Forecast improvements from solar wind data assimilation

Harriet Turner

Department of Meteorology

University of Reading

A thesis presented for the degree of

Doctor of Philosophy

February 2024

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Harriet Turner

Acknowledgements

I would write something deep and philosophical here, but I have used up all my words writing this thesis so there isn't really much left. A Matt Owens quote that has lodged in my brain is that "the deeper you look in the sausage factory, the more disgusted you get" and that couldn't have resonated more with me whilst writing this thesis.

I could not have done this without the people around me. To Matt and Matt, despite you refusing to change your names by deed poll to make it less confusing, I could not have asked for better supervisors. Thank you for answering all my questions and providing unwavering support throughout everything. I'm afraid you're stuck with me for another 3 years.

To my family, thank you for being my rock. Thank you for instilling in me the curiosity and drive needed to complete a PhD. I hope I have made you proud.

To Beth and Izzy, my 2U08 chums. Thank you for the chaos. Thank you for the laughs and thank you for catching me when I fell. Ones like you are hard to come by.

To Charlie, Hannah, Brian and others, you guys are so special to me and have made my time in Meteorology so memorable.

To Lauren, thank you for being the best housemate for 3 years. You showed me that finishing a PhD in space science is possible and I am so glad I got to share that experience with you. I hope we get to share another graduation together.

And to Michael, thank you for teaching me not to submit my thesis half an hour before the four-year deadline. Thank you for bearing with me whilst I have been writing up, I know I was not easy to live with during that time. Thank you for your support. Thank you for everything.

Abstract

Accurate space weather forecasting requires knowledge of the solar wind conditions. Current forecasting methods are initialised with photospheric magnetic field observations but contain no information from observations of the solar wind. Data assimilation (DA) is a method that combines a model of a system with observations to find an optimum estimation of reality. DA has led to large forecast improvements in terrestrial weather forecasting but has been underused in space weather forecasting, especially in the solar wind.

The first part of this thesis investigates the errors present in solar wind forecasts due to the latitudinal separation of spacecraft observations and the forecast location. It is important to understand how the observation locations affect the forecast accuracy to better inform the data assimilation scheme in the future. It was found that up to a separation of 6 degrees in latitude, there was no significant impact on forecast accuracy.

Secondly, the impact of observational age and the presence of CMEs on solar wind speed forecasts produced using DA is examined. This allows us to understand how the forecast error is expected to change across the solar cycle and under different observing configurations. It was found that assimilating multiple sources of observations improves the forecast accuracy. Removing CMEs caused a marginal improvement in forecast accuracy, but it would be difficult to achieve in real time.

The final part of this thesis considers the use of near-real-time (NRT) data in the DA scheme. This is in preparation for it to be used for operational forecasting. The DA still performed well when using NRT data, showing that it is suitable for operational use. Through simulations of similarly configured spacecraft, it was also shown that the future *Vigil* mission would be useful for solar wind DA.

List of Figures

List of Tables

Notation

Find below a description of the major notations used within this thesis.

Notation	Description
θ	Heliographic latitude
ϕ	Heliographic longitude
β	Plasma beta
$\mathbf{B}$	For plasma physics, magnetic field. For data assimilation, observation error covariance matrix
$\mathbf{E}$	Electric field
$\mathbf{F}$	Lorentz force
$\mathcal{H}(\mathbf{x})$	Observation operator
$\mathcal{J}(\mathbf{x})$	Cost function
$\mathbf{J}$	Current density
ρ	Density
P	Pressure
$\mathbf{R}$	Prior error covariance matrix
R_m	Magnetic Reynolds number
R_S	Solar radius $= 695,508$ km
T	Temperature
$\mathbf{y}$	Observations

Acronyms

Find below a list of the acronyms and their meaning used within this thesis.

Acronym	Description
4DVar	4-Dimensional Variational
ACE	Advanced Composition Explorer
AU	Astronomical Unit = 1.496×10^8 km
BFGS	Brandon-Fletcher-Goldfarb-Shanno
BRaVDA	Burger Radius Variational Data Assimilation
CIR	Corotating Interaction Region
CME	Coronal Mass Ejection
DA	Data Assimilation
DSCOVR	Deep Space Climate Observatory
EUHFORIA	European Heliospheric Forecasting Information Asset
ESA	European Space Agency
EUV	Extreme Ultraviolet
GEO	Geostationary Orbit
HCS	Heliospheric Current Sheet
HUX	Heliospheric Upwind Extrapolation
ISS	International Space Station
L1 – 5	Lagrange 1 - 5
MAE	Mean Absolute Error
MAS	Magnetohydrodynamic Algorithm outside a Sphere
MEO	Medium Earth Orbit
MHD	Magnetohydrodynamics
MOSWOC	Met Office Space Weather Operations Centre
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NRT	Near-real-time
PLASTIC	Plasma and Suprathermal Ion Composition
SEE	Single Event Effect
SIR	Stream Interaction Region
STEREO	Solar Terrestrial Relations Observatory
SWEPAM	Solar Wind Electron Proton Alpha Monitor
SWPC	Space Weather Predictions Centre
WSA	Wang-Sheeley-Arge

Contents

It’s gravity, innit.

Mathew J. Owens, 2020

Chapter 1

Thesis overview

1.1 Motivation and aims

Space weather refers to the changing plasma conditions in near-Earth space and can cause impacts on Earth, particularly on our technological infrastructure. Severe space weather can be driven by coronal mass ejections (CMEs), which are eruptions of material and magnetic field from the Sun that can be directed towards Earth. Accurate space weather forecasting, such as that carried out by the UK Met Office Space Weather Operations Centre (MOSWOC), requires knowledge of the solar wind conditions. The solar wind is a constant stream of charged particles that flows off the Sun and is the medium through which CMEs propagate on their way to Earth. Therefore, the solar wind conditions can affect a CMEs speed and arrival time at Earth (Case et al. 2008).

The solar wind is currently forecast using models that are initialised using photospheric magnetic field observations, but are not constrained any further by solar wind observations. Data assimilation (DA) is a technique that combines a model of a system with observations to find an optimum estimation of reality. It has led to large forecast improvements in terrestrial weather forecasting but has been under used in space weather forecasting. The Burger Radius Variational Data Assimilation (BRaVDA, Lang and Owens 2019) scheme has been developed to apply DA to the solar wind with the aim to improve space weather forecasts. This is the scheme that has been used extensively in the work presented here.

This thesis is based around investigating the improvements in solar wind forecasting that DA offers. The main aims of this thesis are:

- To understand some of the errors present in solar wind forecasts
- To verify solar wind DA forecasts under a range of conditions
- To transition to operational solar wind DA

1.2 Structure

The background material relating to the Sun and plasma physics is presented in Chapter 2. Covered here are the structure of the Sun, the solar cycle, basic plasma physics, the solar wind, space weather drivers and space weather effects. This allows for the understanding of the mechanisms driving space weather and the potential impacts on Earth.

Chapter 3 covers the methodology used within this work. The data section describes the instruments used to measure solar wind speed and the space missions that are relevant for this work. Current forecasting methods and data assimilation are covered, with a specific section discussing the Burger Radius Variational Data Assimilation (BRaVDA) scheme.

The work assessing the errors present in solar wind forecasting due to latitudinal separation of spacecraft is presented in Chapter 4. This provides an insight into the errors that could be present in a DA forecast from using less than perfect observation sources.

Chapter 5 investigates the impact of the age of observations and the presence of CMEs on solar wind speed DA forecasts. This allows different conditions and observing configurations to be tested to see how they affect the forecast accuracy.

Finally, the transition towards operational forecasting is presented in Chapter 6. This tests the performance of near-real-time (NRT) data in the DA scheme in preparation for operational forecasting. The future *Vigil* mission is also tested for its possible use for solar wind DA.

The thesis is concluded in Chapter 7, with plans for future work.

Chapter 2

Introduction

2.1 The Sun

The Sun is an ordinary star of mass 2.0×10^{30} kg and radius 7.0×10^{8} m, sitting at an average distance of 1.5×10^{11} m from Earth (Meyer-Vernet 2007). It is primarily made up of plasma, which is an ionised gas that carries an electric charge (Chen 2012). The Sun is split into several layers. The solar interior extends from the core at the centre, through the radiative zone, convection zone and into the photosphere (Kivelson and Russell 1995). The solar atmosphere is made up of the chromosphere and the corona (Figure 2.1). Each layer has its own distinct properties and characteristics; however, here we only give an overview of the photosphere, chromosphere and corona as these are of most importance to the solar wind. Although we only discuss the layers from the photosphere and above, it must be noted that the solar magnetic field is generated by the motion of plasma in the convection zone. The rotation of the Sun and the convective processes within the convection zone drive a dynamo process (e.g. Charbonneau 2020), which generates electric currents and therefore leads to the generation and evolution of the magnetic field.

2.1.1 The photosphere

The photosphere lies at the top of the convection zone and represents the visible surface of the Sun. It is from here that almost all of the light that we receive originates from, with material below being opaque and material above being almost transparent (Kivelson and Russell 1995). It is covered in a granular pattern showing the convection cells from the layer below. These granules generally have a mean diameter of 1000 km (Hanslmeier 2010). Hot material rises from the convection zone, cools in the cooler temperatures of the photosphere, and sinks back down (Vögler et al. 2005). This produces the pattern we see, with the hot material in the bright centre of the granules and the cooler material sinking in the dark regions at the edges, as depicted in Figure 2.2.

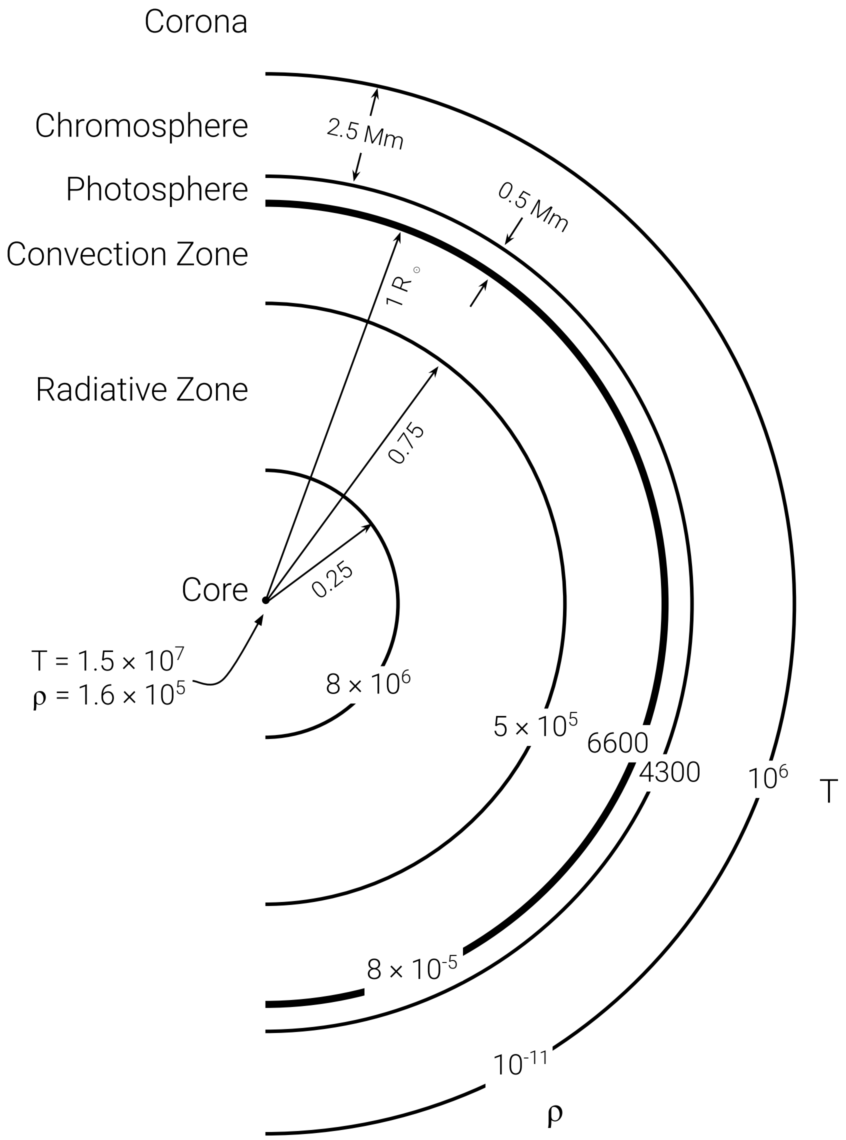


Figure 2.1: Structure of the Sun and each layers associated width, temperature (T, K) and density (ρ, kg m^{-3}) , adapted from Figure 3.2 in Kivelson and Russell (1995).

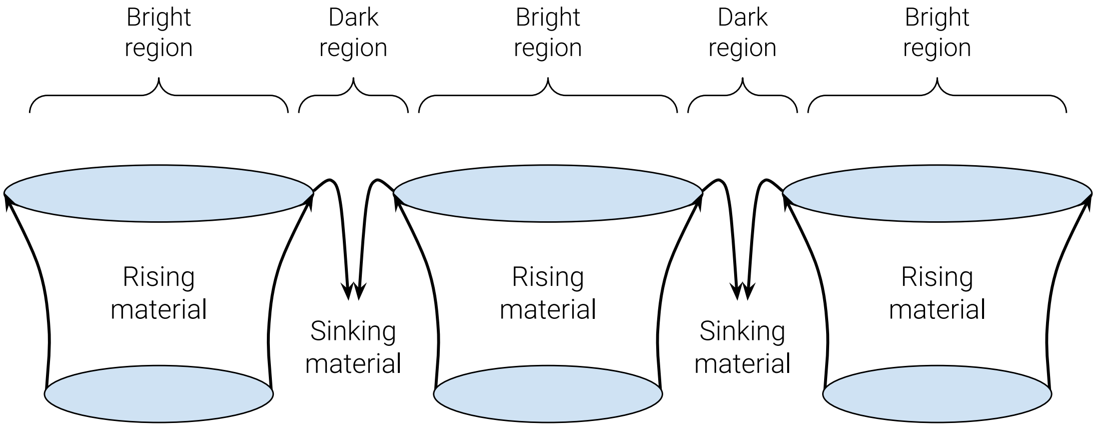


Figure 2.2: Formation of the granular pattern in the photosphere, adapted from Figure 3.4 in Meyer-Vernet (2007).

Darker regions are often seen in the photosphere in two bands centred on the equator. These are regions of intense magnetic field that inhibit convection, thus causing areas that are cooler than the surrounding plasma (Thomas and Weiss 2012). These are known as sunspots and contain a darker umbra surrounded by the filamentary penumbra, as seen in Figure 2.3 (Vaquero 2007). They typically appear in pairs of opposite polarity and the number of sunspots follow the 11-year solar cycle, as will be discussed in Section 2.2.

The photosphere does not rotate as a rigid body. The solar equator is seen to do a full rotation in ~ 25 days, whereas the poles rotate in ~ 33 days (Thompson et al. 1996). This is known as differential rotation, and was first recorded from observing the rotational periods of sunspots at different latitudes.

2.1.2 The chromosphere and solar corona

The chromosphere lies between the photosphere and the corona. Here, the temperature rises from around 4300 K up to temperatures of order 10^6 K in the transition region between the chromosphere and the corona (Kivelson and Russell 1995). The physical processes behind this temperature increase are debated and has led to the "coronal heating problem." The energy required to heat the corona comes from the interaction between plasma flow in the photosphere and the magnetic field (Klimchuk 2006). One theory suggests that the corona is heated through nanoflares, which are small, impulsive releases of energy (e.g. Cargill (2004) and Parker (1988)), covering the Sun. However, other competing theories, such as those suggesting magnetic waves heat the corona (e.g. Van Doorsselaere et al. (2020)), show that the dominant mechanism is not known and may be a combination of multiple theories.

The solar corona is the Sun's outer atmosphere, extending out to a few 10s of

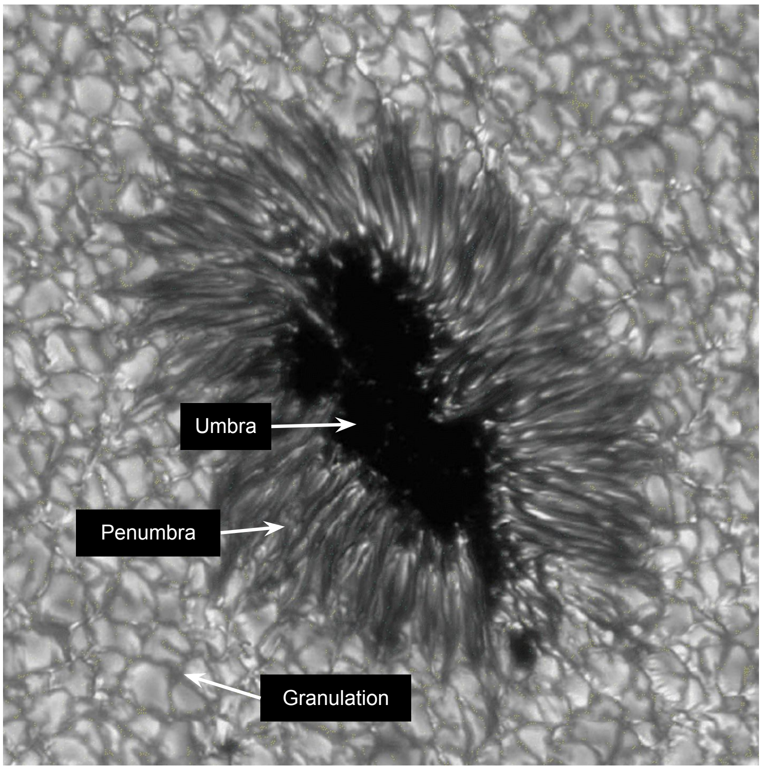


Figure 2.3: Image of a sunspot from Thomas and Weiss (2008) with the darkest umbra in the centre, surrounded by the filamentary penumbra. Granulation can be seen surrounding the sunspot. Image taken in the continuum near 436 nm by the Swedish Solar Telescope on La Palma.

solar radii ($R_S = 695,508$ km). It is optically thin and so can only be viewed using a coronagraph or during a solar eclipse, with the bright light from the Sun being blocked out in both cases as in Figure 2.4. The coronal structure and dynamics are dominated by magnetic pressure, making it a low-beta plasma (meaning that the ratio of magnetic to thermal pressure is low, $P_B >> P_{thermal}$). Therefore, processes within the corona are dominated by the changing coronal magnetic field (Priest 2020, Owens 2020a and references therein). The coronal magnetic field has its footpoints embedded in the photosphere, where they can either be described as ‘open’ or ‘closed’ (Schatten et al. 1969). Open magnetic field has only one footpoint of the magnetic field line rooted in the photosphere, with the line stretching out into the far heliosphere. Although named ‘open’, the other footpoint can either eventually return to the photosphere (possibly reaching the edge of the heliosphere), or may connect to the interstellar magnetic field (Schrijver and Siscoe 2009). However, it is open in the frame of the inner solar system. On the other hand, closed magnetic field has both footpoints rooted in the photosphere, thus creating a local loop. The open and closed magnetic field structure has important implications for the solar wind, as discussed in Section 2.4.

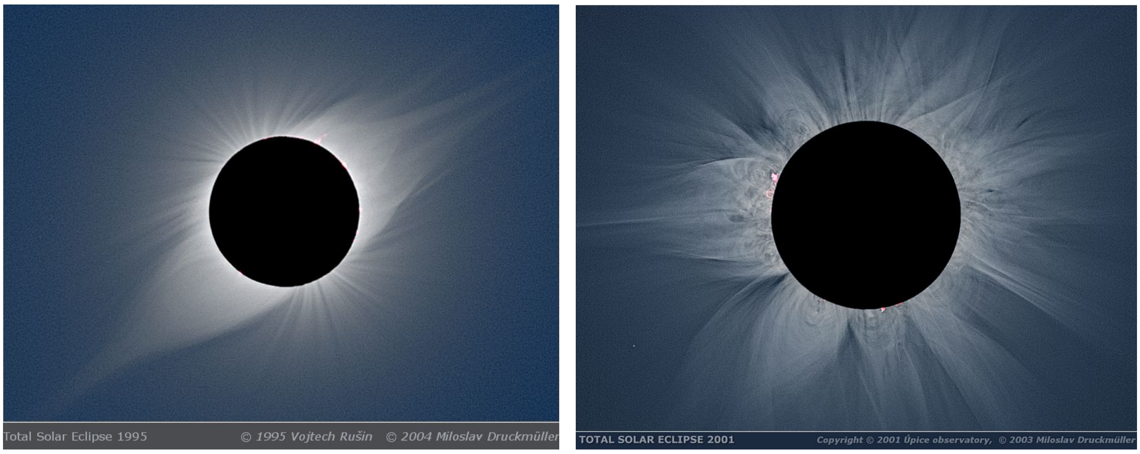


Figure 2.4: Left: image of a solar eclipse at solar minimum, taken on 24/10/1995 in India. Right: image of a solar eclipse at solar maximum, taken on 21/06/2001 in Zambia. Images taken in white light and available at `http://www.zam.fme.vutbr.cz/~druck/eclipse/`.

The corona contains a number of features, including active regions and coronal holes. Active regions are areas of strong magnetic field, with both footpoints rooted in the photosphere (Shibasaki et al. 2011). These footpoints are typically visible as sunspots, often appearing in pairs of opposite magnetic field polarity (Hale and Nicholson 1925). In the corona, active regions are visible as bright loops in extreme ultraviolet (EUV) and in X-ray, as seen in Figure 2.5 (Schrijver and Siscoe 2009).

Coronal holes fulfil two criteria; they are areas of low emission in ultraviolet and X-ray, and are regions of open magnetic field (Cranmer 2009). These serve as the source region of the fast solar wind and are associated with rapid expansion of open magnetic field (Zirker 1977). The low emission of UV and X-rays indicate regions lacking in hot material, which is consistent with the plasma being able to escape into space along open field lines (Wang et al. 1996).

The regions between these features are referred to as the quiet Sun. These are regions of closed magnetic field and are adjacent to coronal holes. Open magnetic field lines from coronal holes expand over quiet Sun regions to fill the volume above the closed magnetic field loops (Gibson 1973).

The structure of the corona varies throughout the solar cycle. Coronal holes are seen to form in polar regions during solar minimum, but will form at a broader range of latitudes during solar maximum (Cranmer 2009). The charged particles of a plasma act as tracers of magnetic field lines; therefore, imaging the solar corona gives an indication of the solar magnetic field. This is explained further in Section 2.3. Figure 2.4 shows two images of the solar corona, taken during a total solar eclipse. For white light images, such as those shown, the brightness is related to electron density, whereby the higher the density, the brighter the image (Cranmer et al. 1999). The left-hand panel shows the Sun during solar minimum. In this

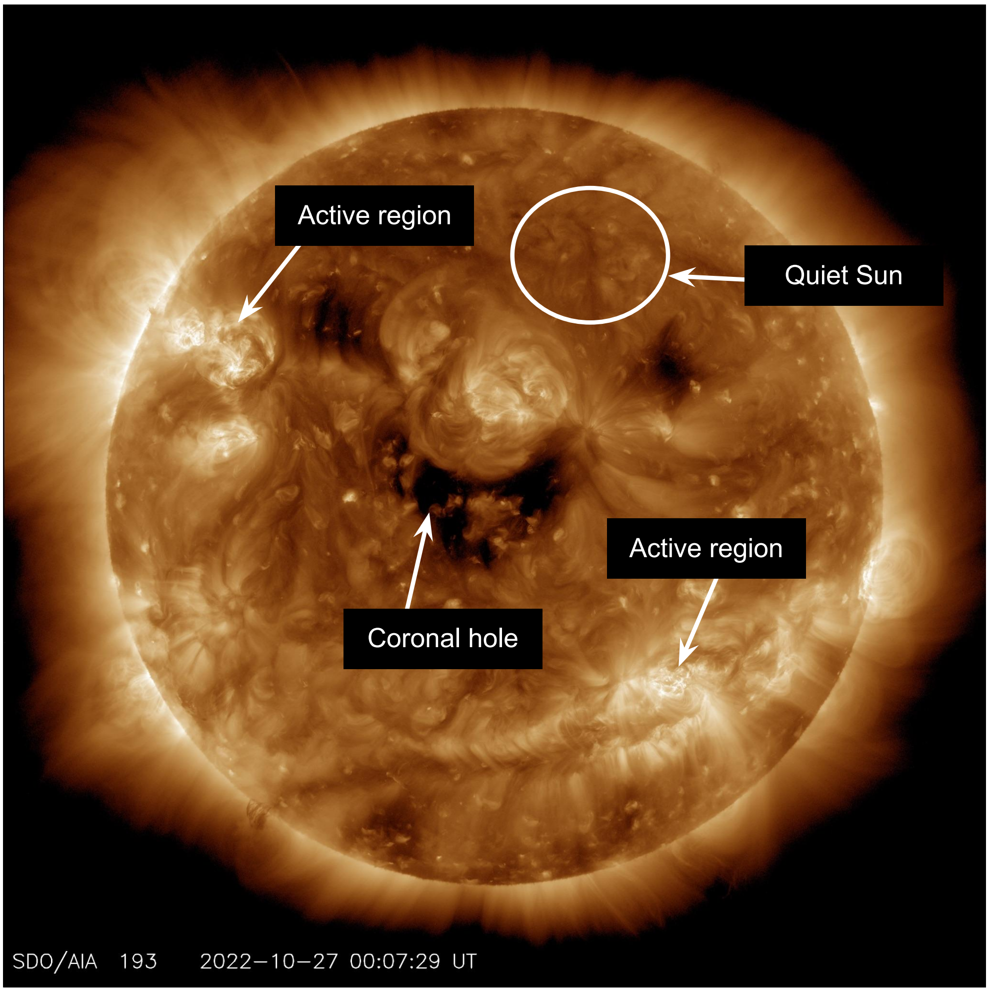


Figure 2.5: Image of the Sun in wavelength 193 Angstrom (19.3 nm, extreme ultraviolet) taken from NASA's Solar Dynamics Observatory (SDO) on 27/10/2022. The bright regions of the image correspond to active regions and the dark regions correspond to coronal holes. Image available for download at `https://sdo.gsfc.nasa.gov/data/aiahmi/`.

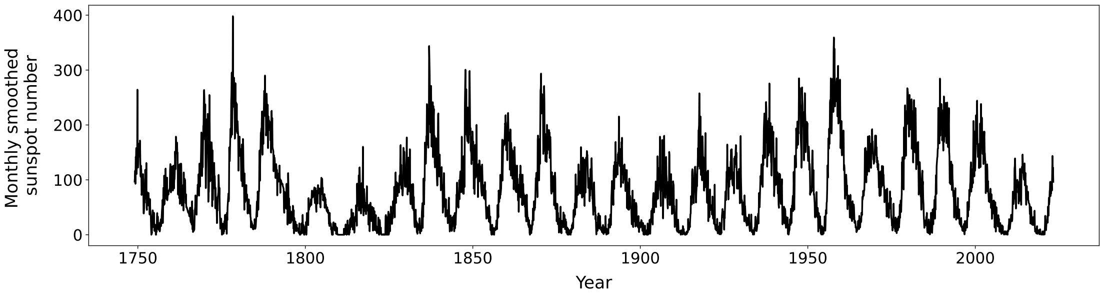


Figure 2.6: Monthly mean sunspot number from January 1749 to April 2023. The sunspot number is smoothed using a tapered-boxcar running mean over a 13-month period, centred on the corresponding month. Data is obtained from SILSO, Royal Observatory of Belgium and is available at `https://www.sidc.be/SILSO/datafiles`.

image, we can see two main streamers coming off the limb of the Sun. These are formed when the flux from coronal holes converge, with large streamers forming when the coronal holes have different polarity (Wang et al. 2000). Many streamers form in the magnetic equatorial region of the Sun, leading to the formation of the 'streamer belt'. The-right hand panel of Figure 2.4 shows the Sun during solar maximum, where the magnetic field is more highly structured and less ordered. Many streamers and pseudostreamers (formed from converging regions of magnetic field of the same polarity) can be seen.

2.2 The solar cycle

The Sun goes through an approximately 11-year activity cycle. Cycles begin at solar minimum, rising to solar maximum, where the Sun's magnetic polarity reverses, and ending at solar minimum (Hathaway 2015). There is an overall approximately 22-year cycle, also known as the Hale Cycle, for the magnetic field to return to its original polarity (Hale and Nicholson 1925). Throughout the 11-year cycle, the number of sunspots on the photosphere changes, as shown in Figure 2.6. The number of sunspots indicates the activity level of the Sun, with large sunspot numbers indicating higher activity levels (Hathaway et al. 2002). Observations of the sunspot number allows us to track the progress through the solar cycle. As the cycle progresses, sunspots form at decreasing latitudes, thus moving closer to the solar equator (Schrijver and Siscoe 2009).

The peak number of sunspots varies between cycles, with the Sun going through particularly quiet periods on decadal and millennial timescales (Meehl et al. 2013). The most notable of these was the Maunder minimum between 1645 and 1715, where there were very few observations of sunspots at all (Usoskin et al. 2015). The Dalton minimum, from 1790 to 1830, was also a period of low solar activity

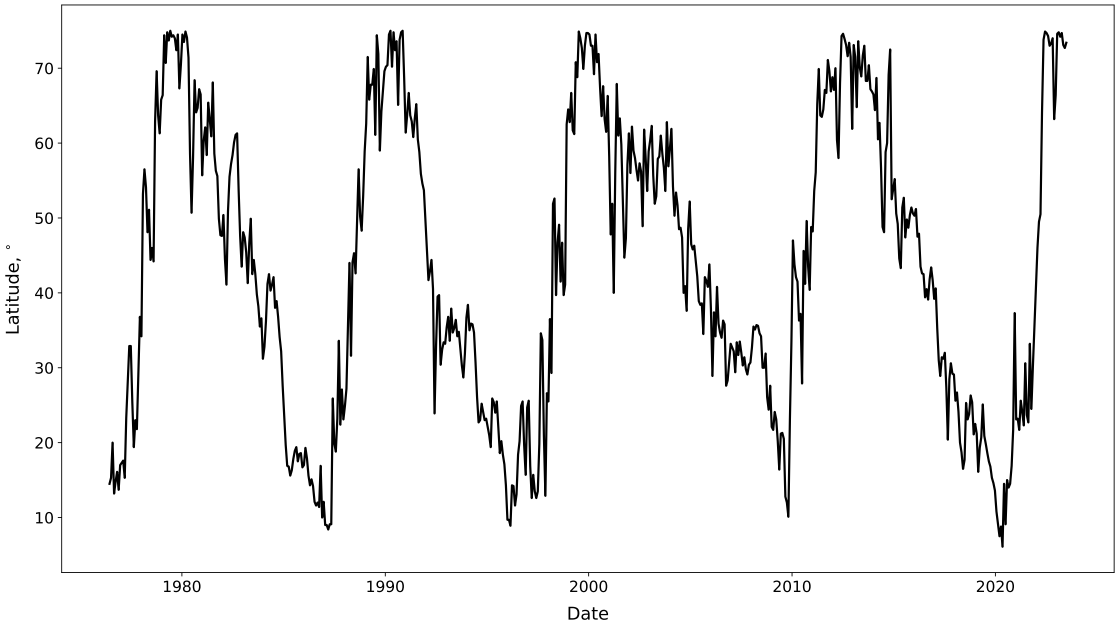


Figure 2.7: Variation of the average latitude of the heliospheric current sheet (HCS) from 1976 to 2023. Data available at `http://wso.stanford.edu/Tilts.html`.

(Silverman and Hayakawa 2021). It is speculated that we are currently entering a period of anomalously quiet activity, with the previous solar cycle (number 24, from 2008 to 2019) only reaching a maximum smoothed monthly mean sunspot number of 146.1 in February 2014, compared to a maximum of 244.3 from the cycle before (data from the Royal Observatory of Belgium). This is following a period of high solar activity, known as grand solar maximum from the 1940s to the early 2000s, which persisted for most of the space age (Lockwood et al. 2009).

Throughout the solar cycle, the magnetic field of the Sun changes. An example of a magnetic field at both solar minimum and solar maximum is shown in Figure 2.9. During solar minimum, the field is approximately dipolar, with a single north and south magnetic pole, approximately lining up with the rotational poles (Hathaway 2015). This can be seen in the left-hand panel of Figure 2.9. There are open magnetic field lines emanating from the north and south poles, with closed magnetic field centred around the equator. This leads to a band of slow solar wind being observed around the Sun's equator and fast solar wind observed towards the poles, as seen in Figure 2.8 (McComas et al. 2003). Conversely, the magnetic field during solar maximum (right-hand panel of Figure 2.9) is more dynamic and does not contain such latitudinal structure seen during solar minimum. As Figure 2.9 shows, the closed magnetic field loops are not confined to the equatorial regions. Due to this, slow solar wind is observed at a greater range of latitudes, as seen in Figure 2.8.

Where open magnetic field lines of opposite polarity converge, the heliospheric

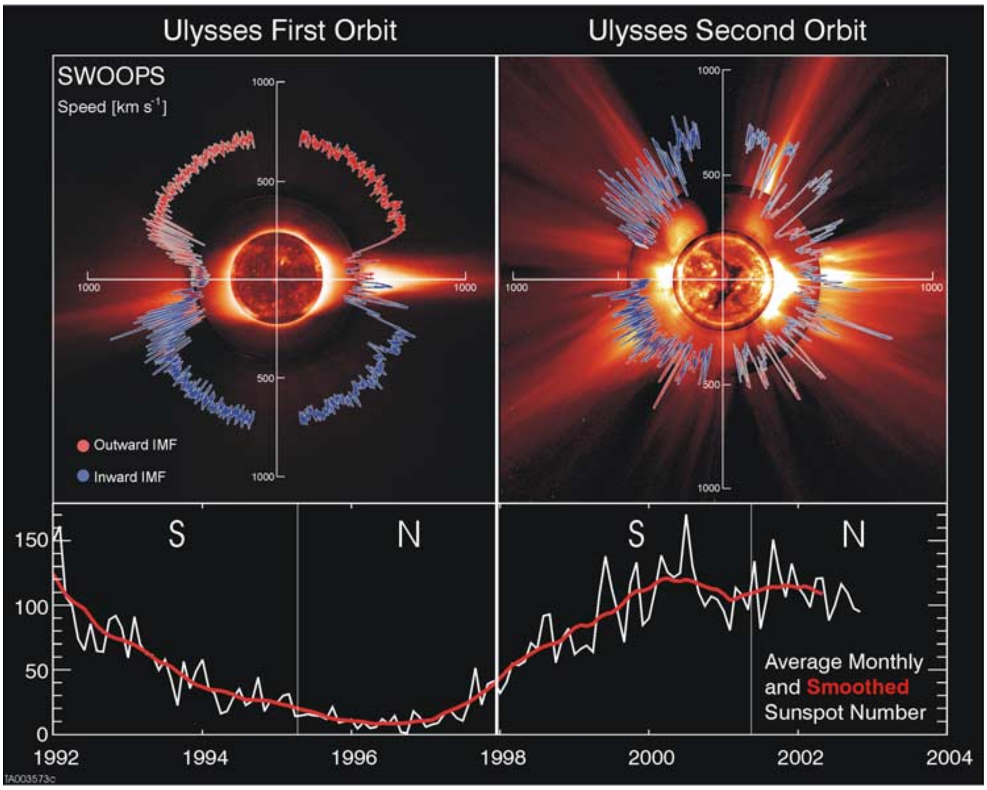


Figure 2.8: Polar plots of solar wind speed as a function of latitude, for declining phase to solar minimum (left) and solar maximum (right). The colour of the line indicates the direction of magnetic field, with red showing outward directed field and blue showing inward directed field. The bottom panel shows the average monthly sunspot number in white and the smoothed sunspot number in red. Figure from McComas et al. (2003).

current sheet (HCS) forms. This is draped throughout the heliosphere, approximately around the solar equator at solar minimum but at a range of higher latitudes during solar maximum (Smith 2001). The changing latitude of the HCS is shown in Figure 2.7. It acts as a magnetic equator separating the positive and negative magnetic field regions and can be seen as the green line in Figure 2.9.

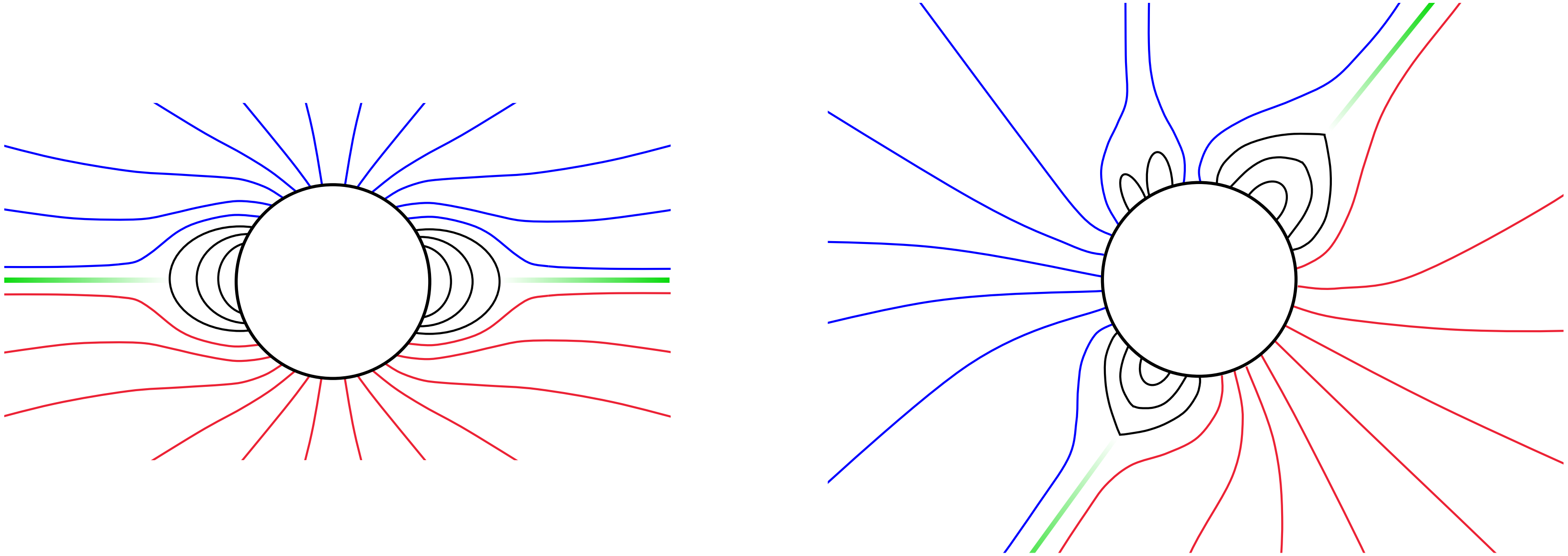


Figure 2.9: Example magnetic field of the Sun in the plane of the sky during solar minimum (left) and solar maximum (right). Blue and red lines show inwards and outwards polarity. Black lines show closed magnetic field loops. The green lines show the heliospheric current sheet. Adapted from Figures 2 and 4 in Owens (2020a).

2.3 Plasma physics

Here we discuss basic principles in plasma physics and why the solar wind can be treated as a fluid.

We can describe the motion of a single particle within electric and magnetic fields using the Lorentz force. Assuming the particle has a negligible effect on the field, this determines how a particle with charge q and velocity $\mathbf{v}$ moves in an electric and magnetic field:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = m\frac{d\mathbf{v}}{dt} \tag{2.1}$$

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field and m is the particle mass. However, solving for every particle within a plasma soon becomes computationally impossible on scales such as the solar wind (Kivelson and Russell 1995). Also, the electric and magnetic field become fixed, rather than responding to the motion of the individual particles (Priest 2012).

The Debye length describes how far the electrostatic effect of a single particle reaches (Verscharen et al. 2019). We call the sphere with a radius equal to the Debye length the Debye sphere, and this is the sphere of influence of the particle and so all particles outside of this sphere are screened from its effects (Xu and Chen 1999). If we are looking at scales greater than the Debye length, it would not make computational sense to solve for each particle individually. Treating a plasma as a fluid allows us to understand its average behaviour, but with no individual particles or information about particle distributions.

To describe a plasma in this way, we use magnetohydrodynamics (MHD). This is based on a set of equations combining the Navier-Stokes equations that describe a fluid and Maxwell's equations for electromagnetism (Sheikholeslami and Ganji 2016). This treats the plasma as a single conducting fluid and this allows us to look at it on spatial scales greater than the Debye length, such as the solar wind, in a more computationally efficient manner. MHD deals with the bulk properties of the plasma, such as bulk speed, temperature and density, effectively averaging over individual particles (Priest 2012). MHD is only valid for length scales larger than the Debye length and for densities greater than one particle per Debye sphere (Kivelson and Russell 1995). If these conditions are not met, the MHD assumptions break down and particles must be treated independently.

The MHD equations relate a plasma's density (ρ), velocity ($\mathbf{v}$), thermodynamic/ kinetic pressure (P) and magnetic field ($\mathbf{B}$) together. Here, we briefly describe the main equations that are relevant for the work within this thesis.

Mass continuity equation

The mass continuity equation describes that, in the absence of sources or sinks, an MHD fluid has a conservation of mass. It states that a change in mass can only be achieved if there is a flow of matter into or out of an element of volume:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{2.2}$$

Momentum equation

The momentum equation describes the motion of an MHD fluid due to all the forces that act upon it. The right hand side of the equation contains any force which may act upon a plasma. Here, we include the force exerted by gas pressure (first term), gravitational forces (second term) and the force exerted by the magnetic field (third term):

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{\nabla P}{\rho} - \frac{G M_S}{r^2} + \mathbf{J} \times \mathbf{B} \tag{2.3}$$

where G is the gravitational constant, M_S is the solar mass, r is the radial distance from the Sun, **J** is the current density and **B** is the magnetic field. Depending on the system being described, this expression can be simplified by approximating many terms on the right hand side to zero.

Induction equation

The induction equation (equation 2.4) relates the velocity and the magnetic field, describing how the magnetic field evolves. This is derived from Ohm's law and Maxwell's equations, with a full derivation found in Schrijver (2015).

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B} \tag{2.4}$$

where μ_0 is the magnetic permeability and σ is the electrical conductivity.

The right hand side is split into two terms; the convection term (first) and the diffusion term (second). The convection term describes how the magnetic field evolves with the plasma flow and the diffusion term describes how the magnetic field moves through the plasma. The ratio between these terms is known as the magnetic Reynolds number, R_m, and this allows for defining which regime a plasma is in. For $R_m << 1$, the diffusion term is dominant, whereas for $R_m >> 1$, the convection term is dominant. The term $\eta = \frac{1}{\mu_0 \sigma}$ is known as the magnetic diffusivity, and due to the high conductivity in space plasma's ($R_m >> 1$), this term is often very small. Therefore, by assuming this term is zero, we obtain:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) \tag{2.5}$$

This is known as ideal-MHD, and gives rise to the frozen-in flux theorem. This states that a fluid can move freely along magnetic field lines, but any motion perpendicular to the field results in the fluid moving the field lines. This allows us to understand the structure of the solar wind, as described in Section 2.4.

2.4 The solar wind

The solar wind is a constant stream of charged particles that continually flows from the Sun, dragging with it the Sun's magnetic field (Meyer-Vernet 2007). It is caused by an outward gas pressure gradient due to the very high temperature in the corona and the cold temperature of interplanetary space. First theorised by Eugene Parker in the 1950s, it was confirmed by later space missions, particularly the Mariner 2 mission to Venus (Parker 1958).

The solar wind is comprised mostly of protons and electrons, with a small proportion of heavier ions (Meyer-Vernet 2007). It is accelerated to supersonic speeds within a few solar radii (see Figure 2.10) and creates the heliosphere, the region of space dominated by the Sun's magnetic field. The dragging out of the Sun's magnetic field by the solar wind is explained by the frozen-in flux theorem (Section 2.3), whereby plasma and magnetic field are inextricably linked. Where frozen-in applies, the ratio of the thermal and magnetic pressure within a plasma, known as the plasma beta (β), determines whether the motion of the plasma impacts the magnetic field or vice versa (Schrijver and Siscoe 2009). For $\beta >> 1$, the plasma motion is dominant over the magnetic field, whereas for $\beta << 1$, the magnetic field is dominant over the plasma motion. In the case of the solar wind, the plasma beta is approximately 1, but the dynamic pressure from the plasma motion dominates over the magnetic and thermal pressure (Kivelson and Russell 1995). This means that the solar wind flow dominates over the Sun's magnetic field (past a few solar radii) and so it is dragged out into space.

The solar wind can be considered 'bi-modal' in nature, with streams of either "fast" or "slow" wind (Coles 1996). Fast wind emanates from regions of open magnetic field at speeds ranging from 400 to 800 km s^{-1} with a density of approximately 3 ions cm^{-3} at 1AU. Conversely, slow wind emerges from regions of closed magnetic field at speeds of 250 to 400 km s^{-1} with a density of 8 ions cm^{-3} at 1AU (Kallenrode 2004). The slow wind is generally confined to the solar equatorial region during solar minimum, but spreads to higher latitudes during solar maximum (McComas et al. 2003. Due to the footpoints of the magnetic field being rooted in the photosphere,

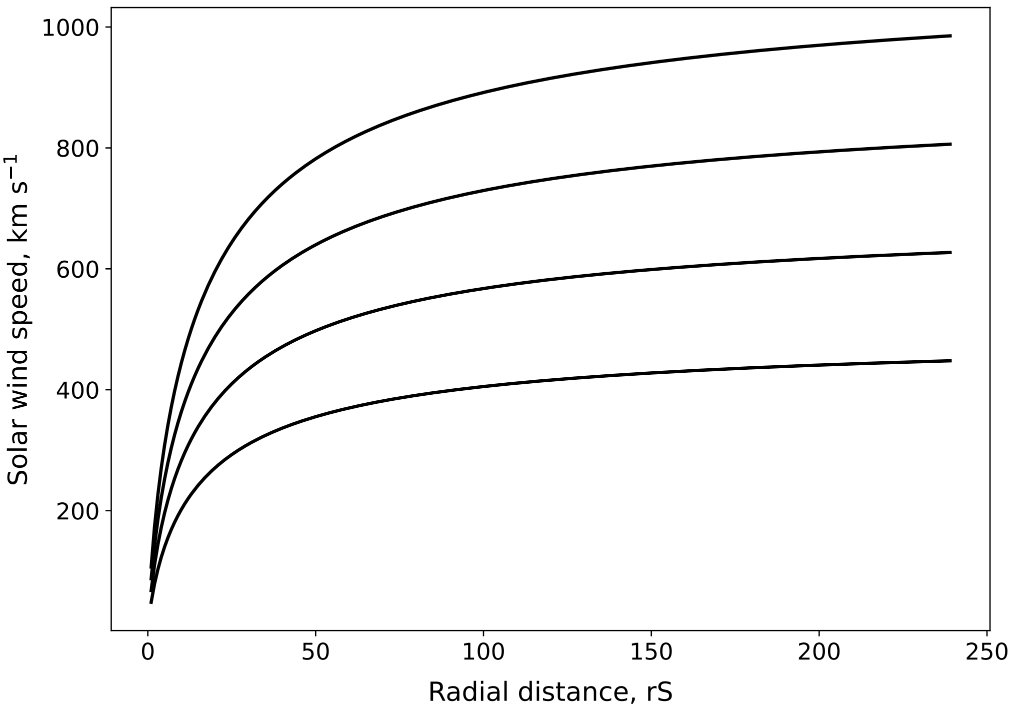


Figure 2.10: Radial variation of the solar wind for a range of speeds at Earth. Profile from the relationship derived in Macneil et al. (2021), based on a power-law fit to the Parker solar wind model by Burlaga (1967) and the acceleration term used by Owens and Riley (2017) in the Heliospheric Upwind eXtrapolation (HUX) model.

the heliospheric magnetic field is pulled into an Archimedean spiral as the Sun rotates (Parker 1958). As viewed from above the ecliptic plane (the plane where the planets lie) although solar wind flow is mostly radial, there will be sectors of fast and slow solar wind rotating, approximately steadily, with the Sun.

The solar wind forms an integral part of space weather, which is the changing conditions in space between the Sun and the Earth. Space weather can cause impacts on Earth itself, through stream interaction regions (SIRs, Section 2.5), but it is also the medium through which the Sun's transient ejecta propagates through (CMEs, Section 2.6). Therefore, knowledge of the background solar wind conditions is vital for accurate space weather forecasting.

2.5 SIRs and CIRs

Where fast solar wind catches up with slow wind, stream interaction regions (SIRs) occur. Due to the resulting compression, these are regions of higher plasma density and stronger magnetic field and can be a recurrent source of space weather

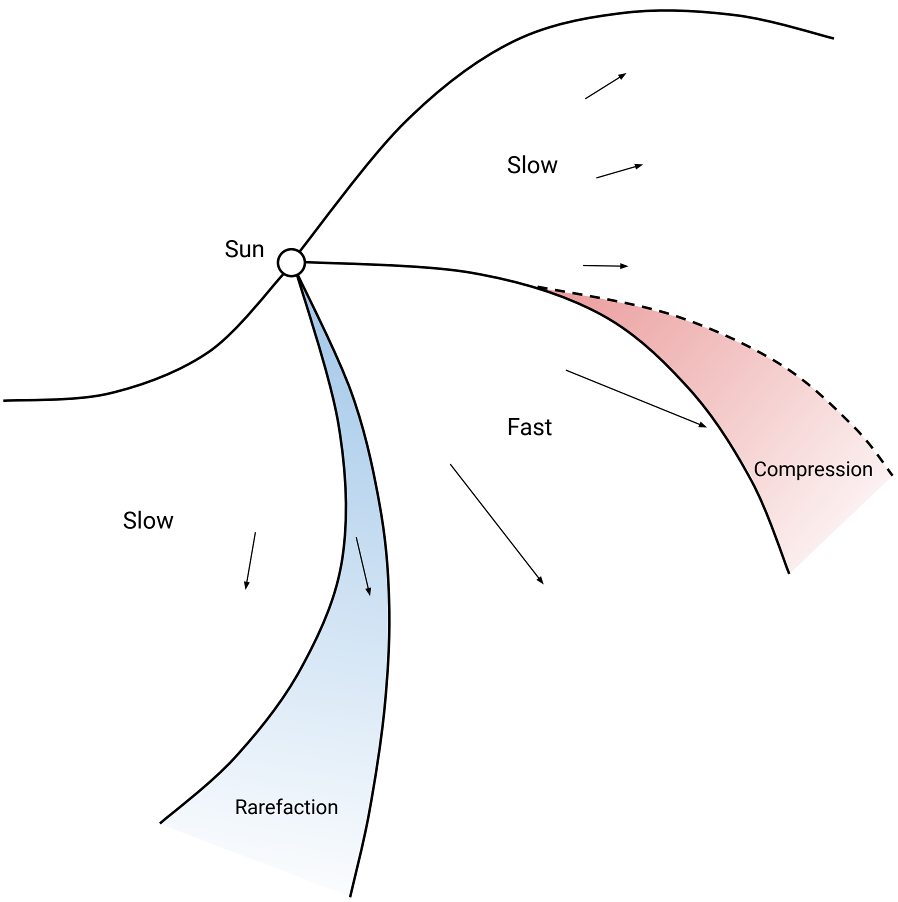


Figure 2.11: Formation of stream interaction regions (SIRs). Two slow solar wind streams are separated by a fast solar wind stream, creating a compression region (red) on the leading edge and a rarefaction region on the trailing edge (blue). The bulk solar wind flow is shown in the black arrows. Adapted from Parker (1965).

