Database of Heliospheric Shock Waves

Method documentation

E. Lumme, T. V. Mäkelä, J. E. Soljento, S. W. Good,
E. K. J. Kilpua, A. Isavnin, and K. Andreeova

September 6, 2024

Contents

1 Data for the analysis

1.1 Downloading and processing of the data

All data were downloaded from the Coordinated Data Analysis Web (CDAWeb). The downloaded parameters are listed in Table 1.

Minor processing of the data included: filtering of the worst spurious spikes from the plasma data, and centering of the time tags to the centers of actual measurement intervals if not done already (the latter was done only to data from the ACE and STEREO spacecraft). More information about the resolution of the parameters, the used coordinate systems, and other spacecraft-related features can be found in Appendix A.

Table 1: Downloaded parameters.

Parameter(s)	Symbol(s)
Interplanetary magnetic field vector and magnitude	$\boldsymbol{B}$, B
Solar wind velocity and bulk speed of protons/ions	$\boldsymbol{V}$, V
Solar wind proton/ion number density	N_{p}
Solar wind proton/ion temperature or the most probable thermal speed	T_{p} or V_{th}
Spacecraft position	$\boldsymbol{X}_{\mathrm{SC}}$

1.2 Data gaps

The number of data points required for determining the conditions upstream and downstream of the shock is specified in Section 2. If significant data gaps occurred the event was excluded from the analysis.

An exception was made with the solar wind velocity component data (both STEREO spacecraft). When only the bulk speed data were available, an assumption of radial solar wind was made. STEREO-B lacked the solar wind velocity component data for about half of the events after 2007. In addition, during the solar conjunction both STEREO spacecraft experienced a partial loss of the velocity component data (and also other data). For STEREO-A the velocity component data are unavailable from around August 2014 to around January 2016. STEREO-B data are not available after September 2014 due to the loss of contact with the spacecraft (for more information see http://stereo-ssc.nascom.nasa.gov/solar_conjunction_science.shtml).

1.3 Resampling of different datasets

Some of the parameters in the database are functions of both the magnetic field ($\boldsymbol{B}$) and plasma variables ($N_{\mathrm{p}}, T_{\mathrm{p}}$). These include the Alfvén speed (v_{A}), sound speed (c_{s}), magnetosonic speed (v_{ms}) and plasma beta (β). When calculating these parameters the magnetic field and plasma data were resampled to the same resolution (if not already done for the downloaded data). When necessary, resampling was achieved by averaging magnetic field data to the resolution of the plasma data. The used resolutions are detailed in Appendix A.1.

2 Identifying the shocks

2.1 Shock candidates

Two techniques for finding the shock candidates have been employed: (1) visual inspection and (2) an automated shock detection algorithm. Visual inspection has been used for finding most of the shock candidates in the database, but in 2016 a permanent transition was made to use the shock detection algorithm for finding shock candidates for the upcoming updates. Further details can be found below and in Appendix A.1.

2.1.1 Visual inspection

In visual inspection daily plots of solar wind plasma and magnetic field parameters were surveyed to identify the shock candidates. The surveyed parameters are given in Table 1. Simultaneous, sudden jumps in the plasma and magnetic field parameters were identified. These jumps had to be significant enough and fulfil the characteristics of either fast forward (FF) or fast reverse shock (FR).

2.1.2 Automatic shock detection algorithm

In the case of the automated shock detection algorithm the novel machine learning algorithm IPSVM was used. The algorithm was trained using all the shocks in the database with spacecraft detection time before October 2015 for the following spacecraft: ACE, Wind, STEREO-A, STEREO-B, Helios-A, Helios-B, and Ulysses. After training the algorithm it was then used to search the shock candidates for Voyager 1, Voyager 2, and DSCOVR spacecraft. The shocks detected by the Wind spacecraft since October 2015 were also found with the aid of IPSVM.

2.2 Determining the shock type

2.2.1 Analysis interval

The upstream and downstream plasma states were determined over a fixed analysis interval (eight minutes, see Table 2). The number of data points within the analysis intervals varies with the data resolution and due to possible data gaps. For Helios, Ulysses, and Voyager, the resolution of the plasma data was occasionally low. In such cases the default length of the analysis interval was extended (up to 30–36 minutes depending on the spacecraft) until enough data points (at least three) were included in both the upstream and downstream intervals.

