

# Solar wind prediction for the Solar Probe Plus orbit

## Near-Sun extrapolations derived from an empirical solar wind model based on Helios and OMNI observations

M. S. Venzmer and V. Bothmer

University of Goettingen, Institute for Astrophysics, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

First draft 9 August 2016; received date; accepted date

### ABSTRACT

**Context.** In view of the planned near-Sun spacecraft mission Solar Probe Plus (SPP) the solar wind environment for its mission duration (2018–2025) and down to its intended closest perihelion (9.86 solar radii) is extrapolated using in situ data. This yet uncharted region is of special interest because the critical Alfvénic and sonic surfaces are believed to lie within and thus still part of the solar wind acceleration as well.

**Aims.** We present an empirical solar wind model for the inner heliosphere which is derived from Helios and OMNI in situ data. The space probes Helios 1 and Helios 2 flew in the 1970s and observed solar wind in the ecliptic within heliocentric distances of 0.29–0.98 au. The OMNI data set consists of multi-spacecraft intercalibrated in situ data from 1 au. This model is used together with sunspot number predictions to estimate the frequency distributions of major solar wind parameters SPP will encounter on its mission.

**Methods.** The model covers the solar wind's magnetic field strength and its plasma parameters proton density, velocity and temperature. Their individual frequency distributions are represented with lognormal functions. In addition, we also consider the velocity distribution's bi-componental shape, consisting of a slower and a faster part. The model accounts for solar activity and for solar distance dependency by shifting of these lognormal distributions. We compile functional relations to solar activity by correlating and fitting the frequency distributions with the sunspot number (SSN), using almost five solar cycles of OMNI data. Further, based on the combined data set from both Helios probes, the parameters' frequency distributions are fitted over solar distance to obtain exponential dependencies. Combining both the solar cycle and the solar distance relations, we get a simple dynamical solar wind model for the inner heliosphere, confined to the ecliptic.

**Results.** The inclusion of SSN predictions and the extrapolation to the SPP perihelion region enables us to estimate the solar wind environment at SPP's planned orbital positions during its mission time.

**Key words.** solar wind – sun: heliosphere – sun: corona

### 1. Introduction

With his theoretical solar wind model Parker (1958) laid the foundations...

Several space missions flew away from Earth's orbit to measure in situ solar wind at different heliocentric distances (e.g. the probes Helios, Voyager, Ulysses...). Voyager 1 is the farthest probe out of the solar system, now covering more than 137 au (cite?). Ulysses was the single probe orbiting the Sun outside the ecliptic and retrieving solar wind measurements from the poles of the Sun. So far, Helios 2 made the nearest in situ solar wind measurements with 0.29 au, followed closely by Helios 1 with 0.31 au. The new mission Solar Probe Plus (SPP), whose planned launch date is in mid 2018, aims to reach further down into the Sun's atmosphere.

The main part of the SPP mission's goals is to gain insight into the coronal heating and solar wind acceleration processes. As these processes occur in the near-Sun region of the corona, SPP will have its closest perihelion at 9.86 solar radii (0.0459 au) (Fox et al. 2015). It will be the first spacecraft flying and measuring in situ solar wind this close to the Sun. Its planned observations range from the perihelion up to 0.25 au, leaving the outer part for data transfer, as Fox et al. (2015) state

in their description of SPP's mission design.

white light imager WISPR... (Vourlidas et al. 2016)

There are many studies analyzing the solar wind dependence on solar activity and heliocentric distance with Helios data (Schwenn (1983); Bougeret et al. (1984); Schwenn (1990), more cites...). They examine average values, additionally distinguishing between slow and fast wind. In this paper we instead look at the variations of the whole frequency distributions, consider solar activity and extrapolate these to the near-Sun region.

### 2. Solar wind environment

The goal of this paper is to estimate the solar wind environment SPP will probe during its orbits. We regard the solar wind primarily as a proton plasma, because protons make up most of it. The average helium abundance is only about 4.5 % and in slow wind at solar cycle minimum even less than 2 % (Feldman et al. 1978; Schwenn 1983; Kasper et al. 2012). Neglecting heavier ions, the electron and mass density can be derived accordingly.

Generally, the characteristic behavior of a plasma is determined by its *density*, *temperature* and *magnetic field strength*. Furthermore, the bulk *velocity* is the parameter which makes the plasma a 'wind'. For our study we define the solar wind environment as the four major solar wind parameters. Quantities like flux densities, mass flux and plasma beta can directly be derived from those four parameters.

In addition, the velocity is the defining parameter of the two types of solar wind. Solar wind mostly consists of and switches between slow and fast wind, whose ratio highly depends on solar activity. These types have different characteristics and their interactions lead to phenomena such as interaction regions. Additionally there also are coronal mass ejections (CMEs) embedded, whose occurrence rate follows solar activity too. Their share varies between almost zero in cycle minimum up to a half in maximum (Richardson & Cane 2012). This study averages over these internal solar wind structures.

One cannot know which specific solar wind type or structure SPP will encounter at a given point in time, so we extrapolate the parameters' probability distributions from existing solar wind measurements.

Our approach is to get an analytical representation of the frequency distributions' shapes, their solar activity dependence and their solar distance scaling. We get the parameters' frequency distributions and solar activity dependence from near-Earth solar wind and sunspot number (SSN) time series with a duration of almost five solar cycles and their distance dependency from solar wind measurements of more than half a solar cycle, covering more than two third of the distance to the Sun (0.29–0.98 au).

From the combination of the obtained frequency distributions, SSN dependence functions and solar distance dependence functions we build a general model representing the solar activity and distance behavior of all four solar wind parameter frequencies.

This general model is then fed with a SSN prediction and extrapolated to SPP's planned orbital positions.

### 3. Frequency distributions

This section looks at the solar wind parameters' frequency distributions which we extract from the in situ OMNI data set. We determine adequate fit function types and evaluate how suitable they are to represent the frequencies' shapes.

#### 3.1. OMNI frequency data

The solar wind parameters are highly variable, due to short-term variations from structures like slow and fast wind streams, interaction regions and CMEs, whose rate and properties depend on solar activity. Hence, for deriving general frequency distributions of the solar wind parameters, averaging over long-term solar wind variations is needed. This requires a distance-independent data set covering multiple solar cycles. The abundance of near-Earth hourly OMNI data is well suited for this task, because it spans almost five solar cycles.

This OMNI 2 data set (King & Papitashvili 2005) combines solar wind plasma and magnetic field data and for this study not important energetic proton fluxes, geomagnetic and solar indices. Because it covers decades, the near-Earth solar wind data is composed of intercalibrated multi-spacecraft data which is time-shifted to the nose of the Earth's bow shock.

The data is obtained from the OMNIWeb interface<sup>1</sup> at NASA's Space Physics Data Facility (SPDF), Goddard Space Flight Center (GSFC). The hourly data of the whole time range up to 2016-12-31 is the basis for the frequency fits. The data starts in 1963-11-27, but the temperature data not before 1965-07-26. The data coverage of the parameters is between 67–74 %, which adds up to about 36–40 years in total. This plethora of data is well suited for our task.

We specify bin sizes considering the individual maximal parameter ranges and the OMNI data precision. Especially for the density and temperature we choose their bin sizes such small that their distributions' peaks can be resolved (the peaks are at their lower end). We set the individual bin sizes to 0.5 nT for the magnetic field strength, 10 km s<sup>-1</sup> for the velocity, 1 cm<sup>-3</sup> for the density and 10 000 K for the temperature.