(Richardson and Cane 2012). With respect to Earth, the Sun completes a full rotation approximately every 27 days. If an SIR persists for more than one solar rotation, it becomes a corotating interaction region (CIR) (Gosling and Pizzo 1999). Where SIRs occur, we get regions of compression and of rarefaction, as seen in Figure 2.11. The region of compression occurs on the leading edge of the fast stream, as the faster moving plasma catches up with the slower plasma. The rarefaction regions occur on the trailing edge of the fast stream, where the faster plasma is moving away from the slower plasma behind it (Parker 1965). SIRs, and fast streams in general, are characterised by a sharp increase in solar wind speed followed by a more gradual decline towards slower speeds (Jian et al. 2006).

SIRs and CIRs can be a source of space weather, with CIRs being a cause of recurrent space weather with a 27 day period. The denser regions and stronger magnetic field act to transfer energy into the Earth's magnetic field, thus causing geomagnetic storms and impacts on the surface (see Section 2.7).

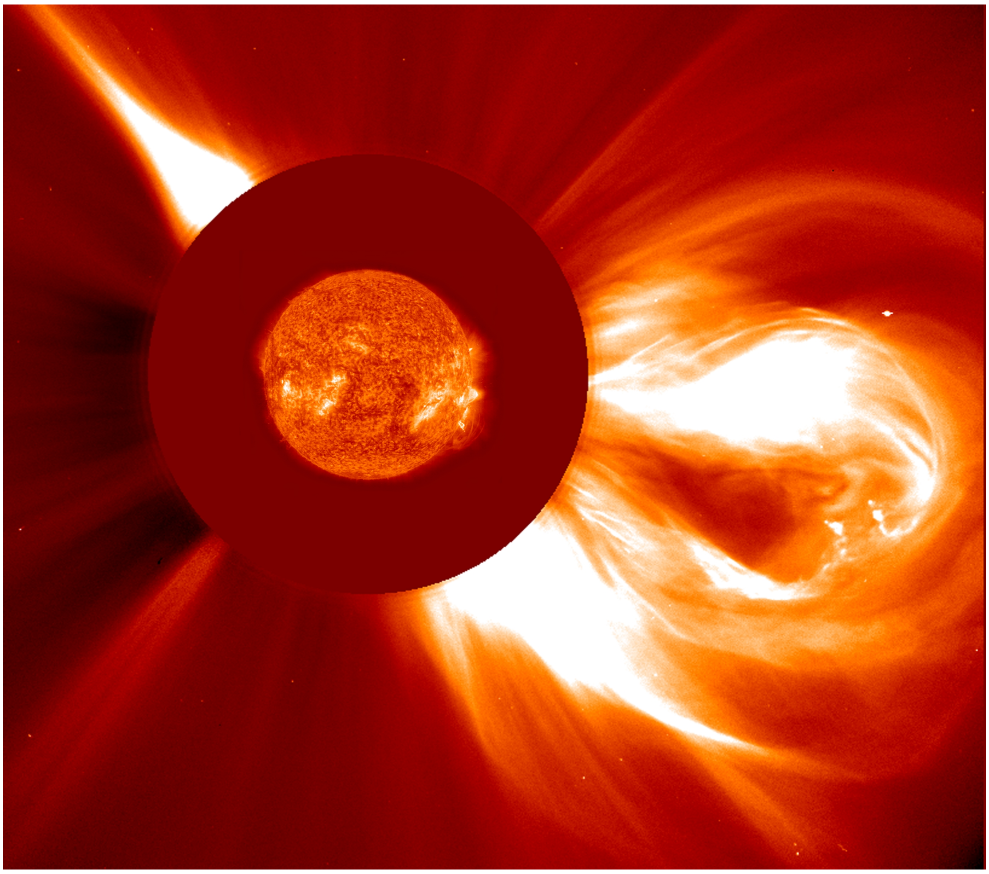


Figure 2.12: Image of a CME taken from the SOHO (Solar Heliospheric Observatory) spacecraft on 02/12/2003. This is a composite image from two instruments, with the central image corresponding to a wavelength of 304 Angstroms showing the upper chromosphere and the outer image taken from a coronagraph. Image courtesy of NASA, found at: `https://soho.nascom.nasa.gov/gallery/images/20031202c2eit304.html`.

2.6 Coronal mass ejections

Coronal mass ejections (CMEs) are transient eruptions of coronal plasma and magnetic field into interplanetary space, as seen in Figure 2.12 (Forbes 2000). Containing up to $10^{12} - 10^{13}$ kg of matter, they can travel in excess of 1000 km s^{-1} and so reach the Earth in a few days (Meyer-Vernet 2007). The occurrence rate of CMEs varies with the solar cycle, with approximately one per week at solar minimum and up to three per day during solar maximum (Hundhausen 1999). CMEs propagate through the solar wind, with the solar wind conditions affecting a CME's speed and arrival time. A CME travelling faster than the surrounding solar wind will tend to be slowed down as it moves through space, whereas a CME travelling slower will be accelerated (Gopalswamy et al. 2001; Vršnak et al. 2013). Therefore, knowing the ambient solar wind conditions allows for better modelling and forecasting of CMEs propagating through interplanetary space.

Figure 2.13 shows the current understood structure of a CME. Here, the CME

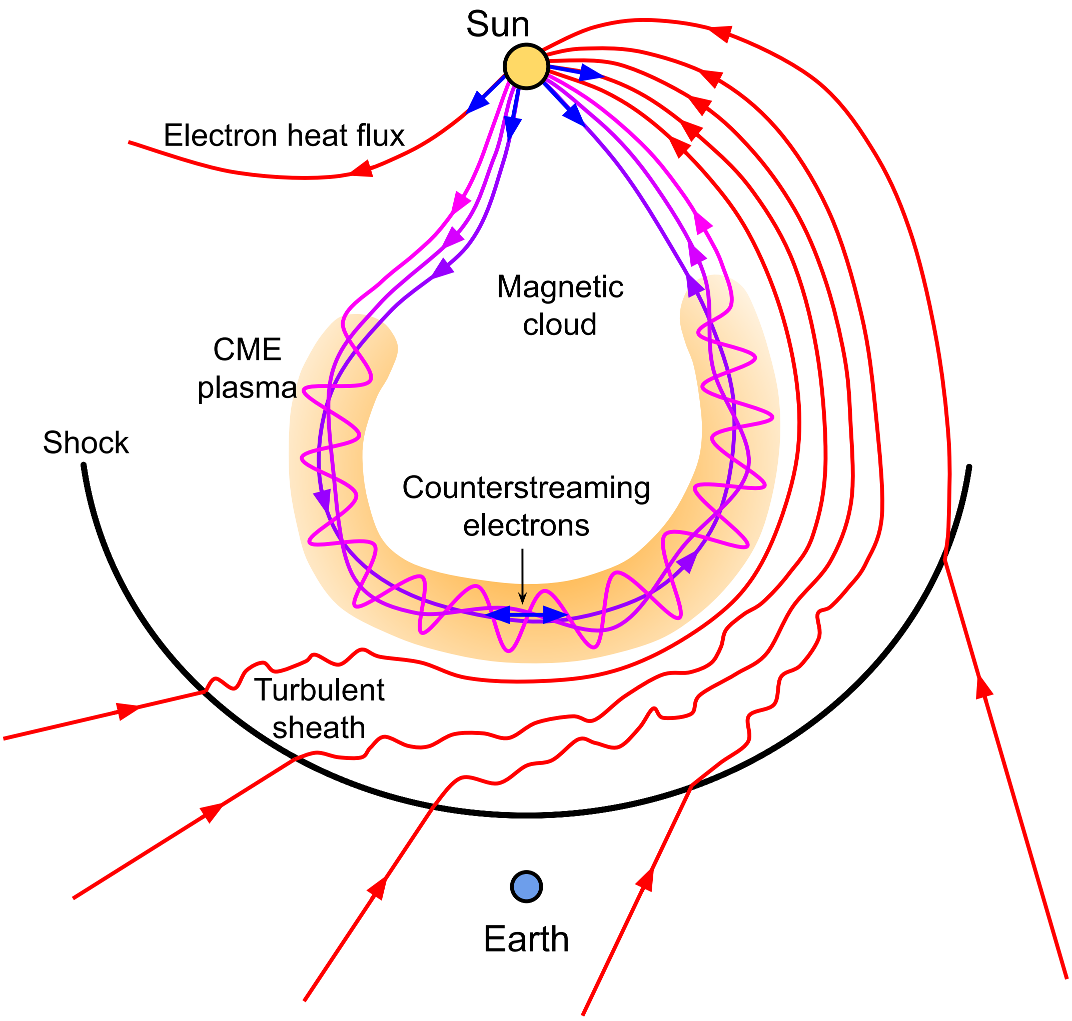


Figure 2.13: Diagram of the structure of a CME. The pink field lines show the magnetic field associated with the CME and the red field lines show the ambient magnetic field. The blue arrows show the field aligned electrons that are accelerated by the CME. The black curve shows the location of the CME shock. Original schematic from Zurbuchen and Richardson (2006) and updated with information from Owens (2016).

propagates radially away from the Sun, preceded by a shock (black curve), which causes turbulence in the heliospheric magnetic field. The shock wave is generated as the CME is moving faster than the surrounding solar wind, and it can accelerate particles ahead of the CME. The CME plasma follows, surrounding the twisted flux rope structure of the CMEs magnetic field. The direction of the CME magnetic field is a major determining factor of the impact the CME has on Earth. If the magnetic field is directed southward (negative in the z-direction when in geocentric solar magnetospheric (GSM) coordinates, i.e. pointing down into the northern magnetic pole of the Earth), then this interacts with the northward directed magnetosphere producing more severe geomagnetic storms.

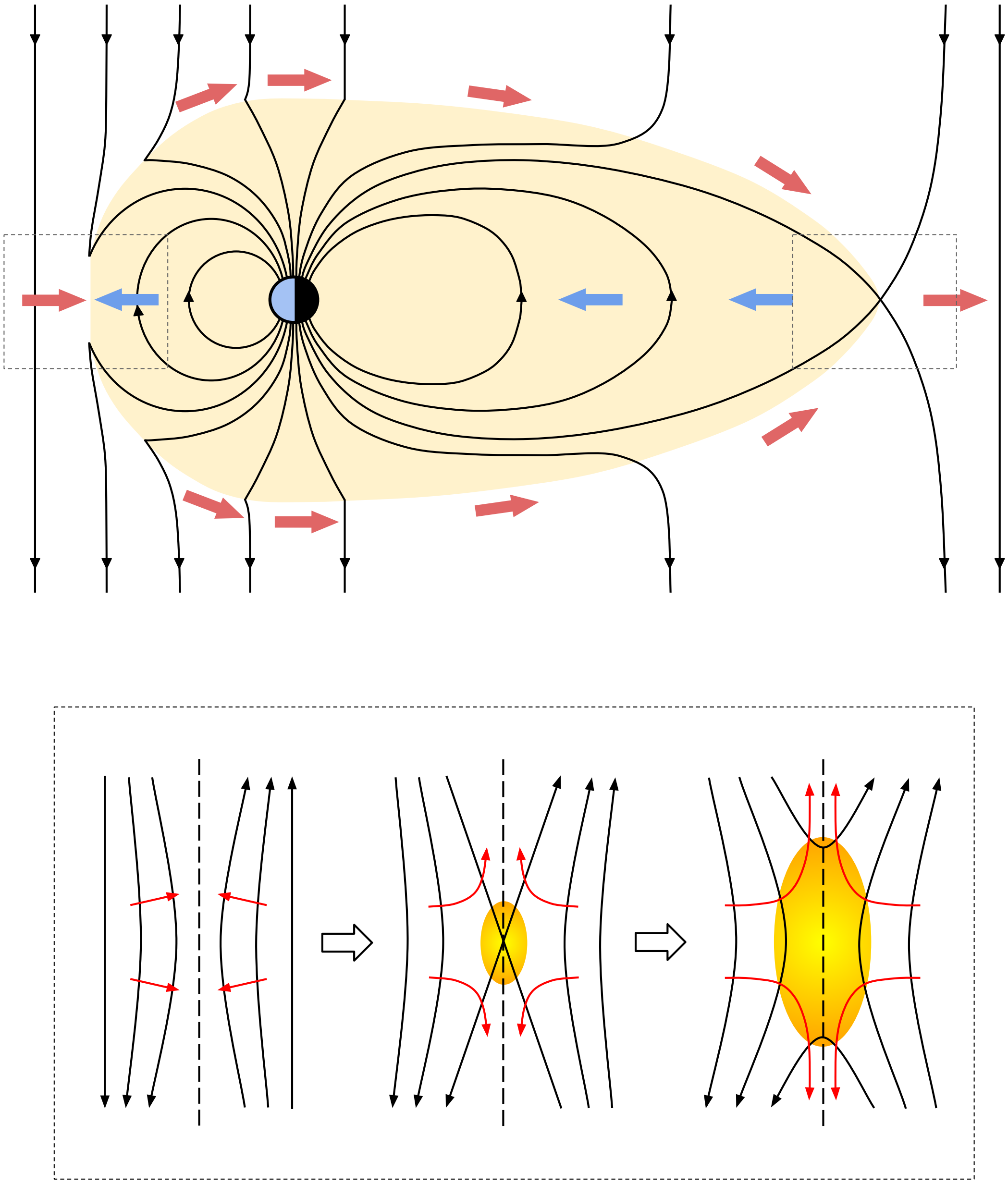


Figure 2.14: Diagram of magnetic reconnection in the Earth's magnetosphere with a close up of the reconnection process. All black arrows show magnetic field lines. Top: the interplanetary magnetic field is directed southwards and the Earth's magnetic field is directed northwards, with the Sun to the left of the Earth. The interplanetary magnetic field lines move around the magnetosphere (light orange region) as per the red arrows, and the Earth's magnetic field moves as per the blue arrows. Bottom: close up diagram of the reconnection process. The red arrows show the movement of plasma into and out of the reconnection region. The orange ovals indicate regions where the plasma has been heated. The top schematic is adapted from Figure 5.4 in Treumann and Baumjohann (1996) and the close up panel of reconnection can be found at `https://wwwmpa.mpa-garching.mpg.de/HIGHLIGHT/2002/highlight0202_e.html`.

This interaction is known as magnetic reconnection. Reconnection occurs when two magnetic field lines with some anti-parallel component (commonly of opposite polarity) come into contact due to plasma flow (Schrijver and Siscoe 2009). They connect in a localised region and reconfigure into a new topology, as shown in Figure 2.14. Magnetic reconnection can only take place if the frozen-in flux theorem, and therefore ideal-MHD, breaks down. This will only occur for low magnetic Reynolds numbers, i.e. $R_m << 1$. If the frozen-in flux condition holds, plasma cannot cross magnetic field lines (Kivelson and Russell 1995). In the local regions that reconnection occurs, the length scales involved are too small for frozen-in to be valid and the diffusion term of equation 2.4 becomes dominant. This means that the magnetic field will diffuse from high to low values, thus allowing reconnection to occur.

When the interplanetary magnetic field is southwards and is presented by the northwards directed magnetic field of the Earth, reconnection occurs. This acts to peel back field lines from the nose of the magnetosphere and into the tail, whereby the field lines reconnect again (Dungey 1961). Reconnection is one source of accelerated particles. The plasma flows bring the field lines together, is locally heated as the magnetic field topology changes. Particles are then accelerated perpendicular to the original flows, as shown by the red arrows in the lower panel of Figure 2.14 (Guo et al. 2015). These particles can be funnelled towards the Earth's poles, creating the aurora.

CMEs are a driver of severe space weather. As well as the large amount of material and magnetic field within the CME, fast CMEs can also generate a shock which can accelerate particles, causing impacts on Earth. Fast CMEs are the main drivers of geomagnetic storms on Earth, and so forecasting them is highly important (Gosling 1993).

2.7 Space weather effects

Although mostly protected by the magnetosphere, space weather can have a large impact on Earth, particularly on our technology-focused way of life. The most extreme space weather events are driven by fast, Earth-directed CMEs. These generate large geomagnetic storms (Section 2.7.1) and can cause a number of effects (Section 2.7.2).

2.7.1 Geomagnetic storms

Geomagnetic storms are a temporary disturbance in the Earth's magnetosphere caused by a transfer of energy from the solar wind, for example by an SIR or CME, into near-Earth space (Gonzalez et al. 1994). The largest geomagnetic storms are driven by CMEs alongside a southward directed interplanetary magnetic field (Gosling 1993).

There are a number of different metrics to categorise a geomagnetic storm, here we use the Dst (Disturbance-storm-time) index. This is an hourly measured, globally averaged value of the strength of the equatorial ring current (Sugiura 1963). The ring current is an electrical current that encircles the Earth and is formed from trapped charged particles in the Earth's magnetic field, caused from the interaction between the solar wind and the magnetosphere (Daglis et al. 1999). The strength of this ring current is related to the magnetic perturbations within the magnetosphere, so an enhanced ring current leads to a decrease in the magnetic field on the Earth's surface, measured using magnetometers. The greater the perturbations in the magnetosphere, the stronger the geomagnetic storm. During normal times, the Dst measures between +20 nT and -20 nT, however during extreme geomagnetic storms the Dst can reach lower than -250 nT (Palacios et al. 2017).

The largest recorded geomagnetic storm occurred in September 1859 and is known as the Carrington event, named after Richard Carrington, a British astronomer who observed the event. The event peaked at a Dst of $-900\,nT$ and is likely to have been caused by a CME (Cliver and Dietrich 2013). The event caused aurora at latitudes as low as central Mexico and caused failures in telegraph systems in Europe and North America.

2.7.2 Impacts on Earth

Aurora

The aurora are the most visible impact of space weather, occurring in the Earth's polar regions. They occur due to precipitating particles in the upper atmosphere colliding with native species, making it glow (Jones 2012). These particles can come from a number of sources, including magnetic reconnection in the magnetosphere tail (Figure 2.14), which are funnelled in towards the poles. This means that the aurora occur in a band circling the magnetic poles, known as the auroral oval. This normally sits between 10 and 20° in latitude, but can expand to lower latitudes during intense geomagnetic storms (Feldstein 2016). The aurora also appear brighter during geomagnetic storms, due to an increased number of particles transferred into the magnetosphere.

Power grid effects

Space weather causes rapid fluctuations of the Earth's magnetic field which can induce an electric field on the surface (Boteler 2001). This can then induce electric currents in the ground, known as geomagnetically induced currents (GICs) (Cannon 2013). These will flow through conducting material, such as power grids, railway lines and long pipelines. Due to magnetic effects from GICs within transformers, these can be shut down and in severe cases, serious damage can occur (Girgis and Vedante 2012). This can lead to local blackouts in moderate cases and national blackouts in severe space weather events. For a Carrington-type space weather event, it is estimated that six grid transformers in England and Wales and seven in Scotland could be damaged and shut down (Cannon 2013).

The largest geomagnetic disturbance recorded in the twentieth century occurred in March 1989, reaching a minimum Dst of $-589\,nT$, and was a result of two CMEs (Boteler 2019). This event led to a number of impacts, most notably on power systems. The Hydro-Québec power station experienced large scale instabilities and eventual system collapse due to an excess of voltage in the system. This resulted in over 9 hours of power outages for some residents and a total cost of \$13.2 million in damages (Bolduc 2002). It is estimated that if such an event were to happen in more current times, this would cost \$2.4 to \$3.4 trillion over one year (Schulte in den Bäumen et al. 2014).

Smaller scale events can cause degradation in equipment and lead to smaller-scale power surges and local black-outs. It was estimated that the impacts of mild to moderate space weather events on the power grid would cost \$1.3 - 2.1 trillion in the EU and US over a century (Schrijver 2015).

Spacecraft effects

Space weather can have a number of impacts on spacecraft, depending on their orbit and design. Extreme events can expose spacecraft to levels of particle flux and radiation that is outside of their design remit. Spacecraft charging, particularly those in geostationary and medium Earth orbits (GEO and MEO respectively) occurs from electrons in the space environment.

Low energy electrons ($<$ 100 keV) can differentially charge the surface of a spacecraft and cause discharges (Cannon 2013). High energy electrons ($>$ 1 MeV) can cause deep dielectric charging, whereby the insulating material between components becomes embedded with these electrons (Baker et al. 1987). This builds up internal electric fields and can lead to a discharge, which can be spontaneous or triggered by, for example, impacts from heavier ions or meteorites (Lai et al. 2018). Such a discharge can cause damage to the insulator and possibly cause the spacecraft to no longer function. If there is sufficient warning of such an electron event, the spacecraft can be powered down to allow the electrons time to diffuse out of the insulating material.

Cosmic rays and high energy ions can cause single event effects (SEEs). These occur due to the deposition of charge in sensitive microelectronics either through direct ionisation or through nuclear interactions (Cannon 2013).

The cumulative radiation dose from the space environment can prematurely age a spacecraft. It can slowly degrade components and electronics and thus weaken a spacecraft, making it more susceptible to damage from an extreme event (Ecoffet 2013). Therefore, a spacecraft may fail before it reaches the end of its planned operational lifetime.

Currents induced in the ionosphere during a space weather event can heat the upper atmosphere, causing it to expand (Yiğit et al. 2016). This can then cause some spacecraft to orbit within the upper atmosphere, thus increasing the drag that they experience. This can lead to orbital decay and even cause spacecraft to burn up, as in the February 2022 Starlink launch. 49 spacecraft were launched into low earth orbit (LEO), in preparation to be raised to an operational altitude once they were confirmed to be working. Due to a CME, combined with the LEO altitude, the spacecraft experienced higher than expected levels of drag, leading to 38 of the spacecraft burning up to a cost of more than $20 million for the spacecraft and $30 million for the wasted launch (Baruah et al. 2024; Lockwood et al. 2023).

Human health effects

While mostly protected by the Earth's magnetosphere and atmosphere, humans operating at higher altitudes are at a higher risk of radiation effects than those on the surface (Wilson 2000). Radiation in space can be in the form of electromagnetic radiation or particles and can have enough energy to be ionising. Therefore, the radiation can damage or destroy living cells or can damage the DNA within a cell. Depending on the type and strength of the radiation, this can cause a reduction in white blood cells, nausea, hair loss, cancer or death in humans (Moldwin 2022).

The crew of aircraft and frequent flyers are exposed to higher radiation doses than those who remain on the surface (Wilson 2000). Radiation increases with both altitude and latitude, therefore trans-polar flights are of higher risk than equatorial flights (Hwang et al. 2010). It is beneficial to fly at higher altitudes, as aircraft have decreased fuel consumption, experience less turbulence and can fly faster (Lovegren and Hansman 2011; Storer et al. 2017). However, this has to be balanced with the radiation doses experienced by crew and passengers. The International Commission on Radiological Protection recommends an effective dose of 20 mSv per year, averaged over 5 years, with no single year exceeding a total of 50 mSv for occupational exposure (ICRP 2007). For the public, this reduces to an effective dose of 1 mSv per year, averaged over 5 years. For a typical flight, radiation exposure ranges from 0.0001 mSv to 0.0644 mSv (Mohler 2003). However, for a solar proton event, whereby there is an emission of high-energy protons emitted from the Sun, this value increases. From a study of 169 solar proton events, for flights at an altitude of 30, 40, 50 and 60 kFt, the radiation dose was found to be 0.10, 0.35, 0.87 and 1.7 mSv per year respectively (Copeland et al. 2008).

Astronauts also have an elevated risk of radiation effects. Those in low Earth orbit (LEO), such as the International Space Station (ISS), are provided some protection from the magnetic field and the atmosphere. Astronauts will routinely wear radiation measuring devices and provide blood samples to measure radiation damage within chromosomes (Cucinotta 2014). The radiation levels are constantly monitored in many locations within the ISS, identifying the most shielded region of the spacecraft for radiation events (Rask et al. 2008). It is also important to monitor the radiation conditions in near-Earth space to time any extra-vehicular activity outside of any high radiation levels to maximise astronaut safety. Astronaut time in space is also limited, with most crew members staying on the ISS for 3 - 6 months to help limit radiation exposure (Cucinotta 2014).

Missions further afield, such as to the moon or to Mars, lose the protection from the magnetic field and the atmosphere. For a mission to Mars, most of the received radiation dose is during the transfer to the planet, as the dose whilst on Mars can be reduced through shielding (Durante 2014).

Communications and GPS failures

Space weather can cause disruption to communications and GPS systems. Many systems use GPS for navigation, including car satnavs, ships and aircraft, and for synchronising timing. GPS signals are transmitted through the ionosphere, which can be disturbed during a space weather event. This acts to change the amplitude and phase of the signal from the satellites, thus making navigation systems lose lock on the satellites and fail (Cannon 2013). These systems are also vulnerable to complete failure of the satellites providing the signal, as described above.

Long distance, high frequency radio communications are used by aircraft to maintain communication with the ground. These signals are transmitted across long distances through reflection off the ionosphere (Gurevich and Tsedilina 1985). During space weather events, the ionosphere can become disturbed leading to a loss of communications. There are procedures in place to follow if an aircraft is in the air and communications are lost; however, plans are lacking when it comes to more severely disturbed environments that would render communications down for a considerable length of time (Cannon 2013).

Chapter 3

Methodology

Here, we cover the necessary information to understand the methods used throughout this thesis. Section 3.1 covers the instrumentation and spacecraft missions that produces the data used in this work. Section 3.2 describes the current state of solar wind modelling and Section 3.3 covers data assimilation (DA). The DA scheme used for the work in thesis is described in Section 3.4.

3.1 Data

Throughout this work, observations of solar wind speed are used from three spacecraft missions. Below, we describe how the plasma detectors on board these spacecraft work and briefly outline the spacecraft missions. The aim for a plasma detector is to measure the particle distribution function, which gives the number of particles per unit volume per unit velocity space, given by $f(x, y, z, t, v_x, v_y, v_z)$. Solar wind speed can then be derived from this, and this is described below. As solar wind speed is the only type of observation used in the work in this thesis, we only consider its derivation here.

3.1.1 Plasma detectors

Faraday Cup

Faraday Cups measure the solar wind by measuring the discharge of ions that are collected within the instrument. Figure 3.1 shows a simplified diagram of a Faraday cup. Ions enter the ‘cup’, which contains an earthed outer shell and an inner shell, whereby they adhere to the inner shell (Brown and Tautfest 1956). The inner is discharged through a resistor with precisely known resistance and this discharge current is measured using a very sensitive ammeter. The minimum energy of the charged particles adhering to the cup is set by the potential difference of the cup. This is caused by the applied voltage, with all particles of energy above a threshold

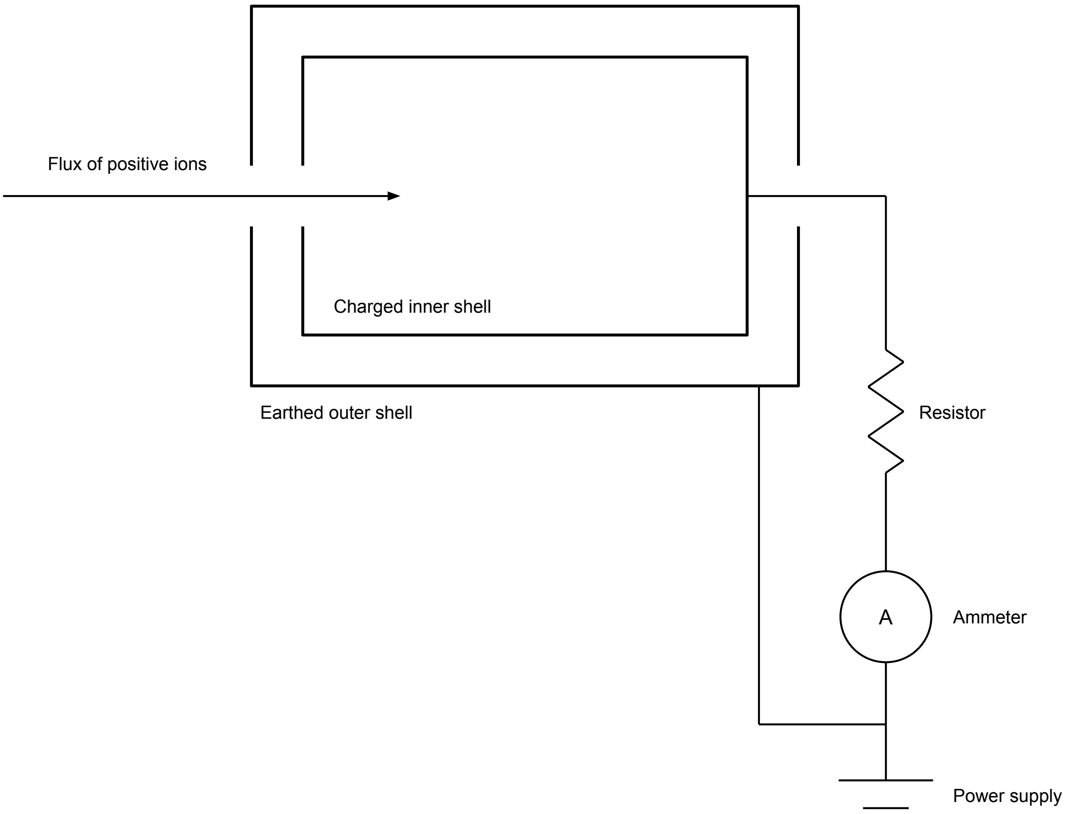


Figure 3.1: Diagram of a Faraday cup. Positive ions enter the cup and adhere to the inner shell, where they discharge through a resistor of known resistance. The current is measured using a sensitive ammeter and the particle flux is determined. A potential difference is applied to the cup, which sets the minimum energy of the entering particles. This potential difference can be cycled through to sample particles of all energies. Image from Lockwood (2024) lecture notes.

potential being able to enter and those below being repelled (Brown and Tautfest 1956). The applied potential difference can be cycled through, to detect particles of all energies and measure the flux of particles. To measure particles in all directions, multiple detectors pointing in different directions can be used, or the instrument can be placed on a spinning spacecraft (Ogilvie et al. 1995). The particle distribution function can be inferred from the flux of particles in the instrument. It is important that the instrument is shielded from electrons as they would reduce the discharge current and the cup is made of correct material so that electrons are not generated when the ions hit the cup.

Electrostatic analyser

Electrostatic analysers only allow particles with a specific and limited energy range to reach the detector. The instrument is comprised of an entrance colimator (a

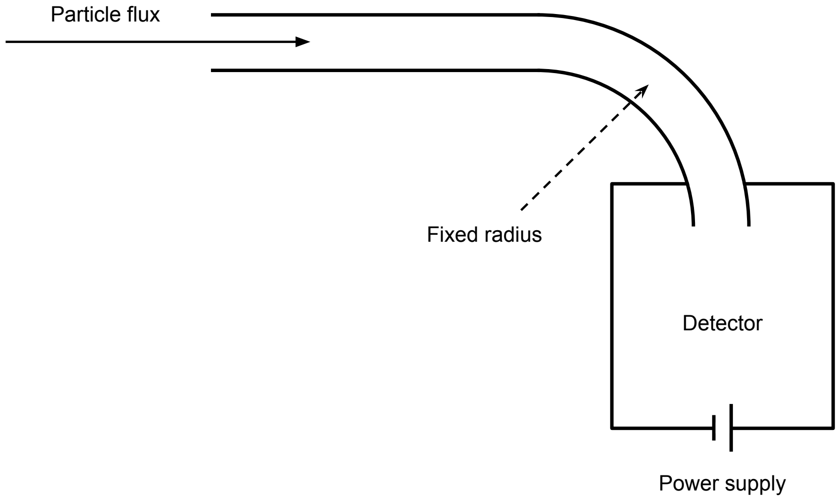


Figure 3.2: Diagram of an electrostatic analyser. Particles enter through a straight channel into a curved channel with a fixed radius and an applied potential difference. The particles paths bend depending on the size of the potential difference, with only certain particle energies bending the precise amount required to make it through to the detector. The particle flux, at a certain energy, is measured by the detector. The applied potential difference can be cycled through to sample the full range of particle energies. Image from Lockwood (2024) lecture notes.

straight channel) and a curved channel with a fixed radius between two plates, to which a potential difference is applied (Barabash et al. 2007). This is shown in Figure 3.2. The potential difference curves the path of the particle due to the creation of an electric field between the plates. The size of the potential difference determines the amount of curvature of the particle path, which is set by the particle's energy. This means that particles with too much energy for a given applied voltage between the plates will not curve enough and not be detected, and particles with too little energy will curve too much (Barabash et al. 2007). The particle flux is then measured by a detector at the end of the curved channel. The potential difference can be cycled through to measure a full spectrum of particle energies and to infer the particle distribution function.

3.1.2 Deriving solar wind speed

To determine the bulk properties of the solar wind, we take the 'moments' of a plasma's particle distribution function. To take a moment, the integral of the distribution function is taken. A 3-dimensional particle distribution function is in both space and velocity space, meaning that we have the number of particles per unit volume (x, y, z) per unit velocity space (v_x, v_y, v_z). The first integral (zeroth moment) under this function gives the particle number, which can be used to find the density. The mean of the velocity distribution (first moment) is the bulk velocity, and its standard deviation (second moment) gives the temperature (Baumjohann and Treumann 2012).

Data measured by a spacecraft can either be processed on board for a real-time data feed or transferred to the ground for more sophisticated processing. Due to spacecraft size and weight, the processing power on board is limited, so deriving solar wind speed from the particle distribution function will not be optimal. This means that the real-time data will often contain data gaps, biases and anomalous results.

The alternative to this is for the full particle distribution function to be downloaded at each time step. The fitting can then be completed on the ground, where the computing power is much greater and a more optimal fit can be achieved. This produces a more complete data set, known as ‘science-level data’. However, this takes considerable time and data may not be available until many days or weeks after it has been measured. Although the science data can be used for verifying methods, the real time data feed must be used for operational forecasting.

3.1.3 Spacecraft missions

STEREO

The Solar Terrestrial Relations Observatory (STEREO) was launched in 2006 and was a twin mission, consisting of the Ahead and Behind (STEREO-A and STEREO-B respectively) spacecraft. They were launched into Earth-like orbits, but with STEREO-A ahead in longitude and STEREO-B behind, separating at a rate of 11° per year from Earth. The mission aimed to investigate CME initiation and to track the propagation of CMEs through interplanetary space (Kaiser et al. 2008). They also offered the opportunity to study the acceleration of energetic particles and the structure of the solar wind. This makes the data from these spacecraft particularly useful for the purposes of this work, as we have measurements of the solar wind speed from the orbital radius of Earth.

The two spacecraft shared an almost identical suite of instruments, including instruments to image the Sun, measure energetic particles, investigate the local plasma and to measure radio bursts. Data from the Plasma and Suprathermal Ion Composition (PLASTIC, see Galvin et al. 2008 for details) instrument were used for the work in this thesis. The instrument makes use of an electrostatic analyser and also has a near-real-time (NRT) data feed known as ‘beacon data’, where solar wind speed is derived using the limited processing power on board the spacecraft. Solar wind speed is also derived on the ground using more computing power, producing a higher quality, but often delayed, dataset known as science data. The derivation of solar wind speed is discussed further in Section 3.1.2.

Both spacecraft passed behind the Sun in 2014, where it was planned to power them down as they would not be able to transmit any data back to Earth. During a test of this system, contact was lost with STEREO-B and could not be regained and so data is only available up to 2014. STEREO-A remains operational to the point of writing, with its current position close to Earth.

ACE

The Advanced Composition Explorer (ACE) was launched in 1997 to the L1 Lagrange point, situated on the Sun-Earth line at approximately 0.99 AU. A spacecraft located at this point can give a short lead time warning of any Earth-directed event of up to one hour. The purpose of the mission was to study the energetic particles within the solar system and to measure the solar wind for space weather forecasting. Despite being in space for almost 26 years, the spacecraft remains functional at the time of writing.

The ACE spacecraft carries numerous instruments to measure the properties of the local plasma and magnetic field. For this work, the Solar Wind Electron, Proton and Alpha Monitor (SWEPAM, see McComas et al. 1998 for more details) was used. This instrument is comprised of an electrostatic analyser to provide measurements of the bulk solar wind properties. It provides both science data and real time data.

DSCOVR

The Deep Space Climate Observatory (DSCOVR) spacecraft was launched in 2015, also to the L1 Lagrange point. Planned as ACE's successor, it was designed to give warning of Earth directed CMEs as well as Earth observation and climate monitoring capabilities.

The spacecraft carries instruments to measure the plasma and magnetic field, to take images of the Earth and to measure irradiances from Earth. For the work used here, the Plasma-Magnetometer instrument (PlasMag, see Burt and Smith 2012 for more details) was used, which contains a Faraday Cup to measure the bulk solar wind properties. The emphasis of the DSCOVR mission was on its real time in-situ solar wind observations. However, there have been some issues with the solar wind speed. When the solar wind density is low, then the speed is more difficult to derive and leads to unphysical, large spikes in solar wind speed (Loto'aniu et al. 2022). This issue is only present in the real time data, as the effect gets removed in the on-ground processed data (as shown in Figure 6.2).

OMNI dataset

The OMNI dataset is constructed from a succession of spacecraft measurements at L1, mostly consisting of observations from the *Wind* (launched in 1994 and monitors the solar wind) spacecraft and from ACE. These observations are intercalibrated and ballistically propagated to the bow shock of the magnetosphere and provide a useful data set for near-Earth solar wind observations (Vokhmyanin et al. 2019).

3.2 Current solar wind forecasting methods

3.2.1 Corotation forecasts

Corotation forecasts use the rotating structure of the solar wind. One rotation of the Sun, with respect to Earth, takes 27.27 days (Bartels 1934). Assuming that the solar wind is steady-state in the reference frame rotating with the Sun, so does not evolve between solar rotations, we can use observations from spacecraft at the L1 point as a forecast. To create a corotation forecast, we assume that the solar wind conditions today are the same as they were 27 days ago. Due to the slowly evolving nature of the solar wind at solar minimum, 27-day corotation produces forecasts of a similar skill to numerical models (Owens et al. 2013). However, at solar maximum, the more dynamic nature of the solar wind means that the steady-state assumption does not hold and 27-day corotation forecasts perform worse.

Corotation is not limited to 27-days. Observations from any spacecraft on the ecliptic plane orbiting close to 1AU can be used for a corotation forecast, with the lead time corresponding to the longitudinal separation between the spacecraft and Earth. It is possible to use observations from a range of radial distances; however, by propagating the solar wind inwards or outwards, then interaction between solar wind streams could become important. 1 degree of separation corresponds to a corotation time of around 1.8 hours. Numerous studies (e.g. Kohutova et al. 2016; Simunac et al. 2009; Thomas et al. 2018; Turner et al. 2021) describe the benefits of corotation from the L5 position. The L5 position is located 60° behind the Earth in longitude, corresponding to a corotation time, and therefore lead time, of approximately 4.5 days. This can provide warning of fast solar wind streams and incoming SIRs. Although this is possible with a 27-day corotation, using observations that are closer to Earth in longitude means that the steady-state assumption is more likely to hold.

As corotation forecasts use a time series of solar wind observations, they cannot be used to predict CME speed and arrival time. Transients would only appear in a time series as an increase in solar wind speed, should a CME be detected at a spacecraft and be travelling faster than the background solar wind. For corotation,

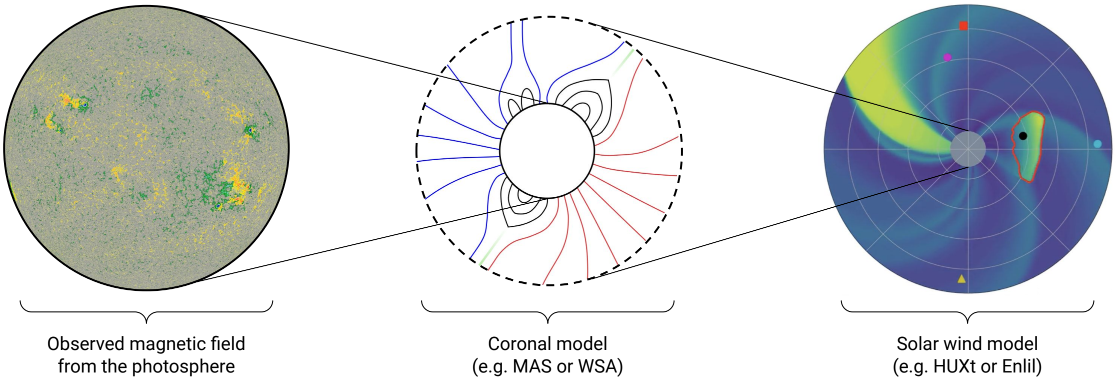


Figure 3.3: Schematic of the coupled coronal and solar wind system. Observations of the photospheric magnetic field are provided as the input into the coronal model, which solves for the corona from $1R_S$ to $20-30R_S$. The output from the coronal model serves as the input into the solar wind model, which propagates the solar wind out to beyond 1 AU. Adapted from Figure 1 in Owens et al. (2014).

this can lead to false alarms in the forecast time series, as the CME motion is purely radial and would not rotate with the Sun.

Corotation forecasts also provide a good opportunity to act as a benchmark for more complex models (Owens et al. 2013). It is often difficult to improve over corotation, particularly at solar minimum, but models that reconstruct the whole domain are still necessary. Corotation offers a solar wind speed forecast for one point, but this does not allow for the propagation of CMEs through the solution. Hence the need for the more complex models described below.

3.2.2 Coronal modelling

For current forecasting methods, the space between the Sun and the Earth is split into two sections, as they are considered to be different regimes (Owens et al. 2014). This is shown in Figure 3.3. The corona is magnetically dominated, whereas the heliosphere is flow dominated. Therefore, it is more efficient to treat them separately with tailored models for each regime. Here, we describe the main methods used to model the inner section; the Sun's corona, from $1R_s$ out to $20-30R_s$.

WSA

The Wang-Sheeley-Arge (WSA) model (Arge et al. 2003; Arge and Pizzo 2000) is based upon a potential field source surface (PFSS) model (Altschuler and Newkirk 1969; Schatten et al. 1969) of the steady-state corona and the Schatten Current Sheet (SCS) model (Schatten 1971).

The PFSS model is for the coronal magnetic field and it makes the assumption that electric currents can be ignored. The field is modelled between two spherical shell boundaries. The inner boundary is the photosphere, where the field is defined using observed magnetograms that are used to create Carrington maps of radial photospheric field. The outer boundary is a spherical 'source surface' where the field is radial and that any closed magnetic field lines are closed before the source surface. I.e. no magnetic field loops will close at radial distances greater than the source surface. The source surface is typically set to a radial distance of $2.5-3.25R_S$ in order to match with observations of the heliospheric magnetic field polarity at 1AU (Riley et al. 2006). This is important as it determines the amount of open flux that extends beyond the source surface. The model also assumes that the magnetic field is dominant over the flow (low β) and so the plasma does not need to be modelled.

In the WSA system, the PFSS output goes into the SCS model. This gives a more realistic magnetic field of the upper corona. To find the solar wind speed at $5R_S$, the following empirical relationship is used (Arge et al. 2003):

$$v(f_s,\theta_b) = 265 + \frac{25}{f_s^{\frac{2}{7}}} \cdot \left[5 - 1.1\exp\left\{1 - \left(\frac{\theta_b}{4}\right)^2\right\}\right]^2 \tag{3.1}$$

This relationship is made up of the flux tube expansion factor (f_s), which describes the radial expansion of magnetic field lines, and the minimum angular separation between an open field foot point and its nearest coronal hole boundary at the photosphere (θ_b). The terms in this relationship can be adjusted for different model versions, different magnetogram sources and for different model resolutions (e.g. McGregor et al. 2008).

Beyond $5R_S$, the WSA model uses a simple kinematic propagation of the solar wind to 1AU. The speed of the solar wind streams are modified depending on whether they are travelling faster or slower than the surrounding plasma (Arge and Pizzo 2000; Riley and Lionello 2011). Therefore, for WSA maps to be used as the inner boundary for more complex heliospheric models, these speeds have to be reduced back to the value they would be at the inner boundary (most commonly $21.5R_S$).

MAS

The Magnetohydrodynamic Algorithm outside a Sphere (MAS) model solves the magnetohydrodynamic (MHD) equations in 3-dimensional spherical coordinates (Linker et al. 1999; Mikić and Linker 1996; Riley et al. 2001). It models the solar corona from $1R_S$ to $30R_S$. The boundary conditions are specified using observations of the photospheric magnetic field, taken from a number of instruments both on the ground and in space, from which the initial conditions are derived. The MHD equations are time-dependent; however, the boundary conditions used in this work are time-independent. After an initial period of run-time, the model will relax to a steady state, which can then be used to drive the heliospheric model.

The MHD equations (Maxwell's equations, mass continuity and the momentum equation) solved by MAS do not form a closed set, i.e. the number of unknowns is greater than the number of equations. Therefore, an equation for energy that describes how density, temperature and pressure vary together is required. The simplest form of this equation is the polytropic assumption, which describes the expansion and contraction of the plasma, including heat transfer. The polytropic equation links together pressure (P) and volume (V):

$$PV^n = C \tag{3.2}$$

where n is the polytropic index, which depends on the situation, and C is a constant. n can take any value, but for situations where the volume is at a constant temperature, $n = 1$. In the MAS model, for the corona, $n \approx 1.05$ (Riley et al. 2001), meaning that it is close to a constant temperature. There are recent versions of MAS that contain more sophisticated energy equations. However, there are no large databases containing simulations from these versions and so they have not been used for this work.

The solar wind speeds given directly from the MAS code are too low and do not show a large enough separation between slow and fast wind (Riley et al. 2001). To that effect, an expression (similar to Equation 3.1) that estimates the solar wind speed depending on the distance to the edge of a coronal hole of the traced back magnetic field line is used (Riley et al. 2015). This distance is related to the flux tube expansion factor as in Equation 3.1. If the field line is deep in the centre of a coronal hole, then there is little radial expansion of the magnetic field. However, if the field line is rooted close to the edge of the coronal hole, then there is much greater radial expansion. The following relation is used in Riley et al. 2001:

$$v_r(\theta_b) = v_{slow} + \frac{1}{2}(v_{fast} - v_{slow}) \left[1 + \tanh \frac{\theta_b - \alpha}{w} \right] x \tag{3.3}$$

where θ_b is the minimum distance to the boundary of the coronal hole, α is the thickness of the slow solar wind band about the equator, w is the width over which the flow increases to coronal hole speeds, $v_{slow} = 250\,\mathrm{km\ s^{-1}}$ and $v_{fast} = 650\,\mathrm{km\ s^{-1}}$.

3.2.3 Heliospheric modelling

The outer section of space to be modelled is the wider heliosphere, from the edge of the corona into interplanetary space. This can be achieved through different methods; here, we describe the reduced physics approach, for which a large amount of the work within this thesis is based upon, and the widely used magnetohydrodynamic approach.

Reduced physics approach

The reduced physics approach bridges the gap between a purely ballistic approach, whereby the solar wind speed at the inner boundary is mapped to different locations within the heliosphere at a constant speed, and the complex magnetohydrodynamic approach. The model is known as the Heliospheric Upwind Extrapolation (HUX). It is based upon a reduced version of the MHD equations that discards the magnetic field, the effects of gravity and gas pressure (Riley and Lionello 2011). From Pizzo (1978), we can write the momentum equation (Equation 2.3) in the corotating reference frame:

$$-\Omega_{rot}\frac{\partial \mathbf{v}}{\partial \phi} + (\mathbf{v}\cdot\nabla)\mathbf{v} = \frac{\nabla P}{\rho} - \frac{GM_S}{r^2} \tag{3.4}$$

where Ω_{rot} is the equatorial angular rotation rate of the Sun, $\mathbf{v}$ is the velocity, ϕ is the azimuth angle, P is the thermal pressure, ρ is the proton mass density, G is the gravitational constant, M_s is the mass of the Sun and r is the radial distance. The time derivative, $\frac{\partial}{\partial t}$, has been replaced by $-\Omega_{rot}\frac{\partial}{\partial \phi}$, since this is equivalent for a time-stationary flow in the corotating reference frame.

For one dimension, ignoring the gas pressure and gravity terms, this simplifies to:

$$-\Omega_{rot}\frac{\partial v_r}{\partial \phi} + v_r\frac{\partial v_r}{\partial r} = 0 \tag{3.5}$$

when only considering velocity in a radial (r) direction.

Burgers' equation describes the transfer of physical quantities (such as particles) within a physical system, through the process of diffusion and convection (Burgers 1940). When viscosity is not involved, we get the inviscid Burgers' equation, which is the form given in Equation 3.5. This can be solved using an upwind scheme, which allows us to propagate a velocity stream from the inner boundary ($30R_S$) to 1AU in steps of Δr. A full explanation of the method can be found in Riley and Lionello (2011), but we find the following expression:

$$v_{i+1,j} = v_{i,j} + \frac{\Delta r\Omega_{rot}}{v_{i,j}}\left(\frac{v_{i,j+1} - v_{i,j}}{\Delta\phi}\right) \tag{3.6}$$

The indices i and j refer to the radial (r) and azimuthal (ϕ) grids respectively.

This scheme does not include any acceleration caused by gas pressure, as this term is neglected from the MHD equations. Although the solar wind is mostly accelerated before the inner boundary (see Figure 2.10), there is still some acceleration before 1AU (Schwenn 1990). To account for this acceleration, the following term is used:

$$v_{acc}(r) = \alpha v_{ro}\left(1 - e^{-\frac{r}{r_h}}\right) \tag{3.7}$$

where v_{ro} is the speed at the inner boundary (r_o), α is the acceleration term, and r_h is the length over which the acceleration occurs. These are found empirically to match polytropic MHD for values of $\alpha = 0.15$ and $r_h = 50r_S$. Therefore, as a function of radius, we can describe the total solar wind speed as $v(r) = v_{ro} + v_{acc}(r)$. Equation 3.7 shows that the speed of the accelerated solar wind is proportional to the speed at the inner boundary, meaning that fast wind will accelerate by a greater total than slow wind.

HUX has proved to produce results similar to those using an MHD model and outperforms the simple ballistic model. One comparison to a heliospheric MHD solution for Carrington Rotation 2068 found a correlation of 0.98 between the MHD solution and the HUX solution (Riley and Lionello 2011). Due to it being reduced-physics in nature, it is extremely efficient to run and only takes seconds on a desktop computer, thus making it suitable for ensemble forecasting. It also provides the opportunity to explore the parameter space and to constrain free parameters that can be used to inform the MHD models.

Due to the time-stationarity of the HUX model, where flows are assumed not to evolve with time in the frame corotating with the Sun, this does not allow for the propagation of CMEs. The Heliospheric Upwind Extrapolation with time-dependency (HUXt) model was developed, retaining the time dependence to allow for boundary conditions that evolve with time. This is more analogous to the MHD models and has again produced similar results (Owens et al. 2020a). Although not used in the main work chapters here, the HUXt model is referenced in Section 7.

