one or another direction.

In Figure 6, we show the DTW cost matrix (heat map) and the alignment of the time series presented in Figure 4(c) and Figure 5(c) when we apply a time window constraint of ±2 days. The DTW score and alignment of the points are the same, as is the warping path. The only thing that changes is the extent of the DTW matrix. We notice that it is now only filled in within a zone along the diagonal. This is because of the ±2 day window restriction that we imposed, which does not allow the alignment of points outside that time window. The subset of the matrix that the warping path is allowed to visit is called a "warping window" or "band" (Ratanamahatana & Keogh 2004). In our case, we implement a band similar to the Sakoe–Chiba band (Sakoe & Chiba 1978). The reason that the DTW score and warping path are the same in this particular example is because the maximum time difference between the two sequences is 1 day. Therefore, our results are not affected by the window constraint.

3.4. Influence of CMEs

During the considered time intervals, 15 CMEs were recorded influencing Earth. Five of them occurred very close to each other, between 2019 May 6 and 2019 May 30 (see



(a) DTW score = 34785.24, Sum(diag) $\approx 92.7 \cdot 10^6$

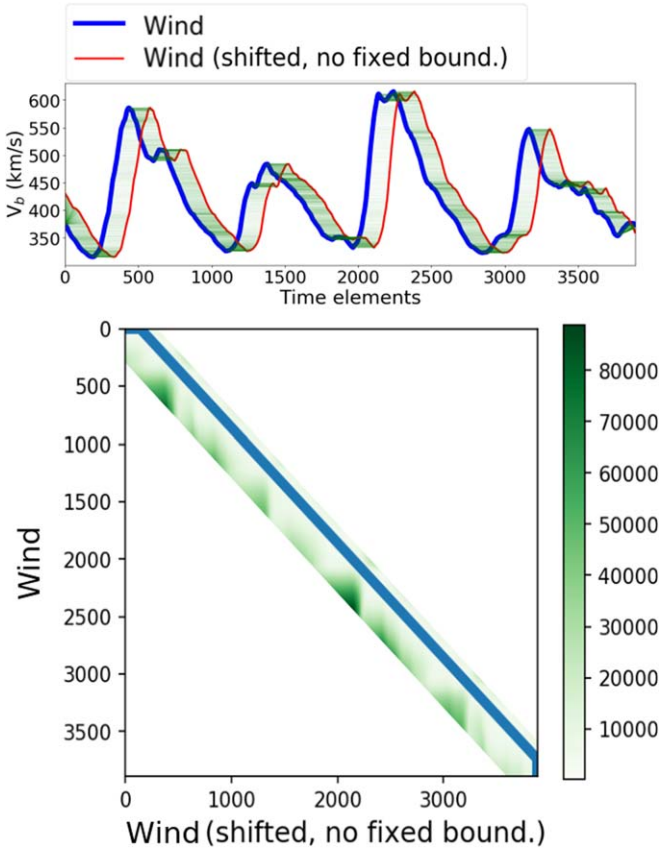
(b) DTW score = 6342.19, Sum(diag) $\approx 36.6 \cdot 10^6$

(c) DTW score = 5242.86, Sum(diag) $\approx 33.7 \cdot 10^6$

(d)DTW score = 4473.49, Sum(diag) $\approx 32.6 \cdot 10^6$

Figure 5. The same as Figure 4, but for time series with different smoothing. (a) A comparison between time series with no smoothing and 12 hr smoothing. In panels (b), (c), and (d), 6 hr, 12 hr, and 24 hr smoothing has been applied to both time series, respectively. The red sequence is always shifted by 1 day compared to the blue one, similar to Figure 4(c).

Cane & Richardson 2003; Richardson & Cane 2010). Even though most of the periods we evaluated did not include any CMEs, for those cases in which CMEs were detected, we ignored them as events and applied DTW normally, as if these structures were not there. As a result, our recommendation for a future user who wants to assess an improved version of EUHFORIA for the same time intervals is to work in the same way that we do in this paper; namely, to ignore the potential CME structures and to apply DTW as if these structures never occurred. Only then will the decrease (or increase) of the DTW score be consistent and comparable to the one calculated based on the current EUHFORIA version, allowing us to track how the change of the model influences the modeling output.



(a) DTW score = 5242.86, Sum(diag) $\approx 33.7 \cdot 10^6$

Figure 6. The same as Figure 4(c) and Figure 5(c), but with a time window constraint of ±2 days.

4. Application of DTW for Assessing the Performance of Solar Wind Time Series

DTW can be applied in two ways for the evaluation of modeled solar wind time series. The first way compares the DTW score of the predicted time series to an ideal and a nonideal (reference) case scenario. The second way quantifies the time and amplitude differences between each point of the time series, as aligned by DTW.

4.1. First Way of Applying DTW: The Sequence Similarity Factor (SSF)

The Wind data for CR 2198 (blue time series in Figure 6) will be considered our observations, while the same data set, shifted by 1 day (red time series in Figure 6), will constitute our predictions. The ideal scenario will be the flawless forecast in which the red time series is identical to the blue one (see Figure 4(a), for which the DTW score = 0 and sum(diag) = 0). If the case we study is not the ideal one, then the DTW score will not be zero and will not have an actual meaning unless it is compared to (a) the ideal scenario and (b) a nonideal (reference) prediction. We define this reference prediction as the mean model of observations, which represents the forecast of the average observed speed for the period of interest. Such a model has no variations in time, and will later prove useful for the direct comparison of the DTW results with traditional metrics (see Section 5).

In Figures 7(a), (b), and (c), we show the application of DTW between observations and (a) the ideal prediction scenario, (b) the mean model, and (c) our predicted data set, respectively. After calculating the DTW scores for each of these cases, we quantify the similarity of the observed and predicted time series. This is done through the sequence similarity factor (SSF), which we define as:

$$\mathrm{SSF} = \frac{\mathrm{DTW}_{\mathrm{score}}(O, M)}{\mathrm{DTW}_{\mathrm{score}}(O, \bar{O})}, \quad \mathrm{SSF} \in [0, \infty], \tag{2}$$