Next, we look for a suitable fit function for the resulting histogram shape of the solar wind parameters' frequency distributions.

#### 3.2. Lognormal fitting

Obviously all possible values for the four parameters are positive. This hints to the supposition that they are lognormally distributed, as many positive natural quantities conform to lognormal distributions. Its probability density function is described by a lognormal function. Therefore we use a lognormal function as the fit function in the process of the least squares regression fitting. The lognormal function

$$W(x) = \frac{1}{\sigma \sqrt{2\pi}x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad (1)$$

depends on the location  $\mu$  and the shape parameter  $\sigma$ . Changes in  $\mu$  affect both the horizontal and vertical scaling of the function whereas  $\sigma$  influences its shape. The distribution's median  $x_{\text{med}}$  and mean  $x_{\text{avg}}$  (average) positions are easier to interpret and can directly be calculated from  $\mu$  and  $\sigma$ :

$$x_{\text{med}} = \exp(\mu) \quad \iff \quad \mu = \ln(x_{\text{med}}), \quad (2)$$

$$x_{\text{avg}} = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad \iff \quad \sigma = \sqrt{2 \ln\left(\frac{x_{\text{avg}}}{x_{\text{med}}}\right)}. \quad (3)$$

It is apparent that the mean is always larger than the median. Replacing the variables  $\mu$  and  $\sigma$  with these relations, the lognormal function (1) becomes

$$W(x) = \frac{1}{2 \sqrt{\pi \ln\left(\frac{x_{\text{avg}}}{x_{\text{med}}}\right)} x} \exp\left(-\frac{\ln^2\left(\frac{x}{x_{\text{med}}}\right)}{4 \ln\left(\frac{x_{\text{avg}}}{x_{\text{med}}}\right)}\right). \quad (4)$$

The values of  $x_{\text{med}}$  and  $x_{\text{avg}}$  obtained from fitting the solar wind frequency distributions are listed in Table 1.

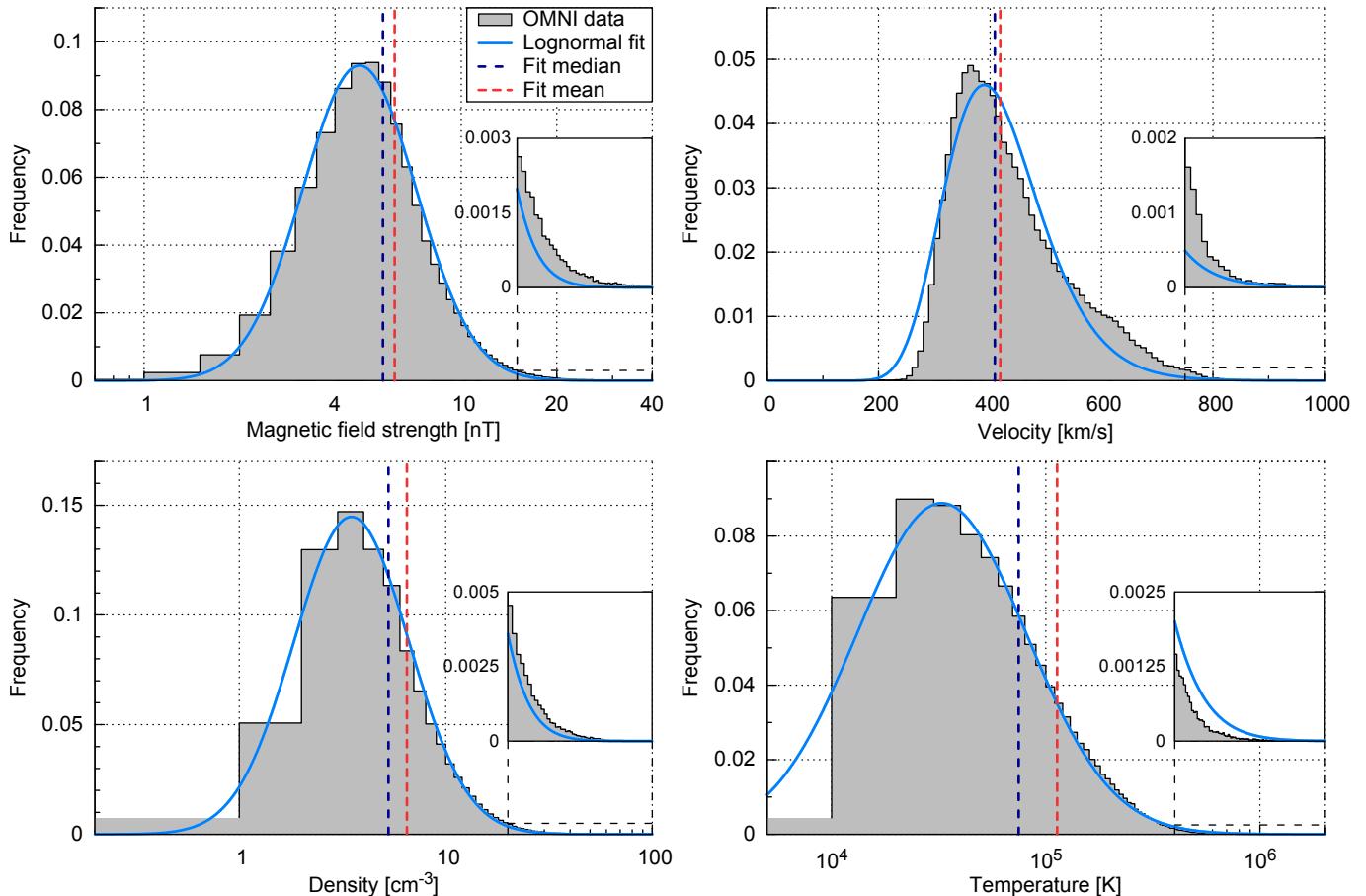
From visual inspection, the resulting curves match well with the shape of the magnetic field strength, density and temperature distributions (Fig. 1). However, for the velocity the fit function seems insufficient in describing its more complex shape, especially at its peak position and the faster end of the distribution. Its sum of absolute residuals (SAR) between data and fit is almost three times larger than those from the other parameters (Table 1). They can be compared, because the area of probability density functions is unity.