Magnetohydrodynamic solar wind modelling

MHD modelling uses our best understanding of the physics of the system to model the solar wind (Riley et al. 2001). Unlike the reduced physics approach, MHD models take into account the magnetic field, incorporate solar wind compression and heating effects, and often allow for fully three-dimensional solar wind interactions. They typically utilise the ideal-MHD equations, and so diffusion is regarded as negligible and reconnection cannot occur, though it is often approximated by numerical diffusion in the model. There are numerous MHD models that have been developed for simulating the solar wind (e.g. Linker et al. 1999; Odstrcil and Pizzo 1999;

Pomoell and Poedts 2018; Tóth et al. 2005).

Here, we briefly describe the Enlil model (Odstrcil 2003), as this is one of the most commonly used MHD models in solar wind forecasting. Enlil is a 3-dimensional time-dependent model of the heliosphere that solves the ideal-MHD equations. The equations for mass, momentum, energy density and magnetic field are solved using an algorithm that solves partial differential equations based on finite differences (see Tóth and Odstrčil (1996) for more details). Enlil covers the domain from $21.5\mathrm{R}_S$ to 1AU, $-60°$ to $60°$ in latitude and $360°$ in longitude. The inner boundary is provided by a coronal solution (e.g. MAS or WSA) and the radial component of the magnetic field is used directly. There is assumed to be no nonradial flow at the inner boundary. The meridional component of the magnetic field (directed from south to north) is assumed to be zero. The azimuthal component (directed from west to east) is derived from the rotation rate of the source surface.

The solar wind is propagated from the inner boundary to Earth, where we get magnetic field and solar wind velocity, temperature and number density. It has the capability to propagate transient structures, such as CMEs, into the heliosphere and so can give speed and arrival time estimates at Earth.

3.2.4 Coupled modelling and operational forecasting

Current state-of-the-art forecasting methods use a coupled system of a coronal model, from 1 to $20 - 30R_s$, and a heliospheric (solar wind) model out to 1AU and beyond. The WSA model coupled to the Enlil model (known as WSA-Enlil) is used operationally at the UK Met Office Space Weather Operations Centre (MOSWOC) and at the US National Oceanic and Atmospheric Administration's (NOAA) Space Weather Prediction Centre (SWPC). They produce CME forecasts by inserting them at the inner boundary of Enlil ($21.5R_s$) and using the cone model (Xie et al. 2004; Zhao et al. 2002). This assumes that the CME is cone shaped with isotropic expansion, radial propagation and constant cone angular width (Wold et al. 2018). A CME will be inserted as a cloud of plasma with a constant velocity, density and temperature with a steady magnetic field (Wold et al. 2018). The forecasts produced by MOSWOC and SWPC are available for viewing by the public (`https://www.metoffice.gov.uk/weather/specialist-forecasts/space-weather` and `https://www.swpc.noaa.gov/products/wsa-enlil-solar-wind-prediction` respectively) and for their industrial clients, such as those in the military, power supply chain or in aviation. Constant model development takes place in response to verification.

Solar wind forecasts are verified against the observations they are trying to predict, often using spacecraft observations from L1. For solar wind speed, there are

Term	Description
State	Vector containing all the variables of interest.
Prior	Best estimate of the state before observations are included (often called background state).
Posterior	State estimate after the observations are included (often called analysis state).
Covariance matrix	Square matrix showing the relation between multiple random variables.
Gaussian	Way of describing a normally distributed random variable.
Model	Computer simulation of the system.
Observations	Measurements taken of the system.
Observation operator	Provides the link between the observations and the model.
Prior error covariance matrix	Describes the uncertainty in the error in the prior.
Observation error covariance matrix	Describes the uncertainty in the error in the observations.

Table 3.1: Basic definitions of the terms frequently used when discussing data assimilation in this thesis.

two avenues for which to verify the forecasts; the accuracy with which enhanced solar wind speed is predicted (e.g. SIRs or CMEs) and the average error on the predicted speed (Reiss et al. 2016). Numerous techniques can be used to verify a solar wind forecast, from simple error metrics such as mean absolute error (e.g. Norquist and Meeks 2010), to more sophisticated methods based on event timings (e.g. Owens et al. 2005). For the work presented in this thesis, we focus on verifying the average solar wind speed using mean absolute error (MAE).

3.3 Data assimilation

Data assimilation (DA) is the process that combines prior knowledge of a system, usually in the form of a model run, with observations of that system to find an optimum estimation of reality. This is shown schematically in Figure 3.4. It has been used extensively in numerical weather prediction, leading to large improvements in forecast accuracy (see Kalnay 2003), but has been underused in space weather forecasting (Innocenti et al. 2011). Table 3.1 shows the definitions of common DA terms used within this thesis.

Current solar wind forecasting makes use of observations to initialise the coronal model, which then feeds into the heliospheric model. This process is based upon measurements of the photospheric magnetic field and is converted into solar wind speed using models that make large assumptions, as discussed in the previous sec-

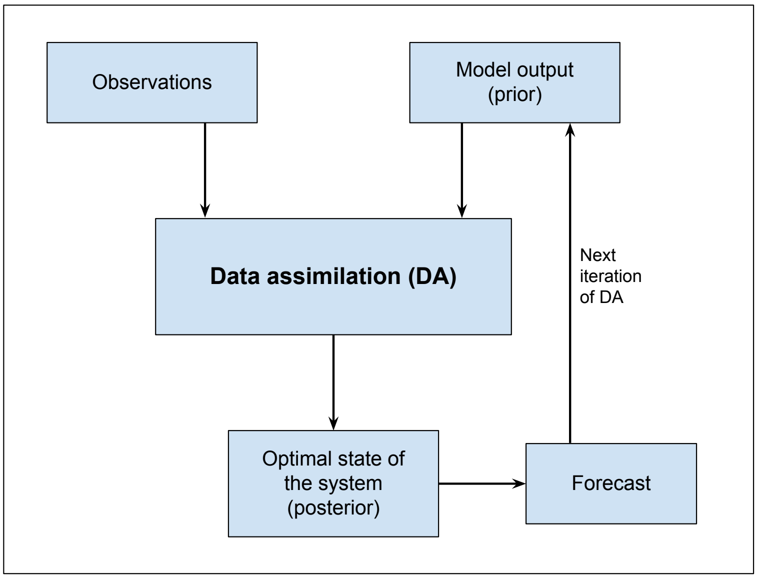


Figure 3.4: Schematic of the data assimilation process.

tion. By introducing information from observations taken of the solar wind, we can update the model inner boundary so that reality is better represented.

The underpinning of DA is based around Bayes' theorem (Bayes 1763), whereby we can use information about event B to update our prior knowledge of the probability of event A. Bayes' theorem is given by:

$$p(A|B) = \frac{p(B|A)\, p(A)}{p(B)} \tag{3.8}$$

where $p(A|B)$ is the conditional probability of event A given that event B is true, also known as the posterior distribution or posterior probability. $p(B|A)$ is the conditional probability of event B happening given that event A is true, it can be interpreted as the likelihood of A given a fixed B. $p(A)$ and $p(B)$ are the probabilities of observing events A and B respectively without any given conditions. $p(A)$ is also known as the prior distribution/ probability and $p(B)$ can be treated as a normalising constant. Therefore, this can be written as:

$$\text{posterior distribution} = \frac{\text{likelihood} \times \text{prior distribution}}{\text{normalising constant}} \tag{3.9}$$

The aim of DA is to find the best estimate of the system that is consistent with the observations, the model state and the system physics. This means that we want to find the posterior probability density function (pdf). This is the distribution that shows the probability of the state $\mathbf{x}$ given the information from the observations $\mathbf{y}$. If we know this pdf, then we can find the most probable state of the system. To find this, we can use Bayes' theorem. Note that Equation 3.8 can be used equivalently for pdfs as for point probabilities. This assumes that we know the prior probability distribution, $p_x(\mathbf{x})$, and the likelihood function, $p_y(\mathbf{y}|\mathbf{x})$. Using Bayes' theorem, we can find the posterior distribution:

$$p_x(\mathbf{x}|\mathbf{y}) = \frac{p_y(\mathbf{y}|\mathbf{x})\, p_x(\mathbf{x})}{p_y(\mathbf{y})} \tag{3.10}$$

here, the term $p_y(\mathbf{y})$ is a normalising constant. The prior distribution ($p_x(\mathbf{x})$) is the pdf of the state before the observations are considered, the likelihood ($p_y(\mathbf{y}|\mathbf{x})$) is the pdf of the observations given that the state is $\mathbf{x}$ and the posterior ($p_x(\mathbf{x}|\mathbf{y})$) is the pdf of the state after the observations have been considered.

We can define the optimum state, $\mathbf{x}$, as the combination of our best guess, the prior $\mathbf{x^b}$, and some random error, ξ:

$$\mathbf{x} = \mathbf{x^b} + \xi \tag{3.11}$$

The random error is taken to have zero mean and covariance matrix $\mathbf{B}$. This is one of the most fundamental components of DA; the prior error covariance matrix. A covariance matrix describes the relation between multiple random variables. For example, in terrestrial weather forecasting, it might show how temperature at one grid point relates to temperature at adjacent grid points. The prior error covariance matrix describes the errors in the prior state and determines how much we should 'trust' the model. It is often not possible to explicitly represent $\mathbf{B}$, especially in high-dimensional systems, as this leads to matrices that are, computationally, too large. Therefore, it is necessary to estimate it (Bannister 2008). This can be done through a number of methods, with the specific method used in this work described in Section 3.4.

When combining the prior ($\mathbf{x}$) with the observations ($\mathbf{y}$), we need to be able to compare the two. Observations will not be taken on the same spatial and temporal grid as the model, and sometimes will not be the correct quantity (for example, satellite radiances to determine surface temperature, e.g. Huang et al. (2019)), so we need to transform and interpolate the model to create "model equivalents" to compare to the observations. The set of operations performed to achieve this is known as the observation operator, $\mathcal{H}$:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \epsilon \tag{3.12}$$

where ϵ is a random variable assumed to have zero mean and covariance matrix $\mathbf{R}$, also known as the observation error covariance matrix. This is another important component of DA as it describes the error in the observations arising from errors in the measuring instrument, errors in the observation operator and respresentivity errors (see Figure 1 in Janjić et al. (2018) for more information). Representivity errors come from poor representation of the observations, for example, in location or differences in resolution (Janjić et al. 2018). The observations are typically assumed to be independent of one another, meaning that **R** is a diagonal matrix. This is often due to computational restrictions and/ or unknown observation error statistics (Waller 2021). Determining the full **R** matrix is often not possible in high-dimensional systems, so estimation is required. This can be achieved through a number of methods, but are not relevant for this thesis. Further information can be found in Waller (2021).

There are many DA methodologies that have been developed, and these can be broadly split into two kinds; sequential and variational. Sequential DA methods take in observations as soon as they are available, propagating the distribution through time, along with its associated error, and updating it with each new observation. Some of these methods are not suitable for high-dimensional systems, such as terrestrial weather forecasting (more than 10^9 dimensions), due to computational cost. Alternatively, variational DA methods intake the observations at the end of a given time window and adjusts the full model trajectory given the observations. This is achieved through minimising a cost function that relates together the errors associated with the model and the observations, so finding the state where the error is a minimum over a whole time window.

Below, we briefly discuss the main sequential methods and then a more in depth description of variational methods (specifically 4DVar), as this method has been used for the work presented in this thesis.

3.3.1 Sequential methods

Sequential DA methods take in observations as soon as they become available. In sequential DA, the prior error covariance evolves with time, as it needs to be updated for every observation that is assimilated. There are numerous methods for sequential DA, with the main ones described below.

Kalman Filter

The Kalman Filter is a technique that ‘nudges’ a model trajectory as new observations come in. It can be applied independently to a system, meaning that it is not model specific. It also allows for propagation of the covariance matrix through time, which is important for flow-dependent systems. However, due to the sequential nature of the Kalman Filter, variational methods offer a better trajectory over the whole assimilation window as they consider all the observations at once.

The Kalman Filter is a two-phase process consisting of a prediction phase and

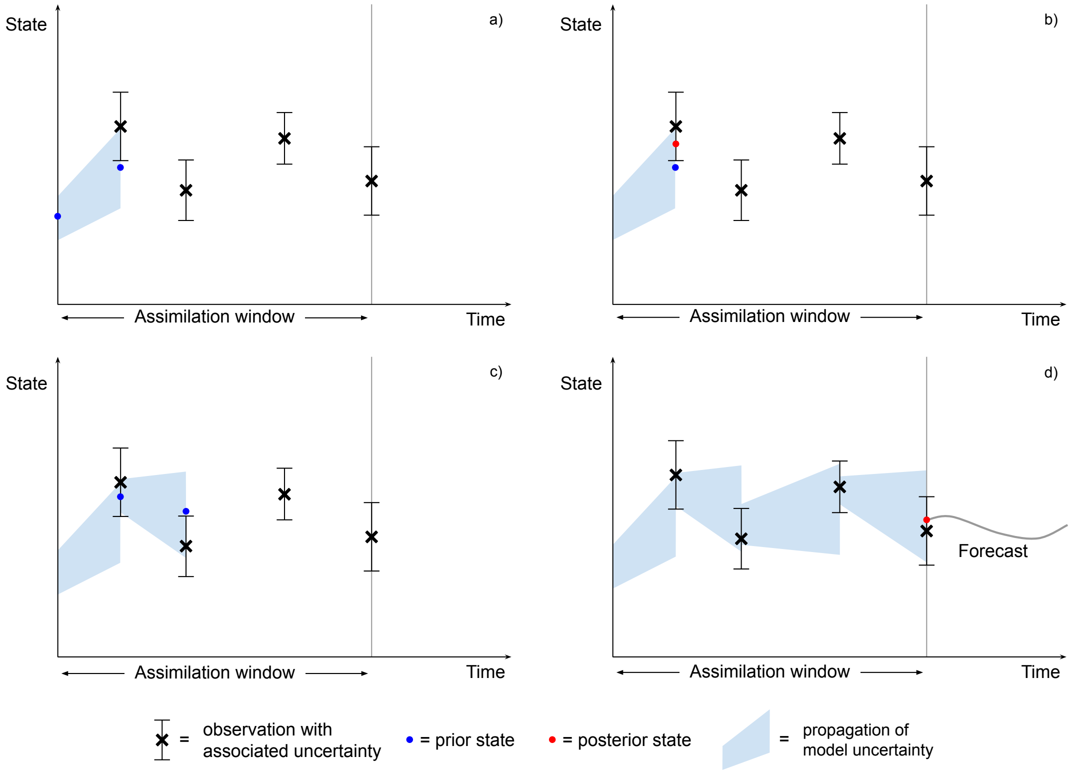


Figure 3.5: Schematic showing the Kalman Filter algorithm. The prior state (left-most blue dot) in panel a is propagated forwards in time to an observation point (black cross) using the model (right-most blue dot) with its associated uncertainty (light blue shaded region). Taking into account the observations uncertainty, the state is nudged towards the observation in panel b (red dot). This is then propagated forwards again in panel c. This is repeated for all observations (panel d) until they have all been assimilated. The estimation of the state can then be used to initialise a forecast.

an update phase. The prediction phase will be based upon a model of the system and maps the state and covariance from one observation to the next. The update phase then 'corrects' the current estimate of the state and the covariance based upon the new observation and the previous model information. This update is weighted depending on the uncertainty in both the observation and the model, with it tending towards the one with the lowest uncertainty. The new state then returns to the prediction phase, where it is mapped forwards again to the new observation time. This process repeats for all observations, until all observations have been processed. At this point, we can use the last estimate of the state to produce a forecast. This process is shown in Figure 3.5.

The Kalman Filter is for linear models only, other methods (such as the Extended Kalman Filter) are used for non-linear systems. Due to the requirement for the covariance matrix to be propagated forwards with each step, the Kalman Filter is only suitable for low-dimensional systems. It would require too much computing

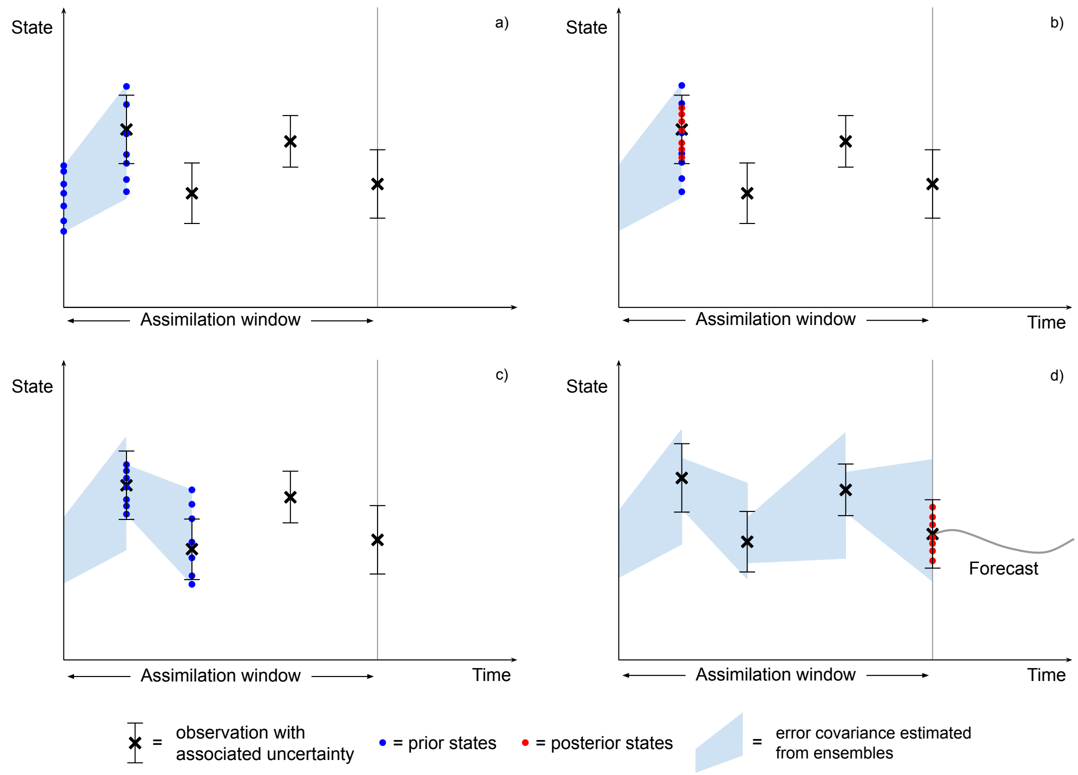


Figure 3.6: Similar to Figure 3.5, but with ensembles used at each step. The model error is estimated from the spread in the ensembles. The ensemble members at the end can be used to create an ensemble forecast, or the ensemble mean can be used to initialise a forecast.

power for systems with a high number of dimensions; hence, the Ensemble Kalman Filter is used for such high-dimensional systems.

Ensemble Kalman Filter

The ensemble Kalman Filter (Figure 3.6) is a Monte Carlo approximation of the Kalman Filter for linear systems. Monte Carlo methods use repeated random sampling to form an estimation of a system, often due to the system being too complex to solve fully. Using this technique avoids the need for the covariance matrix to evolve through time. Instead, an ensemble of states is produced and the covariance is estimated from this sample. We take a selection of initial conditions from the uncertainty in the prior, propagate these forwards to the observation time, then estimate the mean and the covariance from the evolved states.

Once the ensemble members have been evolved using the model, they are then updated given the observations as for the Kalman Filter. Each updated ensemble member is then run forwards again, until all observations have been processed. At the final step, we can either use the ensemble members to create an ensemble forecast

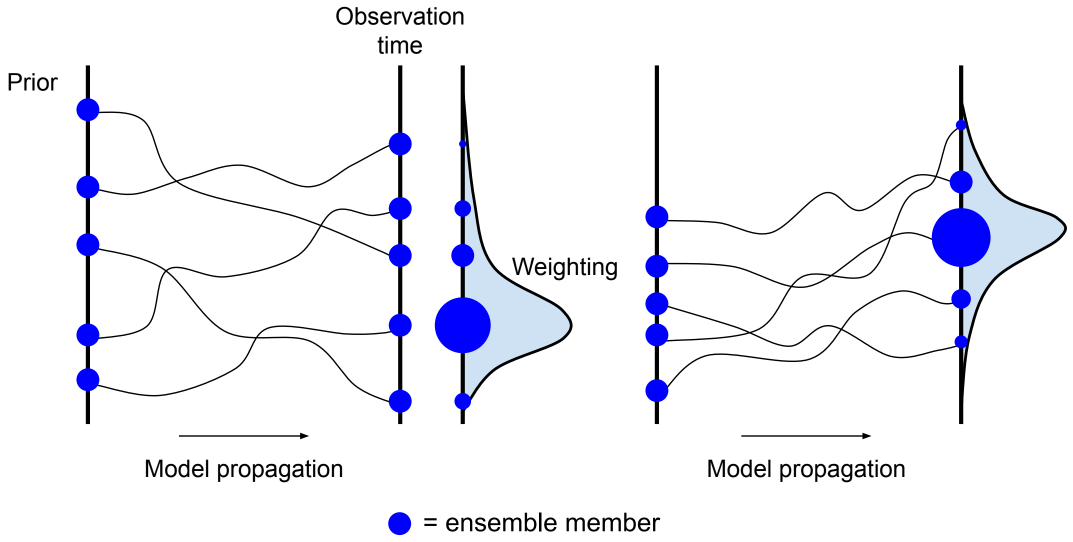


Figure 3.7: Schematic of the particle filter process. An ensemble of particles is propagated forwards from the prior state to an observation time (left) and then weighted depending on how much they agree with the observations (middle). This creates a new ensemble, closer to the observations, which can then be propagated forwards again (right).

or the ensemble mean can be used as the optimal state to generate the forecast. This choice depends on the intended construction of the forecast.

Particle filter

Particle filters aim to find the full posterior distribution. To achieve this, the algorithm takes samples from the prior distribution, which are known as ‘particles’, and weights them. The particle filter differs from the Kalman filter as it does not assume a Gaussian distribution of the error statistics and is suitable for non-linear systems.

The prior is generated by using an ensemble of initial conditions, representing the uncertainty. These conditions are propagated forwards in time to when an observation is available. The ensemble members are then weighted, depending on their agreement with the observations. The closer the ensemble member is to the observation, the higher the weight. This is then used to generate a new ensemble, whose initial conditions are closer to the observations. This is shown schematically in Figure 3.7.

3.3.2 Variational methods

4DVar overview

Variational methods aim to find the mode of the posterior distribution, and so the most probable state given the model and the observations (Le Dimet and Talagrand 1986). Here, we discuss the variational method that considers observations that cover a time window (rather than a single observation at a single time, known as 3DVar), with the assumption that the model contains no errors. This is known as strong-constraint 4DVar and is the method used for the work in this thesis. The alternative, weak-constraint 4DVar, explicitly contains model error (e.g. Fisher et al. 2012; Qin et al. 2007). Weak-constraint 4DVar is more complex than strong constraint and is a technique that has not been used in the work in this thesis, and so will not be further discussed. See Kalnay (2003) for more details.

The aim of 4DVar is to find the initial conditions of the model such that the model state best fits with the observations over the assimilation window. The 'best fit' is the state that has the minimum contribution of error from both the model and the observations. This is achieved using the process shown in Figure 3.8, starting from the model and the observations and finishing with a forecast being produced from the updated model state.

Step 1 of Figure 3.8 shows the introduction of the model and the observations. Both of these are affected by random errors and so should be considered as random variables, which are assumed to be described by a Gaussian distribution. The Gaussian distribution of a random variable is used because it can be fully described by its mean and covariance only, reducing the number of values that have to be stored. This can be used to determine a cost function.

The cost function is calculated in step 2. This is the term that needs to be minimised, as it quantifies the error contribution of both the observations and the prior state. This is described in more detail in the following section. It is often not possible to explicitly find the minimum of the cost function, so we find the gradient as the minimum of the cost function will be found when its gradient is equal to zero. For a high-dimensional system, finding the minimum gradient cannot be found analytically, so we have to rely on iterative methods.

The gradient is found in step 3. There are many methods to find the gradient, but the work in this thesis makes use of the adjoint method (see Errico 1997 for more details). This method is a type of sensitivity analysis that allows the gradient to be found at the initial time, which is then used to inform a minimisation update equation (step 4), leading to a new estimate of the state.

A threshold (step 5) is used to determine whether the process repeats to find a lower cost function and gradient (step 6) or whether it terminates (step 7). This

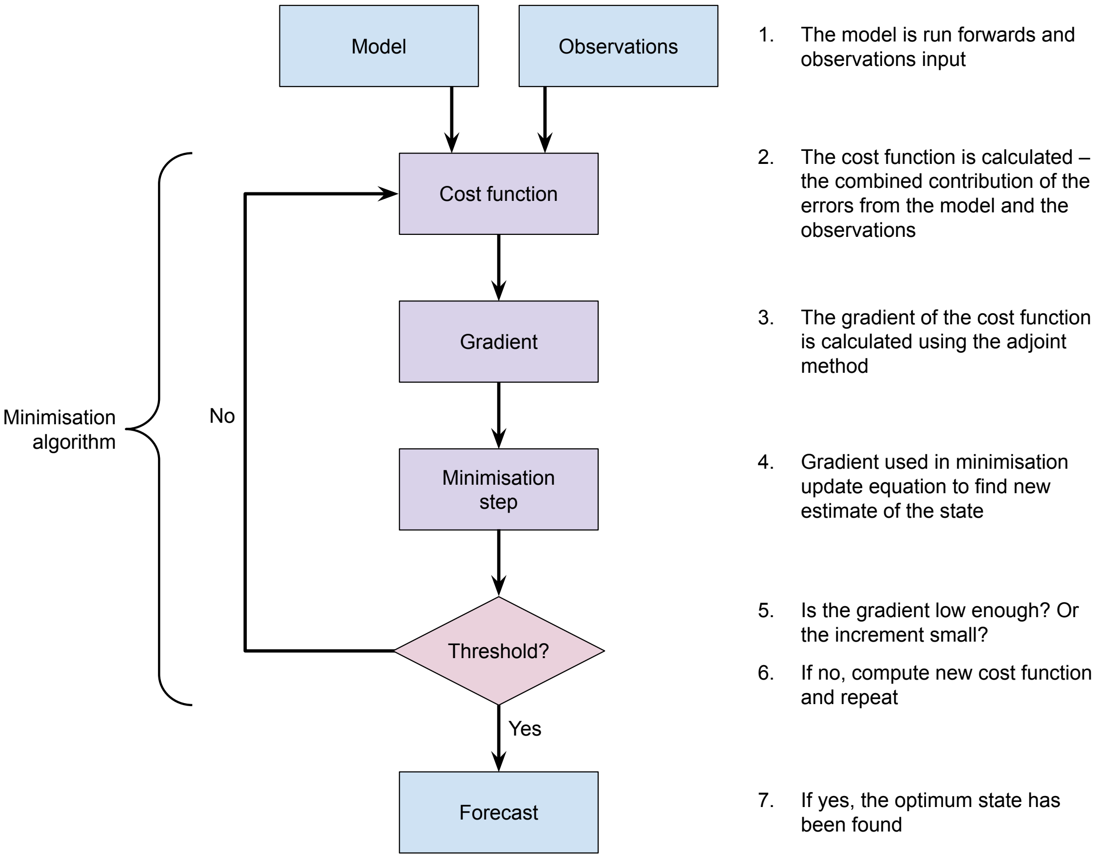


Figure 3.8: Flow chart of the 4DVar data assimilation process.

threshold could be a variety of options; for example, if the gradient is low enough or if the increment between steps is small enough. If the threshold is not met, then steps 2 to 5 are repeated until the threshold is met. Once the process terminates in step 7, a new state with optimum initial conditions has been found, which can then be used to initiate a forecast.

Steps 2 to 5 are called a minimisation algorithm, and a simplified example (with one dimension) can be seen in Figure 3.9. Here, we are aiming to find the minimum value of the cost function, $\mathcal{J}(\mathbf{x})$, for the given state $\mathbf{x}$. Panels a to d show 3 iterations, calculating the gradient at each step and moving to a lower value each time. A typical 4DVar system will compute more iterations, with the system used in this work computing approximately 500 iterations.

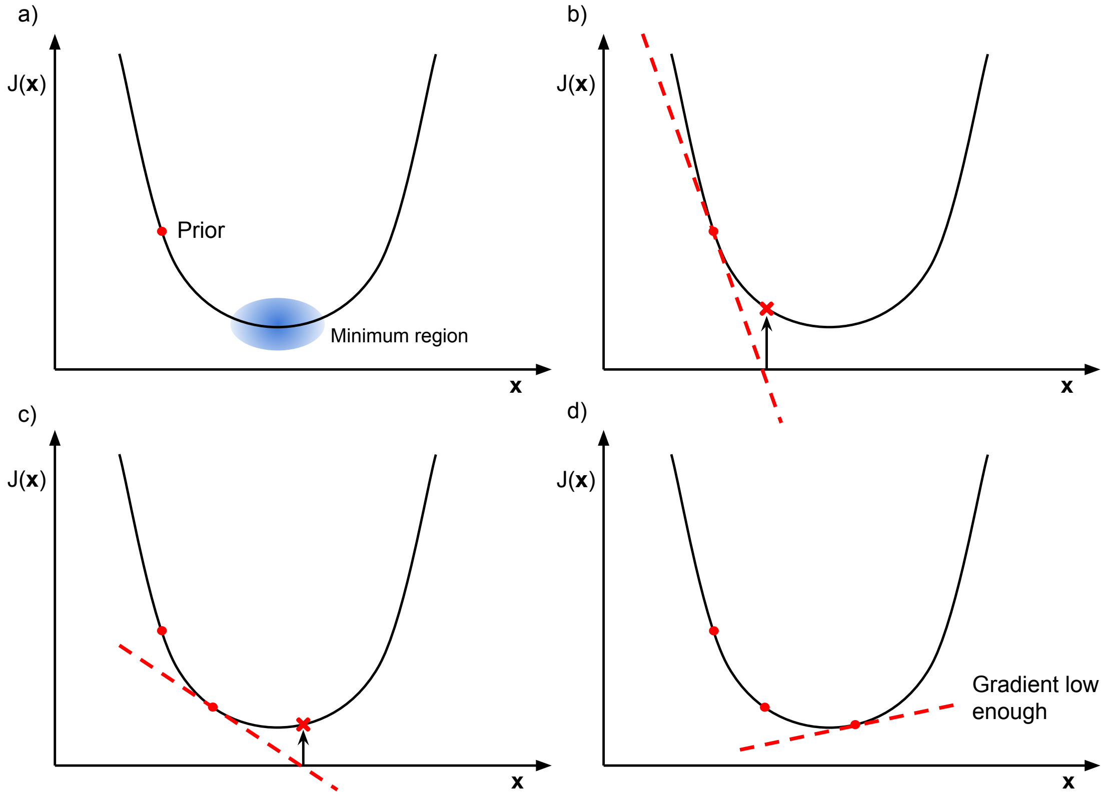


Figure 3.9: Schematic of a one dimensional gradient minimisation algorithm. The value of the cost function ($\mathcal{J}(\mathbf{x})$) is calculated using the prior information in panel a, with the aim of finding the state ($\mathbf{x}$) where $\mathcal{J}(\mathbf{x})$ is at a minimum. Panels b - d show an iterative process where the gradient of the cost function is computed and then a new value of the cost function is found from the new estimation of the **x**. The gradient in panel d is low enough, so the process will terminate and we have a new, optimum estimate of **x**.

4DVar cost function

As described in Section 3.3.2, we can describe the model and the observation using a Gaussian distribution, using only their mean and covariance. A Gaussian distribution is defined as:

$$p(\mathbf{x}) = A \, \exp\left\{ -\frac{1}{2} (\mathbf{x} - \langle\mathbf{x}\rangle)^T \, \mathbf{C}^{-1} \, (\mathbf{x} - \langle\mathbf{x}\rangle) \right\} \tag{3.13}$$

where A is a constant, $\mathbf{x}$ is the random variable, $\mathbf{C}$ is the covariance matrix and $\langle\mathbf{x}\rangle$ is the mean of the distribution. The letter T represents taking the transpose of the vector.

To produce a cost function using this information, we describe the prior ($p(\mathbf{x})$) and the likelihood ($p(\mathbf{y}|\mathbf{x})$) using the Gaussian distribution shown in Equation 3.13.

The mean of the prior is denoted as $\mathbf{x}^{\mathbf{b}}$ with covariance $\mathbf{B}$:

$$p(\mathbf{x}) = A_1 \, \exp\left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{x^b})^T \, \mathbf{B}^{-1} \, (\mathbf{x} - \mathbf{x^b}) \right\} \tag{3.14}$$

The mean of the likelihood is defined using the observation operator, $\mathcal{H}(\mathbf{x})$, with covariance $\mathbf{R}$:

$$p(\mathbf{y}|\mathbf{x}) = A_2 \, \exp\left\{ -\frac{1}{2} (\mathbf{y} - \mathcal{H}(\mathbf{x}))^T \, \mathbf{R}^{-1} \, (\mathbf{y} - \mathcal{H}(\mathbf{x})) \right\} \tag{3.15}$$

Using the relationship from Bayes theorem in Equation 3.10, we find that:

$$p_x(\mathbf{x}|\mathbf{y}) \propto \exp\left\{ -\frac{1}{2} \left[(\mathbf{x} - \mathbf{x^b})^T \, \mathbf{B^{-1}} \, (\mathbf{x} - \mathbf{x^b}) + (\mathbf{y} - \mathcal{H}(\mathbf{x}))^T \, \mathbf{R^{-1}} \, (\mathbf{y} - \mathcal{H}(\mathbf{x})) \right] \right\} \tag{3.16}$$

We wish to maximise the posterior probability, as we are searching for the most probable state given the observations. Due to the exponential in Equation 3.16, this is equivalent to minimising $-\ln p(\mathbf{x}|\mathbf{y}) = \mathcal{J}(\mathbf{x})$, where $\mathcal{J}(\mathbf{x})$ is the cost function:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x_0} - \mathbf{x_0^b} \right)^T \, \mathbf{B_0^{-1}} \, \left(\mathbf{x_0} - \mathbf{x_0^b} \right) + \frac{1}{2} \sum_{i=1}^{N} \left(\mathbf{y_i} - \mathcal{H}_i(\mathbf{x_i}) \right)^T \, \mathbf{R_i}^{-1} \, \left(\mathbf{y_i} - \mathcal{H}_i(\mathbf{x_i}) \right) \tag{3.17}$$

where we are using observations from some time in the interval $[t_0, t_N]$ and the summation is due to having multiple (N) observations. The subscript 0 denotes the state at time t_0 and the subscript i denotes the state at the observation times t_i, $i : 1 \rightarrow N$.

The gradient of $\mathcal{J}(\mathbf{x})$ is found using the adjoint method, as described above. Applying the adjoint to any system is a very complex task and the equations are not shown here as they are not directly relevant for the understanding of this thesis. Errico (1997) gives an in depth discussion of the adjoint method and the precise adjoint equations used for the work in this thesis can be found in Lang and Owens (2019).

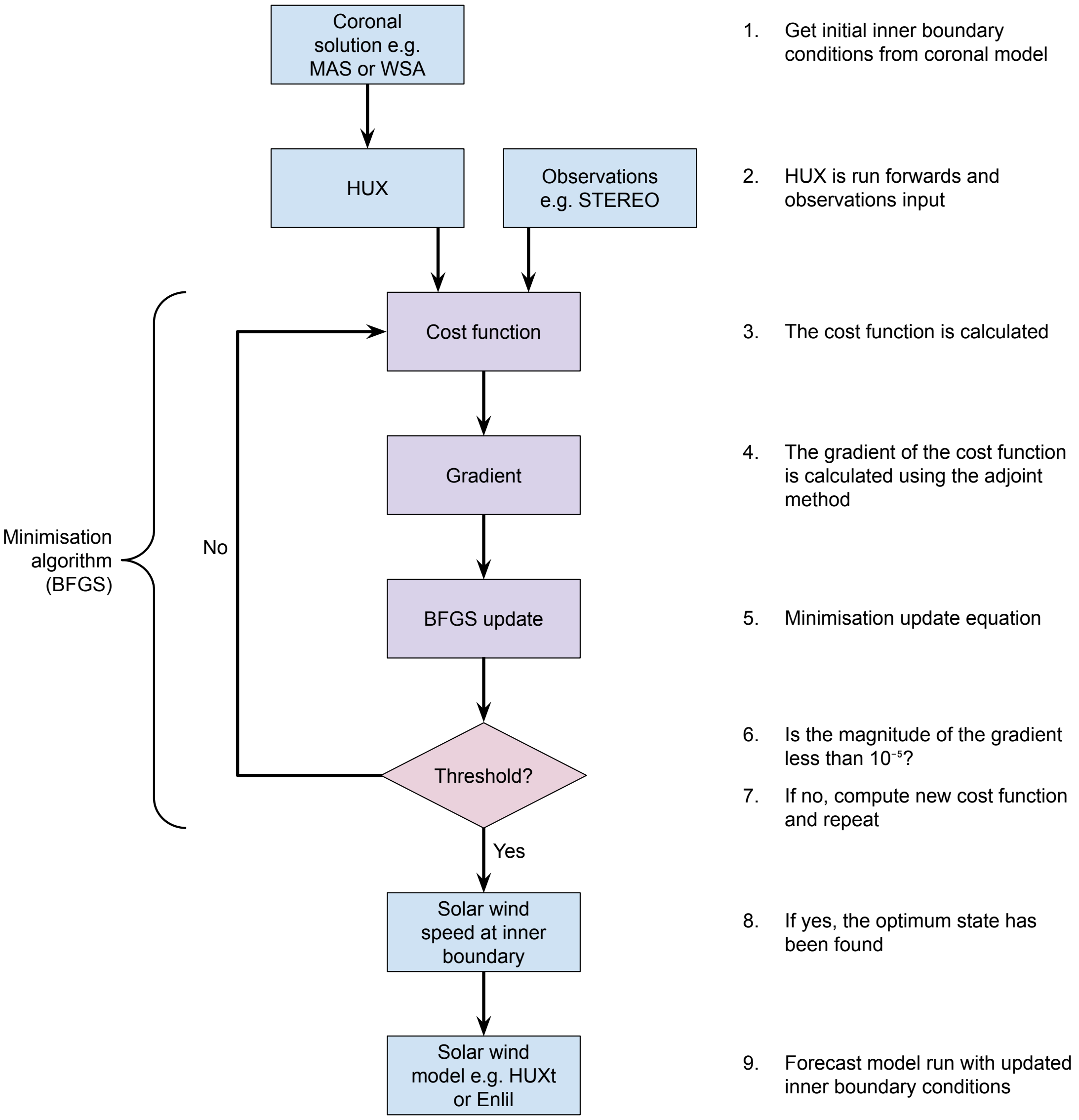


Figure 3.10: Flow chart of the BRaVDA scheme.

3.4 BRaVDA

The Burger Radius Variational Data Assimilation (BRaVDA, Lang and Owens 2019) scheme uses a strong-constraint 4DVar data assimilation method to reconstruct solar wind speed. It is based around the HUX solar wind model, which propagates the solar wind from the inner boundary at $30R_s$ out to Earth (see Section 3.2.3). The scheme can use any observations, as long as they contain information about the conditions at Earth's latitude. Here, we make use of the STEREO mission for observations far from Earth at approximately 1AU, and from the ACE and DSCOVR missions for near-Earth. The aim of BRaVDA is to reproduce the solar wind conditions at Earth's latitude, using prior information from a coronal solution and observations from a similar latitude.

Using the same setup from Owens and Riley (2017), we have a radial resolution of $\Delta r = 1R_s$ and a longitudinal resolution of $\Delta\phi = \frac{360}{128} = 2.81^\circ$. The inner boundary is located at $30R_s$ and the outer boundary can be extended out into the heliosphere; however, for most of BRaVDAs application, this is set just beyond Earth's orbital radius, at approximately $240R_s$, to fully encompass the orbit of Earth and the STEREO spacecraft. Due to the 2-dimensional nature of HUX, BRaVDA solves the solar wind in the plane at Earth's heliographic latitude. This is typically not the solar equator or the ecliptic plane, and so this can introduce a source of representivity error. The impact of latitudinal separation of observations and forecast location is explored in Chapter 4.

The process used in the BRaVDA scheme is shown in Figure 3.10. Step 1 shows a coronal solution, the prior (from WSA or MAS), which is used as the initial conditions for HUX. In step 2, the HUX model is run forwards with the initial conditions and the observations are input into the system. For the BRaVDA scheme, the assimilation window is one whole solar rotation. Therefore, it takes in 27-days worth of observations, at an hourly interval. The cost function is calculated in step 3 and the gradient found using the adjoint method in step 4.

Step 5 shows the minimisation update equation (one step in Figure 3.9) that produces a state with a lower gradient and therefore lower cost function. Here, the method used for minimisation is the Brandon-Fletcher-Goldfarb-Shanno (BFGS) algorithm. This finds the intercept of the gradient with the x-axis, as shown in Figure 3.9. The BFGS was chosen as it is quick, efficient and does not require a second derivative. The threshold set in step 6 is whether the norm of the cost function gradient is less than 10^{-5}. If the threshold is not met (step 7), the process repeats. If the threshold is met, then we have found the state with the optimum initial conditions (step 8). These initial conditions can then be used to initialise a solar wind model, for example HUXt or Enlil, to produce a forecast (step 9).

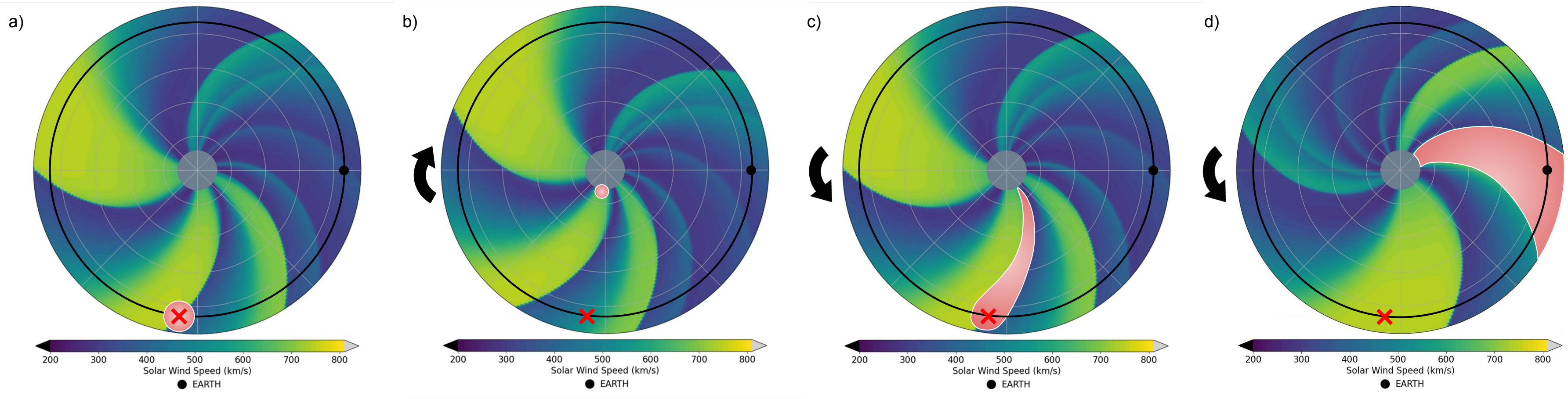


Figure 3.11: Schematic of how the BRaVDA scheme updates the initial conditions of the model given an observation (red cross) at Earth's orbital radius (black circle). The pink shaded region shows the influence of the observation on the model state, i.e. how the information is used to update the initial conditions and then propagate them to Earth's orbital radius. Panel a shows the observation at 1AU, which contains information about the solar wind speed at the inner boundary at a time in the past. Panel b shows the information from that observation being used to update the initial conditions at the inner boundary. The initial conditions are updated given the observations information, and the model is run forwards with this update in panel c. Panel d shows the updated solar wind conditions run into the future for a forecast computed at Earth. Figure recreated from Figure 1 in Lang and Owens (2019) and the background solar wind image is taken from a HUXt simulation.

Figure 3.11 shows this process schematically, with the use of a single observation taken from Earth's orbital radius. Information is taken from the observation and used to update the initial conditions using the process described above. Once the initial conditions have been updated, the solar wind can then be propagated out to Earth using any solar wind model, producing a forecast.

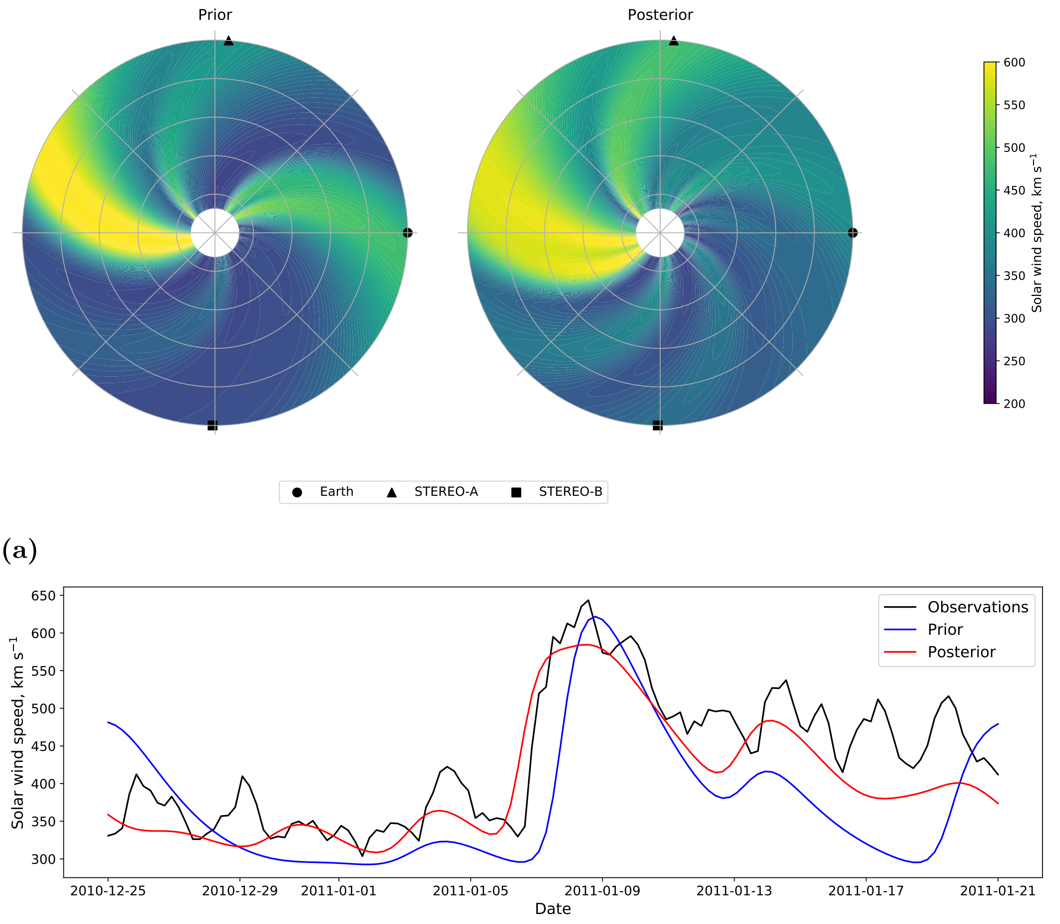
The prior error covariance is generated through producing an ensemble with the prior state. This is done by perturbing the prior, and therefore the model initial conditions, and running the HUX model forwards accordingly. The perturbed initial conditions are found by sampling in latitude about the point where Earth's orbit would be at the inner boundary. As shown in Owens and Riley (2017), the spread of the ensemble members describes the uncertainty in the unperturbed forecast well, hence why it can be used for determining the prior error covariance matrix and then used in the cost function.

BRaVDA uses variational data assimilation as it allows us to map the information in the observation back to the models inner boundary, through the use of the adjoint model. If sequential data assimilation methods were used with observations at 1AU, any update to the model state would be swept away into the heliosphere. This is due to the solar wind flow being radial away from the Sun, and so this would not lead to an improved forecast, as any changes would be made beyond Earth's radius (Lang et al. 2017).

Figure 3.12a shows an example of a prior and posterior state produced using BRaVDA. The prior is taken from a MAS solution and BRaVDA assimilates solar wind speed data from both STEREO spacecraft and from the OMNI dataset. We can see that BRaVDA has introduced more structure into the state and increased the longitudinal width of the fast stream in opposition to the Earth. The moderate stream passing close to Earth has been reduced in magnitude and there are other faint structures introduced.

Figure 3.12b shows the time series of Figure 3.12a at 1AU, with the prior in blue, posterior in red and observations from the OMNI dataset in black. The root mean square error (RMSE) reduces from 90.1 km s^{-1} for the prior to 49.5 km s^{-1} for the posterior, showing that the state after data assimilation is closer to the observations.

The remainder of this thesis is focused around testing the performance of the DA forecasting methods under a range of conditions. Section 4 evaluates the impact of latitudinal separation of observations, aiming to better inform the DA scheme on how to use the observations. Section 5 tests how the number of observation sources affects forecast accuracy. It also assesses how much of an impact removing CMEs from the input and verification time series has on forecast accuracy. Section 6 tests how well we would expect BRaVDA to work operationally and the implications of the future *Vigil* mission might have for future solar wind DA.



(a)

(b)

Figure 3.12: The top panel (a) shows an example of a prior (left) and posterior (right) state in the BRaVDA scheme. This is a reconstruction of the state on 21/01/2011, assimilating data from the OMNI dataset and STEREO spacecraft. In this case, the prior is taken from a MAS solution. The bottom panel (b) shows the time series of solar wind speed at 1AU from panel (a). The prior state is shown in blue, the posterior is shown in red and the observations from the OMNI dataset is given in black.

Chapter 4

The influence of spacecraft latitudinal offset on the accuracy of corotation forecasts

This chapter has been published in the AGU journal *Space Weather* with the following reference:

Turner, H., Owens, M. J., Lang, M. S., & Gonzi, S. (2021). The influence of spacecraft latitudinal offset on the accuracy of corotation forecasts. *Space Weather, 19,* e2021SW002802. `https://doi.org/10.1029/2021SW002802`

Key points

- Solar wind speed corotation forecast error is affected by solar activity and spacecraft longitudinal and latitudinal separation.
- Latitudinal separation has little effect when it is below 6°, but increasing importance above this.
- A period in solar minimum gives a 46% increase in mean absolute error from low to high latitudinal offset for STEREO-A and B corotations.