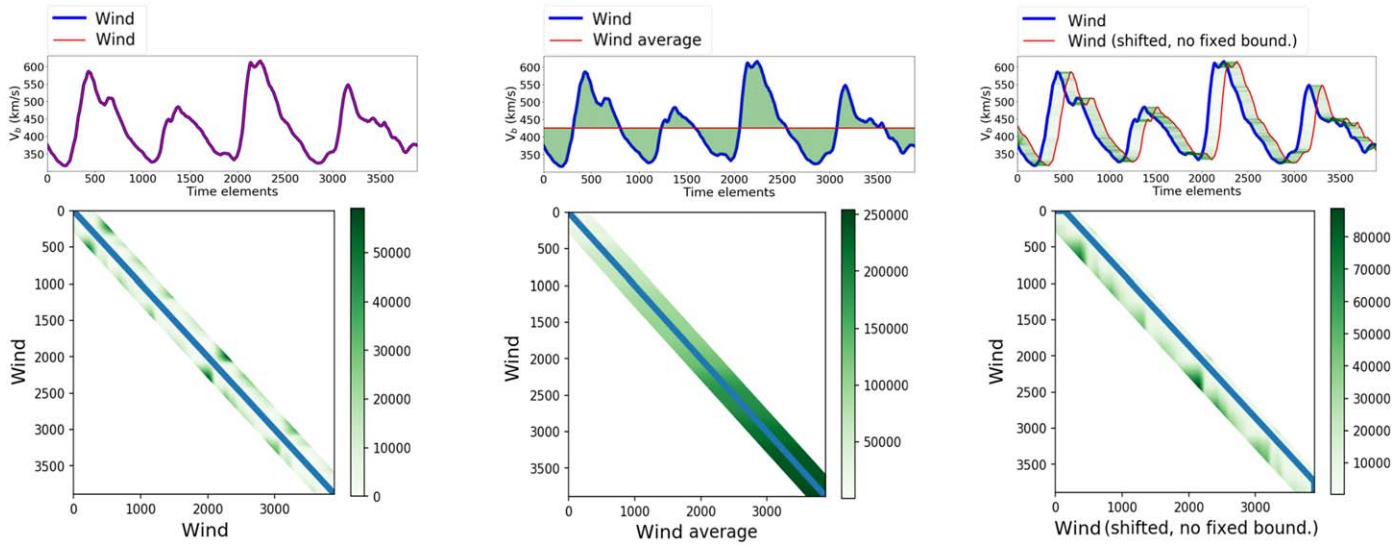
where O, M, and $\bar{O}$ stand for *Observed*, *Modeled*, and *Averaged Observed* data, respectively. The SSF is equal to zero when we have achieved the perfect forecast, and equal to one when our forecast is as bad as a straight average line prediction. In Figure 7(c), the SSF between the observed and predicted time series is 0.021, very close to the perfect scenario of SSF = 0 (Figure 7(a)).

The fact that DTW dynamically warps the sequences in time and, as a result, is able to locate which point from one time series better corresponds to a point from the other time series is a huge advantage compared to other metrics. In their study, Owens et al. (2005) noted that one of the most frequent metrics used, MSE, has a very significant drawback, even if it is a useful tool for a first-order assessment of time series. This drawback comes from the fact that a straight line (Model A, red dotted line in Figure 8) can sometimes give a lower MSE when compared to observations, from a time series that is very similar to observations but shifted in time (Model B, black dashed line in Figure 8). DTW overcomes this problem, opposite to the simple error functions that are completely based on the Euclidean distance measure (see, e.g., the comparison between the DTW scores calculated in Figures 7(b) and (c) and the example shown in Figure 8). Nevertheless, there are still some cases for which the straight average line performs better than our modeled data set. For these cases, it is not the potential shifting in time that causes this discrepancy, as this has already been solved by DTW. On the contrary, the variability in the y-axis, which is sometimes opposite to what is observed in the real data (e.g., the model predicts valleys in place of peaks, and vice versa), is the reason why we get DTW scores that are larger than the ones calculated for the straight-line scenario. As a result, it is reasonable for such cases to obtain an SSF that is higher than one.

4.2. Quantification of Time and Amplitude Differences

Besides the SSF, DTW permits the estimation of time differences and amplitude differences between the points that are best aligned. This important feature is not easily provided by other metrics that can usually quantify only one of these aspects at a time. Figures 9(a) and (b) show histograms of the time and amplitude differences between the aligned points of the sequences presented in Figure 7(c). A maximum Δt of 1 day can be observed in Figure 9(a), which was expected since this is by how much we shifted our data. Figure 9(b) shows a maximum Δv_b of 60 km s^{-1}. The maximum difference in velocity arises at the beginning of the time series, when our predicted data set (red time series) is higher than the observed one (blue time series).

We note that the alignments of the points provided by DTW do contain singularities, i.e., when a point from one sequence can be matched with two or more points from the other



(a) $\mathrm{DTW}_{score}(O, M) = 0$ (b) $\mathrm{DTW}_{score}(O, \bar{O}) = 253850.04$ (c) $\mathrm{DTW}_{score}(O, M) = 5242.86$

Figure 7. Examples of the first approach to evaluate the performance of the predicted solar wind time series compared to observations. (a) The ideal prediction scenario with $\mathrm{DTW}_{\mathrm{score}}(O, M) = 0$. (b) The nonideal/reference case prediction scenario with the maximum DTW score for that specific time interval for which $\mathrm{DTW}_{\mathrm{score}}(O, \bar{O}) = 253{,}850.04$. (c) Our actual observed and predicted time series with $\mathrm{DTW}_{\mathrm{score}}(O, M)$ ranging between 0 and 253,850.04.

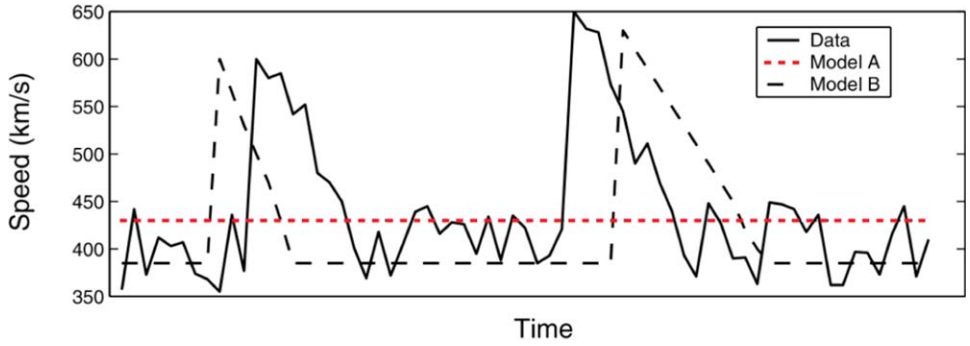


Figure 8. An example of an observed data set (solid line) and two predicted time series (Model A, red dashed line, and Model B, black dashed line). Model A is just a straight line, while Model B is very similar to the observed data but shifted in time. The MSE of Model B is larger than the one calculated for Model A, meaning that a straight-line prediction performs better than a prediction that is very similar to observations, but shifted in time (adapted from Owens et al. 2005).

sequence. Even though singularities are generally assumed to be a drawback of DTW, we find their existence necessary, as they help to ensure the best possible alignment between time series. As a result, the Δt and v_b presented in the histograms are not only relevant to the time and amplitude differences of the points that are uniquely aligned between them, but also to the singularity points. By improving the model and, consequently, the solar wind forecasting, these singularities are minimized, the DTW score becomes lower, and a better agreement with the observations is achieved.

5. Evaluating the Performance of Solar Wind Time Series in EUHFORIA

5.1. SSF versus Traditional Skill Scores: Using the Mean Model as a Reference Model

In this section, we apply the DTW method to the solar wind velocity time series as modeled by EUHFORIA v1.0.4 for the period 2017 November–2019 September. The setup of EUHFORIA is the same as the one presented in Pomoell & Poedts (2018) and Hinterreiter et al. (2019). We first split the considered time interval into smaller periods, as listed in the second column of Table 1. Then, we adopt the smoothing and window constraints (see Sections 3.2 and 3.3, respectively). Finally, we quantify EUHFORIA's performance compared to observations by employing both of the DTW ways mentioned in Section 4. The upper panel of Figure 10 shows the DTW alignment for the time period with the lowest SSF. This corresponds to the time interval between 2019 July 8 and 2019 July 26 (period 20), during which EUHFORIA performed the best compared to all of the other periods we considered in this study. In the same figure, we also present histograms of the time and velocity amplitude difference between the observed and predicted data sets. The time difference between the two sequences is ± 2 days, which is the maximum temporal window we imposed for the alignment. The velocity amplitude differences are the lowest among all of the other periods, with a maximum of 40 km s^{-1}. The application of DTW to the rest of the periods is shown in Figure 13 in the Appendix, and the SSFs are summarized in the fourth column of Table 1.