<sup>1</sup> <http://omniweb.gsfc.nasa.gov/>

**Table 1.** Resulting fit coefficients from the fitting of the lognormal function (4) to the shape of the solar wind parameters' frequency distributions at 1 au. For the velocity also the fit parameters from the double lognormal function (5) are given, as well as the median and mean values of the resulting velocity fit. The mean absolute errors and sums of absolute residuals are shown as well. The values in brackets are the estimated standard deviation of each fit parameter.

| Parameter      | Median<br>$x_{\text{med}}^a$             | Mean<br>$x_{\text{avg}}^a$ | Balance<br>$c$ | MAE                   | SAR<br>[%] |
|----------------|--|----------------------------|----------------|-----------------------|------------|
| Magnetic field | 5.661(16)                                | 6.164(18)                  | –              | $5.51 \times 10^{-4}$ | 6.83       |
| Velocity       | 4.085(19)                                | 4.183(20)                  | –              | $1.80 \times 10^{-3}$ | 18.69      |
| Density        | 5.276(24)                                | 6.484(34)                  | –              | $5.49 \times 10^{-4}$ | 6.48       |
| Temperature    | 7.470(17)                                | 11.301(32)                 | –              | $8.71 \times 10^{-5}$ | 5.78       |
| Velocity       | $W_1$<br>4.89(14)                        | 5.00(14)                   | 0.504(62)      | –                     | –          |
|                | $W_2$<br>5.9089(53)                      | 3.72(20)                   | –              | –                     | –          |
|                | $W_{\text{II}}$<br>4.12(14) <sup>b</sup> | 4.38(14) <sup>b</sup>      | –              | $3.98 \times 10^{-4}$ | 4.20       |

Notes. <sup>(a)</sup> Values in their respective units nT,  $10^2 \text{ km s}^{-1}$ ,  $\text{cm}^{-3}$  and  $10^4 \text{ K}$ . <sup>(b)</sup> Error estimates derived from the individual fit part errors.



**Fig. 1.** Frequency distributions of the four solar wind parameters and their lognormal fits. The histograms have bins of 0.5 nT, 10 km/s, 1  $\text{cm}^{-3}$  and 10 000 K and are based on the hourly OMNI data set. The fit's median and mean values are indicated as well. The insets only have zoomed-in frequency axes, their x-axes stay the same.

To reach a better fit result for the velocity we change the fit function. We do not want to abandon the well-founded application of the lognormal function. However, it is reasonable to assume that the velocity distribution is composed of at least two branches (slow and fast). Therefore a compositional approach promises better fit results, which is why we combine two lognormal functions (4), bearing the disadvantage of more fit variables:

$$W_{\text{II}}(x) = c \cdot W_1(x) + (1 - c) \cdot W_2(x). \quad (5)$$

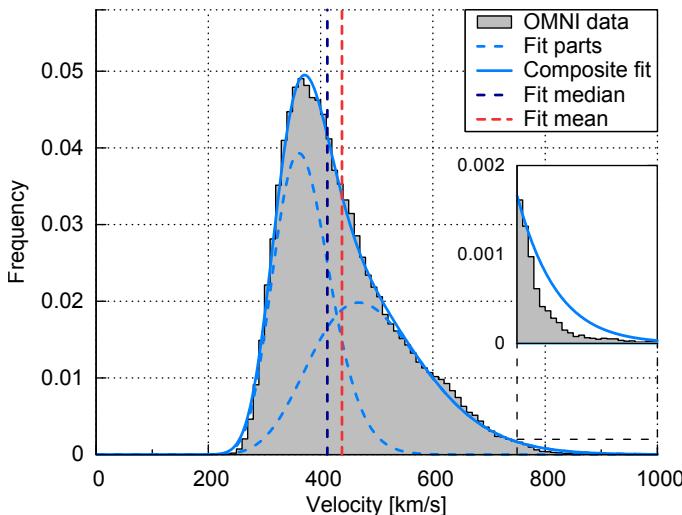
The balancing parameter  $c$  ensures that the resulting function remains normalized as it represents a probability distribution.

The fitting of  $W_{\text{II}}(x)$  to the velocity's frequency distribution gives the values of the now five fit parameters ( $c$ ,  $x_{\text{med},1}$ ,  $x_{\text{avg},1}$ ,  $x_{\text{med},2}$  and  $x_{\text{avg},2}$ ), which are also listed in Table 1 together with the median and mean of the composed distribution, which can be derived via solving

$$\int W_{\text{II}}(x) dx = 0 \quad \text{and} \quad \int x W_{\text{II}}(x) dx = 0 \quad (6)$$

respectively.

As anticipated, this more complex fit function is more accurate in describing the velocity's frequency distribution (see Fig. 2).



**Fig. 2.** Plot of the velocity's frequency distribution (same as in Fig. 1) and its compositional lognormal fit. The fit's median and mean values and its two fit parts are indicated as well. The inset only has a zoomed-in frequency axis, its x-axis stays the same.

That is why we keep using the double lognormal ansatz for the velocity frequency fits in the following sections. In this static model the slow and fast part contribute almost equally ( $c \approx 0.5$ ), which of course is only valid for this kind of long-term average. At different times in a solar cycle their contributions vary strongly.

For the bulk of the solar wind these static lognormal functions describe the parameters' distributions well. This is different for the extreme values (which may also stem from CMEs). The simple lognormal fit models of the magnetic field strength, the velocity and the density underestimate their frequency at the high value tails, whereas the temperature's tail is overestimated (see insets of Fig. 1). The velocity's compositional lognormal fit only slightly overestimates its tail (inset in Fig. 2).

Short-term variations in the solar wind cannot be predicted, but their occurrence rate can. It depends on solar activity, which changes cyclically and thus can be forecasted to a certain degree—at least within a solar cycle.

## 4. Solar activity variations

This section aims to relate changes in the four solar wind parameters to general solar activity. For this we examine their correlations to the yearly sunspot number and determine the lag times with the highest coefficients. Next, we fit lognormal functions to the frequency distributions like before, but implement linear relations to the yearly SSN to shift the distributions. Only for the velocity the approach is different in that its two components are kept fixed and instead their balance is modified with changing SSN.

### 4.1. SSN data

Solar activity is commonly measured via the sunspot number. We want to correlate OMNI in situ measurements with the SSN, but OMNI data are from Earth orbit, causing variations in solar

latitude and distance. To dodge these variations we use yearly OMNI and SSN data.

The international sunspot number (1963–2016) is retrieved from the online catalogue<sup>2</sup> at the World Data Center – Sunspot Index and Long-term Solar Observations (WDC-SILSO), Solar Influences Data Analysis Center (SIDC), Royal Observatory of Belgium (ROB).

seasonal influence contained within yearly frequency distributions

SIDC and SWPC provide limited SSN forecast (a few months)

### 4.2. SSN correlation

Our current interest lies in the correlation of the SSN to the solar wind median value, because it defines the position of the lognormal function. The yearly OMNI parameter medians and the yearly SSN are plotted in Fig. 3.

$$\text{solar activity} = \text{SSN} \neq \text{solar cycle}$$

The solar wind velocity (and its close friends N & T) is known to depend on the state of the solar cycle (cite?), which is why it follows the SSN only indirectly (with time lag). Thus we derive the correlation coefficients for different time lags between solar wind parameters and SSN (see Fig. 3).