Abstract

Knowledge of the ambient solar wind is important for accurate space-weather forecasting. A simple-but-effective method of forecasting near-Earth solar-wind speed is “corotation”, wherein solar-wind structure is assumed to be fixed in the reference frame rotating with the Sun. Under this approximation, observations at a source spacecraft can be rotated to a target location, such as Earth. Forecast accuracy depends upon the rate of solar-wind evolution, longitudinal and latitudinal separation between the source and target, and latitudinal structure in the solar wind itself. The time-evolution rate and latitudinal structure of the solar wind are both strongly influenced by the solar cycle, though in opposing ways. A latitudinal separation (offset) between source and target spacecraft is typically present, introducing an error to corotation forecasts. In this study, we use observations from the STEREO and near-Earth spacecraft to quantify the latitudinal error. Aliasing between the solar cycle and STEREO orbits means that individual contributions to the forecast error are difficult to isolate. However, by considering an 18-month interval near the end of solar minimum, we find that the latitudinal-offset contribution to corotation-forecast error cannot be directly detected for offsets $< 6°$, but is increasingly important as offsets increase. This result can be used to improve solar-wind data assimilation, allowing representivity errors in solar-wind observations to be correctly specified. Furthermore, as the maximum latitudinal offset between L5 and Earth is $\approx 5°$, corotation forecasts from a future L5 spacecraft should not be greatly affected by latitudinal offset.

Plain language summary

Space weather can damage our technologies, from power lines to satellites, as well as pose a hazard to humans in space. Forecasting space weather requires prediction of solar wind, a continuous outflow of material from the Sun, conditions in near-Earth space. A simple way to achieve this is “corotation”, which assumes the structure of the solar wind is unchanging, but simply rotates around with the Sun. Thus solar wind which sweeps past one spacecraft will arrive at Earth some time later. This forecast will be less accurate when the spacecraft is far from Earth, as we need to wait longer for the solar wind to rotate around, during which time its structure may have changed. If the spacecraft is at a different latitude to Earth, it will also create problems for the forecast. In this study, we quantify the contribution of these factors to the forecast error. This knowledge will improve corotation forecasts, but also aid in other, more sophisticated, forecast techniques. By defining how close spacecraft need to be to Earth to sample the same solar wind, it also helps define where future

space weather monitoring spacecraft should be positioned.

4.1 Introduction

Space weather has the potential to damage electricity grids, cause satellite failures, disrupt communications and threaten the health of humans in space (Cannon 2013). The most severe space weather events are driven by transient coronal mass ejections [CMEs] (Gosling 1993). However, prediction of the ambient solar wind is still important for space weather for two reasons. Firstly, the structure of the solar wind impacts the evolution of CMEs through interplanetary space and can determine their arrival time and severity at Earth (Case et al. 2008). Secondly, steady-state structures in the solar wind can also be a driver for space weather events in their own right (e.g. Alves et al. 2006). High speed streams emanating from coronal holes can compress into the slower solar wind and form stream interaction regions (SIRs), which become corotating interaction regions (CIRs) if they persist for several solar rotations (Reiss et al. 2016). These are regions of higher plasma density and magnetic field strength and can cause geomagnetic disturbances at Earth (Richardson and Cane 2012). Thus, for accurate space weather forecasting of both recurrent and non-recurrent events, knowledge of the solar wind conditions near-Earth is required.

The Sun's magnetic field forms an Archimedean spiral due to the field lines remaining rooted at the Sun as it rotates; this shape is known as the "Parker Spiral" (Parker 1958). The solar wind flows almost radially away from the Sun; however, fast and slow structures rotate with the Sun, which has a rotational period of 27-days with respect to Earth. Assuming a steady-state solar-wind structure in the rotating solar frame, this means that the same solar wind conditions will reoccur every 27-days at Earth. Owens et al. (2013) quantified the potential benefit of such 27-day "recurrence" (also called "27-day persistence") forecasts. They can provide useful long-lead time forecasts and a benchmark for more sophisticated models. The same principle can be used with observations from spacecraft distant from Earth, but at the same heliographic latitude, for example, on the ecliptic plane. Applying a time shift to observations at one spacecraft allows them to be used as a forecast for another, further on in the Sun's rotation. These are known as corotation forecasts and the reduction in forecast lead time limits (but does not eliminate) the assumption of a completely steady-state solar-wind structure. They have been shown to be a useful forecasting tool, often outperforming the 27-day recurrence model (Bailey et al. 2020; Kohutova et al. 2016; Simunac et al. 2009; Thomas et al. 2018). For observations at the same heliocentric distance, the required time shift [Δt] is given solely by corotation time [t_C], the time it takes the Sun to rotate by the angle ϕ between the longitudes of the observation (referred to as the source) and the location to be forecast (referred to as the target). For $\phi = 1^\circ$, $t_C = 1.8$ hours. However, spacecraft (and the Earth) orbiting the Sun typically have elliptical orbits, resulting in radial distance variations, as shown in panel 1 in Figure 4.1. This must also be accounted for in the calculation of Δt.

Due for launch in 2027, the European Space Agency (ESA) propose to place a space weather monitor at the L5 Lagrangian point, a gravitational null located 60° behind Earth (Davies 2020). This point would provide a view of the Sun-Earth line, and so could give a side-on view of Earth directed CMEs (Akioka et al. 2005). It would also present the opportunity to use corotation forecasts from L5 to predict the solar wind conditions near-Earth with a lead time of approximately 4.5 days. The accuracy of corotation forecasts from L5 has been investigated using data from the Solar Terrestrial Relations Observatory [STEREO] (Kaiser et al. 2008) mission during specific phases of the operational lifetime (Bailey et al. 2020; Kohutova et al. 2016; Simunac et al. 2009; Thomas et al. 2018). The STEREO mission consisted of two spacecraft, STEREO-A ahead of Earth's orbit and STEREO-B behind, moving away from Earth at a rate of $\sim$ 22.5° per year (Kaiser et al. 2008). Combining these observations with near-Earth data from the Advanced Composition Explorer [ACE] (Stone et al. 1998) or the OMNI data set (Vokhmyanin et al. 2019) provides a number of periods where there are two spacecraft 60° apart in longitude. Simunac et al. (2009) demonstrated that the profiles of solar wind speed are similar when using STEREO-B as a forecast for STEREO-A at $\phi \approx 60^\circ$ in July 2008. Further to this, Kohutova et al. (2016) found that using corotation from STEREO-B to ACE improved the forecast of the B_z component of the interplanetary magnetic field when compared to a 27-day recurrence forecast. Thomas et al. (2018) used different combinations of STEREO and ACE to show the effectiveness of an L5 monitor. It was found that a number of solar wind parameters, including speed, density and temperature, were well predicted using four combinations of spacecraft, with $\phi \approx 60^\circ$, to produce corotation forecasts. The geomagnetic storm time index (Dst) has also been effectively forecasted through corotation from STEREO-B to OMNI, when $\phi \approx 60^\circ$ (Bailey et al. 2020). In both cases, corotation provides an improvement from 27-day recurrence.

Although extensive research has been conducted on the effectiveness of corotation from L5, previous studies have been limited to short periods of time when the spacecraft are separated by 60°. The majority of previous analysis has been around periods of low solar activity (solar minimum). The corotation forecasting method assumes a steady state solar wind; however, in reality the solar wind varies with time and the rate of evolution is linked to the 11-year solar cycle. At solar minimum, the steady state assumption is more valid as solar-wind structure slowly evolves with time. Conversely, at solar maximum the higher activity levels and rapid evolution of solar-wind structure lead to the steady Sun assumption breaking down more readily (Owens et al. 2013). This means that longer corotation times, and therefore longer forecast lead times, are generally expected to be more accurate at solar minimum than at solar maximum.

As the Sun progresses through the solar cycle, the latitudinal structure of the solar wind changes. At solar minimum, there is a slow solar-wind band centred on the heliographic equator, with faster winds emanating from coronal holes at higher latitudes (McComas et al. 2003). This latitudinal ordering breaks down at solar maximum, due to the weaker dipole component of the Sun's magnetic field (Wang and Sheeley Jr 1991). The variation in latitudinal structure is important for corotation forecasts, as spacecraft in the ecliptic plane vary in heliographic latitude $[\theta]$ owing to the 7.25° tilt between the ecliptic plane and the rotational plane of the Sun, as shown in panel 3 of Figure 4.1. This results in a 'latitudinal offset' $[\Delta\theta]$ between the point of observation (the source) and the location where the forecast is required (the target). The term 'latitudinal offset' is used here as the latitudinal separation of the source and target spacecraft. This can introduce a representation error into corotation forecasts that needs to be accounted for in solar-wind data assimilation (Owens et al. 2020b).

As will be demonstrated below, the available observations from the STEREO mission make the effect of $\Delta\theta$ on solar wind speed corotation forecast accuracy difficult to disentangle from solar activity and Δt. In particular, the contribution of latitudinal offset errors to the total corotation forecast error is expected to be significantly higher at solar minimum than solar maximum. Conversely, the contribution from Δt will increase as the corona becomes more dynamic, which is at solar maximum. Finally, as all spacecraft are in the ecliptic plane, large latitudinal spacecraft separation can only occur when there is also large longitudinal separation (and hence large Δt). Thus all contributions to corotation forecast error are interdependent.

Previous study into the error introduced by $\Delta\theta$ in corotation forecasts has been investigated using steady-state model output. Owens et al. (2019) and Owens et al. (2020b) showed that $\Delta\theta$ can have a significant effect on corotation forecast error, especially at solar minimum. It was found that during solar minimum, due to the band of slow solar-wind near the heliographic equator, the solar wind could be considered broadly similar up to $|\Delta\theta| = 3°$. Although this increases to $|\Delta\theta| = 10°$ during solar maximum, there are increasing numbers of transient events and smaller-scale turbulence (Owens et al. 2020b). This shows a breakdown of the latitudinal structure found at solar minimum. Corotation forecast error from L5 to Earth was shown to be up to 80 km s^{-1} at solar minimum purely due to $\Delta\theta$ (Owens et al. 2019), though the peak value occurs around the winter and summer solstices. Averaged over the year, this reduces to around 50 km s^{-1}. Although models have shown that a strong effect of $\Delta\theta$ on wind speed corotation forecast error is expected, Allen et al. (2020) found that the arrival time of SIR/ CIRs does not seem to be affected. This may be the result of compensating errors, as shown below.

The Owens et al. (2019) and Owens et al. (2020b) work was motivated by improving solar-wind data assimilation (DA) capabilities (Lang et al. 2017, 2021). DA combines model output and observational data to find an optimal estimation of reality and is beginning to be used in solar-wind forecasting (Lang and Owens 2019). Quantifying the error from $\Delta\theta$ allows observations to be fully utilised through the specification of more accurate observation error covariances and the removal of any potential biases. Observation errors can arise through poor specification of the observations position, resolution difference between observations and the model, pre-processing errors and many others. These are known as representivity errors, as they are errors that arise from poor representation of the observations, and they are what we wish to specify with greater accuracy in solar wind DA. The interested reader can find more information on representivity errors in Janjić et al. (2018). DA is a step forward in the use of observations for solar wind forecasting as it allows for the observations to be mapped to all longitudes and radial distances, whereas corotation only gives a forecast for a single point. Current DA schemes developed for solar-wind forecasting make use of solar wind speed observations (Lang and Owens 2019; Lang et al. 2017, 2021). Although both corotation and DA can be used for forecasting parameters such as plasma density and magnetic polarity, at present, the DA methods presented in Lang et al. (2017) and Lang and Owens (2019) only use solar wind speed. Therefore, this is the parameter investigated in this study to aid comparison with DA results. Furthermore, the flow dominated regime of the solar wind, coupled with the “frozen-in” flux theorem indicates that if the solar wind speed is not correctly forecast, timing errors will be present in the forecast of other solar wind parameters, such as solar wind density and heliospheric magnetic field.

Due to the previous reliance on model output for analysis of the effect of $\Delta\theta$ on corotation forecast error, it is necessary to investigate whether this is present in observational solar wind data. Magnetohydrodynamic (MHD) models produce a solar wind that is ‘smoother’ than what is observed. This means that an error introduced from $\Delta\theta$ could be more easily detected as there are no transient events or any small-scale turbulence. Furthermore, analysis of the model output was steady-state, and so no time variation of the solar wind was captured. This study uses solar wind data from the STEREO mission and near-Earth to investigate the error introduced from $\Delta\theta$.

Sections 4.2 and 5.2 will detail the data and methods used to produce our analysis, followed by the results in Section 5.3. Finally, we will discuss and draw conclusions in Section 5.4.

4.2 Data

Corotation forecasts require spacecraft at approximately the same heliocentric distance [R] as the intended forecast position (typically Earth). Although radial scaling from Parker Solar Probe to 1AU has been successfully used for high-speed streams (eg. Perrone et al. 2019, Wijsen et al. 2021 and Allen et al. 2021), here we use observations from 1AU. The solar wind captured in this study covers a broad range of solar wind speeds and so as it evolves radially, solar wind streams will interact, and this may not be captured when scaling. Spacecraft at approximately the same heliographic latitude [θ] are also required, but they must be separated in heliographic longitude [ϕ]. Greater ϕ implies longer forecast lead-times, but also decreases the forecast reliability, as will be demonstrated. Furthermore, throughout this analysis, only the radial component of the solar wind is used to allow direct comparison between observation sources.

The twin STEREO (Kaiser et al. 2008) spacecraft, in conjunction with near-Earth observations from the OMNI dataset (Vokhmyanin et al. 2019), provide a unique opportunity to test corotation forecasting and to better understand the factors that contribute to errors. Here, we use 1-hour STEREO plasma data obtained from the PLASTIC instrument (Galvin et al. 2008) and downloaded from CDAWeb (`cdaweb.gsfc.nasa.gov`). One-hour OMNI data are obtained from OMNIWeb (`omniweb.gsfc.nasa.gov`). The heliographic locations of STEREO spacecraft and Earth were obtained from OMNIWeb (`omniweb.gsfc.nasa.gov/coho/helios/heli.html`). We additionally use daily sunspot number from SILSO (*Sunspot Index and Long-term Solar Observations*) as a proxy for solar activity, downloaded from `sidc.be/silso`.

The STEREO spacecraft separate from Earth at a rate of $\approx 22.5°$ per year. Thus from launch in late 2006, they passed behind the Sun in 2014 leading to a data gap in the STEREO-A data from August 2014 to November 2015 and the loss of communication with STEREO-B from August 2014. While both spacecraft orbit the Sun at a distance close to 1 AU, R varies by up to 0.11 AU and must be accounted for when computing corotation forecasts. For completeness, we also take account of Earth's orbital eccentricity, though the associated change in R is small (~ 0.01 AU).

Here, OMNI data are used for near-Earth solar wind observations. Solar wind data from a succession of spacecraft (including Wind and ACE) located at the L1 Lagrange point is propagated to the bowshock of Earth, providing another source of

Source spacecraft	Target spacecraft	Time period
STEREO-A	OMNI	Feb 2007 - Aug 2019
STEREO-A	STEREO-B	Feb 2007 - Aug 2014
STEREO-B	OMNI	Feb 2007 - Aug 2014
STEREO-B	STEREO-A	Feb 2007 - Aug 2014
OMNI	STEREO-A	Feb 2007 - Aug 2019
OMNI	STEREO-B	Feb 2007 - Aug 2014

Table 4.1: The possible corotation forecast configurations using STEREO-A, STEREO-B and OMNI solar wind data, where the data from the source spacecraft is used as a forecast for the target spacecraft. A gap in the STEREO-A data exists between August 2014 and November 2015, affecting the STEREO-A to OMNI and OMNI to STEREO-A corotations.

data for use in corotation forecasts alongside the STEREO spacecraft (Vokhmyanin et al. 2019).

A single spacecraft can be used to provide a corotation forecast for one whole Carrington rotation (approximately 27 days) ahead. Such 27-day recurrence forecasts have already been considered in detail (Owens et al. 2013). The maximum latitude difference for a single in-ecliptic spacecraft over a 27-day period occurs at the equinoxes and reaches a magnitude of approximately 3.5°. As will be demonstrated later (see Figure 4.6), this is insufficient to quantify the latitudinal-offset contribution to corotation forecast error. Thus we focus on pairs of spacecraft. By combining the STEREO-A, B and OMNI spacecraft, there are 6 potential pairs of source (at the position where the solar wind observations are made) and target (at the position of the forecast) spacecraft. These can be seen in Table 4.1.

4.3 Methods

Solar-wind speed corotation forecasts are calculated and tested from combinations of the STEREO and OMNI spacecraft observations of solar wind speed [V]. Throughout this analysis, solar wind speed is the only quantity forecast and the radial component is used. Simunac et al. (2009) and Thomas et al. (2018) describe this process in terms of mapping between Carrington longitudes at pairs of spacecraft. We here use an equivalent description in terms of time, to make explicit a number of assumptions.

Each hourly V observation at the source spacecraft [V_S] is used to produce a forecast [V_F] at the target spacecraft location at a time Δt in the future:

$$V_F(t + \Delta t) = V_S(t) \tag{4.1}$$

where $V_S(t)$ is the observed solar wind speed at the source spacecraft at time t. Δt

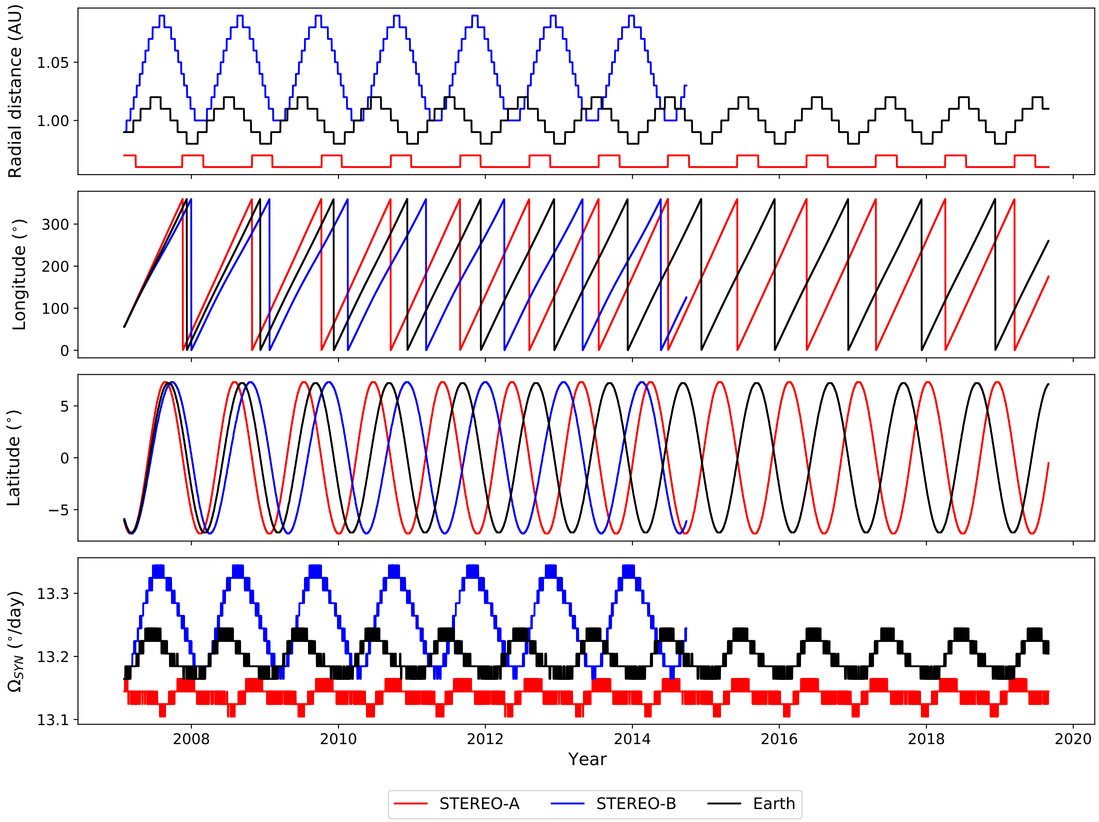


Figure 4.1: Variation of radial distance, heliographic longitude, heliographic latitude and synodic angular speed of STEREO-A (red), STEREO-B (blue) and Earth (black).

is the required time delay for the same solar wind observed at the source location to reach the target location. Consequently, it is also the forecast lead time. Δt consists of two elements: t_R, the time for solar wind to propagate between the source and target radial distances [R_S and R_T respectively] and t_C, the time for solar wind sources to rotate between the source and target longitudes, accounting for spacecraft motion in the inertial frame. This is shown schematically in Figure 4.2.

Figure 4.2a shows a time $t = t_0$, when the target spacecraft has a longitude of ϕ_{T0} and a radial distance of R_T. Similarly, the source spacecraft has a longitude of ϕ_{S0} and a radial distance of R_S. The first step is to ballistically map the source observations to a radial distance of R_T:

$$t_R = \frac{R_T - R_S}{V_S(t_0)} \tag{4.2}$$

where t_R is the radial propagation time. As the propagation in radial distance is purely ballistic, it ignores any stream interaction effects or solar wind acceleration. Thus this approach is only valid for $|R_T - R_S| << R_T$. Given the radial variations of the spacecraft are slow, we assume R_S and R_T are constant over the interval Δt. As Figure 4.1 shows, the spacecraft can be separated by up to approximately 0.1AU in radius. Therefore, depending on the solar wind speed V_S, the radial propagation time can range from 6 to 14 hours (for solar wind speeds of 700 and 300 kms^{-1} respectively). This can be a significant contribution to Δt and therefore it is important to include.

Figure 4.2b shows that during the interval t_R, the target spacecraft continues to move ahead in longitude at an orbital angular speed in the inertial frame of $\Omega_{I,T}$. Note that if $R_S > R_T$, t_R will be negative and spacecraft will move to smaller ϕ during radial propagation. STEREO-A, B and OMNI spacecraft have different average values of Ω_I, which allows the spacecraft to separate over time. For all three spacecraft these values also vary slowly over the year, owing to the slightly elliptical orbits. This is shown in panel 4 of Figure 4.1. We account for this effect by computing Ω_I from the change in ϕ over a 5-day window centred on time t_0. (This window is short enough to allow for the change in $\Omega_{I,F}$ over the year, but large enough to remove numerical noise from taking the time gradient of ϕ_V.) Over the interval Δt, which is typically a few days, it is reasonable to assume a constant value of $\Omega_{I,S}$, which allows us to express the longitude of the target spacecraft as:

$$\phi_T(t_0) = \Omega_{I,T} \cdot t_0 + \phi_{T0} \tag{4.3}$$

After a time t_R, the solar wind structure observed by the source spacecraft at

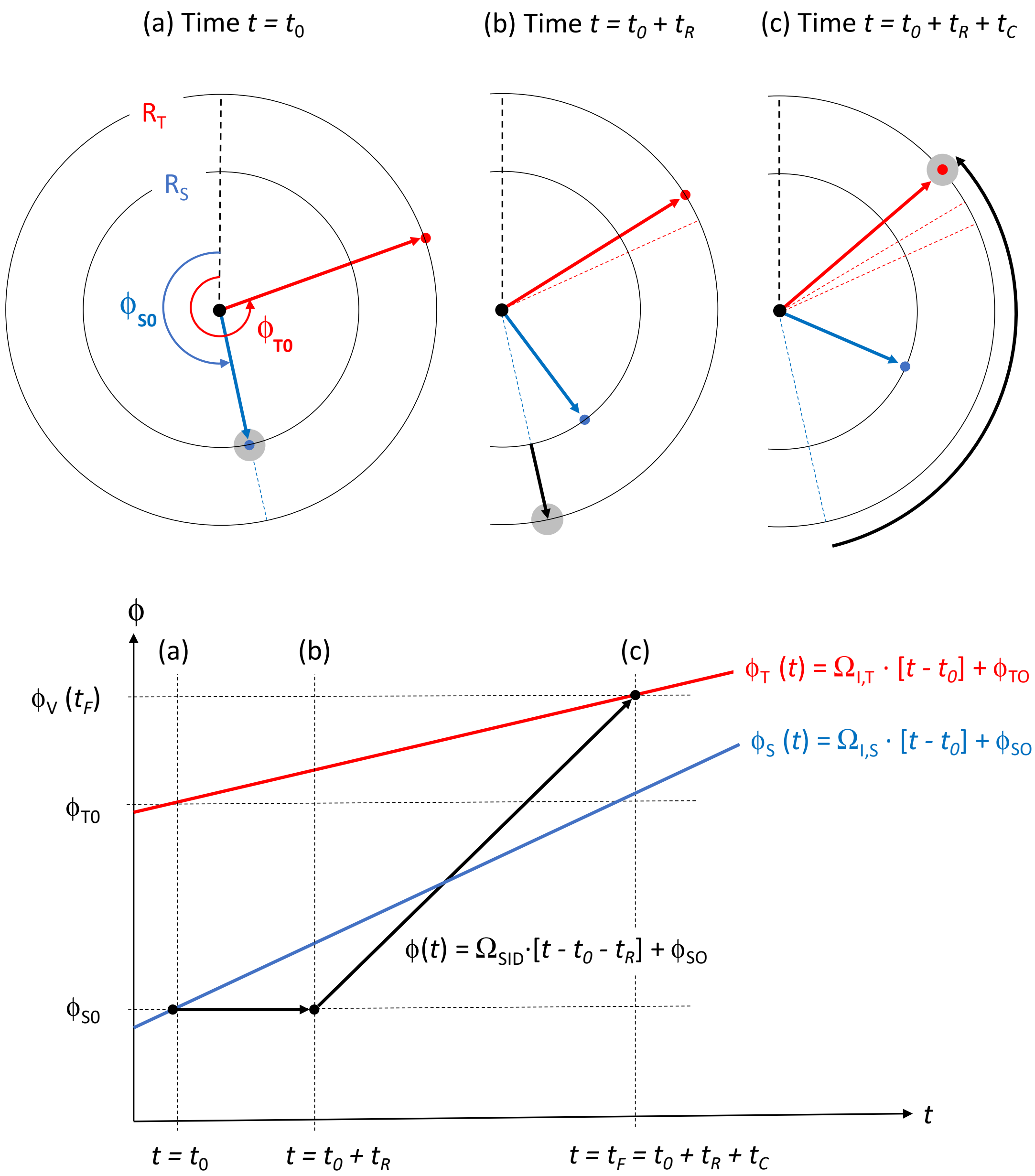


Figure 4.2: A schematic of the corotation forecast in the inertial frame. The grey circle tracks the position of a solar wind structure observed by the source spacecraft at time $t = t_0$. Top: (a) At time $t = t_0$, the target spacecraft (red dot) has a longitude of ϕ_{T0} and radial distance of R_T, while the source spacecraft (blue dot) is at ϕ_{S0} and R_S. (b) Radial propagation from R_S to R_T takes a time t_R (c) Corotation from ϕ_{S0} to the target spacecraft takes time t_C. Bottom: The same steps shown as a time series of ϕ. $\Omega_{I,S}$ and $\Omega_{I,T}$ denotes the orbital angular speed of the source and target spacecraft, respectively, while Ω_{SID} denotes the sidereal rotation speed of the Sun.

$t_0 = 0$ will corotate with the solar wind, meaning its longitude varies as:

$$\phi(t_0) = \Omega_{SID} \cdot [t_0 - t_R] + \phi_{S0} \tag{4.4}$$

where Ω_{SID} is the sidereal rotation speed of the Sun (i.e., 2.86×10^{-6} rad s^{-1}). Thus it will encounter the target spacecraft at a time Δt, where:

$$\Delta t = \frac{1}{\Omega_{SYN,T}}[\phi_{T0} - \phi_{S0} + \Omega_{SID} \cdot t_R] \tag{4.5}$$

where $\Omega_{SYN,T}$ is the synodic orbital angular speed of the target spacecraft, given by $\Omega_{SYN,T} = \Omega_{SID} - \Omega_{I,T}$.

For the STEREO-B corotation forecast of V at STEREO-A's position, Δt is shown as the orange line in Figure 4.3c.

The observed radial solar wind speeds at the target and source spacecraft (V_T and V_S, respectively) are taken from 1-hour resolution data. While the forecast speed, V_F, is computed from hourly V_S data, the t_R term means that the computed forecast speed, V_F, is no longer on a regular 1-hour time step. Thus, V_F is linearly interpolated back to a standard hourly time base for direct comparison with the V_T.

Figure 4.4 shows a stack plot of V at STEREO-A's location. This uses a forecasted solar wind from STEREO-B, V_F in blue, with observations by STEREO-A, V_T in red. V_F is produced from time shifting the STEREO-B observations and V_T is used to verify the forecast. Data have been further averaged to 1-day resolution for clarity. It can be seen that the agreement is extremely good for 2007 and 2008, and becomes gradually worse as time progresses. To quantify the degree of agreement we use the mean absolute error (MAE) between the observed and forecast V:

$$MAE = \frac{1}{N}\sum_{n=1}^{N} |V_T(t_n) - V_F(t_n)| \tag{4.6}$$

where N is the total number of time steps considered. MAE is a point-by-point analysis and thus small timing errors in the forecast can be heavily penalised (Owens 2018). In the case of a corotation forecast, where the solar wind structure is assumed to be merely time lagged between two positions, the general solar-wind structure should be well reproduced. Therefore MAE is an appropriate metric to use as timing errors indicate a time evolution of the solar wind, for which corotation forecasts should be penalised.

The green line in Figure 4.3d shows 27-day averages of MAE for STEREO-B observations used to forecast the solar wind conditions at STEREO-A. The increase in MAE in 2007 through 2009 is consistent with the divergence in the time series seen in Figure 4.4. When using other corotations, such as STEREO-B to OMNI

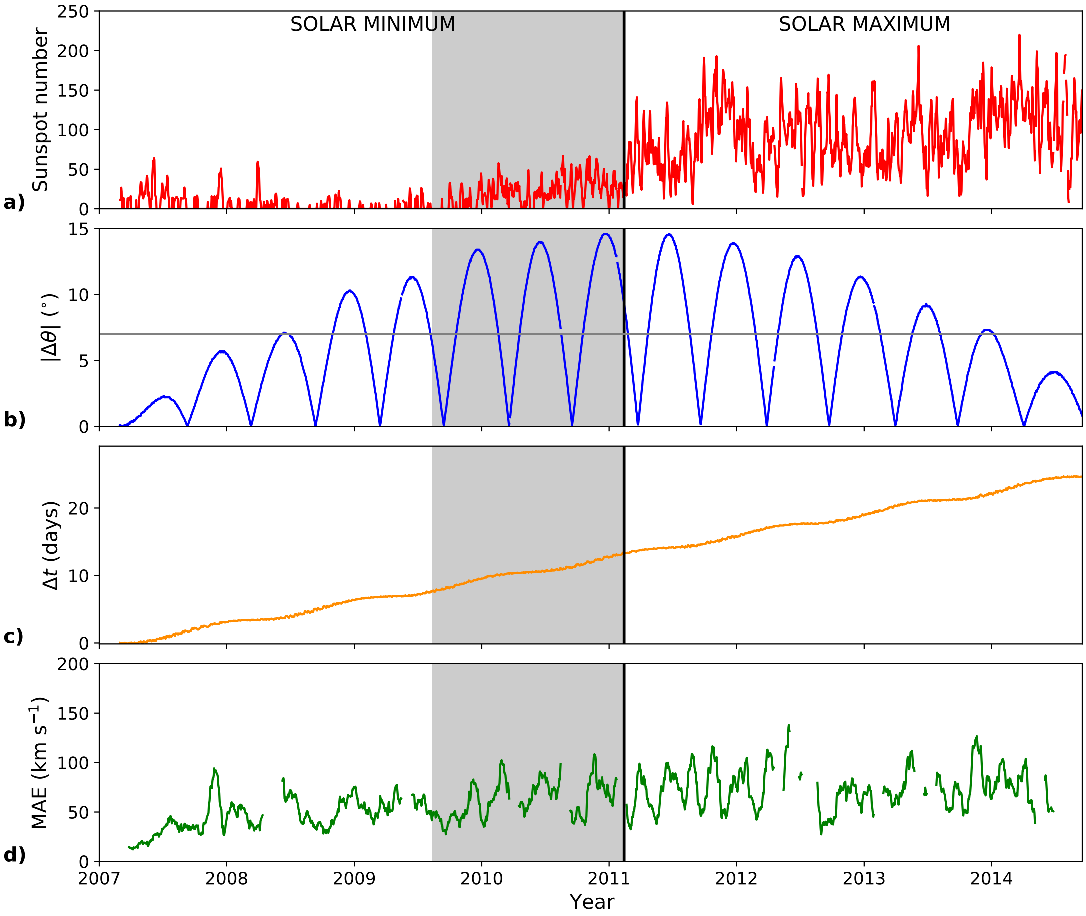


Figure 4.3: Time series at daily resolution for the STEREO-B to A corotation forecast, covering the duration of the STEREO-B lifetime. Here, STEREO-B is the source spacecraft and STEREO-A is the target spacecraft. (a) Sunspot number. (b) $|\Delta\theta|$, the absolute latitude difference between the source and target location. The grey horizontal line indicates 7°, the separation between high and low $|\Delta\theta|$ used in this study. (c) Δt, the forecast lead time. (d) 27-day rolling average of the daily mean absolute error (MAE) for the corotation forecast. In all panels, the vertical black line separates the definitions of solar minimum and solar maximum used in this study (see main text). The grey-shaded region highlights an interval during solar minimum used for further investigation.

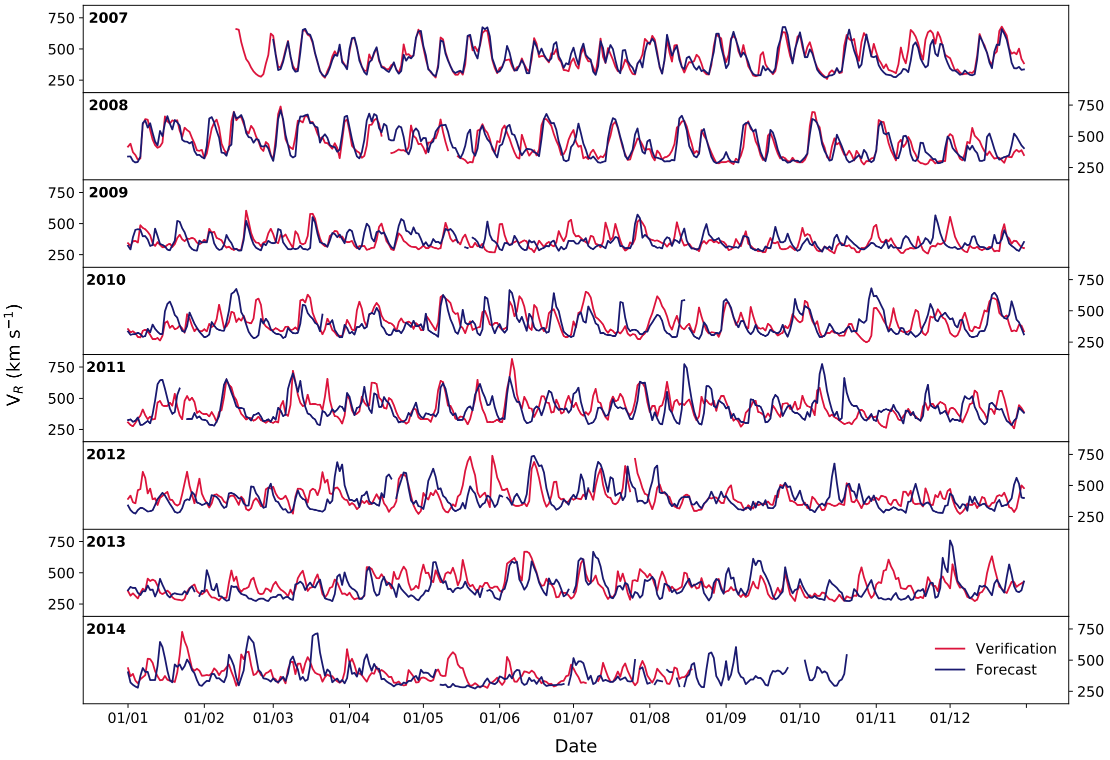


Figure 4.4: Time series of solar wind speed at STEREO-A's location. In red is the observed solar wind speed by STEREO-A. In blue is forecast solar wind speed, V_F, produced from time shifting the STEREO-B observations. Thus, STEREO-A is the target spacecraft providing the observations to verify the forecast, and STEREO-B is the source spacecraft used to produce V_F. Data have been averaged to 1-day resolution for clarity.

and OMNI to STEREO-A, for Figures 4.3 and 4.4, it is found that the plots are qualitatively similar. These are provided in the supporting information for the interested reader. In Section 5.3 we consider the different contributions to this MAE, with a focus on quantifying the role of latitudinal difference between the forecast and verification spacecraft. This is measured as:

$$\Delta\theta(t) = \theta_T(t) - \theta_S(t - \Delta t) \tag{4.7}$$

where θ_T and θ_S are the heliographic latitudes of the target and source spacecraft/positions, respectively. The absolute value of $\Delta\theta(t)$ is shown in Figure 4.3b.

Combining the data from the 6 corotation configurations listed in Table 4.1 allows for the most robust analysis of the variables affecting the forecast error. There are, however, a number of biases in the sampling of this dataset that make it complex to isolate individual factors in corotation forecast error. In particular, the motion of the STEREO spacecraft results in strong aliasing of the $|\Delta\theta|$ and Δt with both each other and the solar cycle, and mean that confounding variables and compensating errors are an ever-present problem.

In order to coarsely isolate the influence of solar activity, the data are split into periods of solar minimum and maximum. Given only a single solar maximum is covered by the STEREO data set and the loss of communication with STEREO-B reducing it further, a simple sunspot number threshold is appropriate. We choose a value of 75 sunspots in the total daily sunspot number time series, as this selects the transition from solar minimum to solar maximum in February 2011, when there is a clear step-change in daily sunspot number. Using the same threshold puts the transition from solar maximum to minimum in March 2016.

We further split the data into periods of high and low $|\Delta\theta|$. Given the maximum $|\Delta\theta|$ available with the STEREO/ OMNI dataset is approximately 14.5°, we use a cut-off of 7°, though this will be further investigated in the remainder of the study.

4.4 Results and discussion

We first consider the effect of forecast lead time, Δt, on forecast accuracy. To minimise the influence of $|\Delta\theta|$, we limit analysis to periods when $|\Delta\theta| < 2^\circ$. Figure 4.5 shows forecast MAE as a function of Δt, for all spacecraft pairings. In general, MAE at solar maximum is higher than at solar minimum for the same Δt, as expected. At solar minimum, there is a trend for increasing MAE with increasing Δt out to around 7 days. However, past 20 days, MAE decreases. $\Delta t > 20$ days is confined exclusively to STEREO-A to STEREO-B forecasts very early in the mission, when the spacecraft were still near Earth. Thus the reduced MAE may actually be the

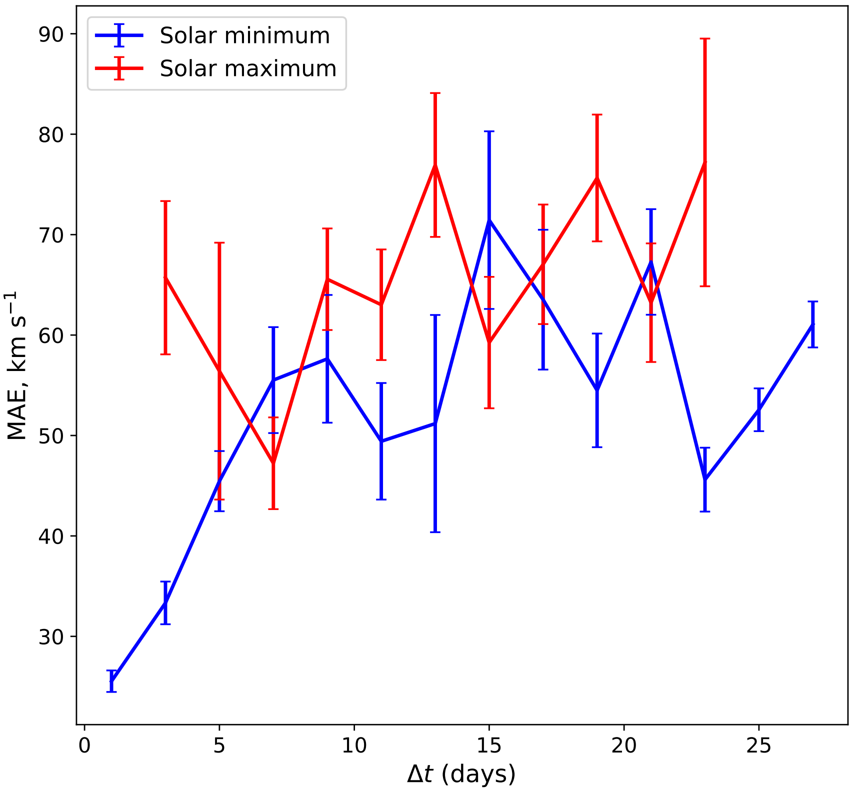


Figure 4.5: Forecast mean absolute error (MAE) as a function of forecast lead time, Δt, for solar minimum and maximum periods. Lines show the mean values, while error bars span one standard error on the mean. All STEREO/OMNI spacecraft pairings are used. Only periods with $|\Delta\theta| < 2°$ are included.

result of particularly quiet solar activity levels at this time. Despite 2007-2010 being classified as solar minimum on the basis of a sunspot number threshold, Figure 4.4 shows a clear difference in the character of the solar wind speed structures between 2007 and 2009, with recurrent fast streams giving way to more persistent slow wind.

Encapsulated within Δt is the effect from radial separation. Analysing this effect, limiting to times when $|\Delta\theta| < 2°$, it is found that there is so obvious trend between MAE and radial separation. This plot is included in the supporting information for the interested reader.

Next we consider the effect of $|\Delta\theta|$. Figure 4.6 shows the differing latitudinal dependencies between solar minimum and solar maximum. At solar maximum, the MAE remains relatively constant at ~ 80 km s^{-1}. At solar minimum, there is a clear correlation between $|\Delta\theta|$ and MAE. However, it is here likely the result of aliasing of $|\Delta\theta|$ with Δt and/or sunspot number. This is shown by Figure 4.7. During the solar minimum interval, splitting the data into high and low $|\Delta\theta|$ does not give equal sampling of either sunspot number or forecast lead time. Thus while MAE is lower for low $|\Delta\theta|$, we cannot rule out the lower mean sunspot number or the existence of lower Δt values as being the causal effect.

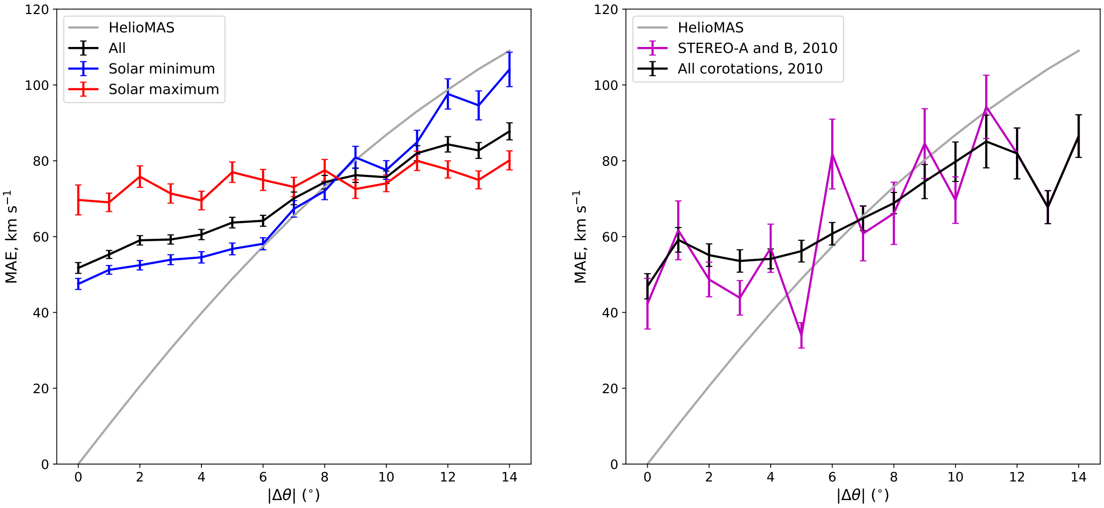


Figure 4.6: Variation of mean absolute error (MAE) with $|\Delta\theta|$. Unless otherwise stated, all spacecraft pairings are included. Lines show the mean values, while error bars span one standard error on the mean. Left: The entire dataset (black), further split into solar maximum (red) and minimum (blue). While there appears to be a correlation between MAE and $|\Delta\theta|$ at solar minimum, this could be a result of aliasing of $|\Delta\theta|$ with Δt. Thus we also show (right) a combination of STEREO-A to B and STEREO-B to A corotations (magenta) and using all corotations (black) for a limited period from August 2009 to February 2011. Both panels show the equivalent modelled MAE at different $|\Delta\theta|$ for steady-state solar wind model solutions (grey). See body text for further description.

For comparison, Figure 4.6 also shows the MAE as a function of $|\Delta\theta|$ from previous modelling studies (Owens et al. 2019, 2020b). In those studies, the "Magnetohydrodynamics Algorithm outside a Sphere" (MAS) global coronal and heliosphere model (Linker et al. 1999; Riley et al. 2012) was constrained by the observed photospheric magnetic field and used to reconstruct the solar wind speed at 1 AU. As the solar wind solutions are steady-state, the MAE obtained is purely a result of $|\Delta\theta|$ and should represent a lower limit to that expected observationally, where transient structures, small-scale turbulence and rapidly-evolving large-scale structures providing additional sources of MAE for corotation forecasts. This agrees somewhat with our findings: At low $|\Delta\theta|$ values ($< 7°$), the observed MAE is higher than the model result, suggesting that $|\Delta\theta|$ is not the dominant contribution to MAE in this parameter range. The model and observations approximately agree in the range $7° \leq |\Delta\theta| \leq 11°$, suggesting that $|\Delta\theta|$ has become a significant contribution to MAE. For $|\Delta\theta| > 11°$, the observational MAE is lower than the model result, suggesting that the difference in solar wind speed with latitude present in the models gets reduced by other processes.

In order to better isolate the effect of $|\Delta\theta|$ in our observational dataset it is necessary to further subdivide it. However, a competing requirement is to retain

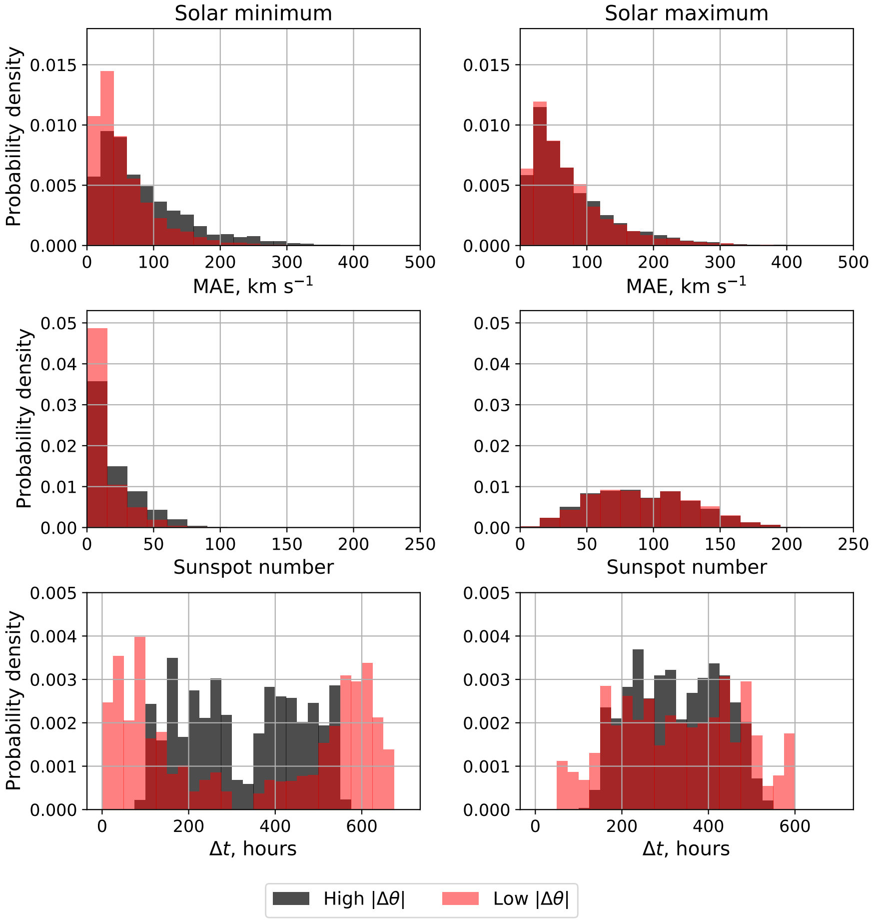


Figure 4.7: Probability density of (top) the mean absolute error (MAE) between the observed and forecast V, (middle) sunspot number and (bottom) Δt. Data have been split into solar minimum (left) and solar maximum (right), and into high (black) and low (red) $|\Delta\theta|$ using a threshold of 7°. All STEREO/ OMNI spacecraft pairings are included. The dark red areas show where there is overlap between high and low $|\Delta\theta|$.

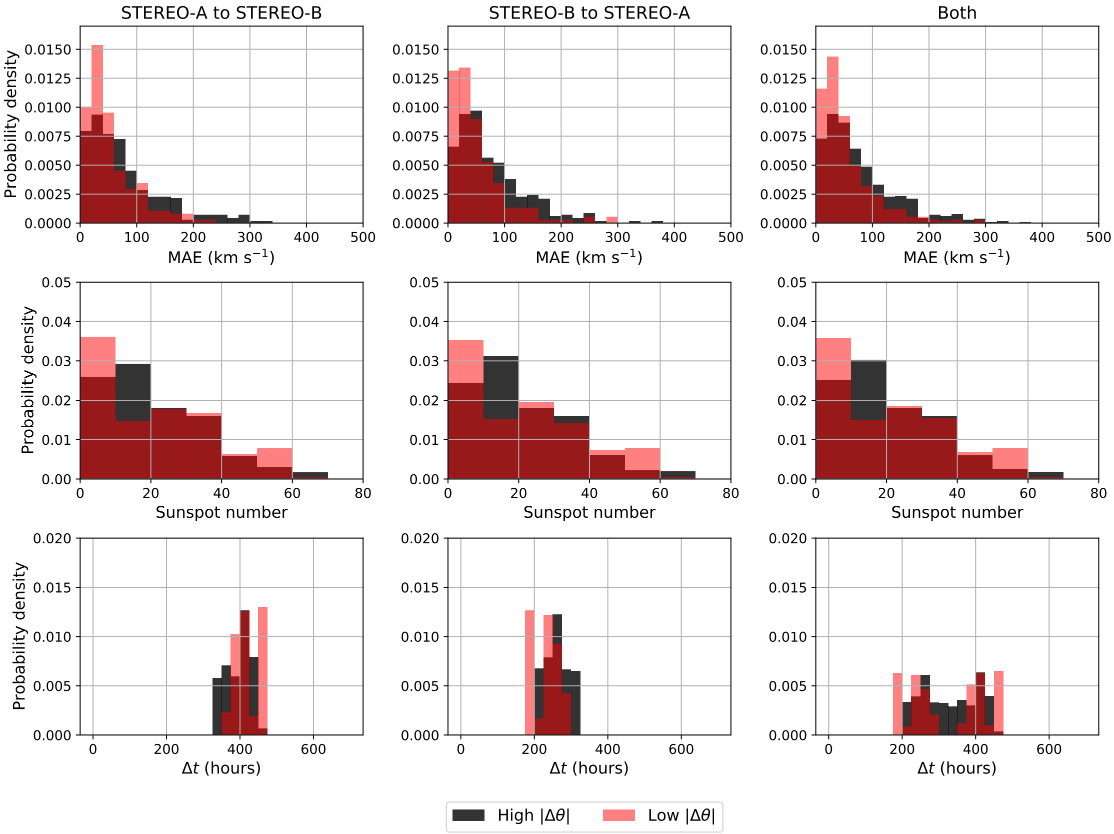


Figure 4.8: Probability density for STEREO-A to B (left) and STEREO-B to A (middle) corotation forecasts limited to the interval August 2009 to February 2011. The right column shows the combined dataset. Rows show (top) the mean absolute error (MAE) between the observed and forecast V, (middle) sunspot number and (bottom) Δt. The dark red areas show where there is overlap between high and low $|\Delta\theta|$.

enough data for meaningful statistical analysis. A compromise of these factors is shown as the grey-shaded region in Figure 4.3, which spans August 2009 to February 2011. This provides a period of time where $|\Delta\theta|$ rises to a large enough value (maximum of 14.9°) but it is still relatively close to solar minimum, when the latitudinal effect is expected. High and low $|\Delta\theta|$ periods are approximately evenly spaced through this period, meaning both contain similar levels of solar activity. Finally, by comparing the STEREO-A to B forecasts with STEREO-B to A forecasts, we can effectively eliminate Δt as contributing factor, as the two combinations have opposing Δt trends.