To prove our point about the significant advantages that DTW offers compared to traditional metrics (see Section 4), in the fifth column of Table 1 we present the results of EUHFORIA's performance compared to observations, as evaluated using a traditional MSE-based skill score metric. This metric is defined

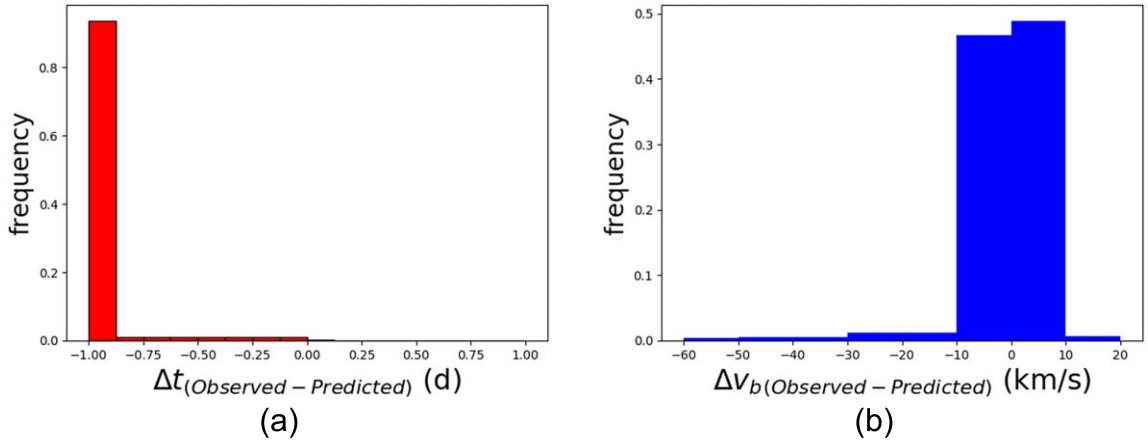


Figure 9. Histograms of the time and velocity amplitude differences between the aligned points, as matched by DTW. (a) Histogram of time differences (Δt in days). (b) Histogram of velocity differences (Δv_b in km s^{-1}).

Table 1
Evaluation of the Performance of the EUHFORIA Solar Wind Time Series Based on the SSFs and Skill Scores for the Individual Periods of Interest

Period	Dates	No. of Elements	SSF_{mean}	Skill $Score_{mean}$	SSF_{27days}	Skill $Score_{27days}$
1	2017-11-06 to 2017-12-03	3889	1.07	2.59	4.33	4.46
2	2017-12-03 to 2017-12-30	3889	**0.70**	**1.74**	2.99	2.00
3	2017-12-30 to 2018-01-18	2737	0.49	0.94	**1.19**	**0.77**
4	2018-01-18 to 2018-03-07	6913	1.43	3.91	2.78	3.19
5	2018-03-07 to 2018-04-17	5905	**0.83**	**1.93**	1.70	1.34
6	2018-04-17 to 2018-05-16	4177	**0.69**	**1.09**	1.93	1.47
7	2018-05-16 to 2018-06-12	3889	**0.66**	**1.12**	1.62	1.30
8	2018-06-12 to 2018-07-29	6769	**0.80**	**1.65**	**1.15**	**0.83**
9	2018-07-29 to 2018-09-02	5040	**0.87**	**1.56**	1.68	1.73
10	2018-09-02 to 2018-09-29	3889	0.42	0.90	**1.03**	**0.72**
11	2018-09-29 to 2018-10-31	4608	0.43	0.78	0.66	0.80
12	2018-10-31 to 2018-11-23	3313	**0.48**	**1.15**	1.90	1.76
13	2018-11-23 to 2018-12-24	4465	0.47	0.92	1.37	1.29
14	2018-12-24 to 2019-01-21	4033	1.07	1.81	1.89	1.07
15	2019-01-21 to 2019-02-19	4177	**0.36**	**1.02**	1.92	2.88
16	2019-02-19 to 2019-03-24	4753	**0.90**	**1.41**	1.69	1.05
17	2019-03-24 to 2019-04-19	3745	1.09	2.73	1.90	1.88
18	2019-04-19 to 2019-06-03	6481	1.17	2.54	1.75	2.07
19	2019-06-03 to 2019-07-08	5041	**0.63**	**1.56**	0.64	0.42
20	2019-07-08 to 2019-07-26	2593	0.15	0.64	0.25	0.54
21	2019-07-26 to 2019-08-22	3889	0.49	0.80	1.52	1.54
22	2019-08-22 to 2019-09-19	4033	**0.86**	**1.06**	2.89	1.37

Note. A 10 minute resolution and a 12 hr smoothing window was adopted. The subscripts "mean" and "27 days" indicate which reference model was employed each time. The bold values correspond to cases at which the SSF and Skill Score provided opposite performance evaluations for the time series of interest.

as:

$$\text{Skill Score} = \frac{\text{MSE}}{\text{MSE}_{\text{ref}}}, \quad \text{Skill Score} \in [0, \infty]. \tag{3}$$

The nominator corresponds to the MSE between observations and EUHFORIA, while the denominator corresponds to the MSE between observations and a reference model. In this section, the reference model will be the mean model. Similar to Equation (2), a Skill Score equal to zero corresponds to the perfect prediction, while a Skill Score equal to one means that the prediction performs the same as the reference model. For a Skill Score higher than one, the reference model performs better than our predictions.

Comparing the SSF and Skill Score values in columns four and five of Table 1, we see that even though both measures have been similarly defined they sometimes provide different assessment results for the same periods (see numbers in bold). Periods 2 and 6 correspond to such controversial examples for which SSF <1, meaning that EUHFORIA performed better than the mean model, but at the same time the Skill Score >1, meaning that EUHFORIA performed worse than the mean model. To clarify the situation, we present the time series for these particular periods in Figure 11. We notice that in both cases EUHFORIA predicted the observations better than the mean model. As a result, we should expect a SSF <1 and a Skill Score <1. Nevertheless, due to the time lag of the predicted time series compared to the Wind data, and the inability of the traditional skill score to capture the overall shape of the sequence, the MSE of the mean model was lower than the MSE of the EUHFORIA time series. This resulted in a Skill Score >1, which does not reflect the actual bad performance of the mean model in terms of forecasting the variability in the solar wind.