For a parameter P, the upstream and downstream values are defined by $P_{\mathrm{up/down}} = \langle P \rangle_{\mathrm{up/down}}$, where the subscripts "up" and "down" refer to the analysis intervals, and the angle brackets indicate a mean over the interval.

Upstream and downstream values of the parameters are determined in the spacecraft's frame of reference. In the case of an FF shock the spacecraft first measures the upstream region of the shock, and in the case of an FR shock the downstream region is detected first.

Consequently, the upstream and downstream analysis intervals ($\Delta t_{\mathrm{up/down}}$) depend on the shock type (see Table 2).

Table 2: Default analysis intervals of FF and FR shocks. For Ulysses and Helios the outer bounds were extended depending on the resolution of the plasma data.

Shock type	**Analysis intervals (default)**	
	Upstream (Δt_{up})	**Downstream (Δt_{down})**
FF	$[t_{\mathrm{shock}} - 9\,\mathrm{min}, t_{\mathrm{shock}} - 1\,\mathrm{min}]$	$[t_{\mathrm{shock}} + 2\,\mathrm{min}, t_{\mathrm{shock}} + 10\,\mathrm{min}]$
FR	$[t_{\mathrm{shock}} + 1\,\mathrm{min}, t_{\mathrm{shock}} + 9\,\mathrm{min}]$	$[t_{\mathrm{shock}} - 10\,\mathrm{min}, t_{\mathrm{shock}} - 2\,\mathrm{min}]$

The upstream and downstream intervals were chosen so that the mean values are taken sufficiently far from the shock ramp. The 1-minute and 2-minute time intervals are excluded from the vicinity of the shock upstream and downstream, respectively.

2.2.2 Shock criteria

In order to be included in the database, the following upstream to downstream jump conditions have to be fulfilled:

$$\frac{B_{\mathrm{down}}}{B_{\mathrm{up}}} \geq 1.2, \qquad \frac{N_{\mathrm{p,down}}}{N_{\mathrm{p,up}}} \geq 1.2, \qquad \frac{T_{\mathrm{p,down}}}{T_{\mathrm{p,up}}} \geq \frac{1}{1.2}.$$

Additionally, the solar wind speed jump has to fulfill the following condition, which depends on the shock type:

$$\begin{array}{cc} \textbf{FF} & \textbf{FR} \\ V_{\mathrm{down}} - V_{\mathrm{up}} \geq 20\,\mathrm{km\,s^{-1}} & V_{\mathrm{up}} - V_{\mathrm{down}} \geq 20\,\mathrm{km\,s^{-1}}. \end{array}$$

The condition differs for FF and FR type shocks due to the opposite order in which the observing spacecraft encounters the upstream and downstream regions of the passing shock.

Note about temperature

The condition of the temperature jump given above is less rigorous due to larger errors in the temperature measurements. Only if the change in the temperature is clearly opposite to the behaviour of the other parameters is the event excluded from the analysis.

3 Output of the analysis

3.1 Output parameters in the database

The output parameters that are provided for each shock in the database are listed in Table 3.

Table 3: Output parameters in the database. An error estimate is provided for parameters written in bold (see Section 3.3).