As can be seen in Figure 4.8, the interval from August 2009 to February 2011 provides approximately equal sunspot number distributions for high and low $|\Delta\theta|$ periods (see also Table 4.2). Δt distributions are also in approximate agreement, particularly when STEREO-A to B and STEREO-B to A forecasts are combined. Furthermore, the difference in the MAE histograms is consistent when the source/target spacecraft, and hence bias in Δt, are switched. Thus we can reasonably conclude

		MAE (km s^{-1})	**Sunspot number**	**Δt (hours)**
STEREO-A to B	**High $\|\Delta\theta\|$**	78.8 ± 3.5	20.1 ± 0.8	395.5 ± 1.8
	Low $\|\Delta\theta\|$	54.0 ± 3.2	19.7 ± 1.3	418.6 ± 2.4
STEREO-B to A	**High $\|\Delta\theta\|$**	76.7 ± 3.1	20.3 ± 0.8	259.7 ± 1.7
	Low $\|\Delta\theta\|$	52.6 ± 3.7	20.2 ± 1.3	235.2 ± 2.4
Both corotations	**High $\|\Delta\theta\|$**	77.7 ± 2.4	20.2 ± 0.6	327.8 ± 2.8
	Low $\|\Delta\theta\|$	53.3 ± 2.4	19.9 ± 0.9	327.1 ± 5.0
All corotations	**High $\|\Delta\theta\|$**	74.0 ± 1.6	21.8 ± 0.4	325.4 ± 3.9
	Low $\|\Delta\theta\|$	56.4 ± 1.1	18.7 ± 0.4	330.2 ± 4.5

Table 4.2: Averages, with associated standard errors, of the corotation forecast MAE, sunspot number and forecast lead time for the STEREO-A to B and STEREO-B to A corotations, for both combined, and for all possible corotations combined (i.e., including OMNI), for the period August 2009 to February 2011.

	MAE (%)	**Sunspot number (%)**	**Δt (%)**
STEREO-A to B	46.0	2.0	-5.5
STEREO-B to A	45.9	0.3	10.4
Both corotations	45.9	1.2	0.2
All corotations	31.4	16.7	-1.5

Table 4.3: Percentage increase (calculated using Equation 4.8) from low to high $|\Delta\theta|$ for average MAE, sunspot number and Δt from Table 4.2. This covers the period August 2009 to February 2011.

that the observed difference in MAE between high and low $|\Delta\theta|$ periods shown here is not the result of aliasing with other effects.

It is apparent from Table 4.2 that there is a distinct difference in the average corotation forecast MAE for high and low $|\Delta\theta|$ times. The average values for MAE are statistically distinct (i.e., differ by far more than one standard error on the mean). Low $|\Delta\theta|$ periods produce a significantly lower average corotation forecast MAE for both STEREO-B to A and A to B corotations, thus Δt is unlikely to be a factor. This is further seen by mean Δt for the combined dataset being the same for high and low $|\Delta\theta|$ times within uncertainties. The same result is found for sunspot number.

Further to this, Table 4.3 shows the percentage increase in MAE, sunspot number and Δt from low to high $|\Delta\theta|$ averages in Table 4.2. Percentage difference is calculated as:

$$\text{\% difference} = \frac{x_{high} - x_{low}}{x_{low}} \times 100 \tag{4.8}$$

where x_{high} and x_{low} is the data at high and low $|\Delta\theta|$ respectively.

The percentage increase in MAE is consistent at around 46% for the two corotations and when both datasets are combined. It is slightly lower when including the OMNI data, as this skews the data set in favour of low $|\Delta\theta|$ occurrence. For sunspot number, we can see that there is very little increase from low to high $|\Delta\theta|$,

	Correlation coefficient	**p-value**
$\boldsymbol{\lvert\Delta\theta\rvert < 6}$	0.40	0.44
$\boldsymbol{\lvert\Delta\theta\rvert \geq 6}$	0.75	0.02

Table 4.4: Linear correlation coefficients and corresponding p-values for MAE and $|\Delta\theta|$, split into $|\Delta\theta| < 6$ and $|\Delta\theta| \geq 6$. This uses all corotations and covers the period from August 2009 to February 2011. The p-value represents the probability of the correlation occurring by chance.

with the percentage increases below 2% for the STEREO-A and B corotations. Although there is a positive and negative percentage increase between Δt from low to high $|\Delta\theta|$, this seems to have little impact on the difference for MAE. As Figure 4.5 shows, for solar minimum, the effect of corotation time on MAE increases to $\Delta t = \sim 7$ days (168 hours), where it remains approximately constant thereafter. This could explain the minimal effect that differing Δt has for high and low $|\Delta\theta|$, as both for both cases, $\Delta t > 8$ days.

Comparing these results with the right-hand panel of Figure 4.6, we can see that the MAE contribution from $|\Delta\theta| < 6$ remains approximately constant, before increasing. To assess this, correlation coefficients of MAE and $|\Delta\theta|$ were calculated for $|\Delta\theta| < 6$ and $|\Delta\theta| \geq 6$. These are shown in Table 4.4. We can see the transition between the two classes of $|\Delta\theta|$, as there is a marked increase in the correlation. The p-value represents the probability of the correlation occurring by chance, so the low p-value for $|\Delta\theta| \geq 6$ indicates that there is a strong and significant correlation, whereas below this, the relationship is not significant.

4.5 Conclusions

Accurate prediction of near-Earth solar wind conditions over the coming hours to days is vital for space weather forecasting and mitigation. By assuming a steady state solar wind, longitudinally-separated observations in or near the ecliptic plane can be used as a forecast for further on in the Sun's rotation. This implicitly assumes that the solar wind structures seen at both locations will be the same. The accuracy of this "corotation forecast" is affected by time evolution of the solar wind and latitudinal separation ($|\Delta\theta|$). The error due to time evolution is itself a function of how rapidly the solar wind structure is evolving and the lead time of the forecast, which is due to the longitudinal and radial separation of the source and target spacecraft. Increased time evolution is approximately a function of solar activity, and this leads to the steady state assumption breaking down more readily, and so longer corotation times being less valid, at solar maximum. $|\Delta\theta|$ can introduce a forecast error through sampling a solar wind at the source spacecraft that is not representative of the target spacecraft. This is most important at solar minimum, where a narrow band of slow solar wind is located near the Sun's equator.

Several studies have shown the effectiveness of corotation forecasts from L5 (Bailey et al. 2020; Kohutova et al. 2016; Simunac et al. 2009; Thomas et al. 2018). These studies used data from periods when combinations of the STEREO and ACE spacecraft were 60° apart in heliographic longitude (ϕ), and showed that they outperformed the 27-day recurrence forecasts. Due to the limitation of periods where there was such a separation in ϕ, previous analysis has mostly been restricted to solar minimum. Here, similar methods have been followed; however, the corotation forecasts have not been limited to separations of 60° and so a wider range of the solar cycle has been sampled in order to fully understand the effect of $|\Delta\theta|$. Only solar wind speed has been considered here, due to its use in data assimilation and its ability to order other solar wind parameters.

We have produced corotation forecasts using solar wind speed data from the STEREO mission and the OMNI dataset. This produces 6 corotation configurations, mostly covering 2007 to 2014. However, a large amount of aliasing exists within this data set. As the STEREO spacecraft separate, the forecast lead time (Δt) increases, as does the maximum $|\Delta\theta|$ and sunspot number. Therefore, it is difficult to ascribe an increase in forecast error to a single factor. At solar minimum, the solar wind is highly structured in latitude, whereas at solar maximum this structure is far more dynamic (e.g. Figure 5 of Owens (2020a)). Thus, the contribution of $|\Delta\theta|$ to corotation forecast error is expected to be a strong function of the solar cycle, with largest contributions at solar minimum and smallest at solar maximum. Conversely, solar-wind structure evolves much more slowly at solar minimum than solar maximum, so the contribution of Δt to corotation forecast error is expected to be largest at solar maximum. Finally, as STEREO and OMNI spacecraft are all in the same orbital plane, the largest $|\Delta\theta|$ values are restricted to times of large longitudinal separation and hence large Δt.

The combined STEREO/ OMNI data sets, however, do provide a wide range of Δt values during both solar minimum and solar maximum even when restricted to $|\Delta\theta| < 2°$. In general, for a given $|\Delta\theta|$ value, corotation forecast mean absolute error (MAE) is higher at solar maximum than solar minimum. If the elevated MAE at solar maximum was the result of increased time-variability of ambient solar-wind structures, we would expect a correlation of MAE with Δt. Instead, MAE is fairly constant across the range of Δt available at solar maximum, suggesting the increased MAE is the result of increased frequency of transient solar-wind structures at this time (Gopalswamy et al. 2009). At solar minimum, MAE increases steadily with Δt up to around 7 days.

To attempt to isolate the $|\Delta\theta|$ contribution, we focused analysis on a period of time where sunspot number and the Δt were fairly constant so that $|\Delta\theta|$ could be isolated. This period covered August 2009 to February 2011. Combining STEREO-A to B and B to A corotation forecasts for this period, the MAE in the corotation forecasts was significantly smaller for low $|\Delta\theta|$ periods (taken to be $< 7°$) than for high $|\Delta\theta|$ periods ($\geq 7°$). The mean sunspot number and Δt values for low and high $|\Delta\theta|$ periods show no significant difference. Thus we can attribute the difference in MAE to latitudinal offset with reasonable confidence.

Looking in more detail at this August 2009 to February 2011 period, there is a strong correlation between $|\Delta\theta|$ and MAE for $|\Delta\theta| < 6°$, but not for $|\Delta\theta| \geq 6°$. At around this same latitudinal separation value, the observed forecast MAE becomes comparable to that expected from heliospheric modelling, where only the latitudinal effect is present (i.e. there is no time evolution and no solar wind transients Owens et al. (2020b)). See the grey curve in Figure 4.6. Thus we suggest that for $|\Delta\theta| < 6°$, the latitudinal offset error in corotation forecasts is present, but is not detectable due to other factors (such as time evolution and CMEs) dominating. For $|\Delta\theta| \geq 6°$, however, the latitudinal offset is the primary source of corotation forecast error and the magnitude is in good agreement with that expected from steady-state solar-wind modelling. These findings are broadly consistent with studies qualitatively looking at SIRs (eg. Jian et al. 2019). This explains why a latitudinal effect is not detectable in analysis of 27-day recurrence forecasts. The maximum change in latitude for a single spacecraft in the ecliptic plane from one Carrington rotation to the next is 3.5°, hence the signal not being present. This also implies that for using observations in data assimilation (DA), if $|\Delta\theta| < 6$, observation errors could be assumed constant, where above this, the observational error would be dependent on $|\Delta\theta|$.

Typically, observations with $|\Delta\theta| < 6$ would be preferable for DA and corotation forecasting, as at this separation, the latitudinal effect is minimised in comparison with other sources of error. This finding has implications on the use of future L5 mission data. L5 reaches a maximum $|\Delta\theta|$ of around 5 degrees with Earth (at times close to the summer and winter solstices), meaning that the latitudinal variation can be largely disregarded. Owens et al. (2019) showed that there is a time-of-year variation in the modelled impact of $|\Delta\theta|$ on MAE. This is not investigated here due to data limitations.

The future space weather monitoring mission to the L5 Lagrange point offers a new opportunity for corotation forecasts for the solar wind. The investigation into the effect of $|\Delta\theta|$ on forecast error here has found that for $|\Delta\theta| < 6$, there is a minimal impact due to other sources of error. However, for $|\Delta\theta| \geq 6$, the error contribution increases and there is a clear relationship between $|\Delta\theta|$ and forecast MAE. Due to the maximum $|\Delta\theta|$ between L5 and Earth being 5 degrees, this result implies that the effect from $|\Delta\theta|$ on the forecast error would be minimal.

Moving forwards, this work can aid the effective use of observations in data assimilation for forecasting the solar wind. It will enable more accurate observation error covariances to be calculated when there is a $|\Delta\theta|$ between observations and Earth. Furthermore, this will allow observational errors that result from $|\Delta\theta|$ to be corrected, ensuring the DA methodologies perform optimally.

Data availability

We are grateful to the Space Physics Data Facility and National Space Science Data Center for OMNI. These data were downloaded frrom the OMNIWeb portal at `https://omniweb.gsfc.nasa.gov/ow.html`, and accessed through the HelioPy module in Python `https://docs.heliopy.org/en/0.6.7/`. STEREO-A and STEREO-B data were downloaded from the CDAWeb Data Explorer portal at `https://cdaweb.gsfc.nasa.gov/cgi-bin/eval2.cgi?dataset=STA_COHO1HR_MERGED_MAG_PLASMA&index=sp_phys`, but accessed through HelioPy. Spacecraft location data were downloaded from `https://omniweb.gsfc.nasa.gov/coho/helios/heli.html`. HelioMAS model output is available from the Predictive Science Inc. (website: `http://www.predsci.com/mhdweb/home.php`).

Acknowledgements

Harriet Turner is funded through SCENARIO grant number NE/S007261/1. Work was part-funded by Science and Technology Facilities Council (STFC) grant numbers ST/R000921/1 and ST/V000497/1, and Natural Environment Research Council (NERC) grant number NE/S010033/1.

4.6 Supporting information

The supporting information included here shows Figures 3 and 4 from the main text, but using alternative corotation configurations. Also included is a figure showing the variation of mean absolute error with radial separation of the source and target spacecraft.

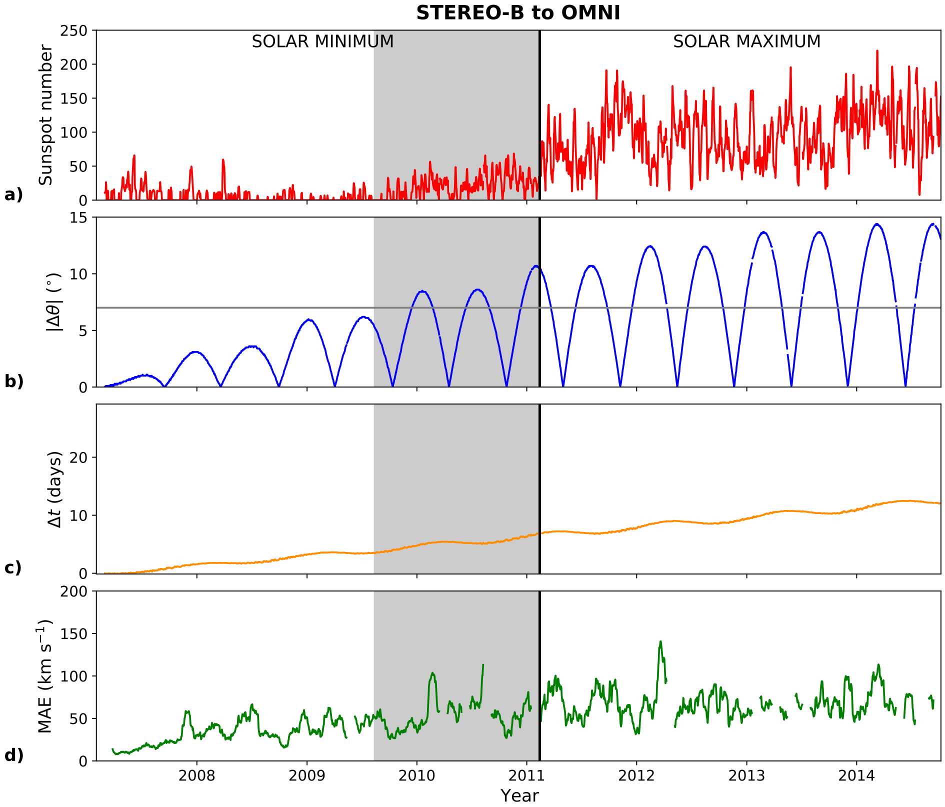


Figure 4.9: Time series at daily resolution for the STEREO-B to OMNI corotation forecast, covering the duration of the STEREO-B lifetime. Here, STEREO-B is the source spacecraft and OMNI is the target. (a) Sunspot number. (b) $|\Delta\theta|$, the absolute latitude difference between the source and target location. The grey horizontal line indicates 7°, the separation between high and low $|\Delta\theta|$ used in this study. (c) Δt, the forecast lead time. (d) 27-day rolling average of the daily mean absolute error (MAE) for the corotation forecast. In all panels, the vertical black line separates the definitions of solar minimum and solar maximum used in this study (see main text). The grey-shaded region highlights an interval during solar minimum used for further investigation.

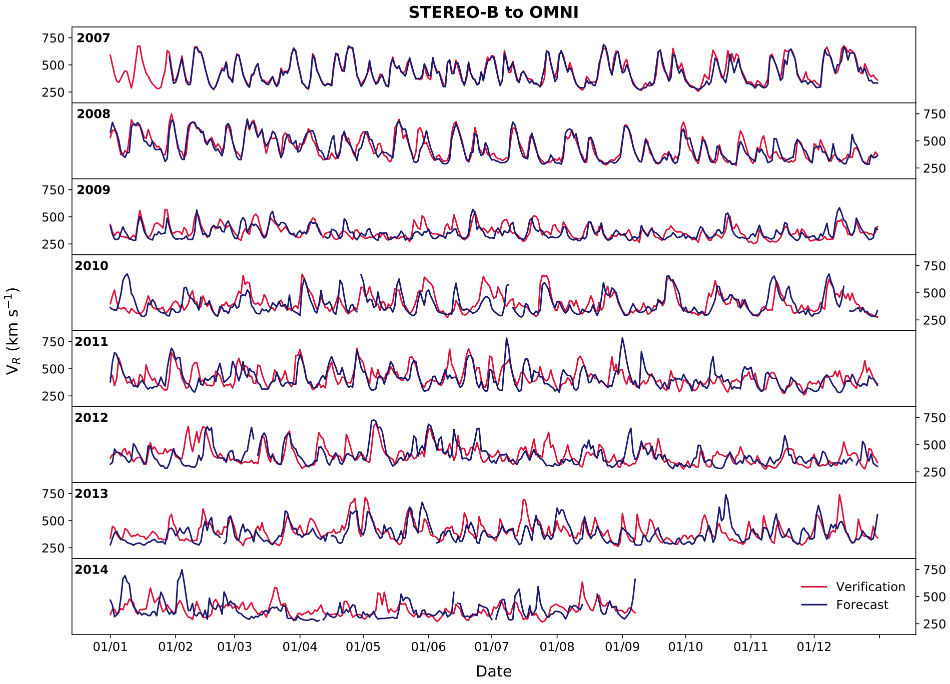


Figure 4.10: Time series of solar wind speed at OMNI's location. In red is the observed solar wind speed by OMNI. In blue is forecast solar wind speed, V_F, produced from time shifting the STEREO-B observations. Thus, OMNI is the target providing the observations to verify the forecast, and STEREO-B is the source spacecraft used to produce V_F. Data have been averaged to 1-day resolution for clarity.

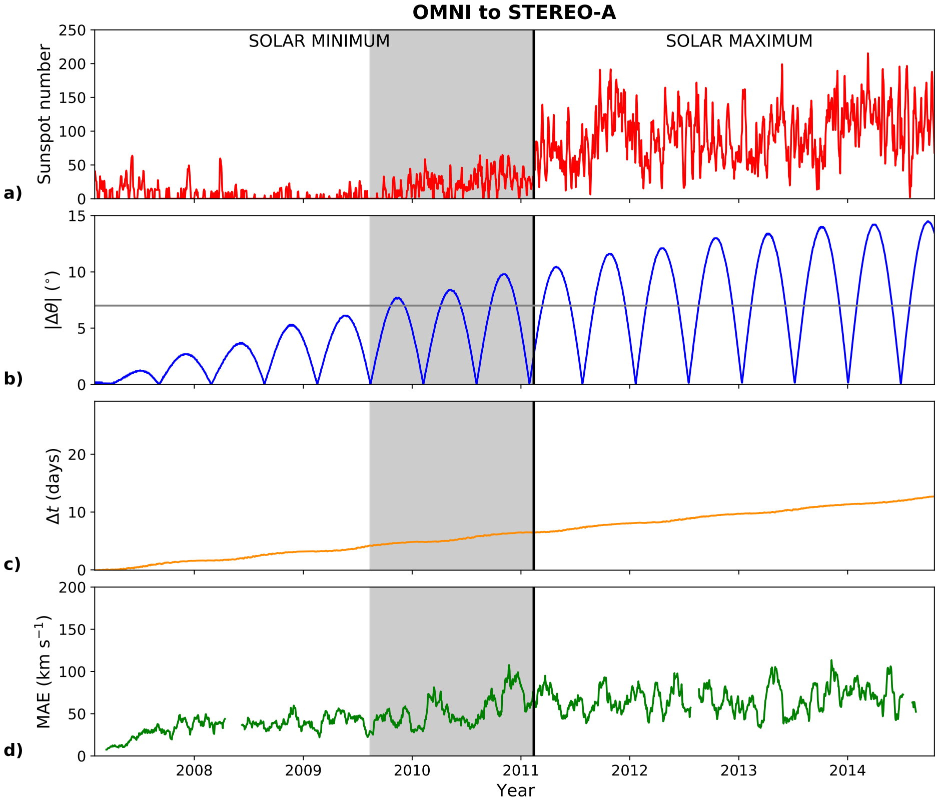


Figure 4.11: As for Figure 4.9, but using the OMNI to STEREO-A corotation.

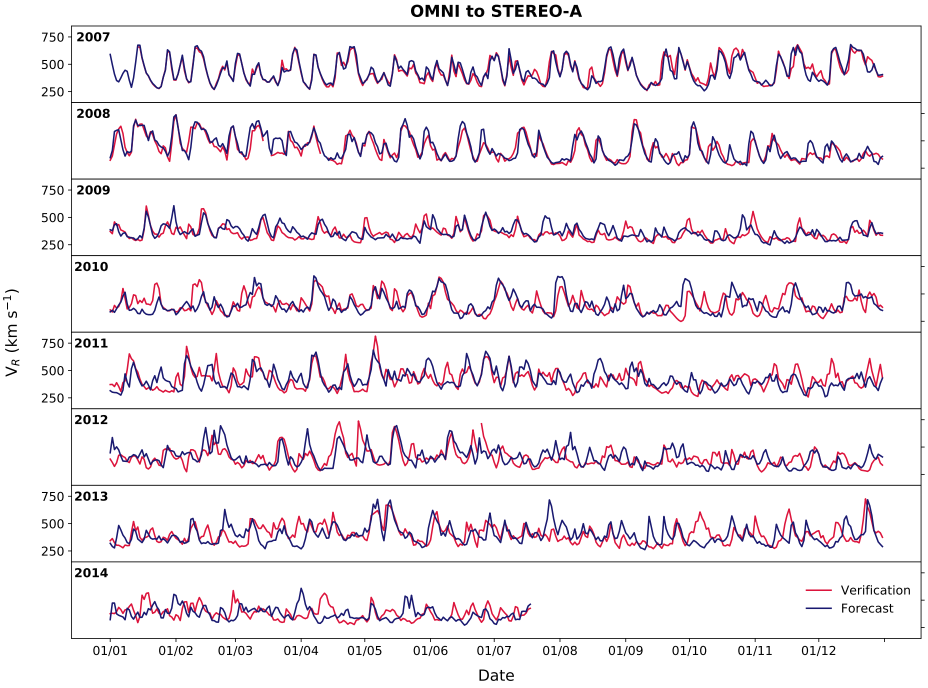


Figure 4.12: As for Figure 4.10, but using the OMNI to STEREO-A corotation.

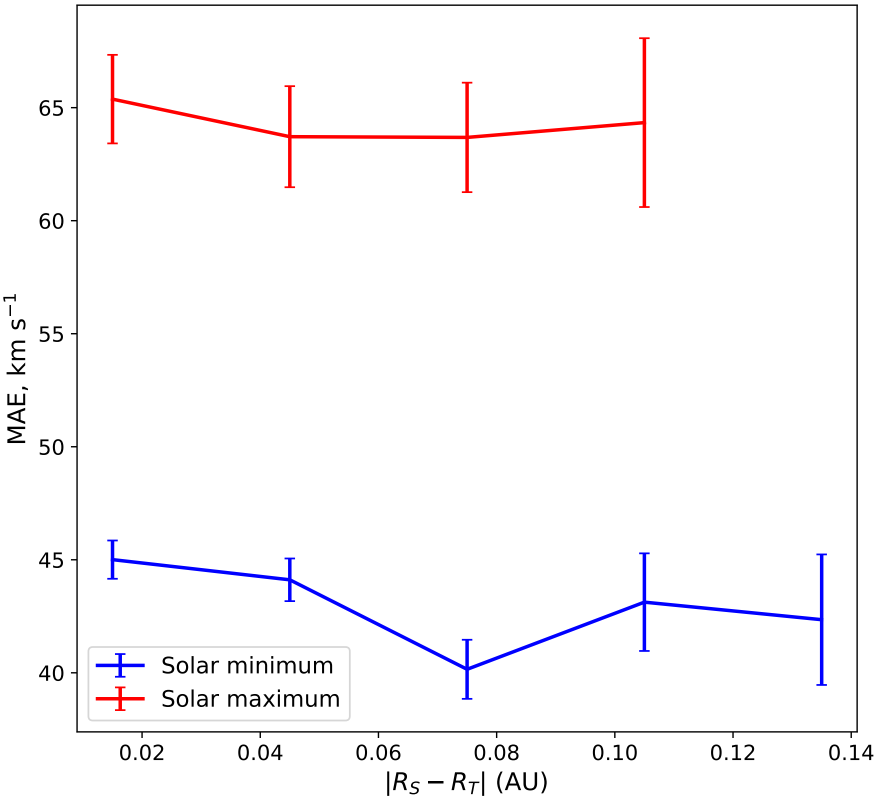


Figure 4.13: Forecast MAE as a function of radial separation between the source and target spacecraft, $|R_S - R_T|$, for solar minimum and maximum periods. Lines show the mean values, while error bars span one standard error on the mean. All STEREO/ OMNI spacecraft pairings are included. Only periods with $|\Delta\theta| < 2^\circ$ are included.

Chapter 5

Quantifying the effect of ICME removal and observation age for in situ solar wind data assimilation

This chapter has been published in the AGU journal *Space Weather* with the following reference:

Turner, H., Owens, M. J., Lang, M. S., Gonzi, S., & Riley, P. (2022). Quantifying the effect of ICME removal and observation age for in situ solar wind data assimilation. *Space Weather, 20,* e2022SW003109. `https://doi.org/10.1029/2022SW003109`

Key points

- Assimilating in situ data from multiple spacecraft provides higher forecast skill than from any one spacecraft individually.
- The age of observations, in terms of time when the required Carrington longitude was last observed, has a large effect on forecast skill.
- Removing ICMEs from the assimilated time series provides a small increase in forecast skill.

Abstract

Accurate space weather forecasting requires advanced knowledge of the solar wind conditions in near-Earth space. Data assimilation (DA) combines model output and observations to find an optimum estimation of reality and has led to large advances in terrestrial weather forecasting. It is now being applied to space weather forecasting. Here, we use solar wind DA with in-situ observations to reconstruct solar wind speed in the ecliptic plane between 30 solar radii and Earth's orbit. This is used to provide solar wind speed hindcasts. Here, we assimilate observations from the Solar Terrestrial Relations Observatory (STEREO) and the near-Earth dataset, OMNI. Analysis of two periods of time, one in solar minimum and one in solar maximum, reveals that assimilating observations from multiple spacecraft provides a more accurate forecast than using any one spacecraft individually. The age of the observations also has a significant impact on forecast error, whereby the mean absolute error (MAE) sharply increases by up to 23% when the forecast lead time first exceeds the corotation time associated with the longitudinal separation between the observing spacecraft and the forecast location. It was also found that removing coronal mass ejections from the DA input and verification time series reduces the forecast MAE by up to 10% as it removes false streams from the forecast time series. This work highlights the importance of an L5 space weather monitoring mission for near-Earth solar wind forecasting and suggests that an additional mission to L4 would further improve future solar wind DA forecasting capabilities.

Plain language summary

The effects of space weather can be damaging to technologies on Earth, potentially causing power outages and posing a hazard to humans in space. Accurate space weather forecasting requires advanced knowledge of the solar wind; a continual outflow of material from the Sun. Data assimilation (DA) is one method used in terrestrial weather forecasting, whereby model results are combined with observations to create an optimum estimation of reality. Here, we use a solar wind DA scheme to create 3 years of forecasts. It is found that assimilating observations from multiple spacecraft produces better forecasts than assimilating observations from a single spacecraft. It was also found that removing large eruptions, known as coronal mass ejections, from the DA input improves forecasts by reducing false alarms.

5.1 Introduction

The changing plasma conditions in the near-Earth space environment is a major component of space weather (Lilensten and Belehaki 2009). It poses a threat to modern life through damaging technology, causing power failures and posing a risk to the health of humans in space (Cannon 2013). For accurate space weather forecasting, advanced knowledge of the solar wind conditions is required. The solar wind is a continual stream of charged particles that flows from the high temperature corona (Parker 1958). The most severe space weather events occur as a result of coronal mass ejections (CMEs), large eruptions of coronal plasma and magnetic field (Webb and Howard 2012). CMEs have to propagate through the ambient solar wind, so it acts to modulate the severity of the CME and its impacts on Earth (Cargill 2004; Case et al. 2008). Stream interaction regions (SIRs) are an inherent feature of the ambient solar wind and are caused by fast streams catching up with slower streams and creating regions of higher plasma density and stronger magnetic field (Gosling and Pizzo 1999; Richardson and Cane 2012). SIRs which persist for more than one solar rotation, are also referred to as corotating interaction regions (CIRs) and provide a source of recurring space weather.

Solar wind forecasting can be achieved through simple empirical methods, such as corotation (Kohutova et al. 2016; Thomas et al. 2018; Turner et al. 2021) and persistence (Owens et al. 2013; Temmer et al. 2018), or through more complex, physics-based approaches such as magnetohydrodynamic (MHD) models (Merkin et al. 2016; Odstrcil 2003; Riley et al. 2001; Tóth et al. 2005). We here focus on improving the latter.

Data assimilation (DA) combines prior information about a system (typically, from a numerical model) with observations to form an optimal estimation of reality, known as the posterior. It has been used extensively in terrestrial weather forecasting, leading to large improvements. It has been attempted in three major areas of space weather forecasting; the photosphere (e.g. Hickmann et al. 2015), the ionosphere (e.g. Elvidge and Angling 2019) and the solar wind (e.g. Lang et al. 2017). Initial experiments in solar wind DA have shown potential for significant improvement in forecast skill (Lang et al. 2017). The BRaVDA (Burger Radius Variational Data Assimilation) methodology developed in Lang and Owens (2019) was subsequently used for producing hindcasts in Lang et al. (2021). BRaVDA uses a variational DA scheme (Le Dimet and Talagrand 1986; Lorenc 1986), with the simplified solar wind model, HUX (Riley and Lionello 2011). The output from BRaVDA was used to initialise a second reduced-physics solar wind propagation model, HUXt (Owens et al. 2020a), though it could equally be used with MHD models too. Lang et al. (2021) showed that whilst the 27-day forecast root mean square error (RMSE) was comparable to that of corotation forecasts, it showed improvement over non-DA forecasts. To further investigate the performance of the BRaVDA scheme and perform a more rigorous analysis, we have increased the hindcast cadence from 27-days to 1-day, as this is how forecasts would be generated if a DA scheme were deployed operationally.

The BRaVDA scheme makes use of in situ observations of near-Earth solar wind conditions from the OMNI dataset (Vokhmyanin et al. 2019), and distant observations from the STEREO (Solar Terrestrial Relations Observatory) mission, which was launched in 2007 (Kaiser et al. 2008). The OMNI dataset uses solar wind observations from a succession of spacecraft located at the L1 Lagrange point on the Sun-Earth line, at approximately 0.99AU. This is mostly comprised of observations from the *Wind* (Lepping et al. 1995; Ogilvie et al. 1995) and ACE (Advanced Composition Explorer; Stone et al. 1998) spacecraft propagated to the bow shock of Earth. The STEREO mission comprised of two spacecraft, STEREO-A and STEREO-B, which were placed into orbit around the Sun at approximately 1AU with STEREO-A ahead of Earth and STEREO-B behind. The spacecraft separate from Earth at approximately 22° per year and they passed behind the Sun in 2014. It was during this time that communication was lost with STEREO-B, and so the data used in this study is limited to STEREO-B's operational lifetime between 2007 and 2014.

The Lagrange points are gravitational nulls whereby the gravity of two large bodies balances the centripetal force of a smaller body. This means that spacecraft located at these positions will remain there, thus reducing the fuel required. There are five Lagrange points, with L4 and L5 positioned 60 degrees ahead and behind Earth in its orbit, respectively. A spacecraft located at either point would provide a near side-on view of the Sun-Earth line and so could provide remote-sensing observations of Earth-directed CMEs. Extensive studies have also shown the potential usefulness of an in situ space weather monitor at L5 [e.g. Akioka et al. (2005), Simunac et al. (2009), and Temmer et al. (2018)] and a mission is set for launch in 2027 (Davies 2020). If this is joined by a space weather monitor at L4 Posner et al. 2021, then these missions will provide additional observations that are useful for solar wind DA, as will be demonstrated in this study.

In this study, two analysis periods are used to assess the accuracy of hindcasts generated using the BRaVDA scheme. The methods used in this study are described in Section 5.2, with BRaVDA methodology described in 5.2.1 and the forecast generation method in 5.2.2. The data assimilation experiments and their results are described in Section 5.3. Finally, we discuss implications and draw conclusions in Section 5.4.

5.2 Methods

5.2.1 BRaVDA scheme

The BRaVDA methodology was developed and extensively described in Lang and Owens (2019). The code is available at: `https://github.com/University-of-Reading-Space-Science/BRaVDA`. Here, we provide only a short overview of the methodology. BRaVDA is a variational DA scheme that incorporates in situ spacecraft observations of solar wind speed into the steady-state "HUX" solar wind model, based on Riley and Lionello (2011). Using the adjoint model of HUX, BRaVDA maps information contained within the in situ observations at 1 AU ($\sim$ 215 solar radii (R_S)) radially inwards to HUX's inner boundary at $30R_S$. This information is then merged with a prior inner-boundary condition through the minimisation of a cost function comprised of the prior and observation errors weighted by their relative uncertainties. By finding the inner boundary condition that minimises this cost function, we find the solar wind speeds with the lowest errors respective of their relative uncertainties. This produces an updated inner boundary (the posterior state) which can be propagated radially outward by any solar wind model. For efficiency, we again use the HUX model, producing an optimal estimate of the true solar wind in the whole model domain, given the observations. The solar wind propagation model used in BRaVDA maps a 2 dimensional solar wind over the heliocentric domain from $30R_S$ to $236R_S$. The outer boundary here is an arbitrary choice, but it ensures that Earth's orbital radius is well within the boundary. We note that a different solar wind propagation model could be used to map the updated solar wind conditions out from $30R_S$ to Earth. For example: Enlil, (Odstrcil 2003); HelioMAS, (Riley et al. 2001); EUHFORIA, (Pomoell and Poedts 2018) or HUXt, (Owens 2020a) could be used. As the solar wind is a highly driven system, we would expect similar improvements to the forecast skill from other solar wind models.

The BRaVDA scheme requires that we define our prior state (our current estimate of the inner boundary condition), the prior error covariance matrix (a measure of the uncertainty present in our prior information) and the observation error covariance matrix (that gives a measure of the uncertainty in our observations relative to the HUX model).

We generate our 'prior' estimate of the solar wind speed at the inner boundary by using archived output of the HelioMAS model (data available from `https://www.predsci.com/portal/home.php`) at $30R_S$. HelioMAS is an MHD model that is initiated using radial magnetic field and solar wind speed derived from the coronal magnetic field topology (Riley et al. 2015) of the MAS (Magnetohydrodynamics

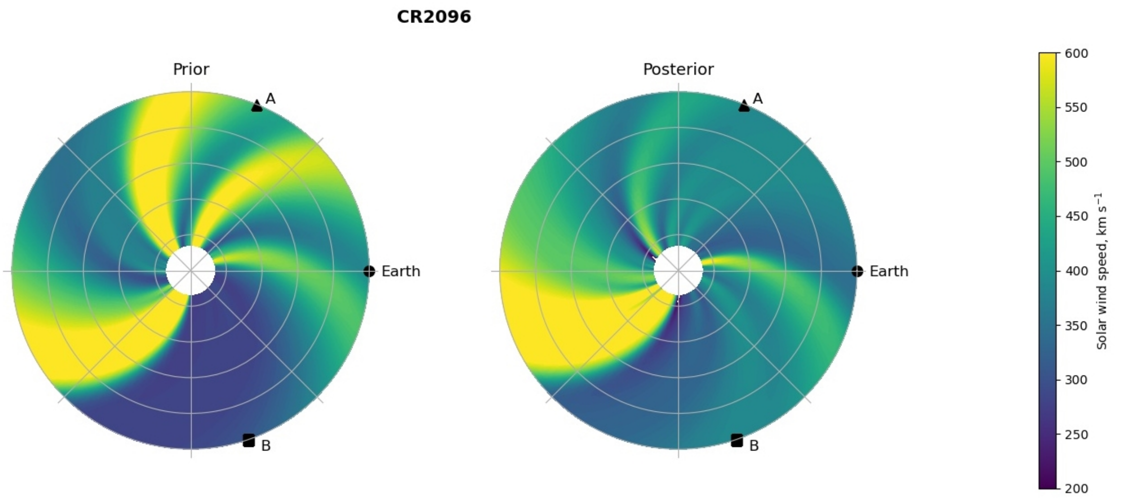


Figure 5.1: Solar wind solution from the HUX model initialised on 22/04/2010 for Carrington Rotation 2096 (22/04/2010 to 19/05/2010). The prior state (left) is that before the in situ data assimilation has taken place and the posterior state (right) is after the data assimilation. Indicated on both panels is the location of STEREO-A (A), Earth and STEREO-B (B) on 22/04/2010.

Around a Sphere) model (Linker et al. 1999) solutions to the observed photospheric magnetic field. This prior state is then propagated out radially to $236R_S$ with the HUX model to generate a prior estimate of the solar wind speed at Earth.

The prior error covariance matrix is estimated from an ensemble of HUX initial conditions (see Lang et al. (2017) for more details) generated by perturbing the HelioMAS $30R_S$ solution in the same manner as Owens and Riley (2017). The observation error covariance contains not only the measurement error, but also representivity errors that arise from the incorrect specification of observations in numerical models (such as errors from assuming the observations are on the model gridpoints, sub-grid processes etc.). An example of such a representivity error in the BRaVDA scheme is the fact that the HUX model is 2-dimensional, meaning that observations are always assumed to be at the heliographic latitude of Earth, whereas in reality observations away from Earth (such as provided by STEREO) may be at other heliographic latitudes. This representivity error is a large unknown at present and an area of ongoing research (Lang et al. 2021; Owens et al. 2020b; Turner et al. 2021). In this study, we use the same observation error covariance matrix as in Lang et al. (2021) to maintain consistency with previous work.

In this study, BRaVDA is run for two time periods; 01/08/2009 to 01/02/2011 and 01/04/2012 to 01/10/2013. These periods are highlighted in Figure 5.3. The earlier period covers the 18-months up to the separation between solar minimum and the rise to solar maximum, as described in Turner et al. (2021), and the later period is around solar maximum. This allows for analysis of solar wind forecasts in

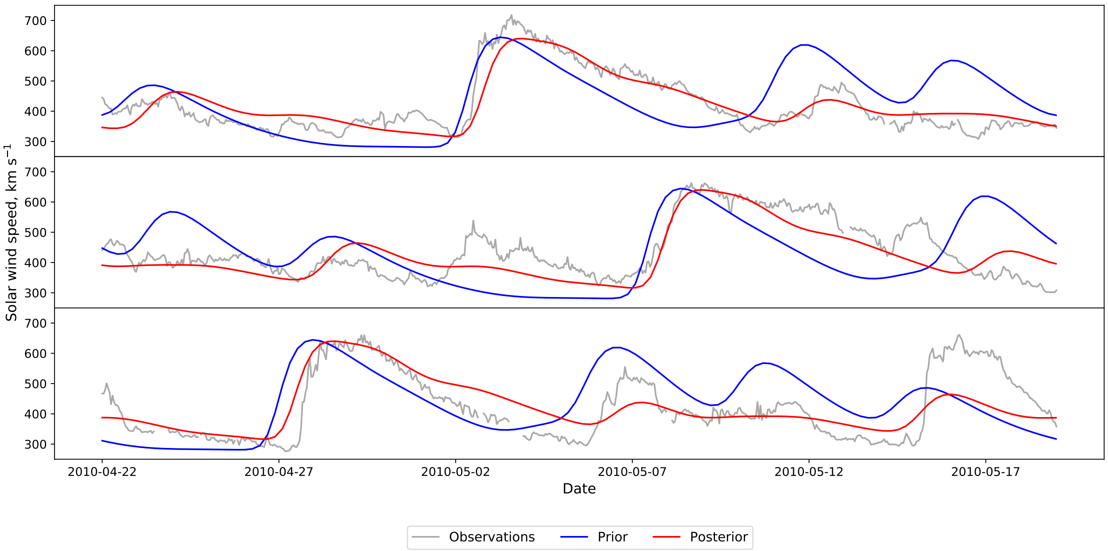


Figure 5.2: Time series at Earth's orbital distance of the solar wind solution for Carrington Rotation 2096 (22/04/2010 to 19/05/2010), as depicted in Figure 5.1. The top panel shows the solution at Earth, the middle panel at STEREO-A and the bottom panel at STEREO-B. The prior state is shown with the blue line, the posterior state with the red line and the grey line shows the observations taken from the respective spacecraft.

both phases of the solar cycle.

BRaVDA is run at daily cadence, using in situ observations which would have been available at the time the forecast is performed, as in a genuine forecast. This expands on the work in Lang et al. (2021), where BRaVDA was run every 27 days. The prior state from HelioMAS, however, is only available as Carrington rotation solutions (i.e. every 27 days). In a true forecasting situation, the prior state would ideally be obtained from daily-updated coronal solutions. However, the Carrington rotation solutions are adequate for our purposes here, as the DA process makes significant changes to the prior state. The likely effect of this is that the accuracy of the prior state is overestimated and the forecast improvement from DA reduced from the value expected in an operational situation.

The output from each BRaVDA run gives a 27-day solar wind reconstruction from $30R_S$ to $215R_S$, as shown in Figure 5.1. By taking the output from $215R_S$, this can be used as a solar wind time series for Earth and the STEREO spacecraft, as shown in Figure 5.2.

5.2.2 Forecast generation

For each (daily) BRaVDA run, observations are assimilated from the previous 27 days up until the time the forecast is made, t_0. By assuming steady-state corotation of the posterior solution, each BRaVDA run can be used to generate a single forecast

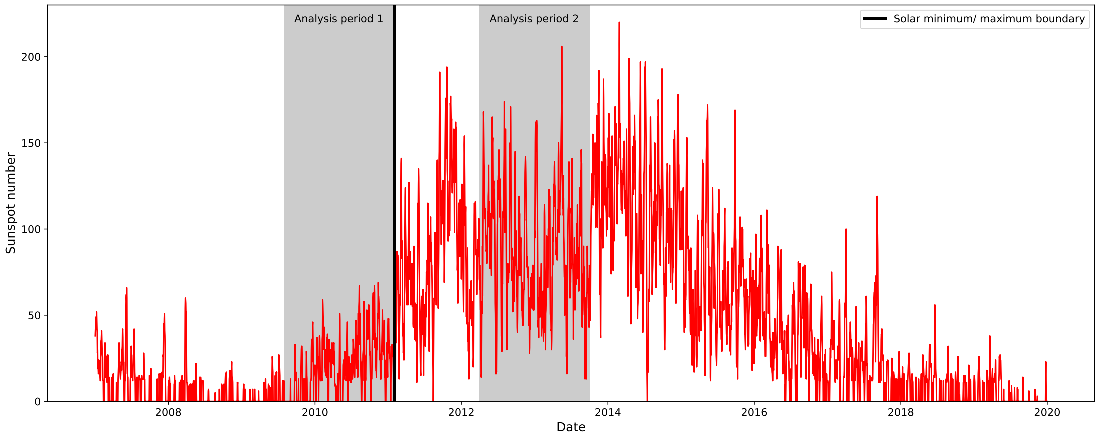


Figure 5.3: Evolution of sunspot number from 2007 to 2020. The analysis periods are highlighted in the grey shaded areas. The black vertical line shows the divide between solar minimum and solar maximum.

Date	Spacecraft corotation time [days]					
	Earth to STA	STA to Earth	STB to Earth	Earth to STB	STB to STA	STA to STB
01/08/2009	4.3	22.7	3.7	23.3	8.0	18.9
01/02/2011	6.5	20.5	6.9	20.1	13.4	13.6
01/04/2012	8.3	18.7	8.9	18.1	17.2	9.8
01/10/2013	11.0	16.0	10.5	16.5	21.5	5.5

Table 5.1: Corotation time for the different spacecraft pairings, taking into account only the spacecraft longitudinal separation. The dates are the starts and ends of the solar minimum and solar maximum intervals.

with a lead-time t_f of 0 to 27 days with respect to t_0. An example is shown in Figure 5.4. The single BRaVDA run produces a single solar wind speed estimate at each forecast lead time from 0 to 27 days. As BRaVDA is run on a daily cadence, a forecast time series for a particular lead time, e.g. $t_f = 5$ days, at a given location can be created by combining forecasts for different t_0. While forecasts can be generated for the whole model domain, from $30R_S$ out to $215R_S$ and for all longitudes, here we only consider the locations of STEREO-A, STEREO-B and OMNI.

For the periods of time used in the analysis here, the spacecraft separation, and therefore corotation time between observation and Earth, changes over the analysis period (see Table 5.1).

5.3 Data assimilation experiments and results

Throughout this study we consider the variation of forecast mean absolute error (MAE) with forecast lead time. Here we describe a number of individual BRaVDA experiments aimed at diagnosing specific aspects of forecast MAE.

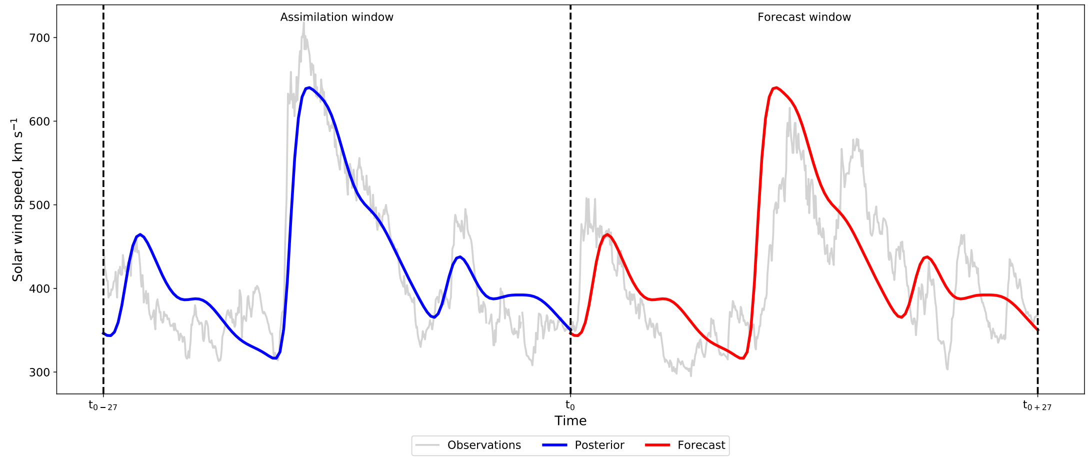


Figure 5.4: Using BRaVDA posterior output (blue) as a solar wind speed forecast (red). t_0 is the start of the forecast window, which is here 27 days long. It is also the end of the assimilation window, wherein observations from the previous 27 days up until t_0 are assimilated. The solar wind observations for this period are shown in grey.

5.3.1 Forecast lead time

We first look at the effect of forecast lead time on MAE. For this, we assimilate observations from all three spacecraft (OMNI, STEREO-A and STEREO-B) and the output from BRaVDA is used to create forecasts at the locations of Earth, STEREO-A and STEREO-B for lead times of 0 to 27 days. This is done for both solar minimum and solar maximum time intervals. Figure 5.5 shows the MAE between the forecast and observed solar wind speed for a range of forecast lead times. As forecast lead time increases, there is a general trend for increasing forecast MAE. The left-hand panel shows the solar minimum interval, where MAE is generally lower than the solar maximum interval, shown in the right-hand panel. This is true for all three forecast locations. Note, however, that spacecraft separation (in solar longitude and latitude) also increases between these two time periods, so the difference in MAE between the solar minimum and maximum intervals cannot necessarily be attributed to the solar cycle. During the solar maximum interval, the MAE for the forecasts at the STEREO spacecraft is higher than for the forecasts at Earth. During this interval, the latitudinal separation between Earth and the STEREO spacecraft reaches a maximum of 14 degrees, compared to 9 degrees for the solar minimum interval (see Figure 1 in Turner et al. 2021). The difference between the forecast skill could be attributed to this greater latitudinal separation, combined with the generally more variable solar wind during solar maximum.

Figure 5.5: Mean absolute error (MAE) of solar wind forecasts as a function of forecast lead time, for the case where all spacecraft observations are assimilated. The forecast at Earth location is shown in black, at STEREO-A in red and at STEREO-B in blue. The solar minimum interval (01/08/2009 to 01/02/2011) is in the left-hand panel and the solar maximum interval (01/04/2012 to 01/10/2013) in the right-hand panel.