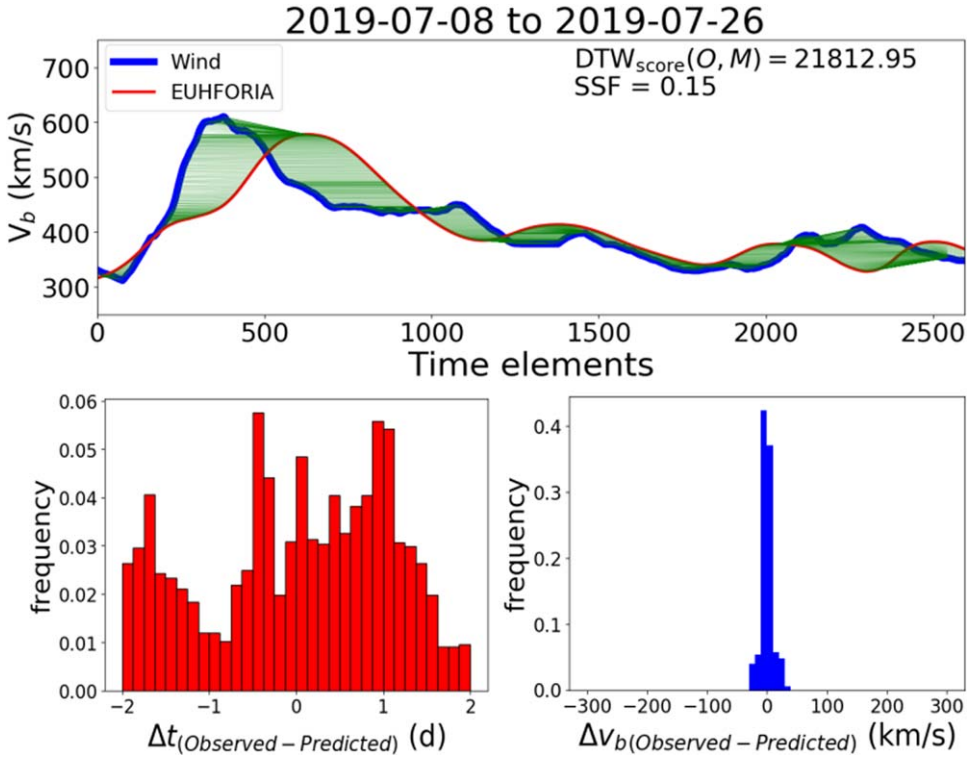


Figure 10. DTW alignment (upper panel) and histograms of the time and velocity differences (lower panels) between 2019 July 8 and 2019 July 26 (period 20), during which EUHFORIA performed the best compared to all of the other considered periods. The time in the *x*-axis of the upper panel corresponds to evenly spaced time elements with a time difference of 10 minutes between each other.

5.2. SSF versus Traditional Skill Scores: Using the 27 Day Persistence Model as a Reference Model

In this section, we perform the same test as in Section 5.1, but we employ a different reference model—the 27 day persistence model (Owens et al. 2013), which is widely used within the space weather community. Based on this model, we assume that the solar wind speeds measured over a full solar rotation predict the future solar rotation as well. The SSF and Skill Score results for the individual periods of interest are summarized in the sixth and seventh columns of Table 1. We notice that in all but three periods (periods 3, 8, and 10) the SSFs and Skill Scores agree with the evaluation of the EUHFORIA time series. Namely, when the SSF <1 and the Skill Score <1, both measures reflect that EUHFORIA performs better compared to the 27 day persistence model, and vice versa.

In Figures 12(a) and (b), we present the time series for two of the three periods for which the SSF and Skill Score showed opposite results. In both cases, the 27 day persistence model predicts the Wind data better than EUHFORIA, not only because of the number of HSSs that it captures, but also because it better reproduces the HSS amplitudes. In Figures 12(c) and (d), we further show two cases for which the SSFs and Skill Scores agree. In the former case, the 27 day persistence model performs better than EUHFORIA, and that is reflected in the SSF and Skill Score numbers. In the latter case, EUHFORIA performs better than the persistence model, with both the SSF and Skill Score numbers being lower than unity.

6. Conclusions and Discussion

In this study, we have introduced an alternative way of assessing the performance of solar wind time series, the so-called dynamic time warping (DTW) technique. Although DTW is not a metric by definition, as it violates the triangular inequality, it acts like one. It obeys the rule of continuity, monotonicity, and the fact that the first and last points of one sequence should be matched with at least the first and last points of the other. It calculates a cumulative cost, the so-called DTW score, which represents the cost of aligning two time series in time when their patterns are similar but differ in time. DTW has already been used in other disciplines, but this is the first time that it has been adapted and applied for the purpose of evaluating the performance of solar wind time series.

We have discussed the benefits and restrictions of the technique, and presented two complementary ways in which DTW can be exploited to assess the solar wind predictions provided by a model. The first way calculates the DTW score between observations and predictions, as well as between observations and a reference model. We define the ratio of these scores as the sequence similarity factor (SSF). This is a skill score that is equal to zero for a perfect forecast, equal to one when the forecast performs the same as the reference model, and higher than one when the model's prediction is even worse than the prediction from the reference model. The second way in which DTW can be exploited is by evaluating the time and amplitude differences between the points aligned by the method. As a result, DTW can be used as a hybrid metric between continuous measurements (such as, e.g., the correlation coefficient) and point-by-point comparisons, by simultaneously assessing time and amplitude differences, a property not often found in traditional metrics.

We then assessed the performance of solar wind predictions from EUHFORIA for an interval of approximately two years (2017 November–2019 September). This interval was first divided into smaller periods for faster and more accurate evaluation. To acknowledge the advantages of DTW and

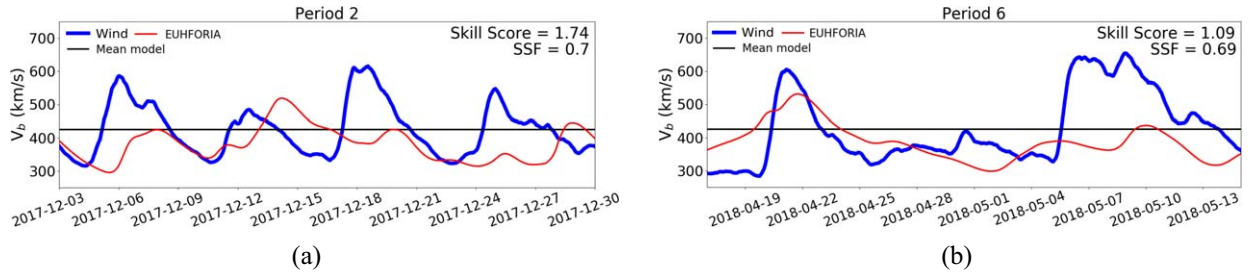


Figure 11. The Wind observations (blue), EUHFORIA output (red), and mean model (black) are shown for periods 2 (panel a) and 6 (panel b). The SSFs and traditional Skill Scores are also presented in the upper right parts of each panel.

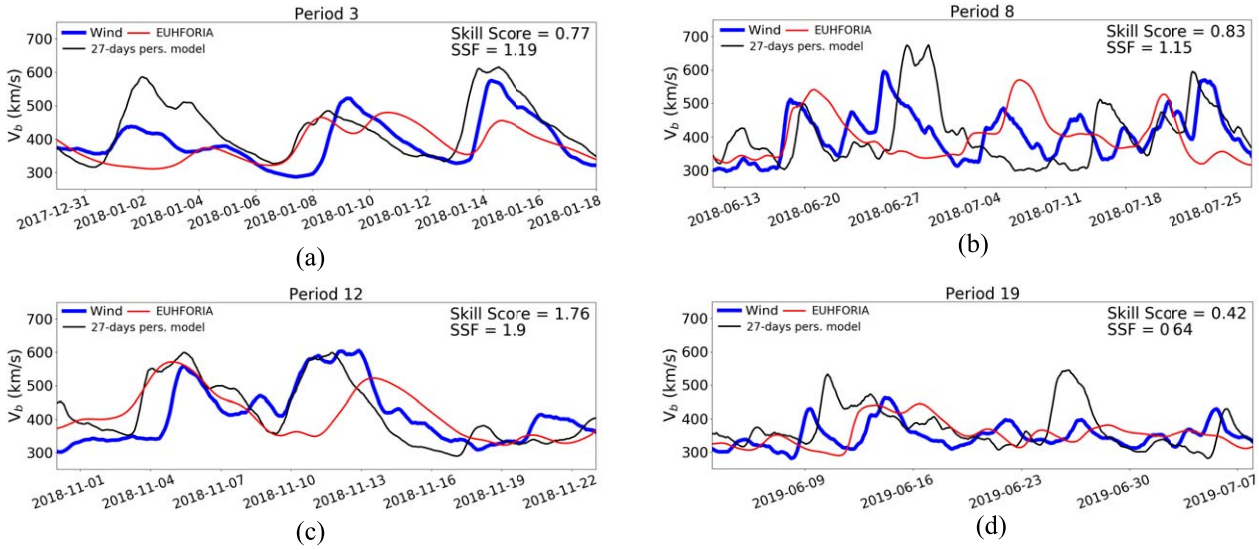


Figure 12. The Wind observations (blue), EUHFORIA output (red), and 27 day persistence model (black) are shown for periods 3 (panel (a)), 8 (panel (b)), 12 (panel (c)), and 19 (panel (d)). The SSFs and Skill Scores are also presented in the upper right parts of the panels.