The highest correlation coefficient for the magnetic field strength is 0.728. It is without lag time and the highest of all solar wind parameters, which is expected because the SSN is directly proportional to the solar surface magnetic flux (cite?). Velocity and temperature have a lag time of 3 years with their maximal correlation coefficients (0.453 and 0.540). The density even 6 years (0.468).

cause: direct/indirect correlation  
cite literature to lag time...

yearly SSN sw model to sw data fits -> SSN dependency (figs)

### 4.3. SSN fitting

To be able to shift the distributions with SSN, add fit a linear SSN dependency to the median

$$x_{\text{med}}(\text{ssn}) = a_{\text{med}} \cdot \text{ssn} + b_{\text{med}}, \quad (7)$$

using a factor to the SSN  $a_{\text{med}}$  with a baseline  $b_{\text{med}}$ . We relate the mean with a scaling factor to the median to keep the new SSN dependency:

$$x_{\text{avg}}(\text{ssn}) = (1 + a_{\text{avg}}) \cdot x_{\text{med}}(\text{ssn}). \quad (8)$$

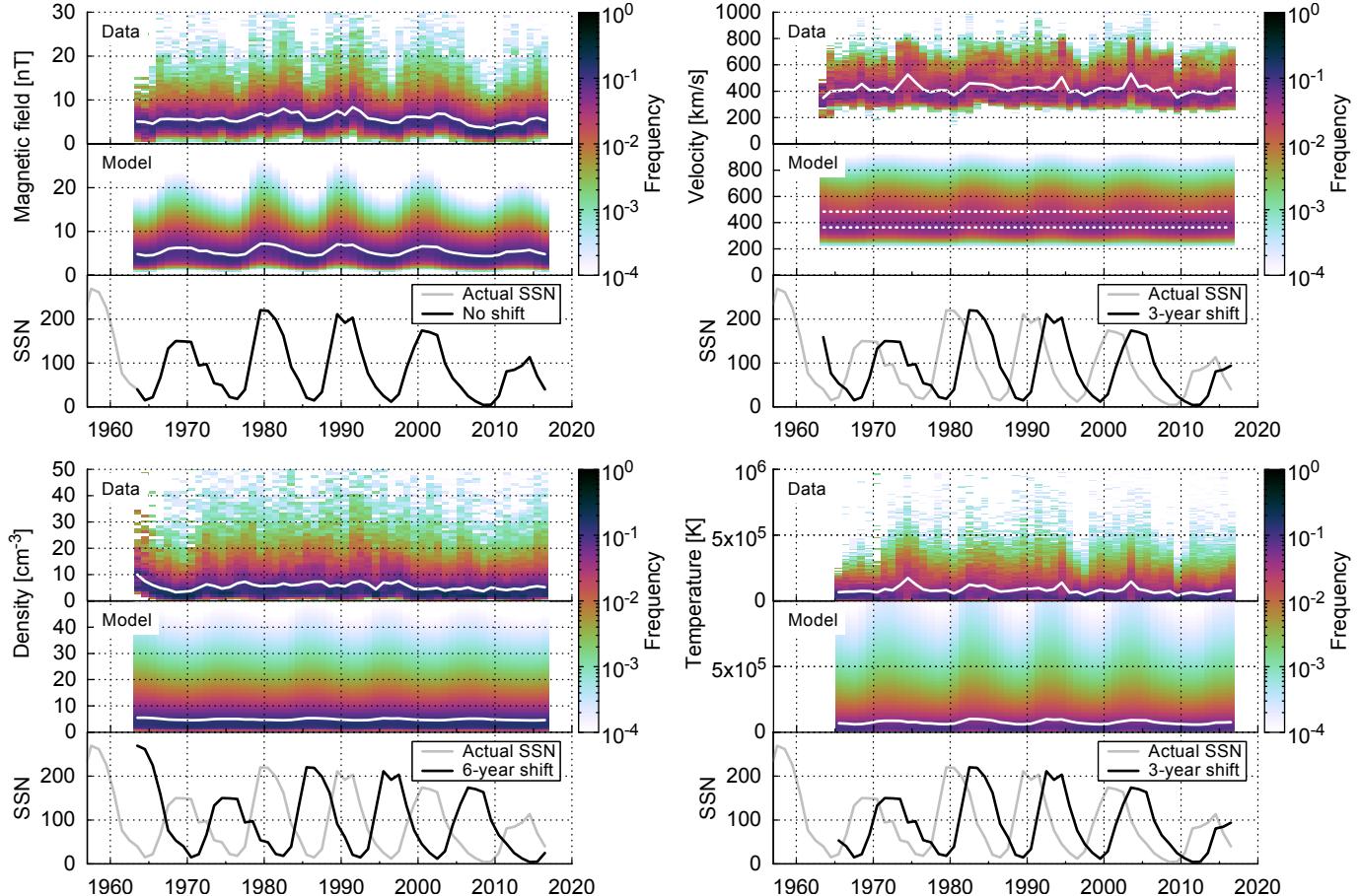
With the implementation of these relations into the lognormal function (4), the new dynamic fit function  $W'(x, \text{ssn})$  is then fitted to the yearly data. The three resulting fit coefficients ( $a_{\text{med}}$ ,  $b_{\text{med}}$  and  $a_{\text{avg}}$ ) are presented in Table 2.

<sup>2</sup> <http://www.sidc.be/silso/>

**Table 2.** Resulting fit coefficients from the data fitting with SSN. For the velocity the fit parameters from the double lognormal fit and their balancing function are given. The values in brackets are the estimated standard deviation of each fit parameter.

| Parameter      | Median                        |                             | Scaling factor $a_{\text{avg}}$ | Mean                       |                             | Balance          |                |
|----------------|-------------------------------|-----------------------------|---------------------------------|----------------------------|-----------------------------|------------------|----------------|
|                | SSN factor $a_{\text{med}}^a$ | Baseline $b_{\text{med}}^a$ |                                 | SSN factor $c_a$           | Baseline $c_b$              | SSN factor $c_b$ | Baseline $c_b$ |
| Magnetic field | $1.309(19) \times 10^{-2}$    | 4.285(17)                   | $8.786(78) \times 10^{-2}$      | —                          | —                           | —                | —              |
| Density        | $3.81(25) \times 10^{-3}$     | 4.495(26)                   | $3.050(27) \times 10^{-1}$      | —                          | —                           | —                | —              |
| Temperature    | $1.974(26) \times 10^{-2}$    | 5.729(19)                   | $6.541(28) \times 10^{-1}$      | —                          | —                           | —                | —              |
| Velocity       | $W_1$                         | —                           | $3.633(12)$                     | $1.008(37) \times 10^{-2}$ | $-1.799(95) \times 10^{-3}$ | $0.638(32)$      |                |
|                | $W_2$                         | —                           | 4.831(81)                       | $2.31(20) \times 10^{-2}$  |                             |                  |                |

Notes. <sup>(a)</sup> Values in their respective units nT,  $10^2 \text{ km s}^{-1}$ ,  $\text{cm}^{-3}$  and  $10^4 \text{ K}$ .



**Fig. 4.** Solar wind parameter data frequencies, lognormal fit models with their median values (white) and the corresponding yearly SSN (grey) over the OMNI time period 1963–2016. The for the models shifted SSN is indicated by a black line. The constant medians of both velocity's lognormal parts are marked with white dots.

In comparison with the data the fit models seem to match with the general trends (Fig. 4).

Again, the velocity gets a special treatment with the double lognormal distribution (5). It is known that slow and fast solar wind stream occurrence rates follow the solar cycle, but their magnitudes stay fairly stable (cite?). This time we keep the two velocity components' positions constant and vary instead their balance with the SSN:

$$c(\text{ssn}) = c_a \cdot \text{ssn} + c_b . \quad (9)$$

The fit result (see Table 2) is a model in which three years after solar cycle minimum (SSN of zero) the slow solar wind has a share of almost two-thirds and decreases further with increasing

SSN (see Fig. 5).