5.3.2 Assimilation of individual spacecraft and age of observations

We now consider the effect of assimilating different combinations of spacecraft. Experiments were carried out assimilating observations from all spacecraft together, as above, and assimilating the spacecraft observations individually. Figure 5.6 shows the MAE variation with forecast lead time for these different experiments. Here, the forecasts at Earth location are in the top row, at STEREO-A in the second row and at STEREO-B in the bottom row. The solar minimum interval is in the left-hand column and solar maximum in the right-hand column.

The most obvious trend is that assimilating all spacecraft produces a forecast with the lower MAE than any one spacecraft individually. This is true at all locations and for all non-zero lead times. The exception is for very long lead-time forecasts ($>$ 20 days) at STEREO-B, where the MAE for assimilating all spacecraft is comparable to assimilating only STEREO-B data.

Overall, there is still a general trend of increasing MAE with lead time. When all spacecraft are assimilated, the MAE increase with lead time is fairly smooth, if not necessarily linear. When looking at the assimilation of individual spacecraft, however, there are clear step changes in MAE with forecast lead time. These can be understood in terms of the corotation time.

When only individual spacecraft are assimilated, the lowest MAE at short, non-

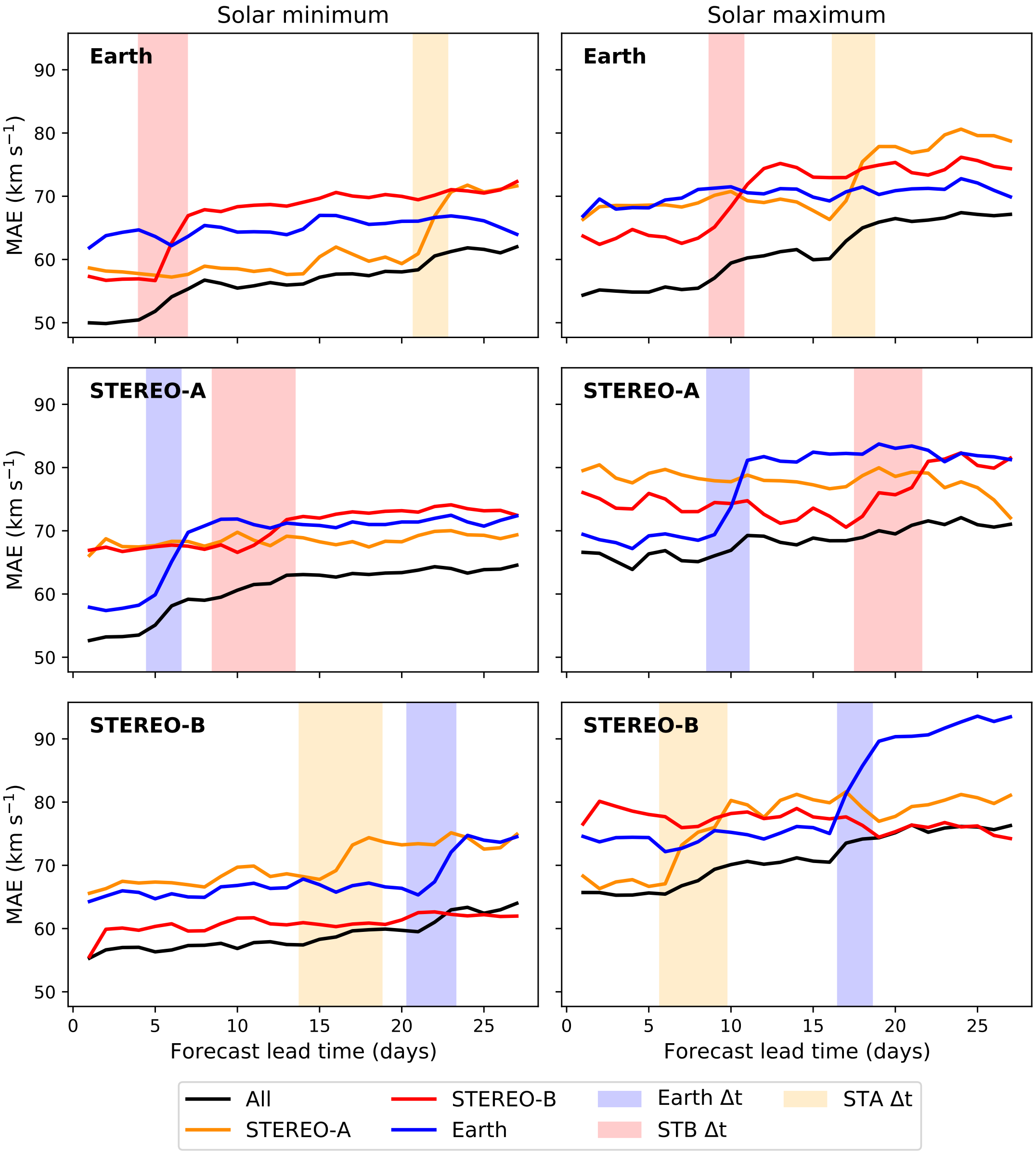


Figure 5.6: Mean absolute error of solar wind forecasts as a function of forecast lead time, for different combinations of assimilated spacecraft (coloured lines, as indicated by the legend) and different forecast locations (rows of panels). The left hand column is the solar minimum interval, from 01/08/2009 to 01/02/2011, and the right hand column is the solar maximum interval, 01/04/2012 to 01/10/2013. The shaded regions show the corotation time from the spacecraft indicated by colour (blue for Earth, yellow for STEREO-A and red for STEREO-B) to the forecast location.

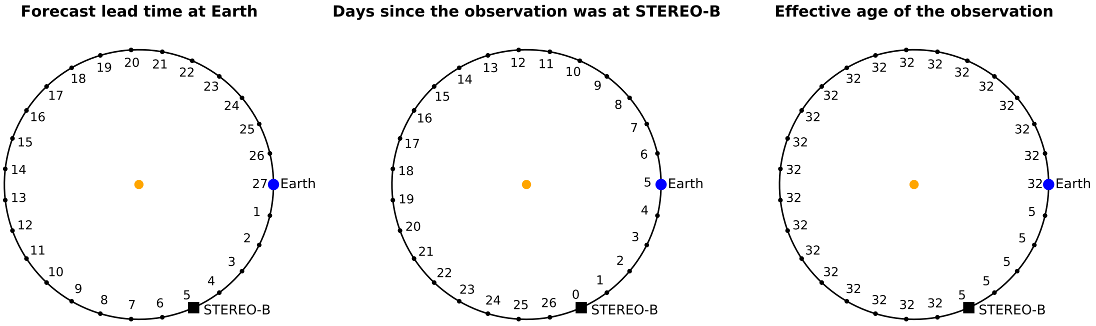


Figure 5.7: Schematics of the ecliptic plane to illustrate the effect of age of observations on solar wind data assimilation. The left-hand panel shows the time (in days) it takes the solar wind to corotate to Earth, which is the potential forecast lead time using observations from that location. The middle panel shows the number of days since solar wind at that position was last observed at STEREO-B. In this example, STEREO-B is approximately 65 degrees behind Earth in Carrington longitude, as occurred in December 2009. The right-hand panel shows the effective age of STEREO-B observations for Earth forecasting, which is the sum of the forecast lead time and the number of days since the observation was observed.

zero forecast lead times (e.g. < 5 days) is obtained when assimilating the spacecraft 'ahead' (in terms of solar rotation, meaning at lower Carrington longitude) of the forecast location. This can be seen in Figure 5.6; for example, the forecast MAE at STEREO-A's location (middle row) when assimilating only near-Earth observations (blue line) is initially below that obtained when assimilating only STEREO-A (orange line) or STEREO-B (red line). This remains the case for forecast lead times out to 5 days during the solar minimum period, and 10 days for the solar maximum period. Between these two time periods, STEREO-A separates from Earth, increasing from 57 to 87 degrees ahead. Thus observations at Earth provide recent information at STEREO-A for up to 5 and 10 days. The other panels show similar transitions are associated with the forecast lead time exceeding the corotation time. This effect can be seen schematically in Figure 5.7.

Assimilated spacecraft	**Forecast location**	**MAE [km/s]**					
		2009-2011			**2012-2013**		
		Before	**After**	**% diff**	**Before**	**After**	**% diff**
OMNI	Earth	-	-	-	-	-	-
	STEREO-A	57.8	71.2	23.2	68.7	82.0	19.4
	STEREO-B	66.1	74.5	12.7	74.5	91.7	23.1
STEREO-A	Earth	58.8	71.2	21.1	68.8	78.7	14.4
	STEREO-A	-	-	-	-	-	-
	STEREO-B	67.6	73.7	9.0	67.3	79.8	18.6
STEREO-B	Earth	57.0	69.6	22.1	63.4	74.2	17.0
	STEREO-A	67.3	73.0	8.5	73.6	81.1	10.2
	STEREO-B	-	-	-	-	-	-

Table 5.2: Forecast mean absolute error (MAE) for different time intervals, different locations and different assimilated spacecraft. Before and After indicate forecasts where the lead time is less than or greater than the minimum corotation time between assimilated spacecraft and forecast location, respectively. Where it is left blank, this is because the lead time never exceeds the corotation time.

Table 5.2 shows the average MAE separated into intervals before and after when the lead time is less than or greater than the minimum corotation time between assimilated spacecraft and forecast location. For all assimilated spacecraft and forecast locations, there is an increase in MAE of between 8.5 and 23.2%, with an average value of 16.6%. Of course, increased lead time is expected to result in increased MAE, regardless of corotation time. But Figure 5.6 shows that this trend is much smaller, typically of the order a few percent.

The large jumps in MAE are due to the age of the observations, whereby once the forecast lead time (t_f) exceeds the corotation time (t_C), observations from the previous Carrington rotation are being used. This means that the effective age of observations (τ) for $t_f > t_C$ is $t_f + 27$ days. For example, when assimilating only STEREO-B and forecasting at Earth, as Table 5.1 shows, the corotation time increases from 3.7 to 6.9 days over the 01/08/2009 to 01/02/2011 period. Therefore, as lead times exceed $\sim$ 5 days, we expect an increase in MAE as τ jumps from $\tau < 6.9$ to $\tau > 30.7$. This can be seen in Figure 5.6, where the red shaded region in the top left plot shows the corotation time for STEREO-B to Earth for that period. This effect can be seen in a number of other situations where single spacecraft are assimilated. When assimilating multiple spacecraft, the abrupt change in age of observation effect is lessened; although there is an increase in the forecast error over the shaded regions of Figure 5.6, the curve is not as steep as when assimilating only the spacecraft associated with that corotation time. Therefore, assimilating multiple spacecraft reduces the effect of the age of observations impacting the forecast MAE.

Figure 5.8 shows experiments where pairs of spacecraft observations are assimilated together (coloured lines) for the different forecast locations compared to when all three spacecraft observations are assimilated (black line). In general, it is preferable to assimilate all three observation sources than any two, though the difference is not as marked as that between three spacecraft and one, as shown in Figure 5.6. This is useful for future mission planning, for example a pairing between the *Vigil* mission (ESA 2022) and a near-Earth monitor. The jump in MAE due to the age of observations seen for single spacecraft in Figure 5.6 is reduced when including another observation source, but not completely removed. For example, the rise in MAE seen when assimilating only STEREO-B for a forecast at Earth (red line) in the top left panel in Figure 5.6 is still largely present when assimilating STEREO-B together with observations for near-Earth (red line) in the top left in Figure 5.8. Comparing that to when all three sources are assimilated (black line in top left in both Figures 5.6 and 5.8), that jump is less severe.

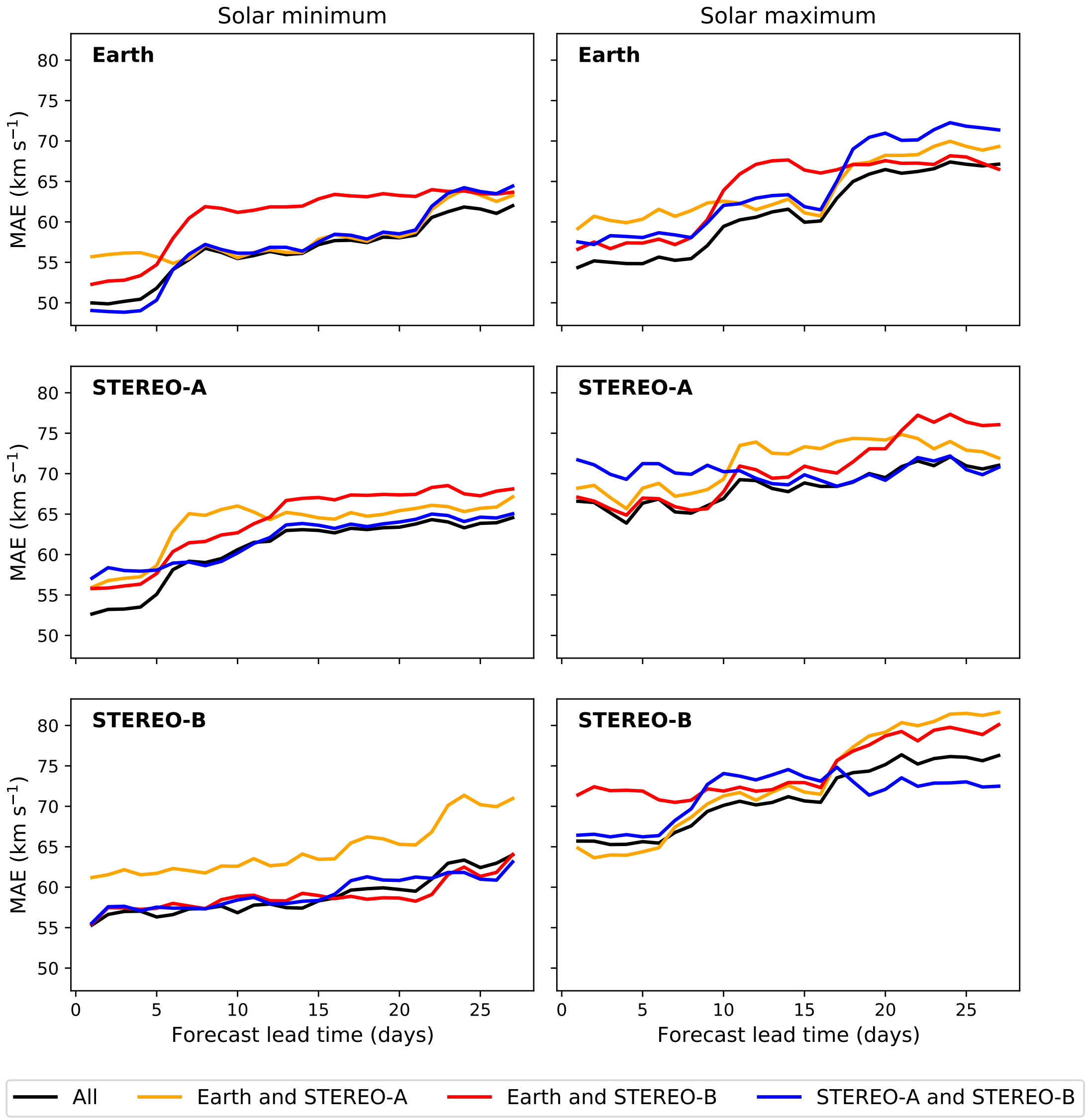


Figure 5.8: Mean absolute error of solar wind forecasts as a function of forecast lead time, for different pairs of assimilated spacecraft observations (coloured lines, as indicated by the legend) and different forecast locations (rows of panels). The left hand column is the solar minimum interval, from 01/08/2009 to 01/02/2011, and the right hand column is the solar maximum interval, 01/04/2012 to 01/10/2013.

5.3.3 Removal of ICMEs

The interplanetary manifestations of CMEs (ICMEs) provide a potential source of error for forecasts using the DA output. For example, if a fast ICME encounters one of the assimilated spacecraft during the assimilation window, the ICME will be reconstructed in the output as if it were time-stationary fast solar wind stream. When this output is subsequently used for forecasting, it would produce a false fast stream in the forecast time series. Therefore, by removing ICMEs from the input time series of all spacecraft, the production of the false streams in the forecast time series can potentially be prevented.

Conversely, if a fast ICME encounters the forecast location during the forecast window, the forecast will miss the transient fast stream, as there will have been no corresponding fast stream in the assimilation window. Thus removing ICMEs from the verification time series allows us to better assess how well BRaVDA is reproducing the ambient solar wind structure, without penalising for missing transient ICME structures, which it is not expected to capture. (However, note that accurate reconstruction of the ambient solar wind is required to produce accurate ICME arrival time forecasts (e.g. Case et al. 2008). Thus we present results of removing ICMEs from the DA-input time series, the verification time series, and both. Times associated with ICMEs are identified using the HELCATS ICME list (Mostl et al. 2022). To ensure all of the disturbance associated with the ICME is removed, we eliminate 24 hours either side of the ICME leading disturbance and ICME trailing edge times. For the purposes of assimilated data, the data gap produced by removing an ICME is then linearly interpolated over to make a complete time series. Qualitatively similar results are obtained by simply removing times identified as ICMEs.

An example of the effect of removing ICMEs is shown in Figure 5.9. A fast ICME (maximum of 720 km s^{-1}) was identified at STEREO-B, seen as an isolated velocity spike between 22/09/2012 to 29/09/2012. This ICME was removed from the STEREO-B input and linearly interpolated to give the green line in the top panel of Figure 5.9. The bottom three panels show 5-day forecasts at Earth, STEREO-A and STEREO-B with (solid lines) and without (dashed lines) the ICME present in the assimilated STEREO-B time series. Red lines show assimilation of STEREO-B data only, while blues lines show assimilation of all three spacecraft data.

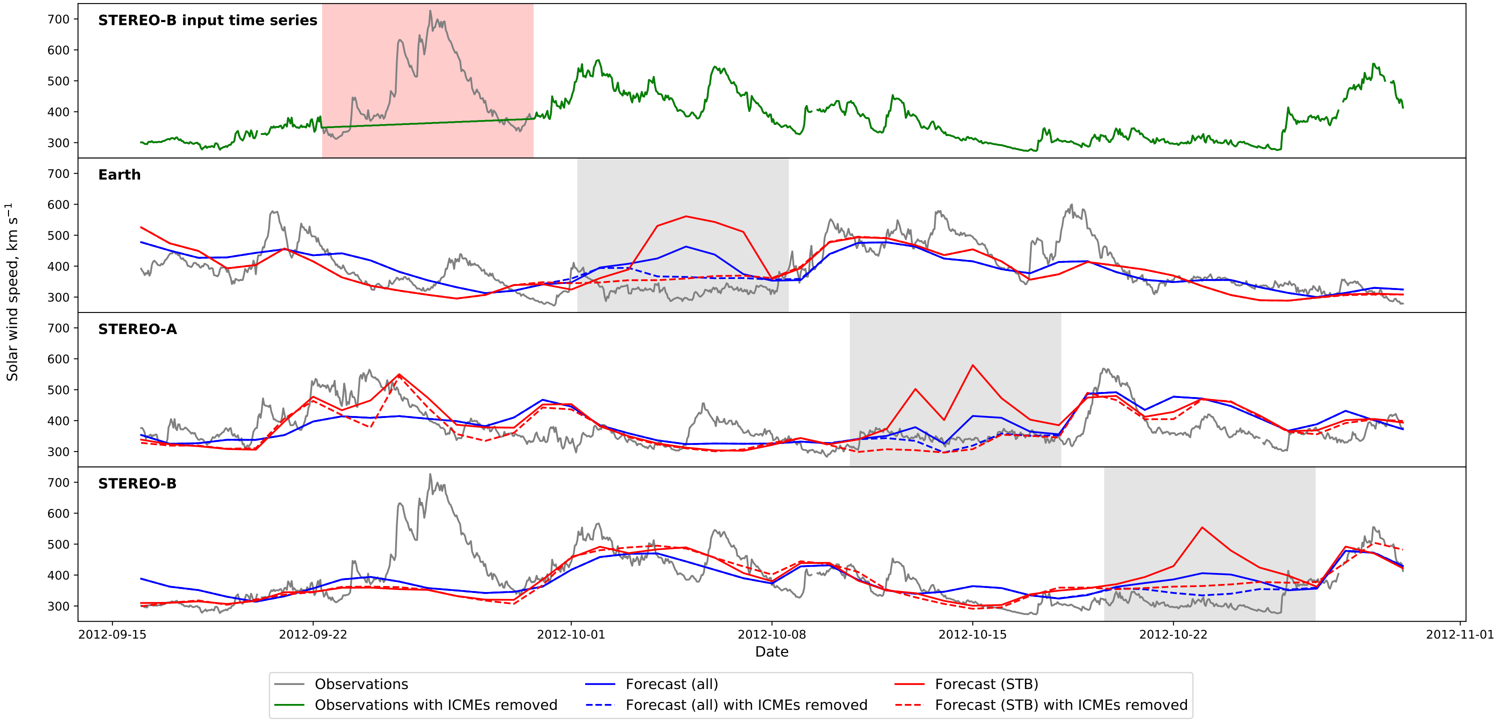


Figure 5.9: Time series of; top panel, the STEREO-B observations with the ICME included (grey) and with it removed (green); second panel, forecast with 5-day lead time for Earth; third panel, forecast with 5-day lead time for STEREO-A; bottom panel, forecast with 5-day lead time for STEREO-B. The blue lines show when all spacecraft observations are assimilated and the red lines show only STEREO-B observations assimilated. The solid line shows when the ICME is included in the STEREO-B input series and the dashed line when the ICME is removed. The red shaded region in the top panel shows the time span of the ICME plus 24 hours either side. The grey shaded regions in panels 2, 3 and 4 show the time span where the forecast time series has been influenced by the removal of the ICME.

ICMEs in input?	ICMEs in verification?	Solar minimum			Solar maximum		
		Earth	STA	STB	Earth	STA	STB
Yes	Yes	56.1	60.3	58.7	60.7	68.1	70.7
	No	55.2	59.8	57.8	60.8	63.1	65.6
No	Yes	55.9	60.4	58.4	60.0	66.9	68.4
	No	55.1	59.6	57.5	59.5	61.9	63.1

Table 5.3: Average forecast MAE (over all lead times) for combinations of removing ICMEs from the DA input time series and the verification time series. The average MAE is shown for both the solar minimum (01/08/2009 to 01/02/2011) and solar maximum (01/04/2012 to 01/10/2013) intervals. Top row: ICMEs are included in both the DA input and verification time series. Second row: ICMEs are included in the DA input time series and removed from the verification time series. Third row: ICMEs are removed from the DA input time series and remain in the verification time series. Bottom row: ICMEs are removed from both the DA input and verification time series.

It can be seen that removing the ICME causes an improvement in the forecast, by removing a false fast stream, a certain time later. The time delay from ICME removal to forecast improvement relates to the corotation time between the spacecraft, which is determined by their longitudinal separation and can be seen in the grey shaded regions in Figure 5.9. At this time, the corotation time from STEREO-B to Earth is approximately 9 days, STEREO-B to STEREO-A is approximately 18 days and the full Carrington rotation to STEREO-B is 27 days. It is after these respective times from the ICME that the improvement is seen in the forecast for that spacecraft. Furthermore, Figure 5.9 shows even without removal of ICMEs from the assimilation data (which may be difficult in real-time operational forecasting), assimilating multiple spacecraft observations reduces the magnitude of the false fast stream that is present in the forecasts following. Comparing the red and blue solid lines, when all spacecraft observations are assimilated, the false stream magnitude is reduced by approximately 150 km s^{-1} in the forecasts.

Figure 5.10 shows the variation of forecast MAE with forecast lead times for combinations of ICME removal from the DA input time series and the verification time series. There are four combinations; ICMEs remaining in both time series, ICMEs removed from the input time series but remaining in the verification time series, ICMEs remaining in the input time series but removed from the verification time series and ICMEs removed from both time series. It can be seen from Figure 5.10 and from the average MAE values in Table 5.3 that, generally, removing ICMEs leads to forecast improvement. For example, for the forecast at Earth, the average MAE across all lead times is 56.1 and 60.7 km s^{-1} for the solar minimum and solar maximum intervals respectively. By removing ICMEs from both the input and verification time series, these are reduced to 55.1 km s^{-1} for solar minimum

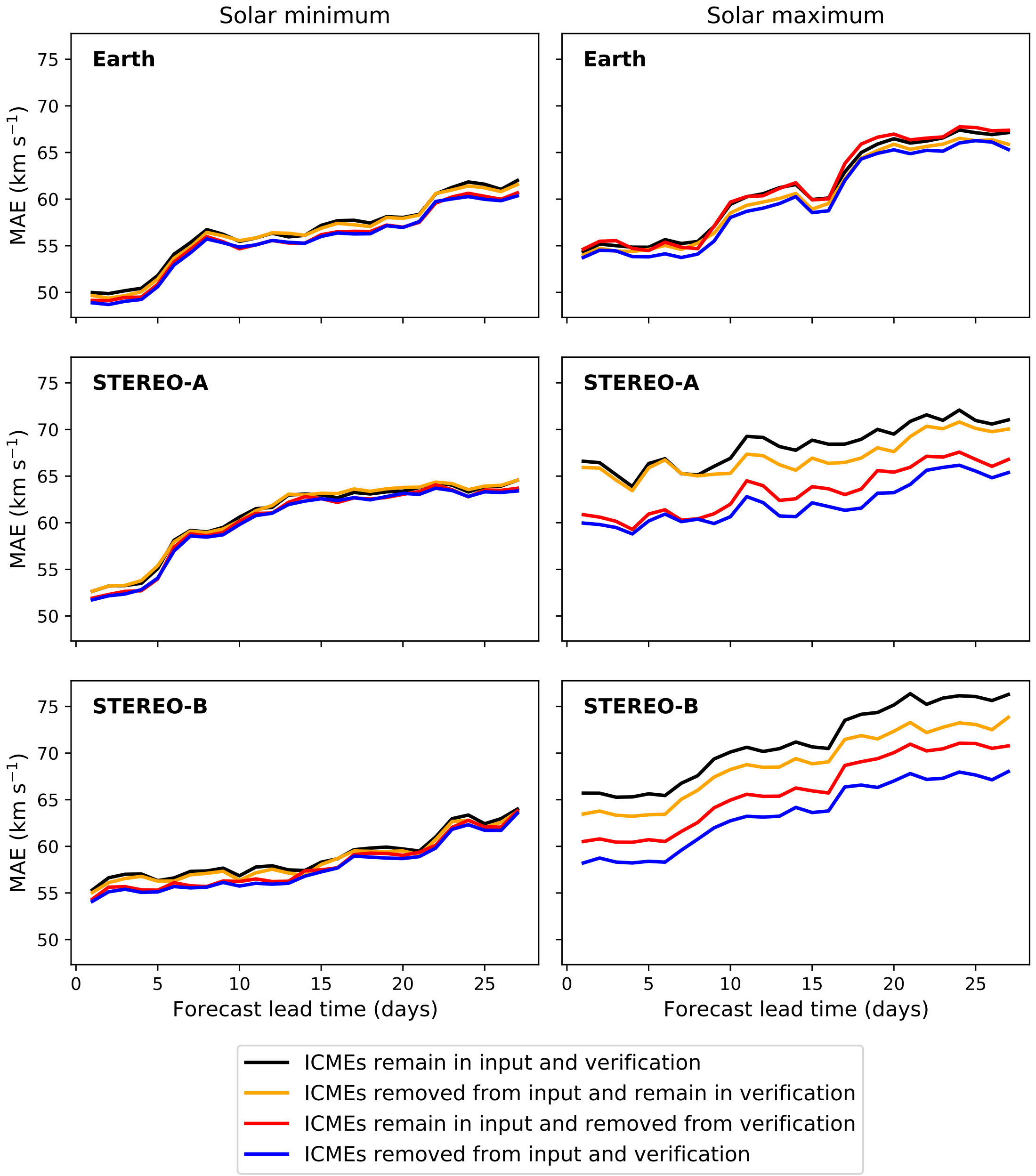


Figure 5.10: Variation of forecast MAE with forecast lead time for different combinations of ICME removal from the input and verification time series. The black line shows the case when the ICMEs remain in both DA input and verification time series. The red line shows the case when ICMEs are removed from the verification time series only. The orange line shows the case when ICMEs are removed from the DA input series only. The blue line shows the case when ICMEs are removed from both the DA input and verification time series. All spacecraft are assimilated in all cases.

ICMEs in input?	ICMEs in verification?	Solar minimum			Solar maximum		
		Earth	STA	STB	Earth	STA	STB
Yes	No	-1.6	-0.8	-1.5	0.2	-7.3	-7.2
No	Yes	-0.4	0.2	-0.5	-1.2	-1.8	-3.3
	No	-1.8	-1.2	-2.0	-2.0	-9.1	-10.7

Table 5.4: Percentage difference of the average MAE for forecasts (over all lead times) with combinations of ICMEs removed from the DA input time series and verification time series compared with the forecasts where ICMEs remain. Where the difference is negative, this indicates the average MAE without ICMEs is smaller than with ICMEs included. Top row: ICMEs remain in the DA input time series and are removed from the verification time series. Middle row: ICMEs are removed from the DA time series and remain in the verification time series. Bottom row: ICMEs are removed from both the DA input and verification time series.

and 59.5 km s^{-1} for solar minimum, which is as percentage difference of -1.8% and -2.0% respectively, as shown in Table 5.4. There is greater improvement in the solar maximum period, particularly at STEREO-A and STEREO-B. This is due to a larger number of ICMEs at this time, with a total number of 72 ICMEs observed during the solar minimum period compared to 138 during the solar maximum period. We further classify fast ICMEs as those with an average proton speed of more than 500 km s^{-1}, as taken from the HELCATS ICME catalogue. 20 fast ICMEs were observed during the solar maximum period and 6 during the solar minimum period. Out of the 20 fast ICMEs during the solar maximum period, 5 of these were at Earth, 7 at STEREO-A and 8 at STEREO-B. Although only a small difference between the spacecraft, this could account for the -2.0% difference at Earth compared with the -9.1% at STEREO-A and -10.7% at STEREO-B. On average, we would not expect such a large difference between the forecasts at the STEREO spacecraft and at Earth. However, here we have a small sample size and so the properties of the individual ICMEs have a greater impact.

5.4 Discussion and conclusions

In this study we have performed a number of solar wind data assimilation experiments to determine how forecast error is expected to vary with a number of different factors. Here, mean absolute error (MAE) was used as the metric to analyse the forecast accuracy. Although MAE is a single metric and it can sometimes mislead (Owens 2018), it is useful here as the individual changes to the DA process that we are testing produce time series that are, for the most part, qualitatively similar. Future work will focus on a more complete analysis of forecast performance.

We have shown that assimilating observations from multiple spacecraft produces a forecast with lower MAE than when assimilating observations from any one single spacecraft. This is despite the fact that observations may not be ideally placed, due to the inclination of the ecliptic plane to the solar equator (Owens et al. 2019, 2020b). Assimilating two sources of well separated observations is preferable to one source, and three sources is preferable to two. However, the difference between one and two sources of observations is greater than two and three. Assimilation of the STEREO spacecraft observations, along with OMNI, has shown to improve forecasts and so the addition of space weather monitors at both L4 and L5 can only aid future development of solar wind DA. This stresses the value of multiple, well separated, space weather monitoring missions, such as at the L5 point (Lagrange mission, Davies 2020), but also the importance of a mission to the L4 point (see Posner et al. 2021).

The "age" of the observations, in terms of the time since a Carrington longitude was last observed by a particular spacecraft, also has a large effect on forecast error. When assimilating data from a single spacecraft, there is a large increase in MAE once the forecast lead time is greater than the corotation time from that spacecraft to the forecast location. This is due to the assumption of steady state conditions becoming increasingly less valid. Although assimilating multiple spacecraft does not completely remove this effect, it is greatly reduced and the discontinuous increases in MAE with forecast lead time are reduced. But further forecast improvement may be obtained by weighting observations by their age. Simple experiments testing this idea with BRaVDA (not shown) have been inconclusive. However, as BRaVDA is based upon HUX (Riley and Lionello 2011), which is not time dependent, it is not easy to explicitly implement this. Future experiments using the time-dependent version of HUX (HUXt, Owens et al. 2020a) are planned.

ICMEs have the potential to introduce false streams into the BRaVDA output, leading to false alarms in the forecasts. As ICMEs are transient events rather than corotating solar wind streams, they are not correctly captured in BRaVDA. If a fast ICME encounters one of the assimilating spacecraft, it will be treated as a fast ambient solar wind stream and assumed to persist at the observed Carrington longitude. Thus a fast stream is subsequently incorrectly forecast. This was demonstrated when considering a single fast ICME encountered by STEREO-B in late September 2012. This ICME was removed from the input observations at STEREO-B and the resulting BRaVDA output was used for forecasting at Earth, STEREO-A and STEREO-B. There was a marked improvement in the 5-day lead time forecast time series at Earth, STEREO-A and STEREO-B, through reduction of a region of forecast high solar wind. This was seen for the cases where only STEREO-B observations were assimilated and when all three spacecraft observations were assimilated. When only the STEREO-B observations were assimilated, the false fast stream appeared larger in the forecast time series than when all spacecraft observations were assimilated, as the ICME was only observed at one spacecraft. Although assimilation of multiple observations caused this false alarm to be of a lower magnitude than the fast ICME in the observations, it did not completely remove it.

More generally applying the ICME removal to the solar minimum and solar maximum time periods showed a general improvement in the forecast accuracy. However, improvements were modest, as all three spacecraft were assimilated, which already reduces the effect of transient ICMEs.

The largest improvement in the solar maximum period at STEREO-A and STEREO-B, due to the larger number of fast ICMEs observed at these spacecraft. It was found that removing ICMEs from the verification time series caused a larger improvement in the forecast MAE than removing them from the DA input time series. This is due to BRaVDA missing fast streams having a larger effect than creating fast streams through misinterpreting an ICME. As an ICME would be observed at only one spacecraft, assimilating multiple observations limits the reconstruction of the ICME as a steady-state solar wind structure, thus reducing the production of a false alarm in the forecast. For BRaVDA to be deployed operationally, an algorithm to automatically detect and remove ICMEs from the real time solar wind data would be required. Some methods based on proton temperature (Cane and Richardson 2003) and iron charge state have been tested with corotation forecasts (Kohutova et al. 2016). Removing ICMEs using these indicators led to an improvement in forecast skill score, so this technique could be applied for operational DA. Developing BRaVDA for operational solar wind forecasting will additionally require the use of real-time spacecraft observations, analysis of which is left for future work.

Data availability

OMNI data were downloaded from the OMNIWeb portal at `https://omniweb.gsfc.nasa.gov/ow.html`. STEREO-A and STEREO-B data were downloaded from the CDAWeb Data Explorer portal at `https://cdaweb.gsfc.nasa.gov/`. Spacecraft location data were downloaded from `https://omniweb.gsfc.nasa.gov/coho/helios/heli.html`. HelioMAS model output is available from the Predictive Science Inc. (website: `http://www.predsci.com/mhdweb/home.php`).

Acknowledgements

Harriet Turner is funded through SCENARIO grant number NE/S007261/1. Work was part-funded by Science and Technology Facilities Council (STFC) grant numbers ST/R000921/1 and ST/V000497/1, and Natural Environment Research Council (NERC) grant number NE/S010033/1.

Chapter 6

Solar wind data assimilation in an operational context: Use of near-real-time data and the forecast value of an L5 monitor

This chapter has been published in the AGU journal *Space Weather* with the following reference:

Turner, H., Lang, M., Owens, M., Smith, A., Riley, P., Marsh, M., & Gonzi, S.(2023). Solar wind data assimilation in an operational context: Use of near-real-time data and the forecast value of an L5 monitor. *Space Weather, 21,* e2023SW003457. `https://doi.org/10.1029/2023SW003457`

Key points

- Solar wind data assimilation needs to perform well with near-real-time data for it to be used operationally for space weather forecasting.
- Despite lower data quality, solar wind speed forecasts based on near-real-time data are comparable to those based on science-level data.
- Assimilation of L1 and L5 data gives forecast error improvement of 15% for lead times up to 5 days over assimilation of only L1 data.

Abstract

For accurate and timely space weather forecasting, advanced knowledge of the ambient solar wind is required, both for its direct impact on the magnetosphere and for accurately forecasting the propagation of coronal mass ejections to Earth. Data assimilation (DA) combines model output and observations to form an optimum estimation of reality. Initial experiments with assimilation of in situ solar wind speed observations suggest the potential for significant improvement in the forecast skill of near-Earth solar wind conditions. However, these experiments have assimilated science-quality observations, rather than near-real-time (NRT) data that would be available to an operational forecast scheme. Here, we assimilate both NRT and science observations from the Solar Terrestrial Relations Observatory (STEREO) and near-Earth observations from the Advanced Composition Explorer (ACE) and Deep Space Climate Observatory (DSCOVR) spacecraft. We show that solar wind speed forecasts using NRT data are comparable to those based on science-level data. This suggests that an operational solar wind DA scheme would provide significant forecast improvement, with reduction in the mean absolute error (MAE) of solar wind speed around 46% over forecasts without DA. With a proposed space weather monitor planned for the L5 Lagrange point, we also quantify the solar wind forecast gain expected from L5 observations alongside existing observations from L1. This is achieved using configurations of the STEREO and L1 spacecraft. There is a 15% improvement for forecast lead times of less than 5 days when observations from L5 are assimilated alongside those from L1, compared to assimilation of L1 observations alone.

Plain Language Summary

Space weather is the conditions of space in the near-Earth environment, and it has the potential to cause a significant impact on modern day life. For accurate space weather forecasting, knowledge of the background solar wind (a continual stream of charged particles flowing from the Sun) conditions is needed. This can be achieved using data assimilation (DA), which combines existing knowledge of the system with observations to form an optimum estimation of reality. Previous solar wind DA experiments have used cleaned-up ‘science-level’ data, which only become available many days after the observations are made. But for forecasting, where a rapid response is important, DA needs to work with near-real-time (NRT) data. NRT data often contains data gaps, biases and noise when compared to the science-level data. Here, we find that using NRT data does not significantly worsen the forecasts, which is promising for DA forecasting. A future space weather

monitoring mission to the L5 Lagrange point (60 degrees behind Earth in longitude) also offers an opportunity for solar wind DA. This is tested using combinations of existing spacecraft observations. Including L5 data alongside observations for Earth improves solar wind forecasting capability for forecasts up to 5 days in the future.

6.1 Introduction

Space weather poses a threat to modern technologies and human health. It can damage satellites, cause communication failures and destroy electricity transformers causing blackouts. It also puts the health of astronauts in space and passengers on high-altitude flights at risk (Cannon 2013). Accurate space weather forecasting requires knowledge of the background solar wind, a continual stream of charged particles and magnetic field that fills the heliosphere (Parker 1958). Stream interaction regions (SIRs) form where fast streams of solar wind catch up with and compress slower streams ahead, leading to regions of higher density and stronger magnetic field (Gosling and Pizzo 1999; Richardson and Cane 2012). These can persist for several solar rotations as corotating interaction regions (CIRs) and can be a source of recurrent space weather. The most severe space weather, namely geomagnetic storms, is driven by coronal mass ejections (CMEs), which are huge eruptions of coronal material and magnetic field from the Sun (Webb and Howard 2012). These propagate through the background solar wind, meaning ambient conditions can impact the CME speed and arrival time at Earth (Cargill 2004; Case et al. 2008; Riley and Ben-Nun 2021). Although severe space weather causes the largest impacts, the effect of mild and moderate space weather also causes a considerable economic impact, with estimates of effects on the power grid over the EU and US costing USD1.3 - 2.1 trillion over a century (Schrijver 2015). With extreme space weather relying less on the background solar wind conditions, the largest improvements in forecasting is expected for mild to moderate space weather events.

Forecasting near-Earth solar wind conditions can be achieved using simple in situ observation-based methods, such as corotation (e.g. Owens et al. 2013; Thomas et al. 2018; Turner et al. 2022), or data driven methods (Riley et al. 2017). These approaches generally do not capture transient solar wind structures, such as CMEs, and only estimate the solar wind at a single point in space. Global solar wind conditions can be forecast on the basis of remote solar observations. Photospheric magnetic field observations are used to constrain semi-empirical (e.g. WSA (Wang-Sheeley-Arge), Arge et al. 2003) and more physics-based (e.g. MAS (Magnetohydrodynamics Around a Sphere), Linker et al. 1999) models of the corona. The solar wind conditions at the top of the corona can then be propagated to Earth (and

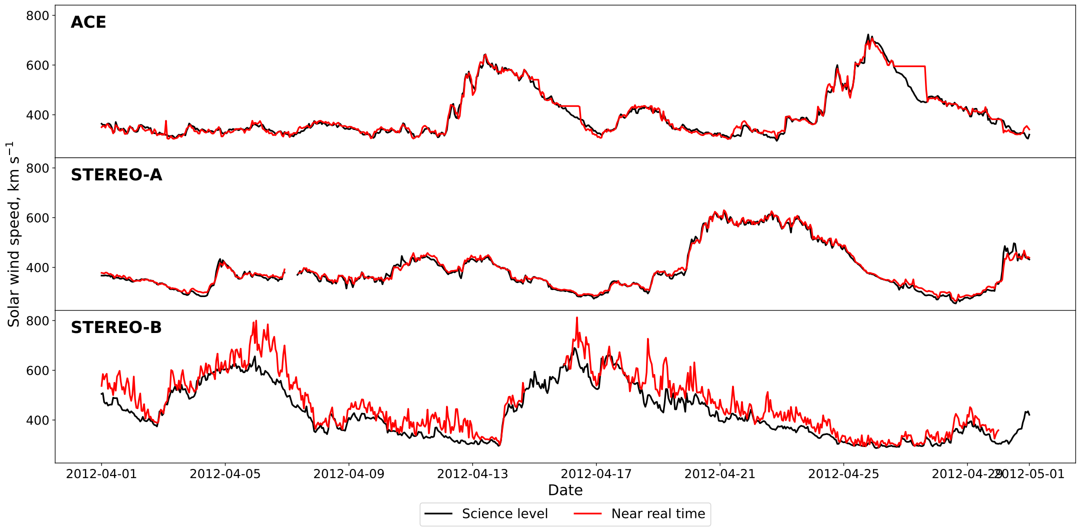


Figure 6.1: Time series of both science-level (black line) and near-real-time (red line) observations from the ACE, STEREO-A and STEREO-B spacecraft; top, middle and bottom respectively. This covers the interval from 2012/04/01 to 2012/05/01. Data are shown at an hourly resolution.

beyond) using solar wind models. This is typically achieved with numerical magnetohydrodynamic (MHD) models (e.g. Merkin et al. 2016; Odstrcil 2003; Riley et al. 2001; Tóth et al. 2005), though reduced-physics approximations can provide a complementary, computationally efficient, approach (HUX (Heliospheric Upwind eXtrapolation), Owens and Riley 2017; Riley and Lionello 2011 and HUXt Owens 2020a). CME-like disturbances can be introduced at the lower boundary of the solar wind model based on the CME characteristics observed in coronagraph observations (Odstrcil et al. 2004; Zhao et al. 2002). Once ambient and CME inner boundary conditions are supplied to the solar wind models, there are no further observational constraints on the model evolution.

Data assimilation (DA) combines model output and observations to form an optimum estimation of reality. It has led to huge improvements in terrestrial weather forecasting (Migliorini and Candy 2019), however has not been fully utilised for solar wind forecasting. The Burger Radius Variational Data Assimilation (BRaVDA) scheme (Lang and Owens 2019) makes use of in situ observations from spacecraft in both near-Earth space and from other locations within the heliosphere. It has been shown to significantly improve the model representation of the ambient solar wind, which is expected to translate to similar forecast gains (Lang et al. 2021). However, all experiments using BRaVDA so far have been carried out using ‘science-level’ data which has been processed on the ground and is often not made available for weeks or months after the observation date. For solar wind DA to be used operationally to produce timely space weather forecasts, it must be able to perform well with near-

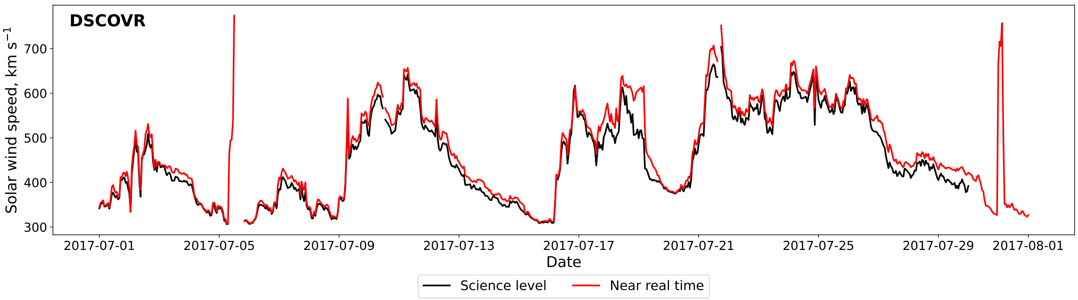


Figure 6.2: Time series of both science-level (black line) and near-real-time (red line) observations from the DSCOVR spacecraft. This covers the interval from 2017/07/01 to 2017/08/01. Data are shown at an hourly resolution.

real-time (NRT) data. NRT data often includes erroneous results, data gaps, and sometimes systematic biases; a lot of which gets corrected in the subsequent data processing stage. Figure 6.1 shows one month of NRT and science-level solar wind speed data from 2012/04/01 to 2012/05/01 for Advanced Composition Explorer (ACE, Stone et al. 1998), Solar Terrestrial Relations Observatory (Kaiser et al. 2008) Ahead (STEREO-A) and Behind (STEREO-B) spacecraft. Similarly, Figure 6.2 shows one month of data from the Deep Space Climate Observatory (DSCOVR, Burt and Smith 2012) spacecraft from 2017/07/01 to 2017/08/01. There are numerous features that show the differing quality between the NRT and science level data; for example, the step changes in the ACE NRT data, increased noise in the STEREO-B NRT data and large spikes and data gaps in the DSCOVR NRT data (Smith et al. 2022). In this study, we assess the performance of the BRaVDA scheme using archived NRT data for three time periods; 2009/08/01 to 2011/02/01, 2012/04/01 to 2013/10/01 and 2017/07/01 to 2019/01/01. The first interval covers the 18 months up to the effective boundary between solar minimum and solar maximum, whereas the second interval is during solar maximum. These were selected for their solar cycle location, whereas the final interval was an arbitrary 18-month period once the DSCOVR spacecraft was operational.

Future deployment of an operational DA scheme would aim to exploit observations from *Vigil* (Luntama et al. 2020), a planned space weather monitoring mission at the L5 Lagrange point, approximately 60 degrees behind Earth in heliospheric longitude. Alongside data from a monitor at L1, e.g. DSCOVR, this could form a framework for solar wind speed forecasting using data assimilation. Using configurations of observations from STEREO and from near-Earth, we can approximate the future pairing of L5 and L1 monitors. Here, we test the performance of BRaVDA using NRT and science-level observations from spacecraft that are separated by approximately 60 degrees in longitude to simulate an operational L5 solar wind monitor. We can then assess what forecast advantage we can expect from a future mission pairing.

The data used in this work are described in Section 6.2 and the methods in Section 6.3. The results and discussion are in Section 6.4 and the conclusions in Section 6.5.

6.2 Data

All data (NRT and science-level) are averaged to an hourly resolution using a boxcar technique with no minimum requirement for the number of data points. This is a good approximation for solar wind speed due to its high autocorrelation (Lockwood et al. 2019).

6.2.1 STEREO data

The STEREO mission was designed to provide a unique viewpoint of ejecta from the Sun and is comprised of two spacecraft; STEREO ahead (STEREO-A) and STEREO behind (STEREO-B) (Kaiser et al. 2008). These were launched into Earth-like orbits in October 2006, where STEREO-A is ahead in Earth's orbit and STEREO-B behind. The spacecraft separate at approximately 22.5° per year and reached opposition to Earth in 2014. During a planned reset of the spacecraft in October 2014 in preparation for opposition, communication with STEREO-B was lost and has not been regained. The STEREO near-real-time (beacon) data is available from `https://stereo-ssc.nascom.nasa.gov/data/beacon/` and science-level data from `https://cdaweb.gsfc.nasa.gov/`. Solar wind speed is measured using the Plasma and Suprathermal Ion Composition (PLASTIC) instrument, which provides in situ solar wind and ion observations (Galvin et al. 2008). The science data is level 2 processed data. The beacon data is provided in a continuous broadcast mode, at 1-minute resolution. For use in BRaVDA, this must be lightly processed so that any unphysical values are removed and the data is on the correct time step. As the input data used in BRaVDA is at an hourly cadence, the NRT data is averaged accordingly. This essentially interpolates over any data gaps that are less than an hour long; if there is a single 1-minute value in an hour interval then this will be taken as representative for that hour. Although this technique would not be suitable for other parameters, such as magnetic field direction, it is expected to be an adequate solution for solar wind speed, which has a long auto-correlation time (Lockwood et al. 2019). The NRT data has a typical latency of less than 10-minutes (Biesecker et al. 2008), which would not cause issues for use operationally, as the DA makes use of hourly averages.

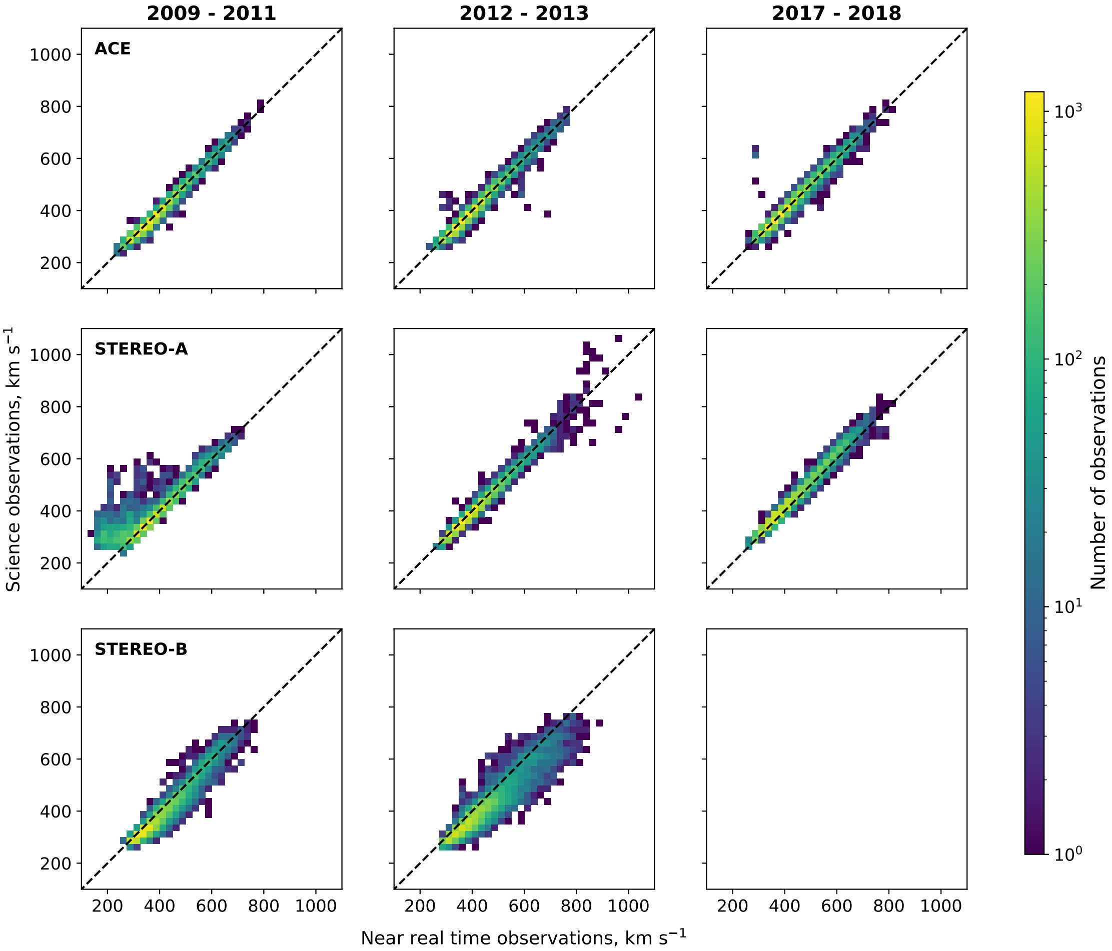


Figure 6.3: Two-dimensional histograms of science-level observations and near-real-time observations for ACE, STEREO-A and STEREO-B (rows) for the periods of time; 2009/08/01 to 2011/02/01 (left-hand column), 2012/04/01 to 2013/10/01 (middle column) and 2017/07/01 to 2019/01/01 (right-hand column). The black dashed line represents $x = y$. The number of observations are shown as a log scale.