understand its differences to traditional skill score metrics, we performed two tests: first, we evaluated our predictions based on the SSF and an MSE-based skill score metric by employing the mean model as a reference model; and second, we repeated the same procedure, but this time we employed the 27 day persistence model as a reference. The former test showed that in 50% of cases (11 out of 22 periods), the SSF and MSE-based skill score yielded opposite results. In particular, the MSE-based skill score indicated that the mean model performed better than the EUHFORIA simulations, even though EUHFORIA reproduced the observations much better. The discrepancy between the two measures arises from the ability of DTW to dynamically warp the sequences in time and locate which point from one sequence better corresponds to a point from another sequence, opposite to conventional Euclidean metrics. Next, employing the 27 day persistence model as the reference model in our study, we concluded that it performed better than EUHFORIA in predicting the observations, for 19 out of 22 periods of interest. In 3 of those 19 cases, the SSF and MSE-based skill score resulted in opposite assessments, with the latter metric providing misleading evaluations for the predictions, due to its inability to capture the overall shape of the time series. Therefore, we prove that DTW can be used as an objective quantification measure for model evaluation. In addition to MSE (and other traditional metrics), it provides more detailed information on the similarities between the profiles of two data sets, thus it should be used in conjunction with other measures to provide the most complete picture of a model's performance.

The use of DTW can also be extended to the evaluation of other solar wind signatures besides velocity, such as the solar wind density, temperature, magnetic field, pressure, etc. Moreover, an extension of DTW, the so-called multidimensional DTW (see Shokoohi-Yekta et al. 2015; Abdullah & Keogh 2016, and references therein), permits the assessment of multiple sequences, i.e., in the case of multidimensional simulations, in which time series from different locations around Earth (or any other point of interest) are considered. This technique enables pattern comparisons between multiple time series at the same time, which is particularly useful for the evaluation of spatial uncertainties during the arrival of an HSS at a measuring satellite. The multidimensional DTW could also be extended to the identification of HSSs by evaluating the various signatures of plasma and magnetic characteristics. For example, during the arrival of an HSS at a particular point of interest, the solar wind density, temperature, pressure, magnetic field, and interplanetary magnetic field polarity should conform to specific patterns and behaviors (see Jian et al. 2006 for more details), which should be recognized by the method. These ideas have not been tested in the frame of the current study, but they constitute promising ideas for the future.

For consistent evaluations of modeled solar wind time series with DTW, we recommend avoiding periods that include potential CME influence at Earth. If this is not possible, the user can still apply the DTW by ignoring the CME signatures. The DTW score will be slightly different then, compared to the case in which there were no CMEs. As mentioned in Section 3.4, 15 CMEs were identified as influencing Earth between 2017 November and 2019 September. For most of the individual periods listed in Table 1, there were no observed CMEs. However, for the cases in which CMEs were detected, we ignored them for the application of DTW. Therefore, a future user who wants to assess an improved version of EUHFORIA for the same time intervals should work in the same way; namely, treat the potential CME structures as if they were not there. Only then will the decrease (or increase) in the DTW score be consistent and comparable to the one calculated using EUHFORIA v1.0.4, allowing us to track how the change of the model influences the modeling output. We note that DTW could be applied separately to CME time series, in terms of the evaluation of the shock, the B_z component, or other in situ parameters.

The authors would like to acknowledge the anonymous referee for the fruitful comments that improved the content of the paper. They would also like to thank Emmanuel Chané for his insights and big contribution to this work, which helped the project to grow. Moreover, they extend their acknowledgments to Eamonn Keogh for the helpful discussions and advice on DTW. E.S. was supported by a PhD grant awarded by the Royal Observatory of Belgium. C.V. is funded by the Research Foundation—Flanders, FWO SB PhD fellowship 11ZZ216N. EUHFORIA is developed as a joint effort between KU Leuven and the University of Helsinki. Work at IRAP was supported by the Centre National de la Recherche Scientifique (CNRS, France), the Centre National d'Etudes Spatiales (CNES, France), and the Université Paul Sabatier (UPS). The validation of solar wind with EUHFORIA is being performed within the BRAIN-be project SWiM (Solar Wind Modeling with EUHFORIA for the new heliospheric missions). This project has received funding from the European Union's Horizon 2020 research and innovation programs under grant agreements No. 870405 (EUHFORIA 2.0) and 870437 (SafeSpace). S.P. acknowledges support from the projects C14/19/089 (C1 project Internal Funds KU Leuven), G.0D07.19N (FWO-Vlaanderen), and SIDC Data Exploitation (ESA Prodex-12). The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government Department EWI.

Software: The DTW code used for this work was based on Senin (2008), Nipun (2020), and Abhishek (2020), and it can be found at https://github.com/SamaraEvangelia/DTW_ForSolarWindEvaluation.

Appendix
DTW Alignments and Histograms between 2017 November and 2019 September

In Figure 13, we present the optimal DTW alignments between the Wind observations and EUHFORIA predictions, as well as histograms of the time and amplitude differences for all of the individual periods listed in Table 1. The green lines show how points from the blue time series are matched with points from the red time series. The DTW score as well as the SSF are shown in the upper right parts of each time series plot. Moreover, the histograms of each period provide a good idea of the minimum and maximum differences in time and velocity between the two sequences.