Parameter(s)	Database symbol(s)	Unit(s)
Date and time		e.g., 11.12.2013 11:22:23
Spacecraft	SC	
Spacecraft position	SC coordinates	Appendix A.2
Coordinate system of the position	SC cs	Appendix A.2
Shock type	Type	FF/FR
Asymptotic values	$\boldsymbol{B}^{\mathrm{up/down}}$, $\boldsymbol{V}^{\mathrm{up/down}}$, $N_{\mathrm{p}}^{\mathrm{up/down}}$, $T_{\mathrm{p}}^{\mathrm{up/down}}$	[nT], [km s^{-1}], [cm^{-3}], [10^4 K]
Downstream-to-upstream ratios	$\frac{B^{\mathrm{down}}}{B^{\mathrm{up}}}$, $\frac{N_{\mathrm{p}}^{\mathrm{down}}}{N_{\mathrm{p}}^{\mathrm{up}}}$, $\frac{T_{\mathrm{p}}^{\mathrm{down}}}{T_{\mathrm{p}}^{\mathrm{up}}}$	
Solar wind speed jump	$\lvert\Delta V\rvert$	[km s^{-1}]
Upstream sound speed	$C_{\mathrm{s}}^{\mathrm{up}}$	[km s^{-1}]
Upstream Alfvén speed	$V_{\mathrm{A}}^{\mathrm{up}}$	[km s^{-1}]
Upstream magnetosonic speed	$V_{\mathrm{ms}}^{\mathrm{up}}$	[km s^{-1}]
Upstream plasma beta	β^{up}	
Shock normal	Normal	
Coordinate system of the normal	Normal cs	
Shock theta	θ_{Bn}	[deg]
Shock speed	V_{sh}	[km s^{-1}]
Alfvén Mach number	M_{A}	
Magnetosonic Mach number	M_{ms}	
Radial solar wind	Radial V_{SW}	yes/no
Length of the analysis interval	Δt_{a}	[min]
Mean resolution of the magnetic field data	Δt_{mag}	[sec]
Mean resolution of the plasma data	Δt_{pla}	[sec]

3.2 Specification of the output parameters

Remarks

- For a parameter P, $P_{\mathrm{up/down}}$ corresponds to the upstream and downstream mean values, respectively, defined in Section 2.2.1.

- The definitions of the sound speed (c_s), Alfvén speed (V_A), magnetosonic speed (V_{ms}), and plasma beta (β) are based on certain assumptions listed in Appendix B.
- When forming the time series for c_s, V_A, V_{ms}, and β, the magnetic field data is resampled to the resolution of the plasma data (see Section 1.3).
- The use of a fixed electron temperature (T_e) may introduce errors to related parameters (C_s^{up}, V_{ms}^{up}, β^{up}, M_{ms}). See Appendix B for further details.
- Inaccuracies in the determination of the shock normal ($\hat{\boldsymbol{n}}$) may introduce errors to related parameters (M_A, M_{ms}, V_{sh}, θ_{Bn}). See Section 3.2.11 for more details.

3.2.1 Spacecraft position

The data point closest to the time of the shock is chosen as the position of the spacecraft.

3.2.2 Coordinate system of the position

See Appendix A.2.

3.2.3 Shock type

See Section 2.

3.2.4 Asymptotic values

Asymptotic values are the mean values calculated over the upstream and downstream regions for: the magnetic field magnitude (B) and its vector components ($\boldsymbol{B}$), the solar wind bulk speed (V) and the velocity components ($\boldsymbol{V}$), the proton/ion number density (N_p), and the temperature of solar wind protons/ions (T_p).

3.2.5 Downstream-to-upstream ratios

The ratios of the magnitudes of the asymptotic values are

$$\frac{B^{down}}{B^{up}}, \quad \frac{N_p^{down}}{N_p^{up}}, \quad \frac{T_p^{down}}{Tp^{up}}.$$

3.2.6 Solar wind (bulk) speed jump

The solar wind speed jump across the shock is defined by

$$|\Delta V| = |V^{down} - V^{up}|.$$

3.2.7 Upstream sound speed

The upstream sound speed is defined by

$$C_{\mathrm{s}}^{\mathrm{up}} = \langle c_{\mathrm{s}} \rangle_{\mathrm{up}} = \left\langle \sqrt{\gamma k_{\mathrm{B}} \frac{T_{\mathrm{p}} + T_{\mathrm{e}}}{m_{\mathrm{p}}}} \right\rangle_{\mathrm{up}},$$

where γ is the polytropic index, k_{B} is the Boltzmann constant, and m_{p} is the mass of a proton.

3.2.8 Upstream Alfvén speed

The upstream Alfvén speed is defined by

$$V_{\mathrm{A}}^{\mathrm{up}} = \langle v_{\mathrm{A}} \rangle_{\mathrm{up}} = \left\langle \frac{B}{\sqrt{\mu_0 N_{\mathrm{p}} m_{\mathrm{p}}}} \right\rangle_{\mathrm{up}},$$

where μ_0 is the permeability of free space.