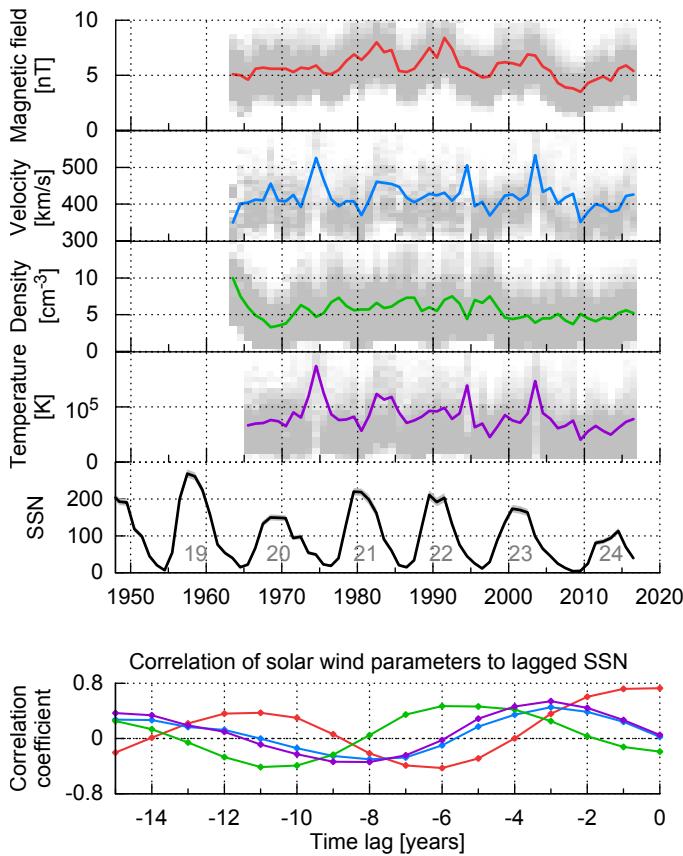
...compare with data

=> there is no specific velocity threshold between slow and fast solar wind types, the velocity ranges of both types overlap.

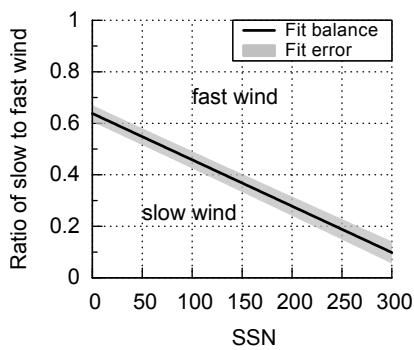
In comparison with the data the fit model roughly matches with the general trend (see Fig. 4).

density modulation with SSN (Schwenn 1983) p. 499

density anticorrelation with SSN (Bougeret et al. 1984) p. 406



**Fig. 3.** Plot of the solar wind parameter yearly medians from OMNI data and the yearly SSN from the SILSO World Data Center (1963–2016) with cycle number (top). Their correlation coefficients with the yearly SSN for time lags back to -15 years (bottom).



**Fig. 5.** Yearly ratio between slow to fast solar wind (threshold  $400 \text{ km s}^{-1}$  over SSN (+)). Fit model ratio depending on the by 3 years lagged SSN (black line). The relation results from the velocity's double lognormal fit.

## 5. Solar distance dependency

In this section we use Helios data to obtain exponential fit functions for the heliocentric distance dependency and evaluate the fits' extrapolation behavior in direction to the Sun. To fit the bulk solar wind distributions' distance dependency we use the frequency fitting method from Sect. 3 on distance-binned Helios data. This results in models comprising of exponentially with distance shifted lognormal functions.

### 5.1. Helios distance data

The Helios probes were the only spacecraft measuring in situ solar wind over large solar distance ranges in the inner heliosphere. We use the combined data from both Helios 1 and Helios 2 probes. Helios 1's (Helios 2's) highly elliptical orbit in the ecliptic covered a solar distance range of 0.31–0.98 au (0.29–0.98 au). Launched during solar cycle minimum, the data of both probes cover the rise to the maximum of cycle 21 (~6.5 years at varying distances).

Helios 1's (Helios 2's) merged hourly data set from the magnetometer and plasma instruments (Rosenbauer et al. 1977) includes ~12.5 orbits (~8 orbits) in the time range 1974-12-10 to 1981-06-14 (1976-01-01 to 1980-03-04). The Helios data was retrieved from the Coordinated Data Analysis Web (CDAWeb) interface at NASA's GSFC/SPDF<sup>3</sup>.

The Helios 1 (Helios 2) magnetometer data coverage is about 43 % (54 %) and amounts to 2.8 years (2.3 years) in total. The plasma data coverage is 76 % (92 %) and amounts to 5.0 years (3.9 years) in total. Thus, the Helios data cover only fractions of a solar cycle and cannot be used for deriving representative time-independent solar wind multi-cycle conditions like the OMNI data can.

Using this data, we also have to keep in mind that its time coverage is unequally distributed over the solar cycle. Dividing the data by the transition from cycle minimum to maximum (mid 1977) and considering the data gap distributions, the Helios data covers about 68 % during cycle minimum whereas during maximum only 38 %.

For calculating the median and mean values at different solar distances the data is binned into 0.01 au bins, which is also the native precision in this data set.

### 5.2. Exponential fitting

An exponential distance behavior is expected from all four parameters (cites?). Therefore we use the exponential function

$$x(r) = a r^b \quad (10)$$

with the solar distance  $r$  for the regression fit of the median and mean. The fits are weighted by data counts per bin. With  $r$  in astronomical units we get the fit coefficients ( $a_{\text{med}}$ ,  $a_{\text{avg}}$ ,  $b_{\text{med}}$  and  $b_{\text{avg}}$ ) as given in Table 3.

The next step is to fit the bulk of the solar wind parameters with lognormal functions. At all considered solar distances the mean of the three plasma parameters is larger than their median (Fig. 6). However, the magnetic field strength's mean crosses the median at 0.339 au (Table 3) and is lower at smaller distances. At the crossing point and below the magnetic field strength can therefore not be described anymore with a lognormal function. For an extrapolation to the SPP perihelion the same happens for the temperature at 0.082 au.

The large velocity's crossing distance and its large error indicate that the median's and mean's distance behavior can be kept identical and so the frequency's shape distance independent.