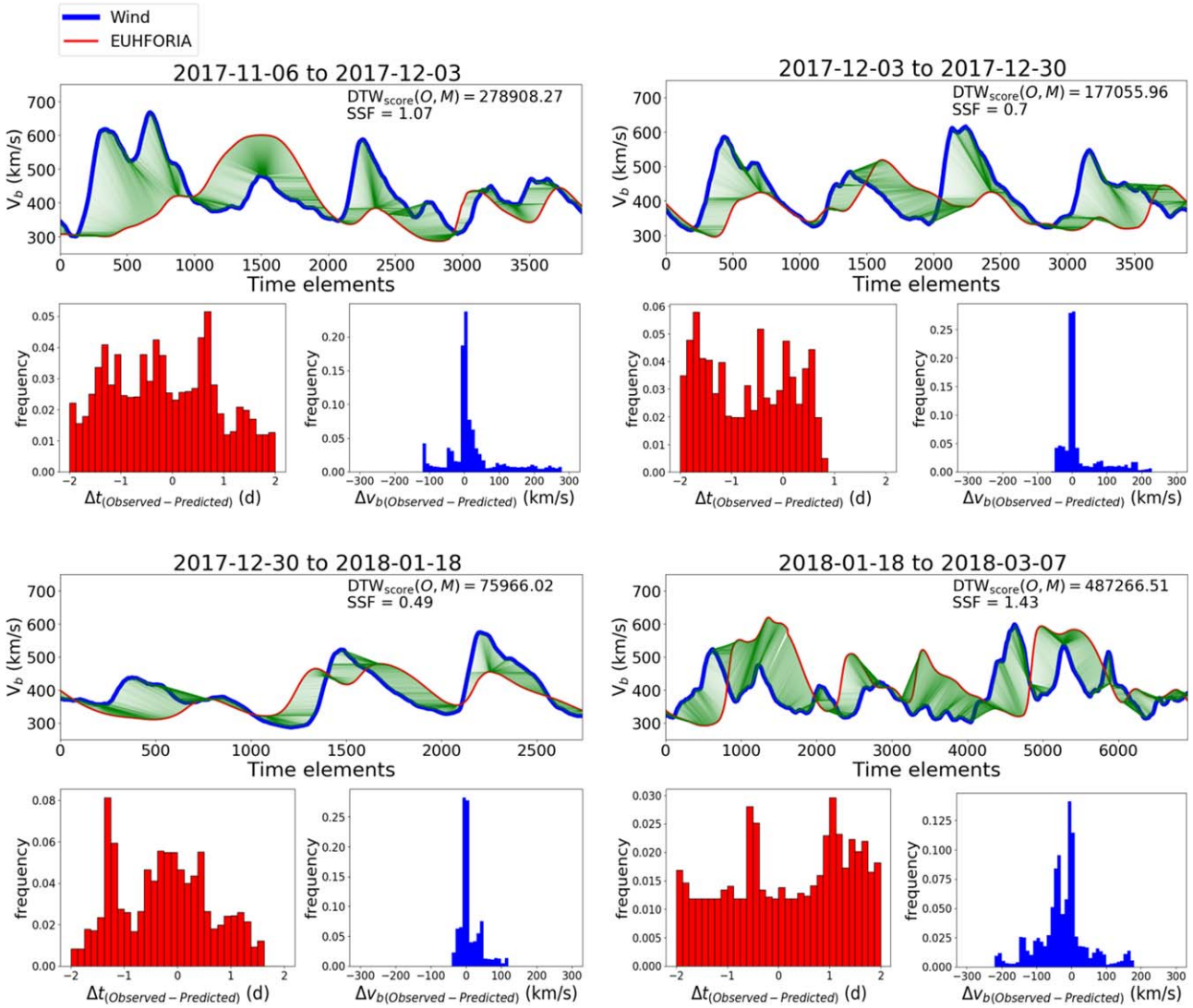


Figure 13. DTW alignments and histograms of the time and velocity differences for the 22 periods under assessment, between 2017 November and 2019 September.

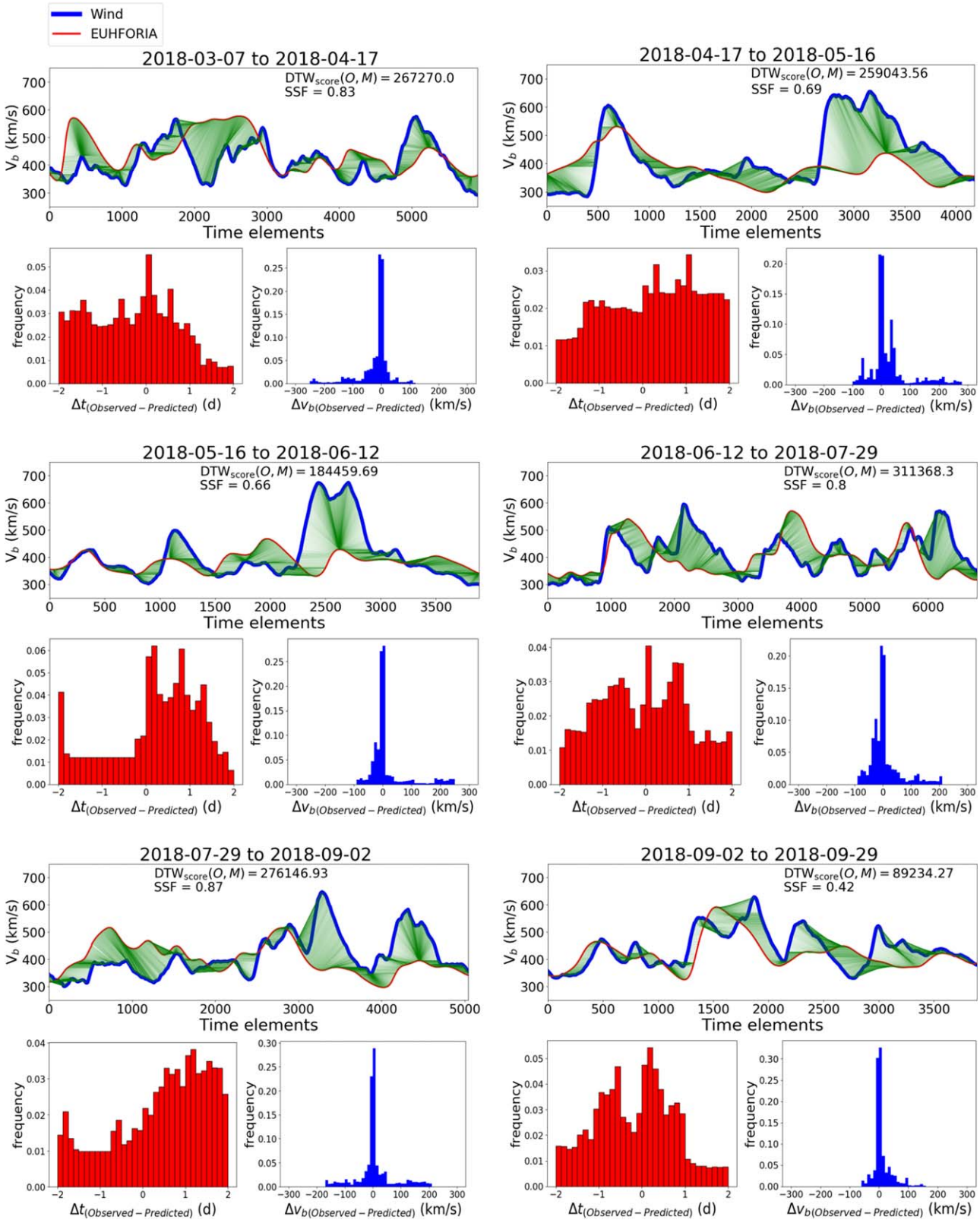


Figure 13. (Continued.)