3.2.9 Upstream magnetosonic speed

The upstream magnetosonic speed is defined by

$$V_{\mathrm{ms}}^{\mathrm{up}} = \langle v_{\mathrm{ms}} \rangle_{\mathrm{up}} = \left\langle \sqrt{v_{\mathrm{A}}^2 + c_{\mathrm{s}}^2} \right\rangle_{\mathrm{up}},$$

where the Alfvén speed, v_{A}, and the sound speed, c_{s}, are defined as above.

3.2.10 Upstream plasma beta

The upstream plasma beta is defined by

$$\beta^{\mathrm{up}} = \langle \beta \rangle_{\mathrm{up}} = \left\langle \frac{2\mu_0 k_{\mathrm{B}} N_{\mathrm{p}} (T_{\mathrm{p}} + T_{\mathrm{e}})}{B^2} \right\rangle_{\mathrm{up}}.$$

3.2.11 Shock normal

The normal vector of the shock front, $\hat{\boldsymbol{n}}$, is calculated using the mixed mode method (method MD3 in Abraham-Shrauner and Yun 1976),

$$\hat{\boldsymbol{n}} = \pm \frac{(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times [(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times (\boldsymbol{V}^{\mathrm{down}} - \boldsymbol{V}^{\mathrm{up}})]}{|(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times [(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times (\boldsymbol{V}^{\mathrm{down}} - \boldsymbol{V}^{\mathrm{up}})]|}.$$

In the case of a data gap in the velocity components the normal is calculated using the magnetic field coplanarity (Colburn and Sonett 1966), according to which

$$\hat{\boldsymbol{n}}_{\mathrm{MC}} = \pm \frac{(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times (\boldsymbol{B}^{\mathrm{down}} \times \boldsymbol{B}^{\mathrm{up}})}{|(\boldsymbol{B}^{\mathrm{down}} - \boldsymbol{B}^{\mathrm{up}}) \times (\boldsymbol{B}^{\mathrm{down}} \times \boldsymbol{B}^{\mathrm{up}})|}.$$

The sign of the normal vector is determined using conditions of the solar wind velocity such that

$$\begin{cases} \boldsymbol{V}^{\mathrm{up}} \cdot \hat{\boldsymbol{n}} & \geq 0, \quad \text{FF-type shocks,} \\ \boldsymbol{V}^{\mathrm{up}} \cdot \hat{\boldsymbol{n}} & \leq 0, \quad \text{FR-type shocks.} \end{cases}$$

Regardless of the method used, there are always several caveats in the determination of the shock normal. Consequently, its values may have significant errors (see, e.g., Schwartz 1998). Therefore, the shock normal and related parameters (M_{A}, M_{ms}, V_{sh}, and θ_{Bn}) should be considered with care.

A particular issue in the determination of the shock normal is the utilisation of fixed upstream and downstream intervals (see Table 2). To obtain correct results, the shock layer itself must be entirely excluded from these intervals. Additionally, the intervals need to correspond to the actual upstream and downstream data points, leaving out disturbances not related to the shock itself. Furthermore, they are required to be long enough to average out turbulence and wave activity. Consequently, the extent and location of the best upstream and downstream analysis intervals vary depending on the shock. The fixed upstream and downstream intervals that are used were chosen to meet the aforementioned conditions for as many shocks as possible. Nevertheless, the used intervals cannot be ideal for all investigated events, which may result in large errors in the shock normals and related parameters.

3.2.12 Coordinate system of the normal

See Appendix A.2.

3.2.13 Shock theta

The angle between the shock normal, $\hat{\boldsymbol{n}}$, and the upstream magnetic field lines is given by

$$\theta_{\mathrm{Bn}} = \frac{180^\circ}{\pi} \arccos\left(\frac{|\boldsymbol{B}^{\mathrm{up}} \cdot \hat{\boldsymbol{n}}|}{\|\boldsymbol{B}^{\mathrm{up}}\| \|\hat{\boldsymbol{n}}\|} \right).$$