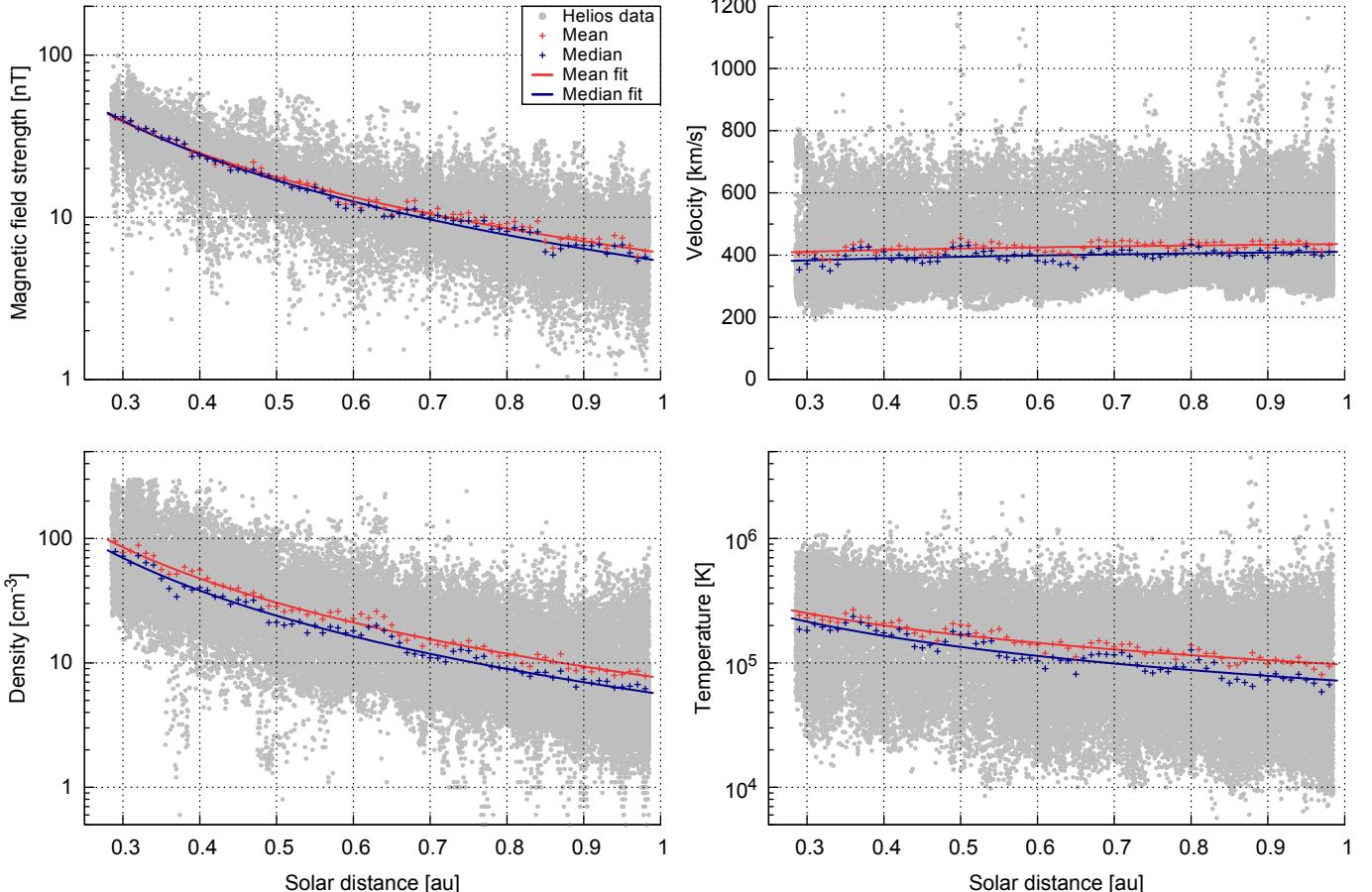
Those crossings limit the possible extrapolation distances with lognormal functions. To circumvent these limitations for all four solar wind parameters we set the exponents  $b_{\text{med}}$  and  $b_{\text{avg}}$  to be identical, avoiding crossing of median and mean. Then

<sup>3</sup> <http://spdf.gsfc.nasa.gov/>

**Table 3.** These are the fit coefficients for the median and mean solar distance dependencies of the four parameters from the combined Helios data set. The errors in brackets are the estimated standard deviations of each fit parameter. The crossing distance is the point where median and mean intersect.

| Parameter      | Median             |                  | Mean               |                  | Crossing distance<br>[au] |
|----------------|--------------------|------------------|--------------------|------------------|---------------------------|
|                | $a_{\text{med}}^a$ | $b_{\text{med}}$ | $a_{\text{avg}}^a$ | $b_{\text{avg}}$ |                           |
| Magnetic field | 5.377(92)          | -1.655(17)       | 6.05(10)           | -1.546(18)       | 0.339(11)                 |
| Velocity       | 4.107(28)          | 0.058(13)        | 4.356(24)          | 0.049(10)        | $0.7(83) \times 10^3$     |
| Density        | 5.61(27)           | -2.093(46)       | 7.57(30)           | -2.010(38)       | 0.027(73)                 |
| Temperature    | 7.14(23)           | -0.913(39)       | 9.67(21)           | -0.792(28)       | 0.082(85)                 |

**Notes.** <sup>(a)</sup> Values in their respective units nT,  $10^2 \text{ km s}^{-1}$ ,  $\text{cm}^{-3}$  and  $10^4 \text{ K}$ .



**Fig. 6.** Helios hourly data plots of the four solar wind parameters over solar distance. The mean and median per 0.01 au data bin and their fit curves are plotted as well. The Helios data has a native distance resolution of 0.01 au. To make the abundance visible in these plots, we added a random distance value of up to  $\pm 0.005$  au.

the distribution's width is constant and scales only exponentially with solar distance. Applying this approximation, we have to accept larger model errors, especially for the magnetic field strength. It also limits the extrapolation accuracy.

### 5.3. Exponential lognormal fitting

To retrieve the frequency distributions for every 0.01 au solar distance bin, we choose the same solar wind parameter binning as with the OMNI data (Sect. 3.1).

As mentioned before, we set the exponents of median and average to be identical. Implementing the exponential distance dependency (10) into the lognormal function (4), we get three fit parameters ( $a_{\text{med}}$ ,  $a_{\text{avg}}$  and the common exponent  $b$ ).

Naturally, we use the double lognormal function (5) for the velocity distribution fit, resulting in  $W''_{\text{H}}(x, r)$ . The additional fit parameters are the balancing parameter  $c$  and from the second lognormal part  $a_{\text{med},2}$  and  $a_{\text{avg},2}$ . The resulting fit coefficients for the four solar wind parameters are presented in Table 4.