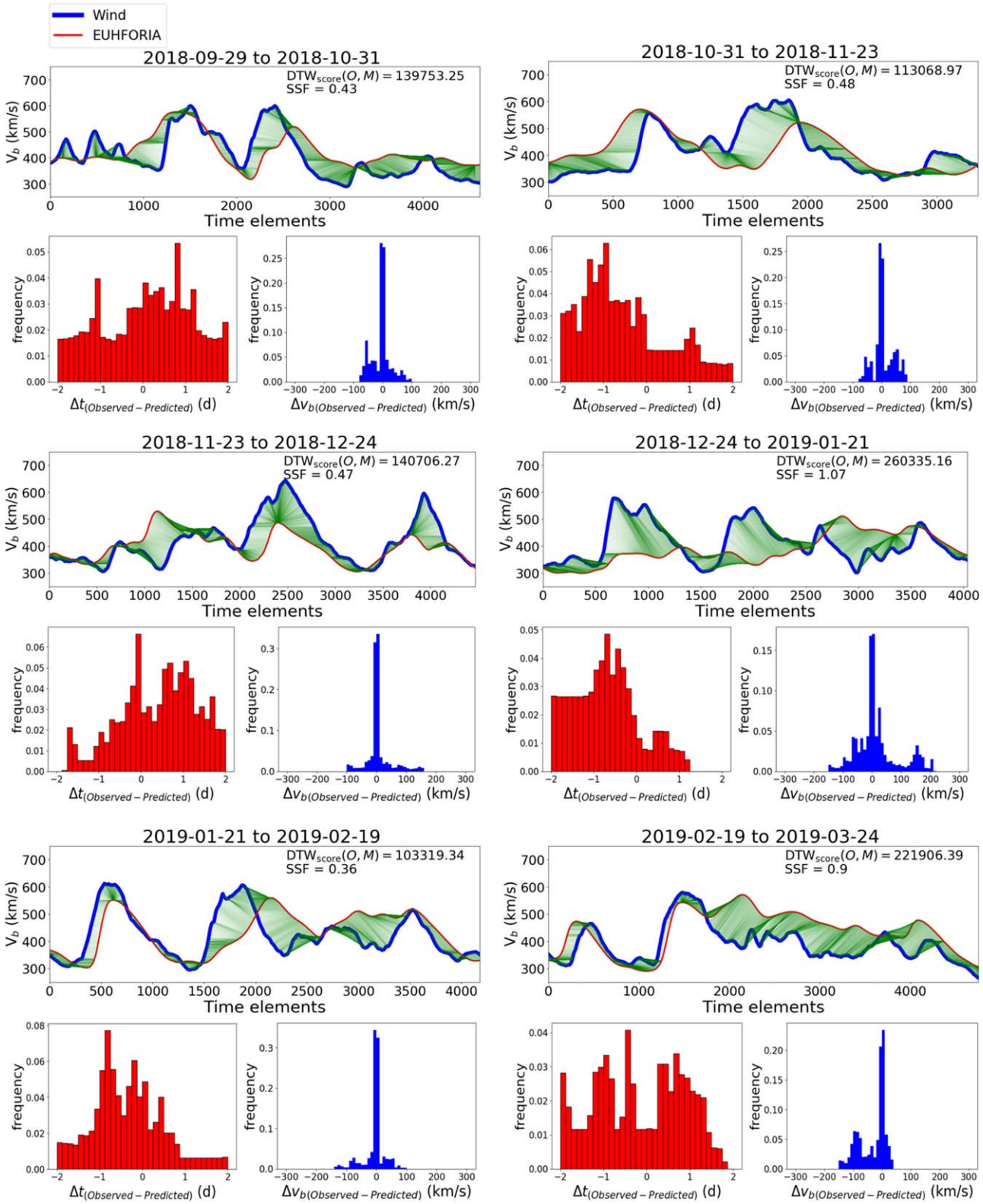


Figure 13. (Continued.)

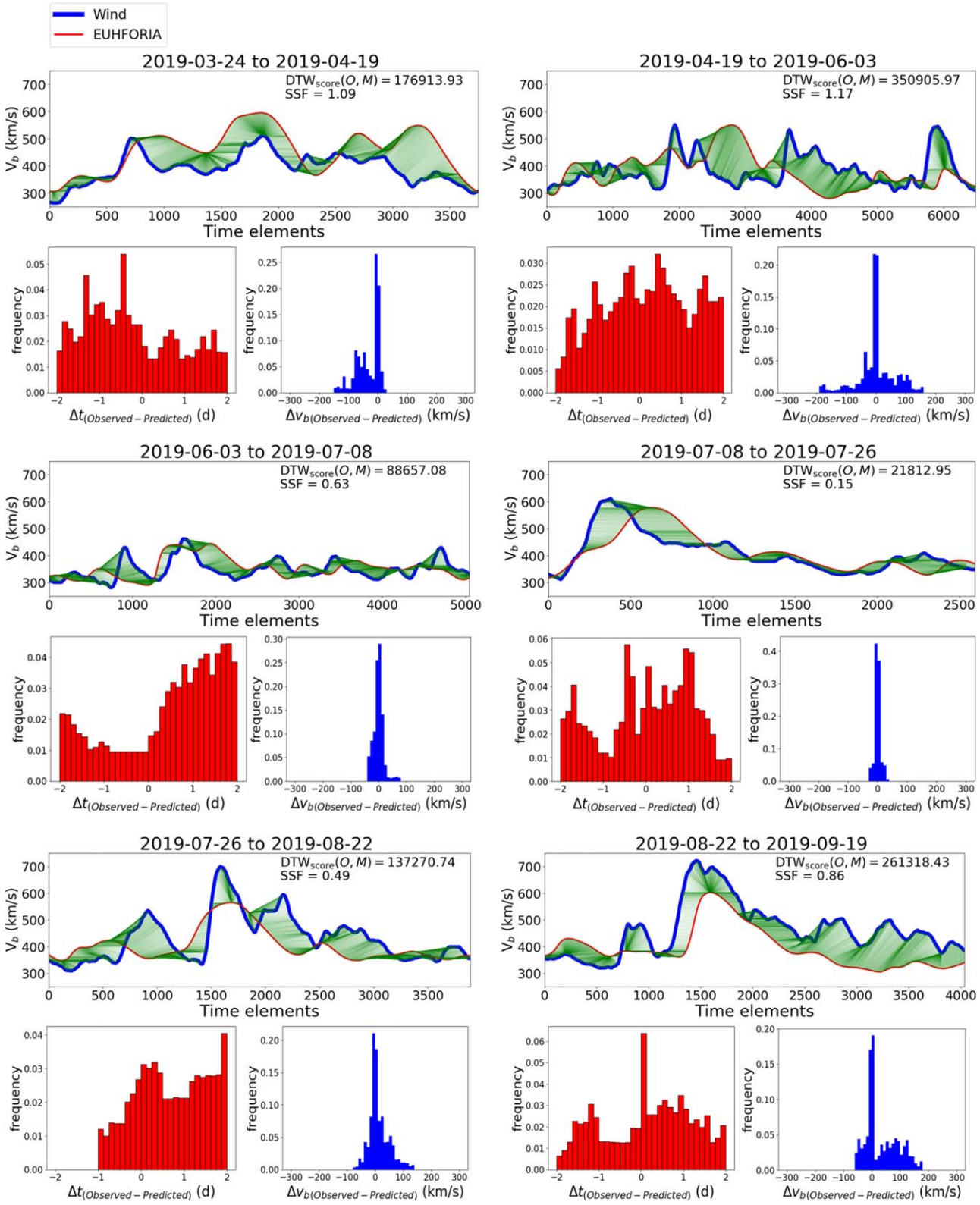


Figure 13. (Continued.)