3.2.14 Shock speed

The shock speed in the spacecraft's frame of reference (see Appendix A.2) is calculated using the mass flux over the shock (Schwartz 1998), i.e.,

$$V_{\mathrm{sh}} = \left| \frac{[\rho_{\mathrm{m}} \boldsymbol{V}]}{[\rho_{\mathrm{m}}]} \cdot \hat{\boldsymbol{n}} \right| = \left| \frac{N_{\mathrm{p}}^{\mathrm{down}} \boldsymbol{V}^{\mathrm{down}} - N_{\mathrm{p}}^{\mathrm{up}} \boldsymbol{V}^{\mathrm{up}}}{N_{\mathrm{p}}^{\mathrm{down}} - N_{\mathrm{p}}^{\mathrm{up}}} \cdot \hat{\boldsymbol{n}} \right|,$$

where $\hat{\boldsymbol{n}}$ is the shock normal.

3.2.15 Alfvén Mach number

In order to calculate the Alfvén Mach number, M_{A}, a Galilean coordinate transformation to the rest frame of the shock must be made (the solar wind velocity transforms $\boldsymbol{V}^{\mathrm{up}} \to \boldsymbol{V}^{\mathrm{up}\prime}$).

As a result (Schwartz 1998):

$$M_{\mathrm{A}} = \frac{|\boldsymbol{V}^{\mathrm{up}\prime} \cdot \hat{\boldsymbol{n}}|}{V_{\mathrm{A}}^{\mathrm{up}}} = \frac{|\boldsymbol{V}^{\mathrm{up}} \cdot \hat{\mathbf{n}} \pm V_{\mathrm{sh}}|}{V_{\mathrm{A}}^{\mathrm{up}}},$$

where $\hat{\boldsymbol{n}}$ is the shock normal, $V_{\mathrm{A}}^{\mathrm{up}}$ is the upstream Alfvén speed, and V_{sh} is the shock speed. The negative sign corresponds to an FF shock, and the positive sign to an FR shock. It is assumed that both shock types propagate away from the Sun in the spacecraft's frame of reference (see Appendix A.2).

3.2.16 Magnetosonic Mach number

Similarly to Alfvén Mach number above, the magnetosonic Mach number is defined by

$$M_{\mathrm{ms}} = \frac{|\boldsymbol{V}^{\mathrm{up}\prime} \cdot \hat{\boldsymbol{n}}|}{V_{\mathrm{ms}}^{\mathrm{up}}} = \frac{|\boldsymbol{V}^{\mathrm{up}} \cdot \hat{\boldsymbol{n}} + V_{\mathrm{sh}}|}{V_{\mathrm{ms}}^{\mathrm{up}}},$$

where $V_{\mathrm{ms}}^{\mathrm{up}}$ is the upstream magnetosonic speed.

3.2.17 Radial solar wind

This parameter indicates whether the solar wind was assumed to be radial in the analysis. It is significant only for STEREO-A and STEREO-B (see Section 1.2). For these spacecraft the assumption is needed when the solar wind velocity component data are missing, but the bulk speed data are available. In such a case

$$\boldsymbol{V} = V\hat{\boldsymbol{e}}_r,$$

where V is the bulk speed and $\hat{\boldsymbol{e}}_r$ is the unit vector pointing radially away from the Sun.

3.2.18 Length of the analysis interval

The length of the analysis interval (upstream and downstream) in minutes. The analysis intervals are defined in Section 2.2.1.

3.2.19 Mean resolution of the magnetic field data

The mean resolution of the magnetic field data over the analysis domain (between the outermost edges of the analysis intervals from the shock). The mean resolution is given in seconds.

3.2.20 Mean resolution of the plasma data

The mean resolution of the plasma data over the analysis domain (between the outermost edges of the analysis intervals from the shock). The mean resolution is given in seconds.

3.3 Error estimates

An error estimate is provided for the parameters written in bold in Table 1. The estimates are based on the sample standard deviations of the mean values calculated over the upstream and downstream intervals (parameters specified in Sections 3.2.4 and 3.2.7–3.2.10). For these mean values, the presented error is the sample standard deviation with the possible addition of the error resulting from the electron temperature estimate (see Appendix B.2). These errors are then propagated to obtain the error estimates for the rest of the output parameters.

NOTE: Due to the simplicity of error propagation, the error estimates may have unrealistically high values in cases where the data have particularly large fluctuations in the analysis intervals. In such cases, extra caution should be taken when interpreting the results of the presented analysis.