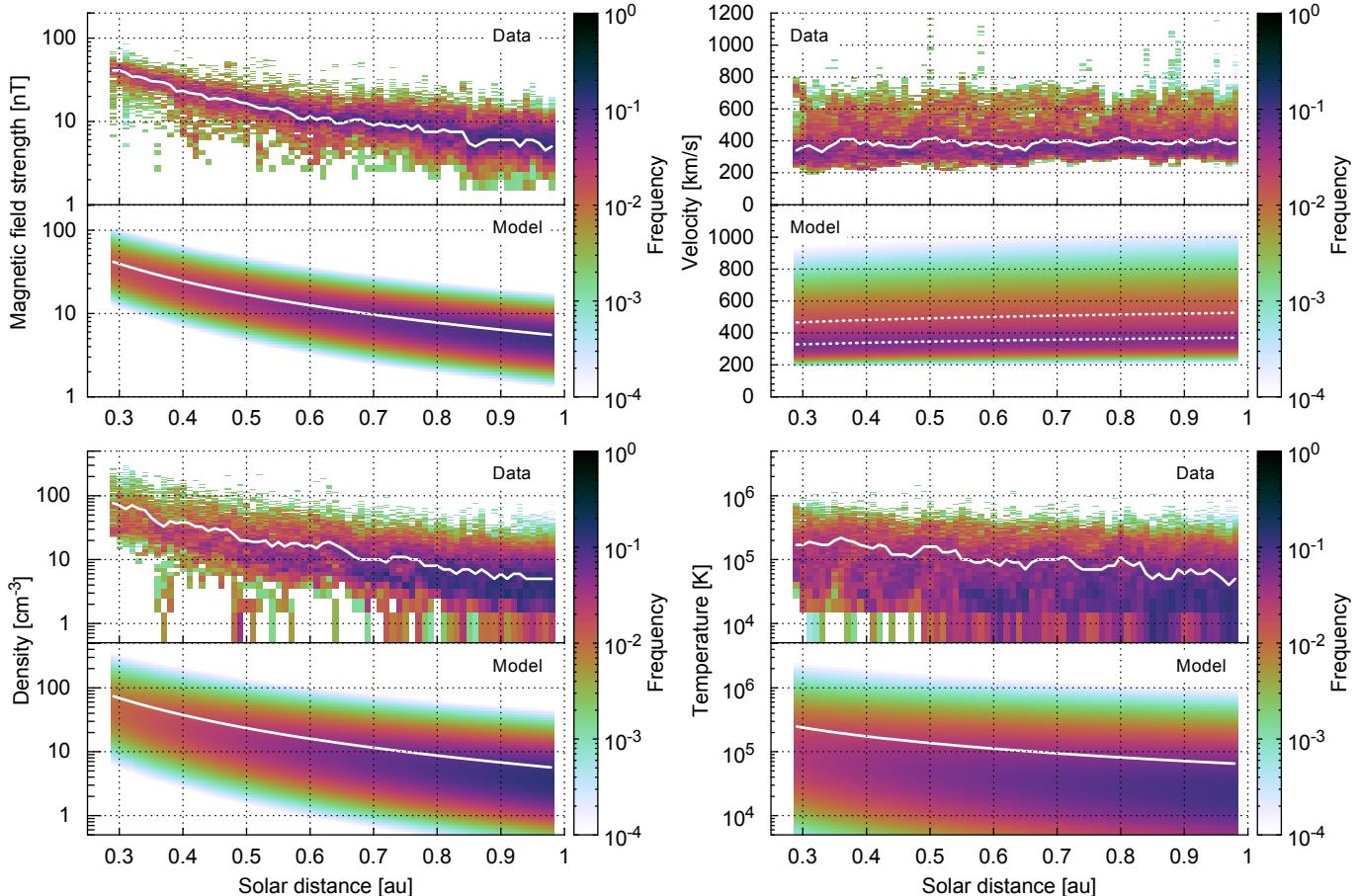
With  $c = 0.557$  the velocity balancing parameter is of an expected value similar to that obtained from the Helios time period (the mean SSN during the Helios period was 59, this corresponds to a  $c(59) = 0.53$ , see Fig. 5).

The fit models seem to resemble the data quite well (Fig. 7). The magnetic field strength frequency is more focused (around 40 nT) at the lower distance boundary than the model's is. This is expected because of our fixed distance independent shape

**Table 4.** These are the resulting fit coefficients with the single lognormal exponential function, respectively double lognormal for the velocity. The errors in brackets are the estimated standard deviations of each fit parameter.

| Parameter      | $c$               | $a_{\text{med}}^a$ | $a_{\text{avg}}^a$    | $b$                   |
|----------------|-------------------|--------------------|-----------------------|-----------------------|
| Magnetic field | –                 | 5.358(25)          | 5.705(28)             | -1.662(11)            |
| Density        | –                 | 5.424(33)          | 6.845(47)             | -2.114(20)            |
| Temperature    | –                 | 6.357(64)          | 10.72(14)             | -1.100(20)            |
| Velocity       | $W_1''$           | 0.557(45)          | 3.707(13)             | 3.748(16)             |
|                | $W_2''$           |                    | 5.26(13)              | 5.42(11)              |
|                | $W_{\text{II}}''$ | –                  | 4.35(13) <sup>b</sup> | 4.77(11) <sup>b</sup> |

**Notes.** <sup>(a)</sup> Values in their respective units nT,  $10^2 \text{ km s}^{-1}$ ,  $\text{cm}^{-3}$  and  $10^4 \text{ K}$ . <sup>(b)</sup> Velocity median and mean 1 au values for the resulting function. Error estimates derived from the individual fit part errors.



**Fig. 7.** Solar wind parameter's frequency distributions over solar distance. Plotted are the binned Helios data and the exponential lognormal fit model (double lognormal for the velocity) with their median values (white).

approximation.

model errors, maximum errors

The extreme values of the velocity and temperature seem to be too high (adjust color scale...)

L-> high value plots... (refer to histogram zoom in Fig. 1)

This model represents the Helios time frame around the rise of solar cycle 21.

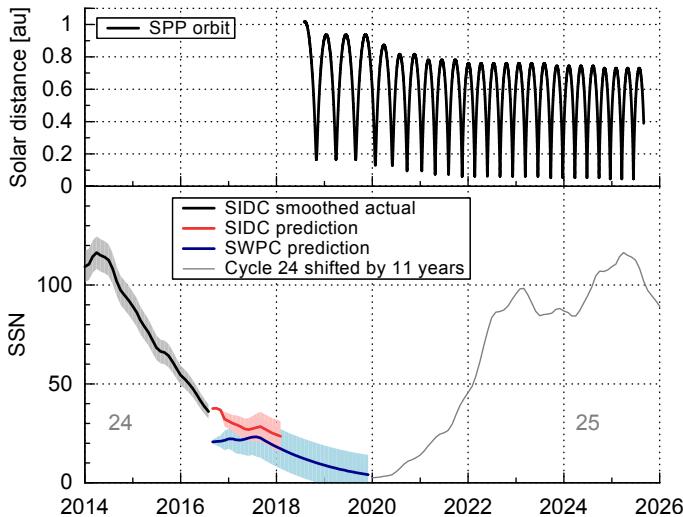
study time/cycle variations

## 6. General solar wind model

Finally, we combine the obtained solar activity and distance dependencies for shifting the frequency distributions. The result will be an empirical solar wind model for the inner heliosphere.

Under the assumption that the exponential falloff laws do not change with time/solar activity, they can be used in general (indications for that: fig. + theory + studies/papers...). assuming the distance scaling laws are time independent, which we deduced...? which is an approximation.

We combine the fit coefficients of the median relation for solar activity dependence (7) with the ones from the exponential



**Fig. 8.** SPP's solar distance during its mission time (top). Consecutive Venus flybys bring its perihelia nearer to the Sun. Actual and predicted SSN (bottom), i.e. 13-month smoothed monthly actual SSN and SIDC prediction, SWPC prediction and simply shifted SSN from previous cycle 24.

distance dependency (10):

$$x_{\text{med}}(ssn, r) = (a_{\text{med}} \cdot ssn + b_{\text{med}}) \cdot r^b. \quad (11)$$

the combined model function  $W''(x, ssn, r)$

...and double lognormal function for the velocity

give model errors sizes/uncertainties (sigmas);  $W''_{\text{err}}(x, ssn, r)$

## 7. Model extrapolation to SPP orbital time and position

To estimate the solar wind environment at SPP's planned orbital positions during its mission time, SSN predictions and extrapolations to the SPP perihelion region are included into the general solar wind model.