The bottom two panels of Figure 6.1 show an example of one month of data from STEREO-A and STEREO-B. The middle panel shows the STEREO-A data, with NRT in red and science data in black, and in general there is a very good agreement between the two time series. However, the STEREO-B NRT data in the bottom panel shows much greater variability in time compared to the science data. The data plotted is at an averaged hourly resolution, meaning that a large amount of noise must have already been filtered out through this averaging. The greater variability is also demonstrated in Figure 6.3, with the STEREO data in the bottom two rows. Here we have 2D histograms of NRT against science observations, with the colour representing the density of observations on a log scale. The three time intervals used in this study are shown; 2009/08/01 to 2011/02/01, 2012/04/01 to 2013/10/01 and 2017/07/01 to 2019/01/01, the choice of which is described in Section 6.4. The STEREO-A NRT data showed periods of low solar wind speed, as shown in the left hand panel of the middle row in Figure 6.3. This is data from the period of time from October 2009 to January 2010, as shown in more detail in Figure 6.4. There is a gradual worsening of the relationship between the NRT and science-level observations, before this is resolved and the relationship returns to lie approximately along $y = x$. Although the cause of this is unknown, it provides a useful test for the DA to see how data quality affects the resulting forecasts. The later two time periods show a good relationship between NRT and science data.

The greater variability in the STEREO-B NRT data shown in Figure 6.1 can also be seen in the greater spread about the $y = x$ line in the bottom row of Figure 6.3. For the intervals shown, the average standard deviation of the difference between the science and NRT observations is 29.1 km s^{-1}, compared to 13.0 and 23.3 km s^{-1} for ACE and STEREO-A respectively. This is due to a known issue with the detector and is present for the whole operational lifetime of STEREO-B. This issue is resolved in the processing of the data on the ground that produces the science-level data.

6.2.2 ACE data

The Advanced Composition Explorer (ACE) was launched in August 1997, with the mission aiming to investigate the composition of solar wind plasma at the L1 Lagrange point. The spacecraft carries a suite of instruments, including the Solar Wind Electron, Proton and Alpha Monitor (SWEPAM) and the Real Time Solar Wind monitoring system (RTSW) (Stone et al. 1998). SWEPAM characterises the bulk flow of the solar wind through measurement of electron and ion distribution functions in 3 dimensions (McComas et al. 1998). This is then available as 1-hour science level 2 data through CDAWeb at `https://cdaweb.gsfc.nasa.gov/`. The RTSW experiment also continually transmits a feed of near-real-time data that can

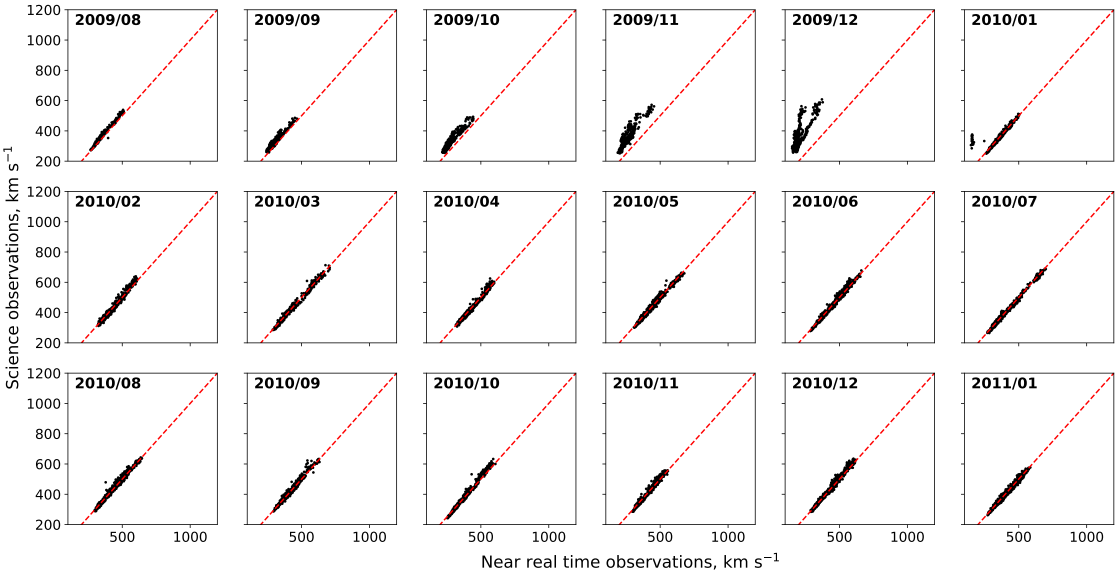


Figure 6.4: Near-real-time solar wind speeds against science-level solar wind speeds for STEREO-A, during the solar minimum interval (2009-2011) further subdivided into monthly intervals.

provide a warning of solar wind conditions to arrive at Earth up to 1 hour later (Stone et al. 1998). This data is available from NASA's Community Coordinated Modelling Centre at `https://ccmc.gsfc.nasa.gov/requests/GetInput/get_ace_K.php`.

The NRT and science-level data from ACE agree very well. As the top panel in Figure 6.1 shows, there are some features where the NRT data is constant and then steps back down to the science data. The cause of this is unknown, however, as Figure 6.3 shows, the observations mostly lie close to the $y = x$ line and so overall there is good agreement.

The NRT data has a typical latency of less than 5 minutes, which is not expected to cause any problems for an operational DA scheme.

6.2.3 DSCOVR data

The Deep Space Climate Observatory (DSCOVR) was launched in February 2015 to the L1 Lagrange point. The mission was launched to succeed ACE and to aid the National Oceans and Atmosphere Administration (NOAA) in real-time monitoring of space weather. For this study, data from the PlasMag instrument was used, which is comprised of a magnetometer, Faraday cup and a top-hat electron electrostatic analyser. Here, we make use of the observations from the Faraday cup, which measures the solar wind velocity, density and temperature. Both the NRT and science-level (level 2) data is available through the DSCOVR Space Weather Data Portal at `https://www.ngdc.noaa.gov/dscovr/portal/index.html#/`. As Figure 6.2 shows, the NRT data shows erroneous spikes in solar wind speed. This is

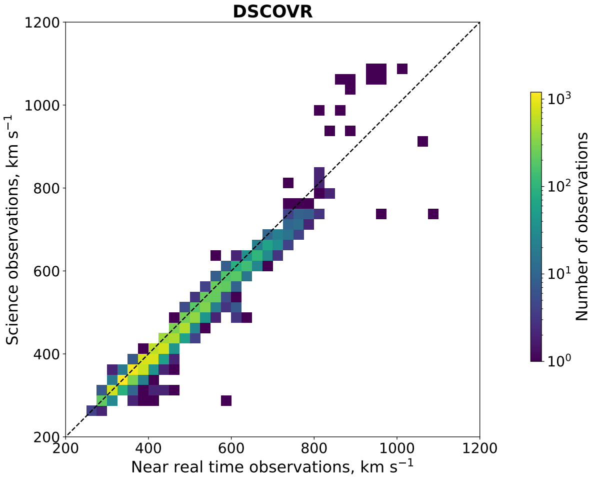


Figure 6.5: Two-dimensional histogram of science-level observations and near-real-time observations for Earth, using data from the DSCOVR spacecraft. This covers the period of time from 2017/07/01 to 2019/01/01. The black dashed line represents $y = x$. The number of observations is shown on a log scale.

due to periods of very low solar wind density, meaning that the Faraday cup cannot accurately measure the solar wind speed (Loto'aniu et al. 2022). There is mostly a good agreement between the NRT and science data (Figure 6.5).

Similarly to ACE, the NRT data latency for DSCOVR is not expected to cause any problems for an operational DA scheme.

6.3 Methods

6.3.1 BRaVDA and forecast generation

A complete description of the BRaVDA methodology can be found in Lang and Owens (2019) and the code is available at `https://zenodo.org/record/7892408#.ZFJ8o3bMK3A`. Here, we provide a brief overview of the scheme. BRaVDA combines in situ solar wind speed observations with the steady-state "HUX" model, based on Riley and Lionello (2011). BRaVDA maps information contained within in situ observations, typically at 1 AU, back to the model's inner boundary at 30 solar radii (R_S), where it is combined with the prior inner boundary condition. This prior is defined using output from the HelioMAS model (Riley et al. 2001) at $30R_S$. These model data are available at `https://www.predsci.com/portal/home.php`. The information is merged through the minimisation of a cost function, which aims to find the optimum compromise between the prior information and the observations, accounting for the uncertainties in both. Once the inner boundary at $30R_S$ is updated, this can then be propagated back out to 1 AU (and beyond) through the use of any solar wind model. For efficiency, HUX is used again for this

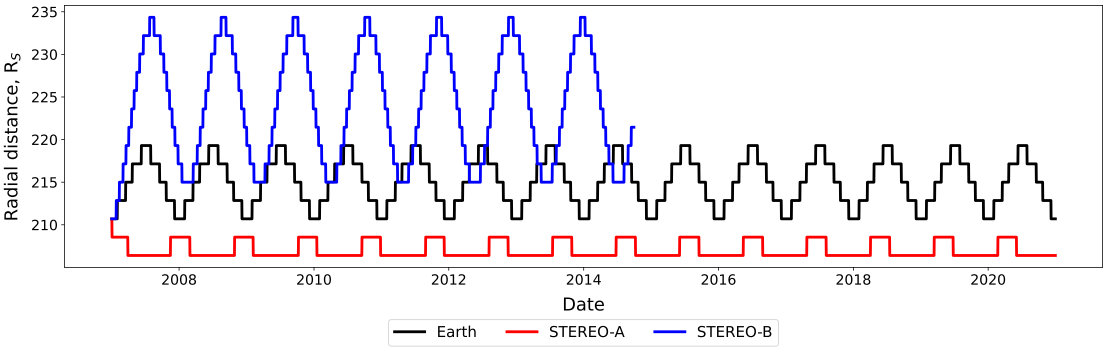


Figure 6.6: Variation of Earth (black), STEREO-A (red) and STEREO-B (blue) in radial distance from the Sun. The y-axis is shown in solar radii (R_S) and covers the time period from 2007 to 2021. Note that contact with STEREO-B was lost in 2014.

stage. This produces an estimate of the solar wind over the 2 dimensional domain from $30R_S$ to the outer boundary, which here is set to $245R_S$, to fully include the orbital radii of all spacecraft considered. The 2D plane considered here is the radius/ longitude plane, located at the solar equator.

Note that previous work using BRaVDA (e.g. Lang et al. 2021; Turner et al. 2022) has made the implicit assumption that the observations made from the STEREO spacecraft were taken from $215R_S$ (1 AU) and the L1 observations are at $213R_S$. In reality, this is not the case. As shown in Figure 6.6, Earth varies from 210 to $219R_S$ over the year, STEREO-A varies from 206 to $208R_S$ and STEREO-B varies from 215 to $234R_S$. These variations are now included into BRaVDA, ensuring that the observations were taken from the correct orbital radius. Due to the highly correlated nature of the solar wind, this radial variation did not have a significant impact on the accuracy of the forecasts, however it is important to be as representative of the system as possible.

Forecasts are generated using the output from BRaVDA in the same way as Turner et al. (2022). (As archived data are used for this work, what we state here are forecasts are actually hindcasts. However, as these hindcasts are used to inform the performance we would expect from forecasts, we retain the use of the word ‘forecast’ for simplicity.) In summary, BRaVDA is run on a daily cadence, which assimilates observations from the previous 27 days to produce a DA solution. Assuming steady state conditions, this can be corotated to produce a forecast for the subsequent 27 days. Here, forecasts are produced from assimilation of NRT and science-level observations, and both are verified against the science-level observations to assess their accuracy.

Effective L5	Effective L1	Start	End
STEREO-B	STEREO-A	02/05/2008	30/08/2008
STEREO-B	Earth	30/07/2009	22/01/2010
Earth	STEREO-A	27/05/2009	06/05/2010
STEREO-A	STEREO-B	25/10/2013	09/02/2014

Table 6.1: Intervals where spacecraft are separated by between 50 and 70 degrees in longitude. These intervals simulate spacecraft at L5 and at L1. In the left and middle panels, the spacecraft are moving away from each other and so the start date indicates the time where they are separated by 50 degrees and the end dates when separated by 70 degrees. In the right panel, the spacecraft are moving towards each other and so the start date is when they are separated by 70 degrees and the end date 50 degrees.

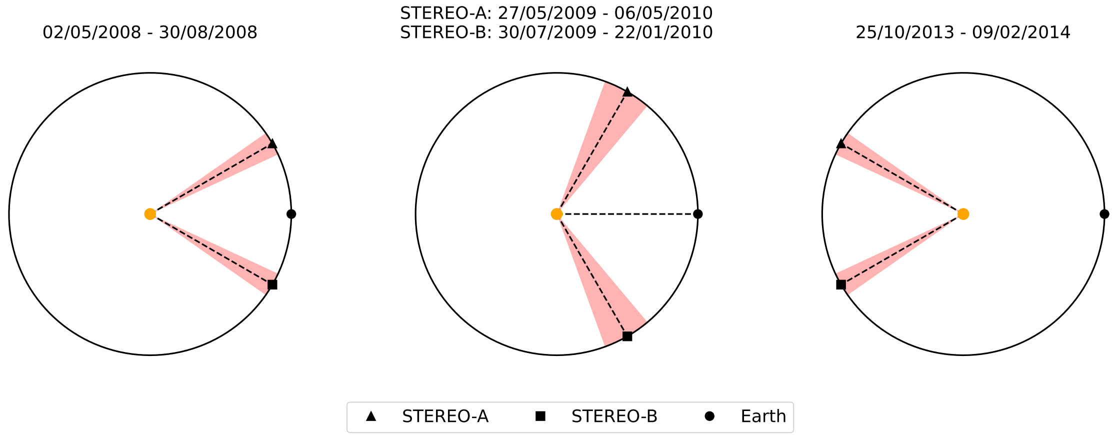


Figure 6.7: Configurations of the spacecraft used in the experiments assessing the possible combination of an L5 and L1 monitor. The red shaded regions show the time where the spacecraft are separated between 50 and 70 degrees. Earth is indicated by the black circle, STEREO-A by the black triangle and STEREO-B by the black square.

6.3.2 L5 experiments

Future deployment of an operational solar wind DA scheme could make use of both observations from near-Earth space (for example, from DSCOVR) and from the planned *Vigil* mission to L5. To test the performance of such a combination, we can use observations from pairs of spacecraft (STEREO-A, STEREO-B and ACE) that are approximately 60 degrees apart in longitude. By using intervals of time where the spacecraft separation is between 50 and 70 degrees, we produce four 'L1-L5' analysis periods. These periods are shown in Table 6.1 and schematically in Figure 6.7. The spacecraft lagging with respect to solar rotation acts as the effective L5 monitor and the spacecraft leading with respect to solar rotation is the effective near-Earth, or L1, monitor. We can then assess the forecast performance at the leading spacecraft, as this would represent a forecast at Earth.

6.4 Results and discussion

Here we conduct a number of experiments to investigate the impact of using near-real-time data on forecasts produced using DA. Here, the science-level observations act as a verification time series for the forecasts to be compared against. The science-level data is also used to produce corotation forecasts, whereby observations are lagged depending on their longitudinal separation from the forecast location. Throughout, we assess the performance of forecasts produced using mean absolute error (MAE) as a function of forecast lead time. As a standard metric, MAE allows for easy comparison of the performance of different forecasts. However, caution must be taken with such "point-by-point" metrics, as they can be misleading with forecasts of markedly different quality, typically over-penalising forecasts with small timing errors and under-penalising forecasts with very low variance (Owens et al. 2005). In this study, the difference between near-real-time and science forecasts is generally expected to be a small quantitative change, rather than leading to a qualitatively different time series. For this reason, MAE is found to generally agree with the assessment gained by visual inspection. However, Section 6.4.1 highlights a case where MAE is inadequate to characterise the forecast performance in isolation.

6.4.1 Assimilation of single and multiple spacecraft observations

We first assimilate observations from a single spacecraft. We have observations from four sources; ACE, STEREO-A, STEREO-B and DSCOVR. Three time intervals are used for analysis; 2009/08/01 to 2011/02/01, 2012/04/01 to 2013/10/01, and 2017/07/01 to 2019/01/01. The first interval is the 18 months effectively covering solar minimum, the second interval occurs during solar maximum and the final interval is an arbitrary 18-month period once the DSCOVR spacecraft was operational. Data from all spacecraft are not available for all time periods, as DSCOVR was only launched in 2015 and communication with STEREO-B was lost in 2014.

Figure 6.8 shows MAE as a function of forecast lead time for experiments assimilating observations from a single spacecraft; ACE observations are assimilated in the top panel, DSCOVR is assimilated in the second panel, STEREO-A in the third panel and STEREO-B in the bottom panel. These are shown for the time intervals where data is available; 2009-to-2011 in the left column, 2012-to-2013 in the middle column and 2017-to-2018 in the right column. Each assimilation experiment is used to produce a forecast at Earth (black lines), a forecast at STEREO-A (red lines) and a forecast at STEREO-B where available (blue lines). Forecasts are verified against the science-level observations at the respective location. Here, and throughout the text, where Earth is used as a forecast verification, this is at the L1 point and so is

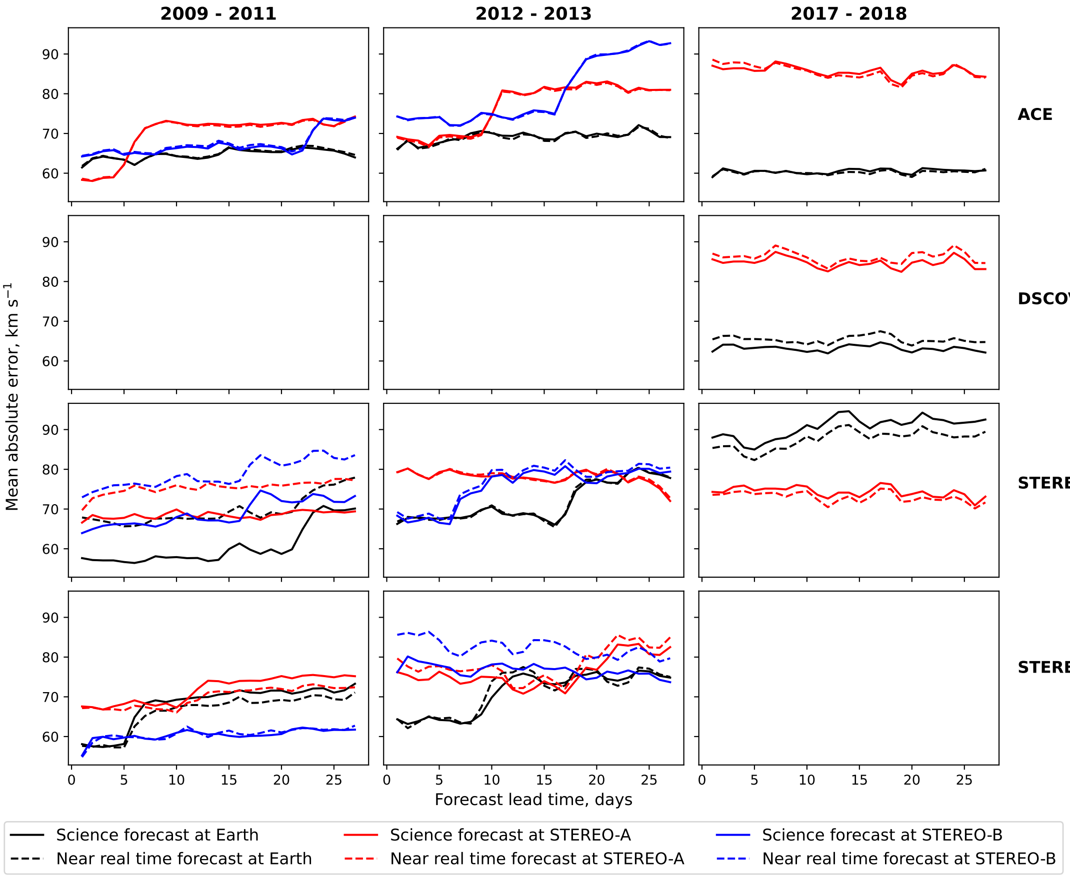


Figure 6.8: Comparison of solar wind speed forecast MAE for experiments assimilating observations from a single spacecraft, shown in the four rows. For each experiment, forecast MAE is shown at three locations; Earth (black lines), STEREO-A (red lines) and STEREO-B (blue lines). The solid lines show forecasts produced using science-level data and the dashed lines show forecasts using near-real-time data. Three time intervals are shown; 2009/08/01 to 2011/02/01, 2012/04/01 to 2013/10/01 and 2017/07/01 to 2019/01/01. Note that due to loss of communication with STEREO-B, there is no data available for the latest interval.

using data from either ACE or DSCOVR, depending on the respective time period. Forecasts produced using science-level data are shown with a solid line and those using NRT data with a dashed line.

As Figure 6.8 shows, in general there is little difference between the real time and science forecasts produced using ACE and DSCOVR data. This means that assimilating these data in an operational setting would still produce forecasts of a similar skill to forecasts produced with science-level data.

There is more difference between forecasts based on NRT and science-level data when assimilating only STEREO data. Due to the issues with the STEREO-A beacon data described in Section 6.2.1 producing a systematic error in the NRT observations, we see a larger difference between the dashed and solid lines for all forecast locations in the 2009-to-2011 panel when assimilating only STEREO-A. This issue is not present in the 2012-to-2013 or 2017-to-2018 data, and we therefore see the NRT and science forecasts producing much more similar results. The forecasts from the STEREO-B real time data in the 2012-to-2013 (approximately solar maximum) interval show greater deviation from the science forecasts than for the 2009-to-2011 (solar minimum) interval.

For the 2009-to-2011 and 2012-to-2013 time intervals, the forecasts assimilating ACE and STEREO data shows the impact from the age of observations, whereby there is a large increase in forecast error when the forecast lead time exceeds the corotation time between the assimilated spacecraft and the forecast location. This is described in more detail in (Turner et al. 2022).

Figure 6.9 shows the simultaneous assimilation of ACE, STEREO-A and STEREO-B science-level (solid lines) and NRT (dashed lines) data, used for forecasts verified at Earth, STEREO-A and STEREO-B (black, red and blue respectively). Also included in this plot is the prior forecast, shown in the dotted line, and the L1 corotation forecast using science-level observations verified at Earth in the light grey shaded region. The prior forecast is the forecast produced from previous available information, before the data assimilation is performed. In this case, the prior forecast is the HelioMAS solution from the photospheric magnetic field that is propagated radially outwards using the HUX solar wind model. The left hand panel is for the 2009-to-2011 interval and the right hand panel for the 2012-to-2013 interval.

Firstly, as Figure 6.9 shows, especially for Earth, it is clear that assimilating either NRT or science level observations offers a significant improvement in forecast skill from the prior state. Secondly, using L1 corotation (also known as recurrence or persistence) as a baseline forecast, whereby we lag the observations by 27 days and use them as a forecast, we also find an improvement over all lead times using DA. For the 2009-to-2011 and 2012-to-2013 intervals for Earth, L1 corotation gives MAEs of 68.9 km s^{-1} and 79.8 km s^{-1} respectively. Using DA also offers improvement

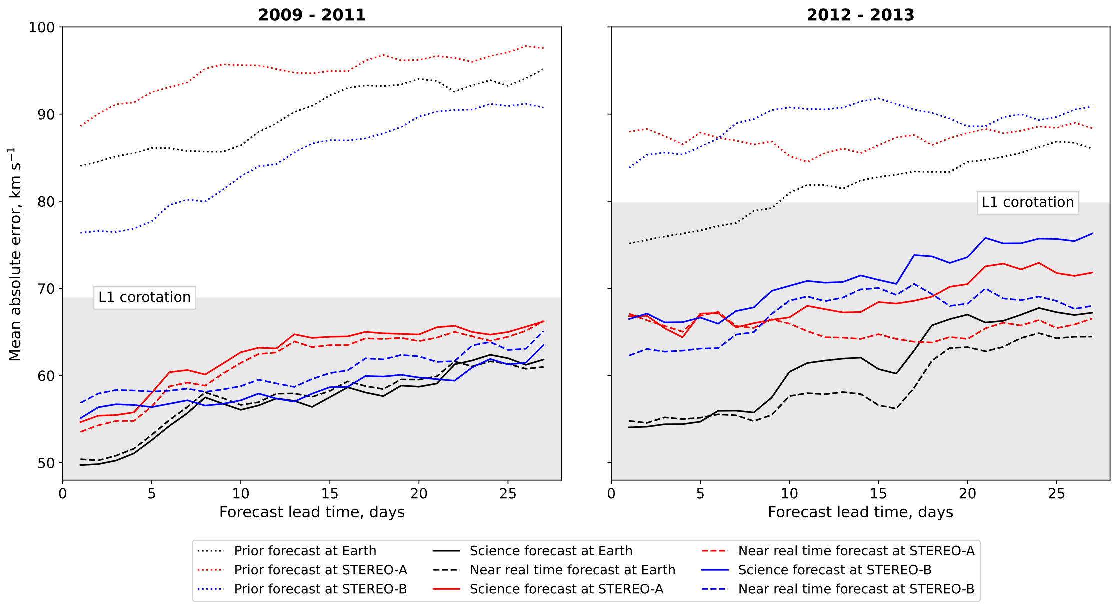


Figure 6.9: Comparison of solar wind speed forecast MAE for experiments assimilating all available observations; near-Earth, STEREO-A and STEREO-B. Forecast MAE is shown at three locations; Earth (black lines), STEREO-A (red lines) and STEREO-B (blue lines). The solid lines show forecasts produced using science-level data, the dashed lines show forecasts using near-real-time data and the dotted lines show forecasts using the prior estimate (i.e. with no DA). L1 corotation forecast error verified at Earth is shown in the light grey shaded region. Two time intervals are shown; 2009/08/01 to 2011/02/01 and 2012/04/01 to 2013/10/01.

over a forecast produced from L1 corotation as it reconstructs the whole domain between the Sun and Earth's orbital radius and provides an updated inner boundary condition that can be used in MHD models. This allows for the propagation of CMEs through the improved background solar wind, something which cannot be achieved through a simple corotation forecast. With CMEs being the main driver of severe space weather, this offers the opportunity to improve their forecasted speed and arrival time.

It can be seen that there is no major difference between the NRT and science forecasts for the earlier interval. Particularly for the 2009 - 2011 interval, it could be expected that the lowest MAE would be seen for forecasts at STEREO-A due to the other observations being closer in longitude behind the spacecraft (with respect to solar rotation). However, it is seen that the lowest MAE are seen for forecasts at Earth. The trends for both Earth and STEREO-A are similar, but there is a systematic offset due to different structures being encountered at the spacecraft over a limited time period. The difference is likely not meaningful due to this reason.

For the 2012-to-2013 interval, from a forecast lead time of approximately 10 days, the forecasts produced using NRT observations appear to perform better than

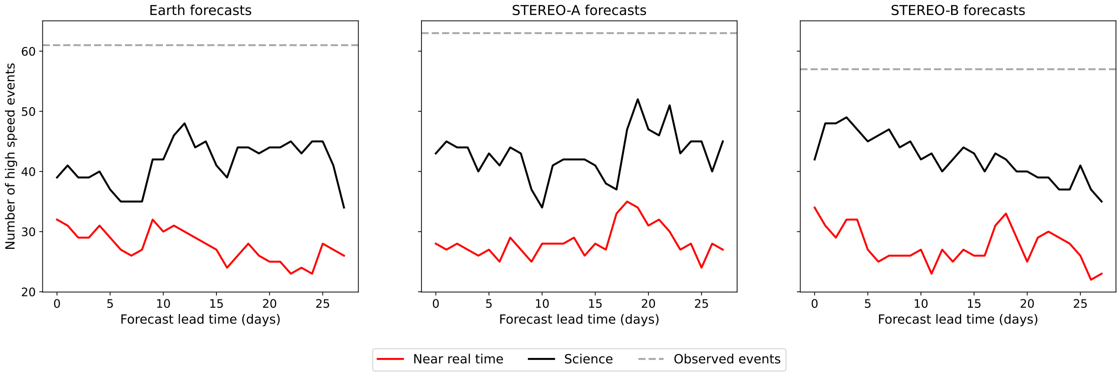


Figure 6.10: Number of forecast fast events (> 500 km s^{-1}) for different lead times for forecasts created using near-real-time and science-level. Observed events are seen in the 1-hour resolution science-level observation time series and shown as the grey-dashed lines. Interval shown; 2012/04/01 to 2013/10/01.

those produced with the science-level observations. As demonstrated below, this improvement comes about due to the NRT-based forecasts producing a 'flatter' solar wind speed time series that doesn't contain the full variability of the observations. Thus, if timing errors are present in both the science-level and NRT-based forecast, the science forecast would suffer greater penalty when assessed by MAE (e.g. Figure 1 of Owens 2018). This is demonstrated in Figure 6.10, where the number of high-speed events in the forecast time series using the science-level observations (black line) is greater than those using NRT observations (red line) for all lead times. Here, we define a high-speed event as having a solar wind speed greater than 500 km s^{-1}. This encapsulates both CMEs and fast solar wind streams. Both science- and NRT-based forecasts underestimate the number of high-speed events compared with observations, as expected as high-speed CMEs are not captured by the steady state data assimilation.

The forecast characteristics can be displayed using a Taylor diagram, as shown in Figure 6.11, which summarises the forecast MAE and linear correlation coefficient with the verification data, as well as the standard deviation of the forecasts. As forecasts improve, they move closer to the observation location, shown as a black star. It can be seen that the NRT and science forecasts group into two areas of roughly equal distance from the ideal forecast, but with the science forecasts having a standard deviation more representative of the observations. We can also see that there is an evolution of forecast MAE as the lead time increases, with the longer lead times producing forecasts with a higher MAE.

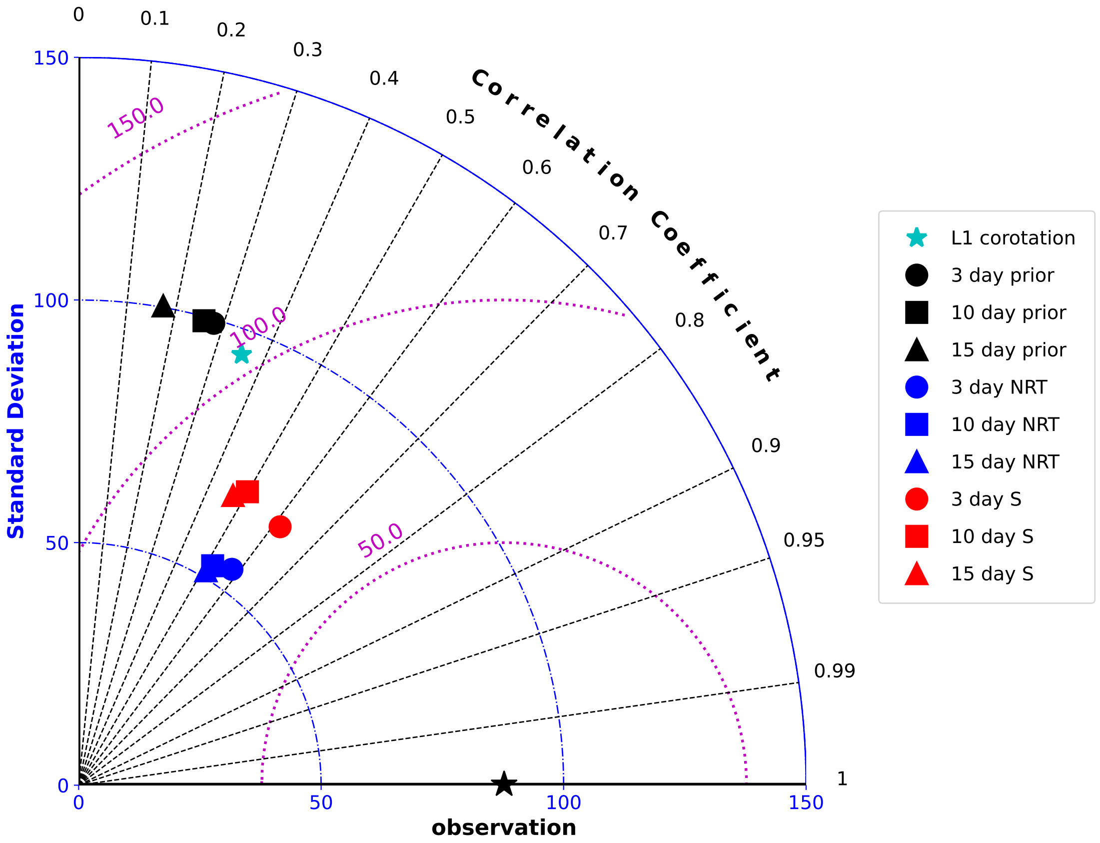


Figure 6.11: Taylor diagram of 3- (circle), 10- (square) and 15- (triangle) day lead-time forecasts of solar wind speed from 2012/04/01 to 2013/10/01 using the prior (black), near-real-time observations (NRT, blue) and science-level observations (S, red). The L1 corotation forecast for this interval, verified at Earth, is shown as the cyan star. Black radial lines show the correlation coefficient between the forecast and the verification values, the blue circular lines show the standard deviation and the purple circular lines show the forecast MAE. The observation metrics are shown with the black star.

6.4.2 L5 experiments

The future *Vigil* mission offers a chance for an operational data assimilation scheme to make routine use of simultaneous L5 and L1 data. To test this scenario, we can use combinations of STEREO and ACE data during specific intervals to mimic such a pairing. The forecast at the effective L1 position can then be assessed, as that would be Earth in an operational setting. Four intervals (Table 6.1) were identified where the spacecraft longitudinal separation was between 50 and 70 degrees, and BRaVDA was run with both NRT and science-level observations. Two sets of experiments were run; assimilating both effective L1 and L5 data and assimilating the effective L1 only. This allows the forecast gains from the L5 mission to be assessed.

Figure 6.12 shows the forecast MAE variation with forecast lead time. The prior is shown in the solid black line, the L1 only assimilated observations in red and the L1 and L5 assimilated observations in blue. The assimilated science data is shown in the solid coloured lines and the NRT data in dashed. Also shown on this plot are the L1 corotation forecast errors verified at the effective L1 spacecraft in the grey shaded region. These forecasts are made using the science-level observations. The forecasts produced using DA show similar forecast errors to L1 corotation, except in panel d), where the error from corotation is similar to that of the prior. The time interval covered in panel d), 2013/10/25 to 2014/02/09, is at approximately solar maximum, whereas the intervals in panels a-c) are in solar minimum. This means that there are likely more CMEs observed during this time, which cannot be captured in corotation forecasts and would therefore lead to a larger forecast error.

We also compare the DA forecasts to those produced using corotation from L5. Due to the separation of the spacecraft, it takes approximately 5 days for the solar wind to corotate round from the effective L5 spacecraft to the effective L1 spacecraft. As a consequence, the forecast lead time is approximately 5 days, thus giving a L5 corotation forecast. This forecast produces a lower MAE than L1 corotation, due to the shorter amount of time through which the solar wind can evolve whilst the Sun rotates from observation point to forecast point. The darker grey shaded region in Figure 6.12 shows the MAE from L5 corotation for each associated time interval. For panel a), the DA outperforms L5 corotation in both instances. For panels b) and d), assimilation of L1 and L5 gives the lowest error, whereas L5 corotation outperforms assimilation of L1 only. In panel c), L5 corotation gives the lowest MAE. Although DA offers no significant improvement over L5 corotation purely through MAE, its advantages come from the reconstruction of the whole domain and updating of the inner boundary condition. This means that it can be used to inform and improve MHD models and also allows for CMEs to be propagated through an updated background solar wind. This could lead to improved CME

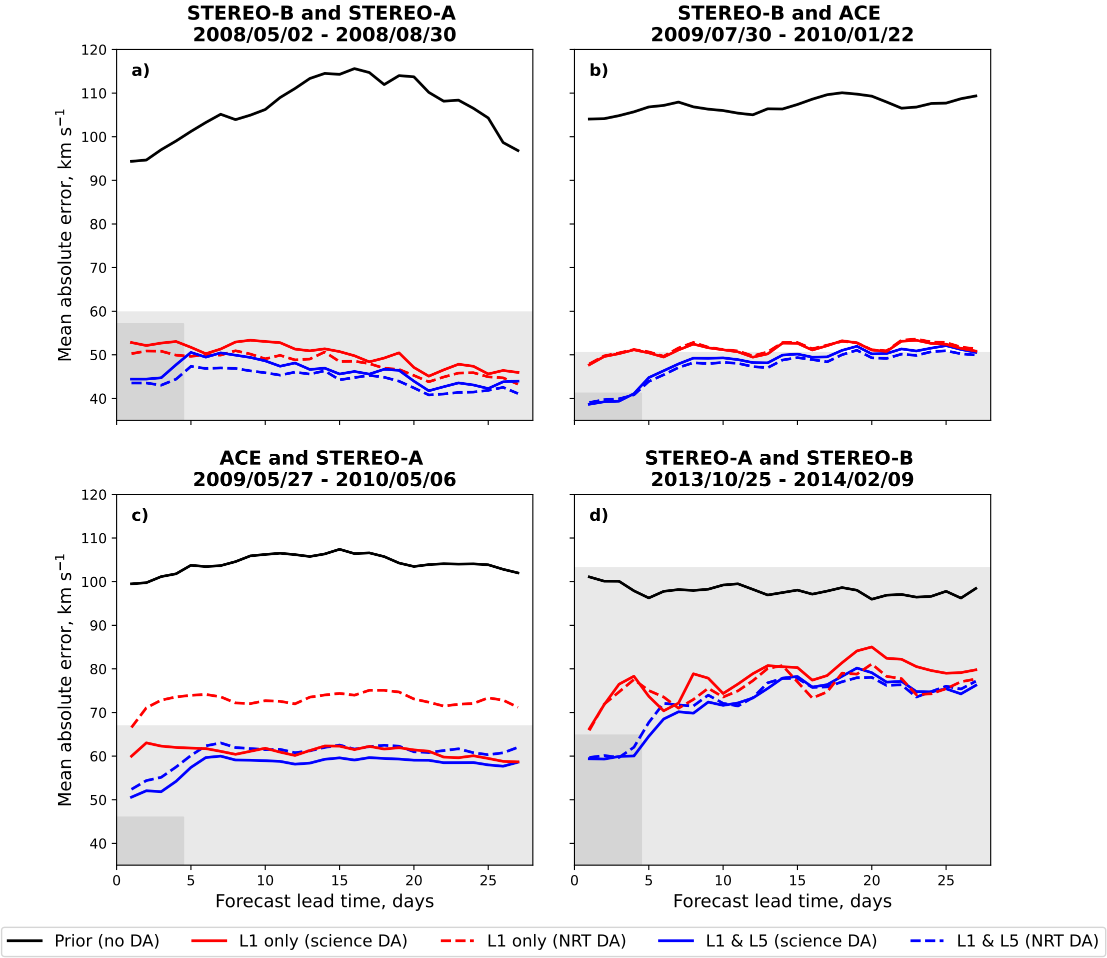


Figure 6.12: (a - d) Solar wind speed forecast MAE for experiments assimilating only effective L1 (red) and both effective L1 and L5 (blue) observations. The solid lines show forecasts produced using science-level observations and the dashed lines using near-real-time (NRT) observations. The black lines show forecasts produced using the prior (i.e., no DA). L1 corotation forecast error verified at the effective L1 spacecraft is shown in the light grey shaded region and the L5 corotation forecast is shown in the dark grey region for each time period, up to its maximum lead-time of 5 days. The panel a) covers the time period 2008/05/02 to 2008/08/30, panel b) covers 2009/07/30 to 2010/01/22, panel c) covers 2009/05/27 to 2010/05/06 and panel d) covers 2013/10/25 to 2014/02/09.

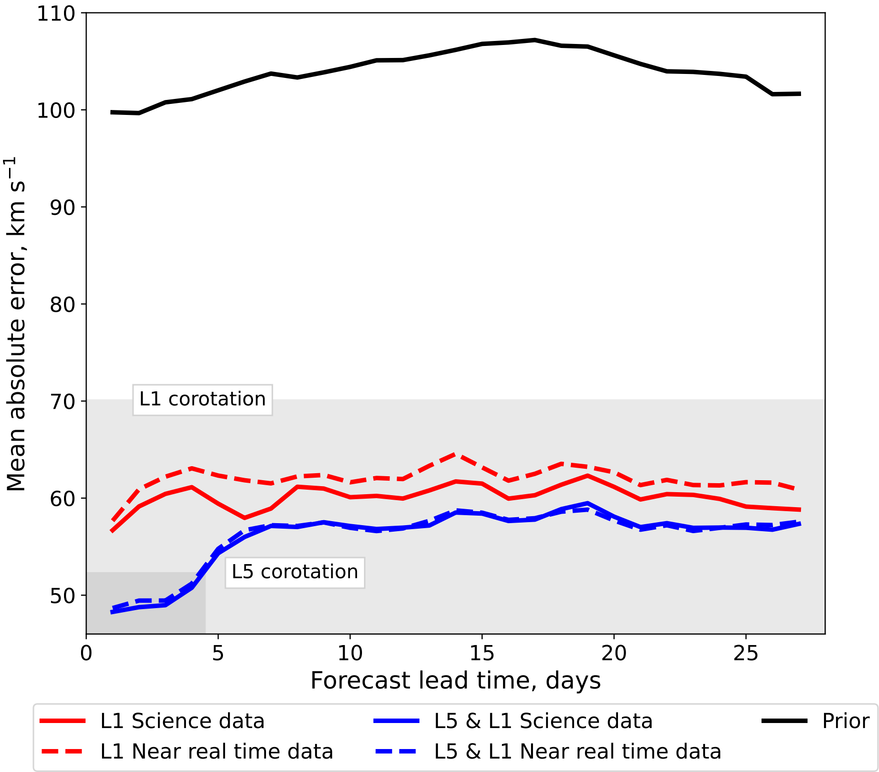


Figure 6.13: Average of the four L5 experiments shown in Figure 6.12. Mean absolute error is shown as a function of forecast lead time. The black line shows the forecast produced using the prior, the red line shows the forecasts from assimilation of only the effective L1 observations and the blue line shows assimilation of both L1 and L5. The coloured solid lines use science-level observations and the dashed lines use near-real-time (NRT) observations. The lighter grey shaded region shows the L1 corotation forecast error and the darker shaded region shows the L5 corotation forecast error, averaged from Figure 6.12.

.

arrival and speed predictions.

The NRT and science-level observations have very similar forecast errors, with no major difference between the solid and the dashed lines. There is one exception; assimilating only STEREO-A NRT as the effective L1. This forecast shows a larger MAE of approximately 10 km s^{-1}, as this interval contains the period of time where there is much lower solar wind speeds in the NRT data when compared with the science-level data, as shown in Figure 6.4.

In general, it can be seen that the assimilation of both L5 and L1 does not offer a large forecast gain for forecast lead times greater than 4-5 days. However; for less than 5 days, the assimilation of L1 and L5 is 9.0 ± 1.1 km s^{-1} lower in MAE. This is because the corotation time associated with 60 degrees of separation is 4.5 days. Thus the effective age of observations increases significantly after around 4 days, as discussed in Turner et al. (2022).

To further summarise these results, we average the four panels in Figure 6.12 to

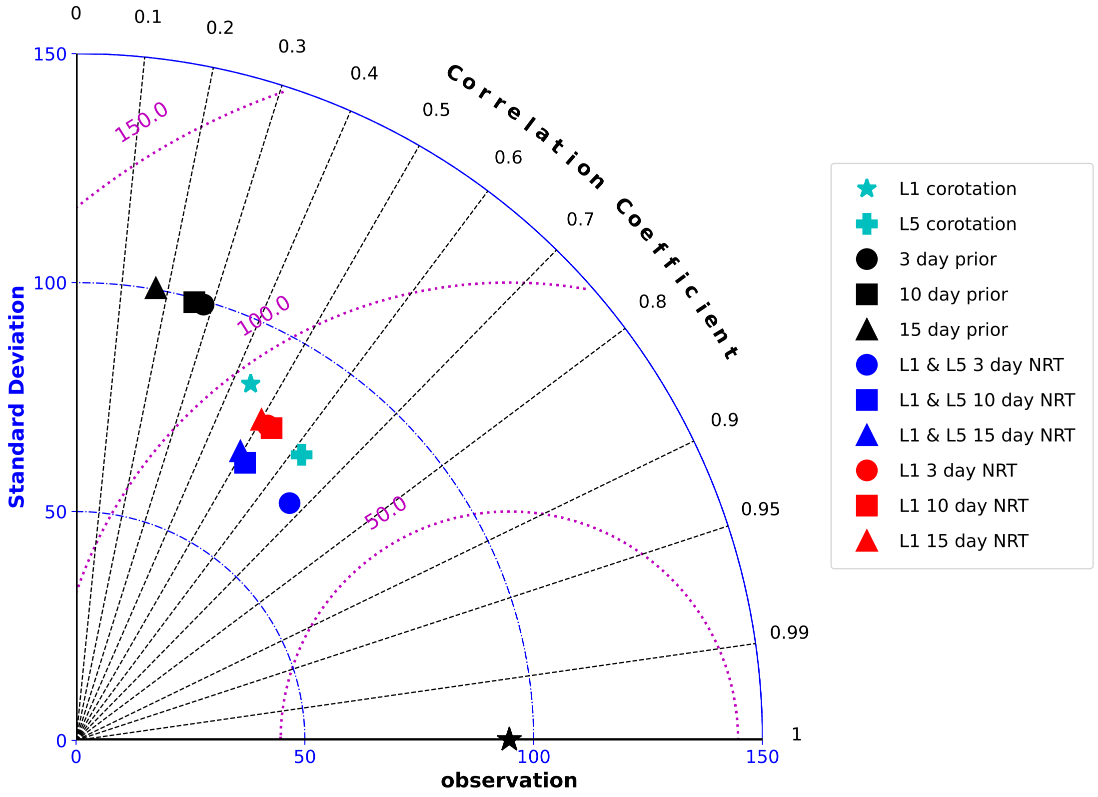


Figure 6.14: Taylor diagram of selected lead time for the prior forecasts (black), L1 NRT only forecasts (red) and L1 and L5 NRT forecasts (blue). 3-day lead time is shown with a circle, 10-day with a square and 15-day with a triangle. The cyan star shows the L1 corotation forecast and the cyan plus shows the L5 corotation forecast, both averaged over the four intervals. The observation metrics are shown with a black star. Note that the red circle is overlaid by the red square.

give Figure 6.13, which shows the improvement in the first 5 days of forecast lead time more clearly. Comparing the assimilation of only L1 and of both L1 and L5 against the forecast using the prior information, we can see significant improvements, with a percentage decrease (absolute difference), averaged over all lead times, of $42.7 \pm 3.3\%$ (44.5 ± 3.5 km s^{-1}) and $46.3 \pm 3.3\%$ (48.2 ± 3.4 km s^{-1}) respectively. Over all lead times, inclusion of L5 in the assimilation provides a $6.2 \pm 1.7\%$ decrease (3.7 ± 1.0 km s^{-1}) in MAE from assimilating only L1. However; in the first five days of forecast lead time, there is a $15.1 \pm 1.8\%$ (9.0 ± 1.1 km s^{-1}) decrease when including L5 data. This is compared to a $4.1 \pm 1.6\%$ (2.5 ± 1.0 km s^{-1}) decrease for lead times greater than 5 days.

As Figure 6.13 shows, assimilation of both the science and NRT observations for both L1 only and L1 and L5 performs better than corotation from L1. Only assimilation of L1 and L5 together performs better than corotation from L5. However, as discussed above, the DA offers improvements over simple corotation due to it updating the whole domain and for allowing the propagation of CMEs through its output.

Figure 6.14 summarises the prior and NRT forecast metrics in a Taylor diagram. The forecasts from the prior information are shown in black, assimilation of L1 and L5 NRT data in blue and only L1 NRT in red. Three lead times are shown; 3 days represented with a circle, 10 days with a square and 15 days with a triangle. The observation metrics are shown with a black star. L1 corotation is shown with the cyan star and L5 corotation with the cyan plus. We can see that assimilating L1 and L5 reduces the variability (standard deviation, blue axis) compared to just L1, so there is not much of an improved forecast for lead time greater than 5 days, despite the lower MAE (purple axis). However; for lead times less than 5 days (the blue circle), despite the correlation and standard deviation remaining similar to the other forecasts, there is a genuine improvement in the MAE when including L5 data.

6.5 Conclusions

In this study we have assessed the performance of the BRaVDA scheme with near-real-time (NRT) observations from the STEREO, ACE and DSCOVR missions. Previous work has been based on the pre-processed, science-level data, but for a solar wind data assimilation scheme to be used operationally it must perform well with NRT data. The forecasts using NRT observations were verified against the science observations, as they are assumed to best represent reality.

Using three test intervals, 2009/08/01 to 2011/02/01 (approximately solar minimum), 2012/04/01 to 2013/10/01 (approximately solar maximum) and 2017/07/01 to 2019/01/01 (interval with DSCOVR availability), BRaVDA was first tested by assimilating individual sources of observations. It was found that for L1 spacecraft (i.e. ACE and DSCOVR), the NRT and science observations produced forecasts with no significant difference, despite there being some quality issues within the input observation time series.