References

Abraham-Shrauner, B. and Yun, S. H. (1976), "Interplanetary shocks seen by Ames Plasma Probe on Pioneer 6 and 7", *Journal of Geophysical Research* **81**, 2097–2102, doi:10.1029/JA081i013p02097.

Colburn, D. S. and Sonett, C. P. (1966), "Discontinuities in the solar wind", *Space Science Reviews* **5**, 439–506, doi:10.1007/BF00240575.

Cranmer, S. R., Matthaeus, W. H., Breech, B. A., and Kasper, J. C. (2009), "Empirical constraints on proton and electron heating in the fast solar wind", *The Astrophysical Journal* **702**, 1604–1614, doi:10.1088/0004-637X/702/2/1604.

Dandouras, I., Barthe, A., and the CIS Team (2009), "User guide to the CIS measurements in the Cluster Active Archive (CAA)", https://caa.esac.esa.int/documents/UG/CAA_EST_UG_CIS_v38.pdf [accessed 26 August 2024].

King, J. and Papitashvili, N. (n.d.), "One min and 5-min solar wind data sets at the Earth's bow shock nose", https://omniweb.gsfc.nasa.gov/html/HROdocum.html [accessed 26 August 2024].

Maksimovic, M., Gary, S. P., and Skoug, R. M. (2000), "Solar wind electron suprathermal strength and temperature gradients: Ulysses observations", *Journal of Geophysical Research* **105**, 18 337–18 350, doi:10.1029/2000JA900039.

NASA (n.d.), "Coordinated Data Analysis Web (CDAWeb)", https://cdaweb.gsfc.nasa.gov/ [accessed 26 August 2024].

Newbury, J. A., Russell, C. T., Phillips, L., and Gary, S. P. (1998), "Electron temperature in the ambient solar wind: Typical properties at a lower bound at 1 AU", *Journal of Geophysical Research* **103**, 9553–9566, doi:10.1029/98JA00067.

Paschmann, G., Fazakerley, A. N., and Schwartz, S. J. (1998), "Moments of plasma velocity distributions", in *Analysis Methods for Multi-Spacecraft Data*, 125–158.

Pilipp, W. G., Miggenrieder, H., Mühläuser, K.-H., Rosenbauer, H., and Schwenn, R. (1990), "Large-scale variations of thermal electron parameters in the solar wind between 0.3 and 1 AU", *Journal of Geophysical Research* **95**, 6305–6329, doi:10.1029/JA095iA05p06305.

Rays of Space (n.d.), "IPSVM: Machine learning algorithm for detection of interplanetary shock waves", https://bitbucket.org/raysofspace/ipsvm/ [accessed 26 August 2024].

Schwartz, S. J. (1998), "Shock and discontinuity normals, Mach numbers, and related parameters", in *Analysis Methods for Multi-Spacecraft Data*, 249–270.

A Spacecraft specification

A.1 Data sources

All the different instruments used as sources of the magnetic field, plasma, and position data of the different spacecraft included in the database, as well as their respective data resolutions, are listed in Table 4, along with some clarifying remarks that are included at the bottom of the table.

Table 4: Data sources for the spacecraft included in the analysis, along with the resolutions of the data products.