ORCID iDs

E. Samara https://orcid.org/0000-0002-7676-9364
B. Laperre https://orcid.org/0000-0001-7218-3561
M. Temmer https://orcid.org/0000-0003-4867-7558
C. Verbeke https://orcid.org/0000-0002-9402-5609
J. Magdalenić https://orcid.org/0000-0003-1169-3722
S. Poedts https://orcid.org/0000-0002-1743-0651

References

Abdullah, M., & Keogh, J. K. 2016, in KDD '16: Proc. of the 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (New York: Association for Computing Machinery), 2129
Abhishek, M. 2020, Time Series Similarity Using Dynamic Time Warping—Explained, https://medium.com/walmartglobaltech/time-series-similarity-using-dynamic-time-warping-explained-9d09119e48ec
Arge, C. N., Luhmann, J. G., Odstrcil, D., Schrijver, C. J., & Li, Y. 2004, JASTP, 66, 1295
Arge, C. N., Odstrcil, D., Pizzo, V. J., & Mayer, L. R. 2003, in AIP Conf. Proc. 679, Solar Wind Ten, ed. M. Velli et al. (Melville, NY: AIP), 190
Arge, C. N., & Pizzo, V. J. 2000, JGR, 105, 10465
Benedikt, L., Kajic, V., Cosker, D., Rosin, P., & Marshall, D. 2008, in Proc. of the British Machine Conf., ed. M. Everingham & C. Needham (Durham: BMVA Press), 107.1
Berndt, D., & Clifford, J. 1994, in AAAIWS'94: Proc. of the 3rd Int. Conf. on Knowledge Discovery and Data Mining (Palo Alto, CA: AAAI Press), 359
Bussy-Virat, C. D., & Ridley, A. J. 2014, SpWea, 12, 337
Cane, H. V., & Richardson, I. G. 2003, JGRA, 108, 1156
Dau, H. A., Keogh, E., Kamgar, K., & Yeh, C. M. 2019, IEEE/CAA J. Autom. Sinica, 6, 1293
Górecki, T., & Luczak, M. 2013, Data Min. Knowl. Discov., 26, 310
Gressl, C., Veronig, A. M., Temmer, M., et al. 2014, SoPh, 289, 1783
Hinterreiter, J., Magdalenic, J., Temmer, M., et al. 2019, SoPh, 294, 170
Itakura, F. 1975, ASAJ, 57, S35
Jackson, B. V., Buffington, A., Cota, L., et al. 2020, FrASS, 7, 76
Jackson, B. V., Hick, P. L., Kojima, M., & Yokobe, A. 1998, JGR, 103, 12049
Jackson, B. V., & Hick, P. P. 2002, SoPh, 211, 345
Jeong, Y.-S., Jeong, M. K., & Omitaomu, O. A. 2011, PatRe, 44, 2231
Jian, L., Russell, C. T., Luhmann, J. G., & Skoug, R. M. 2006, SoPh, 239, 337
Jian, L. K., MacNeice, P. J., Taktakishvili, A., et al. 2015, SpWea, 13, 316
Keogh, E. J., & Pazzani, M. J. 1998, KDD-98 Proc. (Palo Alto, CA: AAAI Press), 239, http://www.aaai.org/Library/KDD/1998/kdd98-041.php
Keogh, E. J., & Pazzani, M. J. 2001, in Proc. of the 2001 SIAM Int. Conf. on Data Mining (SDM) (Philadelphia, PA: Society for Industrial and Applied Mathematics), 1
Kruskall, J. B., & Liberman, M. 1983, Time Warps, String Edits and Macromolecules: The Theory and Practice of String Comparison (Reading, MA: Addison-Wesley)
Kulbacki, M., & Bak, A. 2002, in Intelligent Information Systems 2002, ed. M. A. Kłopotek, S. T. Wierzchoń, & M. Michalewicz (Heidelberg: Physica-Verlag HD), 217
Laperre, B., Amaya, J., & Lapenta, G. 2020, FrASS, 7, 39
Lemen, J. R., Title, A. M., Akin, D. J., et al. 2012, SoPh, 275, 17
Linker, J. A., Caplan, R. M., Downs, C., et al. 2016, JPhCS, 719, 012012
Lionello, R., Linker, J. A., & Mikić, Z. 2003, in AIP Conf. Proc. 679, Solar Wind Ten, ed. M. Velli et al. (Melville, NY: AIP), 222
MacNeice, P. 2009, SpWea, 7, S06004
MacNeice, P., Jian, L. K., Antiochos, S. K., et al. 2018, SpWea, 16, 1644
Meng, X., van der Holst, B., Tóth, G., & Gombosi, T. I. 2015, MNRAS, 454, 3697
Mikić, Z., Linker, J. A., Schnack, D. D., Lionello, R., & Tarditi, A. 1999, PhPl, 6, 2217
Müller, M. 2007, Information Retrieval for Music and Motion (Berlin: Springer), 69
Myers, C. S., Rabiner, L. R., & Rosenberg, A. E. 1981, BSTJ, 60, 303
Nipun, B. 2020, Programatically Understanding Dynamic Time Warping (DTW), https://nipunbatra.github.io/blog/ml/2014/05/01/dtw.html
Odstrčil, D., & Pizzo, V. J. 1999, JGR, 104, 483
Ogilvie, K. W., Chornay, D. J., Fritzenreiter, R. J., et al. 1995, SSRv, 71, 55
Owens, M. J. 2018, SpWea, 16, 1847
Owens, M. J., Arge, C. N., Spence, H. E., & Pembroke, A. 2005, JGRA, 110, A12105
Owens, M. J., Challen, R., Methven, J., Henley, E., & Jackson, D. R. 2013, SpWea, 11, 225
Owens, M. J., Riley, P., & Horbury, T. S. 2017, SoPh, 292, 69
Owens, M. J., Spence, H. E., McGregor, S., et al. 2008, SpWea, 6, S08001
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Pinto, R. F., & Rouillard, A. P. 2017, ApJ, 838, 89
Pomoell, J., & Poedts, S. 2018, JSWSC, 8, A35
Ratanamahatana, C. A., & Keogh, E. 2004, in Proc. of the 2004 SIAM Int. Conf. on Data Mining (SDM) (Philadelphia, PA: Society for Industrial and Applied Mathematics), 11
Reiss, M. A., Temmer, M., Veronig, A. M., et al. 2016, SpWea, 14, 495
Richardson, I. G., & Cane, H. V. 2010, SoPh, 264, 189
Riley, P., Ben-Nun, M., Linker, J. A., Owens, M. J., & Horbury, T. S. 2017, SpWea, 15, 526
Riley, P., Lionello, R., Linker, J. A., et al. 2011, SoPh, 274, 361
Sakoe, H., & Chiba, S. 1978, ITASS, 26, 43
Samara, E., Pinto, R. F., Magdalenić, J., et al. 2021, A&A, 648, A35
Senin, P. 2008, Dynamic Time Warping Algorithm Review, https://csdl.ics.hawaii.edu/techreports/2008/08-04/08-04.pdf
Shokoohi-Yekta, M., Wang, J., & Keogh, E. 2015, in Proc. 2015 SIAM Int. Conf. on Data Mining (Philadelphia, PA: Society for Industrial and Applied Mathematics), 289
Temmer, M., Hinterreiter, J., & Reiss, M. A. 2018, JSWSC, 8, A18
Vidal, E., Casacuberta, F., Benedi, J. M., Lloret, M. J., & Rulot, H. 1988, Speech Commun., 7, 67
Vršnak, B., Temmer, M., & Veronig, A. M. 2007, SoPh, 240, 315
Wong, T., & Wong, M. H. 2003, in Seventh Int. Database Engineering and Applications Symp. 2003 (Piscataway, NJ: IEEE), 139
Zhang, C., Fanaee Tork, H., & Thoresen, M. 2021, Data Min. Knowl. Discov., 35, 1760
Zhang, Z., Tavenard, R., Bailly, A., et al. 2017, Inform. Sci., 393, 91