The NRT STEREO observations were found to be more problematic. In the NRT STEREO-A observations, a period of approximately three months at the end of 2009 had anomalously low NRT values compared to the science-level data. This problem gradually worsened over the three months before the NRT values returned close to the science-level observations in 2010/01. The effect of this was seen in the comparison between the DA-forecasts produced using the NRT and science observations, whereby the NRT forecasts have a greater MAE of approximately 10 km s^{-1}. This problem does not occur in the later two periods, showing that the quality of the observations needs to be continually assessed so that issues can be addressed in a timely manner. From a straight comparison between NRT and science data, it is not obvious what will cause a problem in the assimilation. So it is important to periodically assess the forecast quality by checking previous NRT forecasts against newly made forecasts using science-level data once it is available.

The STEREO-B NRT observations contain a large amount of noise (i.e. high frequency variations) at roughly the hour timescale compared to the science-level observations. As a result, the STEREO-B NRT data produces an inferior forecast in regards to MAE at the position of STEREO-B itself. At other spacecraft locations, however, there is little difference between NRT and science-level forecasts. The reasons for this difference are not obvious, but may be due to the specific solar wind conditions due to these relatively short intervals. However, as the STEREO-B example shows, despite a slight worsening of the forecast error in one instance, the DA copes well with random errors. In the case of STEREO-B, these were large and of the order of 50 km s^{-1} on an hour timescale. Comparing this to the systematic error seen in a few months of the STEREO-A data, we see that this produces a systematic error in the forecast. This is due to the assumption of a non-biased prior and observations in the formation of the data assimilation framework in general. This is well-known and accounted for in numerical weather prediction (Dee, Uppala, et al. 2008; Dee 2005), and bias correction methodologies have been developed. This is an area of active research, which this study falls within, and will be improved in future versions. DA can be used to correct and identify biases in input data, whereas corotation cannot.

BRaVDA was also tested with assimilation of multiple spacecraft observations from ACE, STEREO-A and STEREO-B, for both science and NRT. It was found that assimilation of both science and NRT observations performed better that the prior forecasts (i.e. without DA). Comparing these against a benchmark forecast of L1 corotation, we also see an improvement when using DA. As DA updates the entire domain, rather than a single point forecast that is produced from corotation, its solution can be used to initialise MHD models and allows for the propagation of CMEs through its output. This is not possible using corotation, thus the DA forecast model framework adds significant value to solar wind forecasting.

The future mission to the L5 Lagrange point, *Vigil*, offers the possibility of an operational DA scheme utilising routine NRT data from two vantage points. It is hoped that this will lead to large improvements in solar wind forecasting, but has not been tested from a DA perspective. For this purpose, we used BRaVDA with pairs of the STEREO spacecraft and ACE when they were separated in longitude between 50 and 70 degrees. The forecast was assessed at the effective L1 spacecraft (i.e. 50-70 degrees ahead with respect to solar rotation) to mimic a forecast at Earth. It was found that the NRT observations produce forecasts that are not significantly different to those created with the science-level observation. When these four intervals are averaged together, there is very little difference between the NRT and science forecasts. However, there is a significant improvement when compared to an example of a prior forecast. There is an average improvement of 46.3 (±3.3)%, showing that DA could offer large improvements to solar wind speed forecasting.

The assimilation of effective L1 and L5 observations was compared against assimilation of effective L1 only. We find improvement from L1 corotation, for both L1 only and L1 and L5, and a similar forecast error to L5 corotation for similar lead times for L1 and L5. As stated above, the DA offers value over corotation as it allows for the whole domain to be updated and for the propagation of CMEs. Although including the L5 observations did not provide a large improvement over L1 only for forecast lead times of more than 5 days, it did offer a 15.1 (±1.8)% decrease in forecast MAE for lead times less than 5 days. This lead time is of great interest for space weather forecasting, and so the future mission to L5 could be a step forward for solar wind forecasting capability, if solar wind DA is used operationally to exploit these observations.

Data availability

STEREO science data were downloaded from the CDAWeb Data Explorer portal at `https://cdaweb.gsfc.nasa.gov/` and STEREO NRT data from `https://stereo-ssc.nascom.nasa.gov/data/beacon/`. ACE science data were also downloaded from CDAWeb and the NRT data from NASA's Community Coordinated Modelling Centre at `https://ccmc.gsfc.nasa.gov/requests/GetInput/get_ace_K.php`. Both DSCOVR science and NRT data were downloaded from the DSCOVR Space Weather Data Portal at `https://www.ngdc.noaa.gov/dscovr/portal/index.html#/`. The code for BRaVDA is available at `https://zenodo.org/record/7892408#.ZFJ8o3bMK3A`. HelioMAS output can be found on the Predictive Science website at `https://www.predsci.com/portal/home.php`.

Acknowledgements

We would like to thank the providers of the data for this project. Harriet Turner is funded through SCENARIO Grant No. NE/S007261/1. Mathew Owens is funded through STFC grants ST/V000497/1 and ST/V00235X/1.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In the three studies presented in this thesis, we have investigated the errors present in solar wind speed forecasts in preparation for operational use. The first study diagnosed the errors in corotation forecasts due to latitudinal separation of spacecraft. These errors are expected to be present in any forecast that has to assume that the solar wind is time stationary, such as when using data assimilation (DA). In the second study, errors due to observation age and the presence of CMEs in forecasts using DA were investigated. The final study looked at the comparison between science data and real time data for DA forecasts in preparation for using the scheme operationally. In this chapter, we summarise each study and discuss their key conclusions and how the work links together to lead into the future work.

In Chapter 4, we produced corotation forecasts using solar wind speed data from the STEREO mission and the OMNI data set (near-Earth) to find the influence of latitudinal separation between spacecraft on forecast error. It was discovered that a large amount of aliasing existed between longitudinal separation, latitudinal separation and the solar cycle in the forecast data sets. This was produced from the orbital period of STEREO meaning that the longitudinal and latitudinal separation increased with each other as the solar cycle moved towards solar maximum. It proved difficult to isolate the effect of a single factor; however, by selecting periods of time where both longitudinal separation and sunspot number were approximately constant, the effect of just latitudinal separation could be studied. It was found that when the latitudinal separation is less than 6°, the error it introduces is not detectable above the noise and contribution due to other factors. However, when latitudinal separation is more than 6°, this is the primary source of corotation forecast error. For 27-day corotation forecasts, the latitudinal separation stays well below 6° at all times, hence why a latitudinal dependence is not detectable in the forecast error. This finding also implies that for DA forecasts, observations with a latitudinal separation of less than 6° would be preferable. Furthermore, corotation forecasts provide a good benchmark to compare forecasts against, which is useful when determining the accuracy of more complex methods.

The work in Chapter 5 was focused on a number of DA experiments using BRaVDA to determine how forecast error is expected to change with the solar cycle and under different observing configurations. It was found that assimilating observations from multiple spacecraft sources produces a lower forecast error than assimilating from a single spacecraft. It showed the value of multiple, well separated space weather monitoring missions, such as the future *Vigil* mission (ESA 2022). It was also found that the age of the observations had a large impact on forecast error. When only assimilating observations from a single spacecraft, when the forecast lead time exceeds the spacecraft corotation time, there is a large increase in forecast error. This is due to the steady state assumption becoming less valid with time. Assimilating multiple spacecraft sources separated in longitude reduces the increase in error at specific lead times associated with the age of the observations. The effect of CMEs on the forecast error was also investigated. It was found that including CMEs in the input time series of the DA forecasts introduced false fast streams into the forecast time series. Removing them did improve the forecast error marginally; however, this would be difficult to implement in real time. Assimilation of multiple sources of observations acted to reduce the impact of CMEs on the forecast error without the need for identification and removal.

The final study in Chapter 6 investigated the use of near-real-time (NRT) data in the DA scheme in preparation for its use operationally. Previous work has used pre-processed, science level data but for the DA to be used in an operational context, it would need to perform well with NRT data. The experiments used archived NRT data from the STEREO, ACE and DSCOVR space missions. The experiments assimilating NRT data from L1 (i.e. ACE or DSCOVR) produced forecasts with no significant difference to the forecasts using the science data, despite some data quality issues in those input data. However, the forecasts from only assimilating STEREO NRT data contained more issues, such as biases. It was found that the DA scheme coped well with random errors, but systematic biases in the input time series led to systematic errors in the forecast time series. There are techniques developed for bias correction within DA for numerical weather prediction, and so these would need to be included in future versions for solar wind forecasts. Combinations of spacecraft in a configuration similar to an L5 and L1 pairing was also tested. This was to test whether a future pairing between the *Vigil* spacecraft and an L1 monitor would be useful for operational DA. It was found that there was a significant improvement from just using the prior (i.e. no DA) to when assimilating the NRT data. Comparing the NRT DA forecasts against corotation, it was found that the

DA outperformed 27-day corotation. When comparing against L5 corotation, there was an improvement for forecast lead times of less than 5 days when using DA. Using DA also improves over corotation as it updates the entire domain and allows for CMEs to be propagated through the solution when using a time-dependent model. Corotation, on the other hand, is only for a single point in space. This showed that the Vigil mission would be a step forward for solar wind forecasting capability, if the observations are used for solar wind DA.

The work in this thesis has gradually led to the use of the solar wind DA scheme operationally. In the first study, we attempted to find the errors present in corotation forecasts to inform the DA and also to set a benchmark to compare future forecasts against. The second study showed the benefit of assimilating observations from multiple spacecraft, which could be used to inform future experiments and future spacecraft missions. The final study tested the performance of the DA with real time data, which is necessary for use in operational forecasting. Current operational forecasting methods, such as WSA-Enlil used at both SWPC and the UK Met Office, are not constrained by observations beyond the magnetograms used to initialise the coronal model. As shown in this thesis, DA offers the opportunity to better constrain solar wind models to produce more accurate forecasts.

7.2 Future work

The work in this thesis naturally leads on to the implementation of an operational solar wind data assimilation scheme. Chapters 4 and 5 were focused on diagnosing and quantifying the errors that might be present in a solar wind forecast produced using DA. This led on to testing the BRaVDA scheme with archived real time data to see how it would cope with the issues present in the data. The forecasts produced using the real time data still improved the forecast when compared to the prior. This showed that BRaVDA would be suitable for use operationally.

Previous experiments using BRaVDA have made use of MAS data for the prior. The UK Met Office use a coupled version of WSA-Enlil for their operational forecasting. Therefore, if the DA were to be used in the Met Office, it would need to be modified so that it can use the WSA output. This method has been in development and is in its rudimentary stages, running routinely every 6-hours.

WSA output can be used as the prior for DA and is available from the Met Office, updated every 2-hours. To produce the prior, a 500 member ensemble of perturbed WSA conditions is produced. This is a random perturbation, drawn from a Gaussian distribution, in longitude, latitude and rotation of the original WSA output. The unperturbed WSA state is used as the prior and the WSA ensemble is

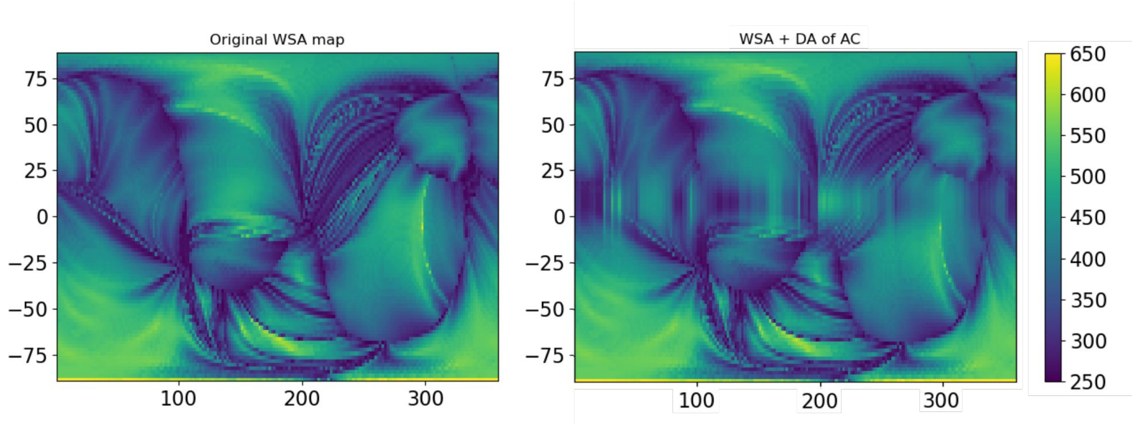


Figure 7.1: The left panel shows the original WSA map, from 1000 28/11/2023. The left panel shows the same map having been modified using the posterior.

used for the estimation of the prior error convariance.

BRaVDA is run using this set up, using real time observations from the DSCOVR and STEREO-A spacecraft. This is downloaded through a web scraper from `https://www.ngdc.noaa.gov/dscovr/data/` and `https://stereo-ssc.nascom.nasa.gov/data/beacon/ahead/plastic/` for DSCOVR and STEREO-A respectively. This must be formatted into an hourly cadence to be ingested into BRaVDA. As described in Section 3.4, the DA is run to produce new initial conditions (posterior), containing information from the observations.

The posterior is then used to modify the original WSA map. BRaVDA is a 2-D solution at Earth's latitude, whereas the WSA map covers all solar latitudes. It is therefore necessary to blend the two solutions together. At present, we apply the BRaVDA posterior values with a Gaussian weighting about Earth's latitude. By using a standard deviation of 10 degrees, the BRaVDA solution is effectively limited to influencing only a small latitudinal range about Earth's latitude.

Figure 7.1 shows an example of an original WSA map and the same map that has been modified using the posterior. The updated region can clearly be seen in the equatorial section of the right hand plot. These initial conditions can then be used in any solar wind model, but initial experiments have used the HUXt solar wind model (Owens et al. 2020a).

Using the updated WSA map in HUXt allows it to be propagated out to Earth's orbit and run forwards in time to produce a forecast. Combining this with CME parameters, also from the Met Office, this can be used in real time to give both ambient solar wind forecasts, and speed and arrival times of any Earth directed CMEs. This is a valuable tool and is the first of its kind to produce operational solar wind forecasts using in-situ DA.

The initial parts of the operational framework are in place; however, future work will include the testing and verifying of this system. After several months of the forecast being produced every 6 hours, there should be an archive of genuine forecasts to compare against observations. This will allow for the assessment of whether the ambient solar wind solution is being improved compared to the prior.

Furthermore, a catalogue of CME speed and arrival times will be required and compared against the actual values to assess whether the modification to the ambient solar wind has led to an improvement in CME forecast accuracy. Using a CME scoreboard, such as the Community Coordinated Modeling Centre (CCMC) Scoreboard (`https://kauai.ccmc.gsfc.nasa.gov/CMEscoreboard/`), will allow for the results to be compared against other models.

The current method used to blend the WSA map with the BRaVDA posterior solution has not gone through rigorous testing. The effect of changing the standard deviation of the Gaussian weighting can be investigated to find the optimum range of influence the DA solution has. Further investigation into better methods to combine these two sources of information can also be carried out.

Further tests of the ambient solar wind modification can be taken by running the updated WSA maps in a full-MHD model. The most likely candidate for this would be Enlil, as used by the Met Office. This is a project that could be run jointly with the Met Office. If improvements in the forecast accuracy are produced, then this could lead to BRaVDA being implemented operationally at the Met Office.

Work will need to be done to make BRaVDA more suitable for operational use and to improve accessibility for the scientific community. Although the code is already open source, the process of installation and use needs to be streamlined. At the least, the documentation needs to be improved and a few worked examples in a Jupyter notebook need to be included.

A lot of this work will be carried out by me, during a 3-year postdoctoral research position at the University of Reading, funded through the UK governments Natural Environment Research Council (NERC).

Bibliography

Akioka, M. et al. (2005). “The L5 mission for space weather forecasting”. *Advances in Space Research* 35.1, pp. 65–69 (cit. on pp. 58, 86).

Allen, R. et al. (2020). “Predictive capabilities and limitations of stream interaction region observations at different solar longitudes”. *Space Weather* 18.4, e2019SW002437 (cit. on p. 60).

Allen, R. et al. (2021). “Radial evolution of a CIR: Observations from a nearly radially aligned event between Parker Solar Probe and STEREO-A”. *Geophysical Research Letters* 48.3, e2020GL091376 (cit. on p. 61).

Altschuler, M. D. and G. Newkirk (1969). “Magnetic fields and the structure of the solar corona: I: Methods of calculating coronal fields”. *Solar Physics* 9, pp. 131–149 (cit. on p. 33).

Alves, M. V., E. Echer, and W. D. Gonzalez (2006). “Geoeffectiveness of corotating interaction regions as measured by Dst index”. *Journal of Geophysical Research: Space Physics* 111.A7 (cit. on p. 57).

Arge, C. N. et al. (2003). “Improved method for specifying solar wind speed near the Sun”. In: *AIP conference proceedings.* Vol. 679. 1. American Institute of Physics, pp. 190–193 (cit. on pp. 33, 34, 109).

Arge, C. and V. Pizzo (2000). “Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates”. *Journal of Geophysical Research: Space Physics* 105.A5, pp. 10465–10479 (cit. on pp. 33, 34).

Bailey, R. et al. (2020). “Prediction of d st during solar minimum using in situ measurements at l5”. *Space Weather* 18.5, e2019SW002424 (cit. on pp. 57, 58, 76).

Baker, D. et al. (1987). “Deep dielectric charging effects due to high-energy electrons in Earth’s outer magnetosphere”. *Journal of Electrostatics* 20.1, pp. 3–19 (cit. on p. 23).

Bannister, R. N. (2008). “A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances”. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 134.637, pp. 1951–1970 (cit. on p. 41).

Barabash, S. et al. (2007). “The analyser of space plasmas and energetic atoms (ASPERA-4) for the Venus Express mission”. *Planetary and Space Science* 55.12, pp. 1772–1792 (cit. on p. 29).

Bartels, J. (1934). “Twenty-seven day recurrences in terrestrial-magnetic and solar activity, 1923-1933”. *Terrestrial Magnetism and Atmospheric Electricity (Journal of Geophysical Research)* 39.3, p. 201 (cit. on p. 32).

Baruah, Y. et al. (2024). “The loss of Starlink satellites in February 2022: How moderate geomagnetic storms can adversely affect assets in low-earth orbit”. *Space Weather* 22.4, e2023SW003716 (cit. on p. 24).

Baumjohann, W. and R. A. Treumann (2012). *Basic space plasma physics.* World Scientific (cit. on p. 30).

Bayes, T. (1763). “LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S”. *Philosophical transactions of the Royal Society of London* 53, pp. 370–418 (cit. on p. 40).

Biesecker, D., D. Webb, and O. St Cyr (2008). “STEREO space weather and the space weather beacon”. *The stereo mission*, pp. 45–65 (cit. on p. 112).

Bolduc, L. (2002). “GIC observations and studies in the Hydro-Québec power system”. *Journal of Atmospheric and Solar-Terrestrial Physics* 64.16, pp. 1793–1802 (cit. on p. 23).

Boteler, D. H. (2019). “A 21st century view of the March 1989 magnetic storm”. *Space Weather* 17.10, pp. 1427–1441 (cit. on p. 23).

Boteler, D. (2001). “Space weather effects on power systems”. *Geophysical Monograph-American Geophysical Union* 125, pp. 347–352 (cit. on p. 22).

Brown, K. and G. Tautfest (1956). “Faraday-cup monitors for high-energy electron beams”. *Review of Scientific Instruments* 27.9, pp. 696–702 (cit. on pp. 27, 28).

Burgers, J. (1940). “Application of a Model System to Illustrate Some Points of the Statistical Theory of Free Turbulence”. In: *Acad. Sci. Amsterdam.* Vol. 43. 1, pp. 2–12 (cit. on p. 36).

Burlaga, L. F. (1967). “Anisotropic diffusion of solar cosmic rays”. *Journal of Geophysical Research* 72.17, pp. 4449–4466 (cit. on p. 16).

Burt, J. and B. Smith (2012). “Deep space climate observatory: The DSCOVR mission”. In: *2012 ieee aerospace conference.* IEEE, pp. 1–13 (cit. on pp. 31, 111).

Cane, H. and I. Richardson (2003). “Interplanetary coronal mass ejections in the near-Earth solar wind during 1996–2002”. *Journal of Geophysical Research: Space Physics* 108.A4 (cit. on p. 104).

Cannon, P. (2013). *Extreme space weather: impacts on engineered systems and infrastructures* (cit. on pp. 23–25, 57, 85, 109).

Cargill, P. J. (2004). “On the aerodynamic drag force acting on interplanetary coronal mass ejections”. *Solar Physics* 221, pp. 135–149 (cit. on pp. 5, 85, 109).

Case, A. et al. (2008). “Ambient solar wind’s effect on ICME transit times”. *Geophysical Research Letters* 35.15 (cit. on pp. 1, 57, 85, 98, 109).

Charbonneau, P. (2020). “Dynamo models of the solar cycle”. *Living Reviews in Solar Physics* 17, pp. 1–104 (cit. on p. 3).

Chen, F. F. (2012). *Introduction to plasma physics.* Springer Science & Business Media (cit. on p. 3).

Cliver, E. W. and W. F. Dietrich (2013). "The 1859 space weather event revisited: limits of extreme activity". *Journal of Space Weather and Space Climate* 3, A31 (cit. on p. 22).

Coles, W. (1996). "A bimodal model of the solar wind speed". In: *International Astronomical Union Colloquium.* Vol. 154. Cambridge University Press, pp. 87–96 (cit. on p. 15).

Copeland, K. et al. (2008). "Cosmic radiation exposure of aircraft occupants on simulated high-latitude flights during solar proton events from 1 January 1986 through 1 January 2008". *Advances in Space Research* 42.6, pp. 1008–1029 (cit. on p. 25).

Cranmer, S. et al. (1999). "An empirical model of a polar coronal hole at solar minimum". *The Astrophysical Journal* 511.1, p. 481 (cit. on p. 7).

Cranmer, S. R. (2009). "Coronal holes". *Living Reviews in Solar Physics* 6, pp. 1–66 (cit. on p. 7).

Cucinotta, F. A. (2014). "Space radiation risks for astronauts on multiple International Space Station missions". *PloS one* 9.4, e96099 (cit. on p. 25).

Daglis, I. A. et al. (1999). "The terrestrial ring current: Origin, formation, and decay". *Reviews of Geophysics* 37.4, pp. 407–438 (cit. on p. 22).

Davies, J. (2020). "The COR and HI instruments for the Lagrange L5 mission". *The COR and HI instruments for the Lagrange L5 mission* (cit. on pp. 58, 86, 103).

Dee, D., S. Uppala, et al. (2008). "Variational bias correction in ERA-Interim" (cit. on p. 130).

Dee, D. P. (2005). "Bias and data assimilation". *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.613, pp. 3323–3343 (cit. on p. 130).

Dungey, J. W. (1961). "Interplanetary magnetic field and the auroral zones". *Physical Review Letters* 6.2, p. 47 (cit. on p. 21).

Durante, M. (2014). "Space radiation protection: destination Mars". *Life sciences in space research* 1, pp. 2–9 (cit. on p. 25).

Ecoffet, R. (2013). "Overview of in-orbit radiation induced spacecraft anomalies". *IEEE Transactions on Nuclear Science* 60.3, pp. 1791–1815 (cit. on p. 24).

Elvidge, S. and M. J. Angling (2019). "Using the local ensemble transform Kalman filter for upper atmospheric modelling". *Journal of Space Weather and Space Climate* 9, A30 (cit. on p. 85).

Errico, R. M. (1997). "What is an adjoint model?" *Bulletin of the American Meteorological Society* 78.11, pp. 2577–2592 (cit. on pp. 46, 49).

ESA (2022). *Introducing ESA Vigil: Earth's devoted solar defender.* URL: `https://www.esa.int/Space_Safety/Vigil` (cit. on pp. 96, 134).

Feldstein, Y. I. (2016). "The discovery and the first studies of the auroral oval: A review". *Geomagnetism and Aeronomy* 56, pp. 129–142 (cit. on p. 22).

Fisher, M. et al. (2012). *Weak-constraint and long-window 4D-Var.* ECMWF Reading, UK (cit. on p. 46).

Forbes, T. (2000). “A review on the genesis of coronal mass ejections”. *Journal of Geophysical Research: Space Physics* 105.A10, pp. 23153–23165 (cit. on p. 18).

Galvin, A. B. et al. (2008). “The Plasma and Suprathermal Ion Composition (PLASTIC) investigation on the STEREO observatories”. *The STEREO Mission*, pp. 437–486 (cit. on pp. 30, 61, 112).

Gibson, E. G. (1973). *The quiet sun.* Vol. 303. National Aeronautics, Space Administration, Scientific, and Technical ... (cit. on p. 7).

Girgis, R. and K. Vedante (2012). “Effects of GIC on power transformers and power systems”. In: *PES T&D 2012.* IEEE, pp. 1–8 (cit. on p. 23).

Gonzalez, W. et al. (1994). “What is a geomagnetic storm?” *Journal of Geophysical Research: Space Physics* 99.A4, pp. 5771–5792 (cit. on p. 21).

Gopalswamy, N. et al. (2009). “The soho/lasco cme catalog”. *Earth, Moon, and Planets* 104, pp. 295–313 (cit. on p. 76).

Gopalswamy, N. et al. (2001). “Predicting the 1-AU arrival times of coronal mass ejections”. *Journal of Geophysical Research: Space Physics* 106.A12, pp. 29207–29217 (cit. on p. 18).

Gosling, J. T. (1993). “The solar flare myth”. *Journal of Geophysical Research: Space Physics* 98.A11, pp. 18937–18949 (cit. on pp. 21, 22, 57).

Gosling, J. and V. Pizzo (1999). “Formation and evolution of corotating interaction regions and their three dimensional structure”. *Space Science Reviews* 89.1-2, pp. 21–52 (cit. on pp. 17, 85, 109).

Guo, F. et al. (2015). “Particle acceleration and plasma dynamics during magnetic reconnection in the magnetically dominated regime”. *The Astrophysical Journal* 806.2, p. 167 (cit. on p. 21).

Gurevich, A. V. and E. E. Tsedilina (1985). *Long distance propagation of HF radio waves.* Springer (cit. on p. 25).

Hale, G. E. and S. B. Nicholson (1925). “The law of sun-spot polarity”. *Astrophysical Journal, vol. 62, p. 270* 62, p. 270 (cit. on pp. 7, 9).

Hanslmeier, A. (2010). *The sun and space weather.* Springer, pp. 233–249 (cit. on p. 3).

Hathaway, D. H. (2015). “The solar cycle”. *Living reviews in solar physics* 12, pp. 1–87 (cit. on pp. 9, 10).

Hathaway, D. H., R. M. Wilson, and E. J. Reichmann (2002). “Group sunspot numbers: Sunspot cycle characteristics”. *Solar Physics* 211, pp. 357–370 (cit. on p. 9).

Hickmann, K. S. et al. (2015). “Data assimilation in the ADAPT photospheric flux transport model”. *Solar Physics* 290, pp. 1105–1118 (cit. on p. 85).

Huang, G. et al. (2019). “Estimating surface solar irradiance from satellites: Past, present, and future perspectives”. *Remote Sensing of Environment* 233, p. 111371 (cit. on p. 41).

Hundhausen, A. (1999). *The many faces of the Sun: A summary of the results from NASA's Solar maximum mission.* Springer (cit. on p. 18).

Hwang, J.-A. et al. (2010). "Space radiation measurement on the polar route onboard the Korean commercial flights". *Journal of Astronomy and Space Sciences* 27.1, pp. 43–54 (cit. on p. 24).

ICRP (2007). *Radiation Protection and Safety of Radiation Sources : International Basic Safety Standards General Safety Requirements Part 3* (cit. on p. 24).

Innocenti, M. E. et al. (2011). "Improved forecasts of solar wind parameters using the Kalman filter". *Space Weather* 9.10 (cit. on p. 39).

Janjić, T. et al. (2018). "On the representation error in data assimilation". *Quarterly Journal of the Royal Meteorological Society* 144.713, pp. 1257–1278 (cit. on pp. 42, 60).

Jian, L. et al. (2006). "Properties of stream interactions at one AU during 1995–2004". *Solar Physics* 239, pp. 337–392 (cit. on p. 17).

Jian, L. et al. (2019). "Solar Terrestrial Relations Observatory (STEREO) observations of stream interaction regions in 2007–2016: Relationship with heliospheric current sheets, solar cycle variations, and dual observations". *Solar physics* 294.3, p. 31 (cit. on p. 77).

Jones, A. V. (2012). *Aurora.* Vol. 9. Springer Science & Business Media (cit. on p. 22).

Kaiser, M. et al. (2008). "The STEREO mission: An introduction". *Space Science Reviews* 136, pp. 5–16 (cit. on pp. 30, 58, 61, 86, 111, 112).

Kallenrode, M.-B. (2004). *Space physics: an introduction to plasmas and particles in the heliosphere and magnetospheres.* Springer Science & Business Media (cit. on p. 15).

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability.* Cambridge university press (cit. on pp. 39, 46).

Kivelson, M. and C. Russell (1995). *Introduction to space physics.* Cambridge university press (cit. on pp. 3–5, 13, 15, 21).

Klimchuk, J. A. (2006). "On solving the coronal heating problem". *Solar Physics* 234, pp. 41–77 (cit. on p. 5).

Kohutova, P. et al. (2016). "Improving solar wind persistence forecasts: Removing transient space weather events, and using observations away from the Sun-Earth line". *Space Weather* 14.10, pp. 802–818 (cit. on pp. 32, 57, 58, 76, 85, 104).

Lai, S. T. et al. (2018). "Deep dielectric charging and spacecraft anomalies". *Extreme events in geospace*, pp. 419–432 (cit. on p. 23).

Lang, M. and M. J. Owens (2019). "A variational approach to data assimilation in the solar wind". *Space Weather* 17.1, pp. 59–83 (cit. on pp. 1, 49, 51, 52, 60, 85, 87, 110, 116).

Lang, M. et al. (2017). "Data assimilation in the solar wind: Challenges and first results". *Space Weather* 15.11, pp. 1490–1510 (cit. on pp. 53, 60, 85, 88).

Lang, M. et al. (2021). "Improving solar wind forecasting using data assimilation". *Space Weather* 19.7, e2020SW002698 (cit. on pp. 60, 85, 88, 89, 110, 117).

Le Dimet, F.-X. and O. Talagrand (1986). “Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects”. *Tellus A: Dynamic Meteorology and Oceanography* 38.2, pp. 97–110 (cit. on pp. 46, 85).

Lepping, R. et al. (1995). “The WIND magnetic field investigation”. *Space Science Reviews* 71, pp. 207–229 (cit. on p. 86).

Lilensten, J. and A. Belehaki (2009). “Developing the scientific basis for monitoring, modelling and predicting space weather”. *Acta Geophysica* 57.1, p. 1 (cit. on p. 85).

Linker, J. A. et al. (1999). “Magnetohydrodynamic modeling of the solar corona during Whole Sun Month”. *Journal of Geophysical Research: Space Physics* 104.A5, pp. 9809–9830 (cit. on pp. 34, 37, 71, 88, 109).

Lockwood, M., A. Rouillard, and I. Finch (2009). “The rise and fall of open solar flux during the current grand solar maximum”. *The Astrophysical Journal* 700.2, p. 937 (cit. on p. 10).

Lockwood, M., M. J. Owens, and L. Barnard (2023). “Universal time variations in the magnetosphere and the effect of cme arrival time: Analysis of the february 2022 event that led to the loss of starlink satellites”. *Journal of Geophysical Research: Space Physics* 128.3, e2022JA031177 (cit. on p. 24).

Lockwood, M. (2024). “Lecture notes on plasma detectors” (cit. on pp. 28, 29).

Lockwood, M. et al. (2019). “The development of a space climatology: 1. Solar wind magnetosphere coupling as a function of timescale and the effect of data gaps”. *Space Weather* 17.1, pp. 133–156 (cit. on p. 112).

Lorenc, A. C. (1986). “Analysis methods for numerical weather prediction”. *Quarterly Journal of the Royal Meteorological Society* 112.474, pp. 1177–1194 (cit. on p. 85).

Loto’aniu, P. T. et al. (2022). “Validation of the DSCOVR spacecraft mission space weather solar wind products”. *Space Weather* 20.10, e2022SW003085 (cit. on pp. 31, 116).

Lovegren, J. and R. J. Hansman (2011). *Estimation of potential aircraft fuel burn reduction in cruise via speed and altitude optimization strategies.* Tech. rep. (cit. on p. 24).

Luntama, J.-P., S. Kraft, and A. Glover (2020). “Esa Lagrange mission for enhanced space weather monitoring”. In: *100th American Meteorological Society Annual Meeting.* AMS (cit. on p. 111).

Macneil, A. R. et al. (2021). “A statistical evaluation of ballistic backmapping for the slow solar wind: the interplay of solar wind acceleration and corotation”. *Monthly Notices of the Royal Astronomical Society* 509.2, pp. 2390–2403 (cit. on p. 16).

McComas, D. et al. (1998). “Solar wind electron proton alpha monitor (SWEPAM) for the Advanced Composition Explorer”. *The Advanced Composition Explorer Mission*, pp. 563–612 (cit. on pp. 31, 114).

McComas, D. et al. (2003). “The three-dimensional solar wind around solar maximum”. *Geophysical research letters* 30.10 (cit. on pp. 10, 11, 15, 59).

McGregor, S. et al. (2008). "Analysis of the magnetic field discontinuity at the potential field source surface and Schatten Current Sheet interface in the Wang–Sheeley–Arge model". *Journal of Geophysical Research: Space Physics* 113.A8 (cit. on p. 34).

Meehl, G. A., J. M. Arblaster, and D. R. Marsh (2013). "Could a future "Grand Solar Minimum" like the Maunder Minimum stop global warming?" *Geophysical Research Letters* 40.9, pp. 1789–1793 (cit. on p. 9).

Merkin, V. et al. (2016). "Time-dependent magnetohydrodynamic simulations of the inner heliosphere". *Journal of Geophysical Research: Space Physics* 121.4, pp. 2866–2890 (cit. on pp. 85, 110).

Meyer-Vernet, N. (2007). *Basics of the solar wind.* Cambridge University Press (cit. on pp. 3, 5, 15, 18).

Migliorini, S. and B. Candy (2019). "All-sky satellite data assimilation of microwave temperature sounding channels at the Met Office". *Quarterly Journal of the Royal Meteorological Society* 145.719, pp. 867–883 (cit. on p. 110).

Mikić, Z. and J. A. Linker (1996). "The large-scale structure of the solar corona and inner heliosphere". In: *AIP Conference Proceedings.* Vol. 382. 1. American Institute of Physics, pp. 104–107 (cit. on p. 34).

Mohler, S. (2003). "Galactic radiation exposure during commercial flights: Is there a risk?" *CMAJ* 168.9, pp. 1157–1158 (cit. on p. 25).

Moldwin, M. (2022). *An introduction to space weather.* Cambridge University Press (cit. on p. 24).

Mostl, C. et al. (2022). *HELCATS WP4 catalogue.* URL: `https://%20www.helcats-fp7.eu/catalogues/wp4%7B%5C_%7Dicmecat.html.%2010.6084/m9.figshare.4588315.v1` (cit. on p. 98).

Norquist, D. C. and W. C. Meeks (2010). "A comparative verification of forecasts from two operational solar wind models". *Space Weather* 8.12 (cit. on p. 39).

Odstrcil, D. (2003). "Modeling 3-D solar wind structure". *Advances in Space Research* 32.4, pp. 497–506 (cit. on pp. 38, 85, 87, 110).

Odstrcil, D., P. Riley, and X. Zhao (2004). "Numerical simulation of the 12 May 1997 interplanetary CME event". *Journal of Geophysical Research: Space Physics* 109.A2 (cit. on p. 110).

Odstrcil, D. and V. J. Pizzo (1999). "Distortion of the interplanetary magnetic field by three-dimensional propagation of coronal mass ejections in a structured solar wind". *Journal of Geophysical Research: Space Physics* 104.A12, pp. 28225–28239 (cit. on p. 37).

Ogilvie, K. et al. (1995). "SWE, a comprehensive plasma instrument for the Wind spacecraft". *Space Science Reviews* 71, pp. 55–77 (cit. on pp. 28, 86).

Owens, M. (2020a). "Solar-wind structure". *Oxford Research Encyclopedia of Physics* (cit. on pp. 6, 12, 76, 87, 110).

Owens, M. et al. (2020a). “A computationally efficient, time-dependent model of the solar wind for use as a surrogate to three-dimensional numerical magnetohydrodynamic simulations”. *Solar Physics* 295.3, p. 43 (cit. on pp. 37, 85, 103, 136).

Owens, M. J. (2018). “Time-window approaches to space-weather forecast metrics: A solar wind case study”. *Space Weather* 16.11, pp. 1847–1861 (cit. on pp. 66, 102, 123).

Owens, M. J. and P. Riley (2017). “Probabilistic solar wind forecasting using large ensembles of near-Sun conditions with a simple one-dimensional “upwind” scheme”. *Space Weather* 15.11, pp. 1461–1474 (cit. on pp. 16, 51, 53, 88, 110).

Owens, M. J. et al. (2014). “Ensemble downscaling in coupled solar wind-magnetosphere modeling for space weather forecasting”. *Space Weather* 12.6, pp. 395–405 (cit. on p. 33).

Owens, M. J. et al. (2019). “Near-Earth solar wind forecasting using corotation from L5: The error introduced by heliographic latitude offset”. *Space Weather* 17.7, pp. 1105–1113 (cit. on pp. 59, 60, 71, 77, 103).

Owens, M. J. et al. (2020b). “Quantifying the latitudinal representivity of in situ solar wind observations”. *Journal of Space Weather and Space Climate* 10, p. 8 (cit. on pp. 59, 60, 71, 77, 88, 103).

Owens, M. J. (2020b). personal communication.

Owens, M. J. et al. (2005). “An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model”. *Journal of Geophysical Research: Space Physics* 110.A12 (cit. on pp. 39, 119).

Owens, M. J. et al. (2013). “A 27 day persistence model of near-Earth solar wind conditions: A long lead-time forecast and a benchmark for dynamical models”. *Space Weather* 11.5, pp. 225–236 (cit. on pp. 32, 33, 57, 59, 62, 85, 109).

Owens, M. (2016). “Do the legs of magnetic clouds contain twisted flux-rope magnetic fields?” *The Astrophysical Journal* 818.2, p. 197 (cit. on p. 19).

Palacios, J. et al. (2017). “Defining scale thresholds for geomagnetic storms through statistics”. *Natural Hazards and Earth System Sciences Discussions*, pp. 1–19 (cit. on p. 22).

Parker, E. (1958). “Dynamics of the interplanetary gas and magnetic fields.” *Astrophysical Journal, vol. 128, p. 664* 128, p. 664 (cit. on pp. 15, 16, 57, 85, 109).

— (1965). “Dynamical theory of the solar wind”. *Space Science Reviews* 4.5-6, pp. 666–708 (cit. on p. 17).

Parker, E. N. (1988). “Nanoflares and the solar X-ray corona”. *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 330, July 1, 1988, p. 474-479.* 330, pp. 474–479 (cit. on p. 5).

Perrone, D. et al. (2019). “Radial evolution of the solar wind in pure high-speed streams: HELIOS revised observations”. *Monthly Notices of the royal astronomical society* 483.3, pp. 3730–3737 (cit. on p. 61).

Pizzo, V. (1978). “A three-dimensional model of corotating streams in the solar wind, 1. Theoretical foundations”. *Journal of Geophysical Research: Space Physics* 83.A12, pp. 5563–5572 (cit. on p. 36).

Pomoell, J. and S. Poedts (2018). "EUHFORIA: European heliospheric forecasting information asset". *Journal of Space Weather and Space Climate* 8, A35 (cit. on pp. 37, 87).

Posner, A. et al. (2021). "A multi-purpose heliophysics L4 mission". *Space Weather* 19.9, e2021SW002777 (cit. on pp. 86, 103).

Priest, E. (2012). *Solar magnetohydrodynamics.* Vol. 21. Springer Science & Business Media (cit. on p. 13).

Priest, E. (2020). "Solar Physics: Overview". *Oxford Research Encyclopedia of Physics* (cit. on p. 6).

Qin, J. et al. (2007). "A weak-constraint-based data assimilation scheme for estimating surface turbulent fluxes". *IEEE Geoscience and Remote Sensing Letters* 4.4, pp. 649–653 (cit. on p. 46).

Rask, J. et al. (2008). "Space faring: the radiation challenge". *Nasa, Module* 3.8, p. 9 (cit. on p. 25).

Reiss, M. A. et al. (2016). "Verification of high-speed solar wind stream forecasts using operational solar wind models". *Space Weather* 14.7, pp. 495–510 (cit. on pp. 39, 57).

Richardson, I. G. and H. V. Cane (2012). "Solar wind drivers of geomagnetic storms during more than four solar cycles". *Journal of Space Weather and Space Climate* 2, A01 (cit. on pp. 17, 57, 85, 109).

Riley, P. and R. Lionello (2011). "Mapping solar wind streams from the Sun to 1 AU: A comparison of techniques". *Solar Physics* 270, pp. 575–592 (cit. on pp. 34, 36, 37, 85, 87, 103, 110, 116).

Riley, P. and M. Ben-Nun (2021). "On the sources and sizes of uncertainty in predicting the arrival time of interplanetary coronal mass ejections using global MHD models". *Space Weather* 19.6, e2021SW002775 (cit. on p. 109).

Riley, P., J. Linker, and Z. Mikić (2001). "An empirically-driven global MHD model of the solar corona and inner heliosphere". *Journal of Geophysical Research: Space Physics* 106.A8, pp. 15889–15901 (cit. on pp. 34, 35, 37, 85, 87, 110, 116).

Riley, P., J. A. Linker, and C. N. Arge (2015). "On the role played by magnetic expansion factor in the prediction of solar wind speed". *Space Weather* 13.3, pp. 154–169 (cit. on pp. 35, 87).

Riley, P. et al. (2006). "A comparison between global solar magnetohydrodynamic and potential field source surface model results". *The Astrophysical Journal* 653.2, p. 1510 (cit. on p. 34).

Riley, P. et al. (2012). "Corotating interaction regions during the recent solar minimum: The power and limitations of global MHD modeling". *Journal of Atmospheric and Solar-Terrestrial Physics* 83, pp. 1–10 (cit. on p. 71).

Riley, P. et al. (2017). "Forecasting the properties of the solar wind using simple pattern recognition". *Space Weather* 15.3, pp. 526–540 (cit. on p. 109).

Schatten, K. H. (1971). *Current sheet magnetic model for the solar corona.* Tech. rep. (cit. on p. 33).

Schatten, K. H., J. M. Wilcox, and N. F. Ness (1969). “A model of interplanetary and coronal magnetic fields”. *Solar Physics* 6, pp. 442–455 (cit. on pp. 6, 33).

Schrijver, C. J. (2015). “Socio-economic hazards and impacts of space weather: The important range between mild and extreme”. *Space Weather* 13.9, pp. 524–528 (cit. on pp. 14, 23, 109).

Schrijver, C. J. and G. L. Siscoe (2009). *Heliophysics: Plasma physics of the local cosmos.* Cambridge University Press (cit. on pp. 6, 7, 9, 15, 21).

Schulte in den Bäumen, H. et al. (2014). “How severe space weather can disrupt global supply chains”. *Natural Hazards and Earth System Sciences* 14.10, pp. 2749–2759 (cit. on p. 23).

Schwenn, R. (1990). “Large-scale structure of the interplanetary medium”. *Physics of the inner heliosphere I: large-scale phenomena*, pp. 99–181 (cit. on p. 37).

Sheikholeslami, M. and D. D. Ganji (2016). *External magnetic field effects on hydrothermal treatment of nanofluid: numerical and analytical studies.* William Andrew (cit. on p. 13).

Shibasaki, K., C. Alissandrakis, and S. Pohjolainen (2011). “Radio emission of the quiet Sun and active regions (invited review)”. *Solar Physics* 273, pp. 309–337 (cit. on p. 7).

Silverman, S. M. and H. Hayakawa (2021). “The Dalton minimum and John Dalton’s auroral observations”. *Journal of Space Weather and Space Climate* 11, p. 17 (cit. on p. 10).

Simunac, K. D. et al. (2009). “In situ observations from STEREO/PLASTIC: a test for L5 space weather monitors”. In: *Annales Geophysicae.* Vol. 27. 10. Copernicus Publications Göttingen, Germany, pp. 3805–3809 (cit. on pp. 32, 57, 58, 62, 76, 86).

Smith, A. et al. (2022). “On the considerations of using near real time data for space weather hazard forecasting”. *Space Weather* 20.7, e2022SW003098 (cit. on p. 111).

Smith, E. J. (2001). “The heliospheric current sheet”. *Journal of Geophysical Research: Space Physics* 106.A8, pp. 15819–15831 (cit. on p. 11).

Stone, E. C. et al. (1998). “The advanced composition explorer”. *Space Science Reviews* 86, pp. 1–22 (cit. on pp. 58, 86, 111, 114, 115).

Storer, L. N., P. D. Williams, and M. M. Joshi (2017). “Global response of clear-air turbulence to climate change”. *Geophysical Research Letters* 44.19, pp. 9976–9984 (cit. on p. 24).

Sugiura, M. (1963). *Hourly values of equatorial Dst for the IGY.* Tech. rep. (cit. on p. 22).

Temmer, M., J. Hinterreiter, and M. A. Reiss (2018). “Coronal hole evolution from multi-viewpoint data as input for a STEREO solar wind speed persistence model”. *Journal of Space Weather and Space Climate* 8, A18 (cit. on pp. 85, 86).

Thomas, J. H. and N. O. Weiss (2008). *Sunspots and starspots* (cit. on p. 6).

— (2012). *Sunspots: Theory and observations.* Vol. 375. Springer Science & Business Media (cit. on p. 5).

Thomas, S. et al. (2018). “Evaluating the skill of forecasts of the near-earth solar wind using a space weather monitor at L5”. *Space Weather* 16.7, pp. 814–828 (cit. on pp. 32, 57, 58, 62, 76, 85, 109).

Thompson, M. et al. (1996). “Differential rotation and dynamics of the solar interior”. *Science* 272.5266, pp. 1300–1305 (cit. on p. 5).

Tóth, G. and D. Odstrčil (1996). “Comparison of some flux corrected transport and total variation diminishing numerical schemes for hydrodynamic and magnetohydrodynamic problems”. *Journal of Computational Physics* 128.1, pp. 82–100 (cit. on p. 38).

Tóth, G. et al. (2005). “Space Weather Modeling Framework: A new tool for the space science community”. *Journal of Geophysical Research: Space Physics* 110.A12 (cit. on pp. 38, 85, 110).

Treumann, R. and W. Baumjohann (1996). *Basic plasma physics.* Imperial College Press (cit. on p. 20).

Turner, H. et al. (2021). “The influence of spacecraft latitudinal offset on the accuracy of corotation forecasts”. *Space Weather* 19.8, e2021SW002802 (cit. on pp. 32, 85, 88, 91).

Turner, H. et al. (2022). “Quantifying the effect of ICME removal and observation age for in situ solar wind data assimilation”. *Space Weather* 20.8, e2022SW003109 (cit. on pp. 109, 117, 121, 127).

Usoskin, I. G. et al. (2015). “The Maunder minimum (1645–1715) was indeed a grand minimum: A reassessment of multiple datasets”. *Astronomy & Astrophysics* 581, A95 (cit. on p. 9).

Van Doorsselaere, T. et al. (2020). “Coronal heating by MHD waves”. *Space Science Reviews* 216, pp. 1–40 (cit. on p. 5).

Vaquero, J. M. (2007). “Historical sunspot observations: a review”. *Advances in Space Research* 40.7, pp. 929–941 (cit. on p. 5).

Verscharen, D., K. G. Klein, and B. A. Maruca (2019). “The multi-scale nature of the solar wind”. *Living Reviews in Solar Physics* 16.1, p. 5 (cit. on p. 13).

Vögler, A. et al. (2005). “Simulations of magneto-convection in the solar photosphere-Equations, methods, and results of the MURaM code”. *Astronomy & Astrophysics* 429.1, pp. 335–351 (cit. on p. 3).

Vokhmyanin, M., N. Stepanov, and V. Sergeev (2019). “On the evaluation of data quality in the OMNI interplanetary magnetic field database”. *Space Weather* 17.3, pp. 476–486 (cit. on pp. 32, 58, 61, 62, 86).

Vršnak, B. et al. (2013). “Propagation of interplanetary coronal mass ejections: The drag-based model”. *Solar physics* 285, pp. 295–315 (cit. on p. 18).

Waller, J. (2021). “Estimating the full observation error covariance matrix” (cit. on p. 42).

Wang, Y.-M., S. H. Hawley, and N. R. Sheeley Jr (1996). “The magnetic nature of coronal holes”. *Science* 271.5248, pp. 464–469 (cit. on p. 7).

Wang, Y.-M. and N. Sheeley Jr (1991). “Magnetic flux transport and the sun’s dipole moment-New twists to the Babcock-Leighton model”. *Astrophysical Journal, Part 1* 375, pp. 761–770 (cit. on p. 59).

Wang, Y.-M. et al. (2000). “The dynamical nature of coronal streamers”. *Journal of Geophysical Research: Space Physics* 105.A11, pp. 25133–25142 (cit. on p. 9).

Webb, D. F. and T. A. Howard (2012). “Coronal mass ejections: Observations”. *Living Reviews in Solar Physics* 9.1, pp. 1–83 (cit. on pp. 85, 109).

Wijsen, N. et al. (2021). “A self-consistent simulation of proton acceleration and transport near a high-speed solar wind stream”. *The Astrophysical Journal Letters* 908.2, p. L26 (cit. on p. 61).

Wilson, J. W. (2000). “Overview of radiation environments and human exposures”. *Health Physics* 79.5, pp. 470–494 (cit. on p. 24).

Wold, A. M. et al. (2018). “Verification of real-time WSA- ENLIL+ Cone simulations of CME arrival-time at the CCMC from 2010 to 2016”. *Journal of Space Weather and Space Climate* 8, A17 (cit. on p. 38).

Xie, H., L. Ofman, and G. Lawrence (2004). “Cone model for halo CMEs: Application to space weather forecasting”. *Journal of Geophysical Research: Space Physics* 109.A3 (cit. on p. 38).

Xu, Y. and Y. P. Chen (1999). “The nonlinear properties of Debye sphere”. *Physica Scripta* 60.2, p. 176 (cit. on p. 13).

Yiğit, E. et al. (2016). “A review of vertical coupling in the Atmosphere–Ionosphere system: Effects of waves, sudden stratospheric warmings, space weather, and of solar activity”. *Journal of Atmospheric and Solar-Terrestrial Physics* 141, pp. 1–12 (cit. on p. 24).

Zhao, X., S. Plunkett, and W. Liu (2002). “Determination of geometrical and kinematical properties of halo coronal mass ejections using the cone model”. *Journal of Geophysical Research: Space Physics* 107.A8, SSH–13 (cit. on pp. 38, 110).

Zirker, J. B. (1977). “Coronal holes and high-speed wind streams”. *Reviews of Geophysics* 15.3, pp. 257–269 (cit. on p. 7).

Zurbuchen, T. H. and I. G. Richardson (2006). *In-situ solar wind and magnetic field signatures of interplanetary coronal mass ejections.* Springer (cit. on p. 19).