Spacecraft	Parameter(s)	Instrument	Resolution
ACE	$B, \boldsymbol{B}$	MAG	16 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	SWEPAM	64 s
	$\boldsymbol{X}_\mathrm{SC}$		16 s (or 64 s)*
Wind[1]	$B, \boldsymbol{B}$	MFI	3 s
	$N_\mathrm{p}, V^\dagger, \boldsymbol{V}^\dagger, T_\mathrm{p}{}^\ddagger$,	SWE	90 s
	$\boldsymbol{X}_\mathrm{SC}$		60 s (or 90 s or 10 min)*
STEREO	$B, \boldsymbol{B}$	IMPACT/MAG	0.125 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	PLASTIC	60 s
	$\boldsymbol{X}_\mathrm{SC}$		60 min
Ulysses	$B, \boldsymbol{B}$	VHM	1 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	SWOOPS	4–8 min
	$\boldsymbol{X}_\mathrm{SC}$		60 min
Helios	$B, \boldsymbol{B}$	E3 (E2)§	6 s ($\geq$ 40.5 s)
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	E1	$\geq$ 40.5 s
	$\boldsymbol{X}_\mathrm{SC}$		$\geq$ 40.5 s
Cluster 1 and 3	$B, \boldsymbol{B}$	FGM	4 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	CIS/HIA	4 s
	$\boldsymbol{X}_\mathrm{SC}$		4 s
Cluster 4	$B, \boldsymbol{B}$	FGM	4 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	CIS/CODIF	4 s
	$\boldsymbol{X}_\mathrm{SC}$		4 s
DSCOVR[2]	$B, \boldsymbol{B}$	PlasMag	1 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	PlasMag	60 s
	$\boldsymbol{X}_\mathrm{SC}$		60 s
Voyager[2]	$B, \boldsymbol{B}$	MAG	1.92 s¶
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	PLS	12–192 s
	$\boldsymbol{X}_\mathrm{SC}$		48 s

Continued on the next page.

Table 4: Data sources for the spacecraft included in the analysis, along with the resolutions of the data products. (continued)

Spacecraft	Parameter(s)	Instrument	Resolution
OMNI[2]	$B, \boldsymbol{B}$	OMNI_HRO_1MIN	60 s
	$N_\mathrm{p}, V, \boldsymbol{V}, T_\mathrm{p}$,	OMNI_HRO_1MIN	60 s
	$\boldsymbol{X}_\mathrm{SC}$		60 s

[1] Shock candidates for Wind data with detection time before October 2015 were found using visual inspection of the data. Shocks detected after this were found with the aid of the automated shock detection algorithm IPSVM (see Section 2.1 for details).

[2] Shock candidates for the Voyager and DSCOVR spacecraft as well as OMNI data were found using the IPSVM algorithm (see Section 2.1 for details).

[*] The position data, $\boldsymbol{X}_\mathrm{SC}$, were downloaded from the source with the best resolution. In the case of ACE and Wind there were multiple sources with different resolutions, and in the case of Wind, there were also data of predicted orbit available. If one source had a data gap, other sources were used if possible.

[†] In the case of Wind, the solar wind velocity, $\boldsymbol{V}$, and the bulk speed, V, were determined as the ion velocity and the ion bulk speed.

[‡] In the case of Wind, temperature data was not available as such, but T_p was determined using the most probable thermal speed of the solar wind, V_th, from which $T_\mathrm{p} = m_\mathrm{p} V_\mathrm{th}^2/(2k_\mathrm{B})$ (this result is based on the assumptions listed in Appendix B).

[§] The magnetic field data of the Helios spacecraft were collected preferably from the instrument with the best resolution (E3). When these data were unavailable, a secondary instrument (E2) was utilised.

[¶] The Voyager magnetic field data has frequent data gaps that reduce the resolution.

Notes about Cluster events

For Cluster, only shocks detected when the spacecraft was in the solar wind and the plasma experiment CIS was in solar wind mode were included. The CIS experiment has two instruments, HIA and CODIF, of which HIA is best suited to solar wind measurements (Dandouras et al. 2009). For most of the shocks detected between 2001–2011, Cluster 1 and 3 had sufficient HIA data for the analysis. Cluster 4 was used as a backup when both Cluster 1 and 3 had data gaps, and as primary spacecraft after 2011, when the CIS instrument on board Cluster 3 was shut down and the functionality of CIS in Cluster 1 became restricted. Since the CIS/HIA instrument of Cluster 4 is switched off, plasma data was taken from the CIS/CODIF instrument, which is not ideally suited to solar wind measurements. Therefore, the parameters dependent on the plasma measurements should be considered only as crude approximations for events detected by Cluster 4.

Notes about OMNI events

The OMNI data is created from Wind observations artificially shifted to the nose of the Earth's bow shock (King and Papitashvili n.d.). The IPSVM algorithm was used to find shock candidates in the OMNI data (see Section 2.1.2). Since the algorithm was trained using the shocks directly observed by various spacecraft, and further analysis of the shocks (see Sections 2.2 and 3) was also designed for such direct observations, the use of the artificially created OMNI data in scheme presented here is problematic. Consequently, the OMNI analysis resulted in a significantly fewer number of shocks than were found from the Wind spacecraft data. The user should therefore be careful when drawing conclusions about the existence of shocks based on the OMNI results in the database, and should always check the corresponding Wind and DSCOVR data and results.

A.2 The Coordinate Systems and Reference Frames

The spacecraft's frame of reference and its coordinate system are used for: the magnetic field vector, $\boldsymbol{B}$, the velocity vector, $\boldsymbol{V}$ (also the bulk speed V), and the normal vector of the shock, $\hat{\boldsymbol{n}}$. The position of the spacecraft may have a different frame and coordinates. The different coordinate systems used for the different spacecraft's frames of reference and their position vectors are listed in Table 5.

Table 5: Coordinate systems used in the database.

Spacecraft	Coordinate system of the spacecraft's frame of reference	Coordinate system of the position vector
ACE	GSE	GSE
Wind	GSE	GSE
STEREO	RTN	HGI
Ulysses	RTN	HGI
Helios	HGIRTN	HGI
Cluster	GSE	GSE
DSCOVR	GSE	GSE
Voyager	RTN	HGI
OMNI	GSE	GSE
PSP	RTN	HGI
SOLO	RTN	HGI

B Assumptions about the solar wind plasma

B.1 General assumptions

The following general assumptions have been made when conducting the shock analysis presented in this document:

- The solar wind plasma is assumed to be an ideal gas, such that $\gamma = 5/3$ and $p_\alpha = n_\alpha k_\mathrm{B} T_\alpha$, where α is the particle species.
- The adiabatic equation of state, $\frac{\mathrm{d}}{\mathrm{d}t}(p\rho_\mathrm{m}^{-\gamma}) = 0$, is assumed to hold.
- Other ions apart from protons are neglected.
- The mass of electrons is neglected, i.e., it is assumed that $m_\mathrm{e}/m_\mathrm{p} \ll 1$.

B.2 Assumptions about the solar wind electrons

The following assumptions have been made regarding the treatment of solar wind electrons:'

- The bulk speeds of electrons and protons are assumed to be equal, $V_\mathrm{e} = V_\mathrm{p}$.
- The temperature of electrons is assumed to depend only on the radial distance from the Sun according to

$$T_\mathrm{e} = T_\mathrm{e}(R_\mathrm{au}) = 146\,277 \times R_\mathrm{au}^{-0.664}\,\mathrm{K}$$

  where R_au is the radial distance from the Sun given in astronomical units (au).

The assumption of equal bulk speed for electrons and protons ($V_\mathrm{e} = V_\mathrm{p} = V$) results in the additivity of pressures (Paschmann et al. 1998). In the isotropic case with the ideal gas assumption

$$p = p_\mathrm{p} + p_\mathrm{e} = k_\mathrm{B} N_\mathrm{p}(T_\mathrm{p} + T_\mathrm{e}).$$

Electron temperature measurements were not generally available and the constant value at a given radial distance from the Sun given by the expression above, was used instead. The variations of T_e with the radial heliospheric distance from the Sun (required for Helios, Ulysses, and Voyager) are typically expressed as a power law $T_\mathrm{e}(R_\mathrm{au}) \sim R_\mathrm{au}^k$ (Maksimovic et al. 2000). The index k, however, varies considerably between different studies. A power law fit, similar to Cranmer et al. (2009), was made. Electron temperature measurements made by the Helios and Ulysses spacecraft at radial distances ranging from 0.29 au to 5.41 au were used. Suitable Helios data points were taken from Pilipp et al. (1990), and all available Ulysses data were collected from CDAWeb. The power law was fitted to these data points so that both spacecraft had equal weight in the fit. This gives the electron temperature at 1 au as being 146 277 K, which is, within error bounds, consistent with Newbury et al. (1998). This constant value was used for the ACE, Wind, STEREO, Cluster, and DSCOVR spacecraft (thus neglecting their small deviations from the radial distance of 1 au) as well as for the analysis of OMNI data.

NOTE: The analysis results of shocks observed prior to August 2015 are based on a rigid electron-to-proton temperature ratio of $T_\mathrm{e} = 2T_\mathrm{p}$.