SPP orbit description, Venus flybys, ecliptic, (Fox et al. 2015), see Fig. 8

choosing SSN prediction source as input (SIDC + SWPC + previous cycle), (see Fig. 8)

plot of SPP's estimated solar wind environment from predicted SSN; zoom into first and nearest perihelion (Fig. 9)

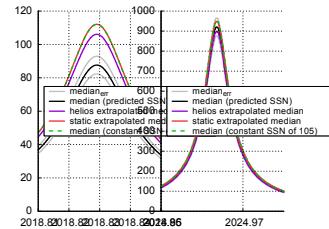
### 7.1. Near-Sun extrapolation behavior/effects

The extrapolation distance is only about one third of the model range, but as the parameters follow exponential change, one has to look at the logarithmic distance which is indeed one and a half times the model range.

extrapolation errors...

velocity:

in direction to the Sun is at  $\sim 2.5$  Rs the source surface (Schatten1969)



**Fig. 9.** Estimated B-field median during the first and nearest perihelia in 2018 and 2024.

sonic and Alfvénic critical point positions (see Sittler & Guhathakurta (1999))

Alfvénic critical surface i.e. source surface (see Fox before 2.1)

approaching these regions, acceleration plays a role

The near-Sun (SPP perihelion) solar wind velocity is expected to be slower than our model estimates, because the position of the source (Alfvénic critical) surface is predicted to lie between 15–30 Rs (Schatten1969, Sittler1999, Exarhos2000, Katsikas2010, Goelzer2014; choose references...), up to which the solar wind is believed to be accelerated.

density:

Leblanc et al. (1998) -> electron density model derived from type III radio burst observations

The density distance dependency scales with  $r^{-2}$ , but steepens below 10 Rs with  $r^{-6}$  (Leblanc et al. 1998).

magnetic field and temperature:  
crossing distance effect

### 7.2. Model requirements and limits

model requirements:

- modeling the frequency distribution with sufficient accuracy -> error size (changing distribution shape with distance)
- modeling the distance dependency with sufficient accuracy -> error size (different scaling law at smaller distances?)
- possibility to extrapolate model down to 0.0459 au

model boundaries (spherical coordinates):

- mean solar wind of the Helios time range -> error size (determine deviation from average)
- heliocentric distance range 0.29–0.98 au

- rotational symmetry

- confined to ecliptic

model constrictions:

- solar distance dependency function
- frequency distribution function
- neglected influence from heliolatitude variation

### 7.3. Model validity and error sources

validity and estimation of error size outside of valid model range...  
derive heliocentric distance depending error...

list simplifications/approximations...

error estimation for general model and extreme value tendencies

error sources:

- extrapolation
- lognormal model
- SSN variance

all estimates outside these boundaries are extrapolations with large uncertainties. see Sect. 7.3

discuss high value zoom figures

Generally, the use of hourly data instead of higher resolution data (less than 1 min) should not have significant influence on the results. The minutely OMNI 2 data's frequency distributions are almost congruent and differ only slightly at their extreme ends (we tested that...).

The many hourly Helios data points which contain only a few measurements, contribute with a larger scatter to the frequency distributions (Schwenn 1983, p. 491).

The OMNI 2 data is time-shifted to the nose of the Earth's bow shock. This leads to yearly solar distance variations of > 2 % (cite?) as the Earth orbits the Sun. Furthermore, its orbit within the ecliptic leads to a yearly variation of  $\pm 7.2^\circ$  in heliospheric latitude.

The solar wind parameters vary with solar distance as well as with latitudinal separation from the heliospheric current sheet (HCS) (cite?).

The HCS's position in latitude is highly variable around the solar equator (Schwenn 1990, p. 127 ff.?).

- over year (seasonal/monthly) -> expect them to be less than 5 %

Sun-Earth distance over the course of the year.

In the year 2017 Earth's perihelion was on 5 January with a distance of -1.67 % from 1 au (Horizons On-Line Ephemeris System<sup>4</sup>, Solar System Dynamics Group, Jet Propulsion Laboratory).

The cosine approximation

$$r_E(t) = 1 - 0.0167 \cdot \cos\left(2\pi\left(t - 2017 - \frac{5}{365}\right)\right), \quad (12)$$

with  $t$  in years, suffices for our accuracy requirements.

seasonal variation function:

$$X_{\text{avg}}(t) = a r_E(t)^b$$

HGI coordinates; latitude range  $-7.25^\circ$  to  $7.25^\circ$   
latitude variation (see Schwenn1990 p. 127)

OMNI data is not corrected for solar distace. only available for 1-/5-min data

## 8. Results and discussion

list of results:

- low velocity at 0.0459 au
- empirical solar wind model for inner heliosphere
- application validity of lognormal distributions
- > B inversion of frequency distribution
- > magnetic field distribution's with distance increasing high value tail -> source are compression regions (why with density no increase?); look into Parker1958's B-field formula...

Comparison with existing near-Sun models (renew extrapolation figure to 1-column)

Wang et al. (2000), sources of slow solar wind + IMF regulation mechanism + blobs; compare with our slow V lognormal part

Schwenn (1983, 1990), who derived the distance dependencies for both Helios spacecraft separately ( $v_{H1}(r) \propto r^{0.083}$  and  $v_{H2}(r) \propto r^{0.036}$ ), average or median? fits well...  
Helios results, radial gradients see Schwenn (1990) p. 155

The electron density distance function, which Bougeret et al. (1984) derived from Helios data and normed to the 1976 1 au density ( $n(r) = 6.14 r^{-2.10(4)} \text{ cm}^{-3}$ ) also fits well...

- Sheeley et al. (1997) -> LASCO coronagraph observed speed profile of coronal features tracing the slow solar wind, 2–30 Rs
- parabolic eq. (2) and exponential eq. (3)
- sonic point 5–6 Rs
- slow solar wind origin 3–4 Rs

The model of Sittler & Guhathakurta (1999) is based on Skylab coronagraph and Ulysses data. it is a 2d semiempirical MHD model (time static)  
L-> compare with their radial density function eq. (18a); B-field function eq. (19a+b); V(r) eq. (24); T(r)...

Balogh et al. (1999) p. 162 ff (origin and formation of CIRs in inner heliosphere with Helios data; latitude V dependence)  
Balogh2009 (HMF review + inner heliosheath)  
Patsourakos2016 (near-Sun B-field of CMEs)  
Aschwanden2004, p. 29

Parker1963

individual velocity part discussion -> there is no specific velocity threshold between slow and fast solar wind types, the velocity ranges of both types overlap.

Sanchez-Diaz et al. (2016) (very slow solar wind)

Not only the slowest wind but also the fastest wind is expected to converge to the average speed (Sanchez-Diaz2016 p. 2835, using MHD-model -> very slow solar wind is continuation of slow wind) (because of interaction).

The ratio of both varies with solar activity, e.g. 3 years after maximum, polar coronal holes are observed to often have equatorial extensions (cite?). see and use Bougeret et al. (1984)

<sup>4</sup> <http://ssd.jpl.nasa.gov/horizons.cgi>

p. 498...

larger influx from higher latitudes (see figure b))

varying shape with distance is indicator for internal physical processes (mixing/turbulence...)

In most studies the density distance dependence is assumed to scale with  $r^{-2}$  (cites), assuming a constant velocity.

## 9. Conclusions

Further investigations should be done into structure extrapolations...

Further questions:

nearer to the Sun (at and below the source surface) the solar wind expansion in the ecliptic should be less spherical but more circular due to the influx from higher latitudes. => density exponent > -2  
see Li2011 Fig. 1

*Acknowledgements.* This work was done within the scope of the Coronagraphic German and US SolarProbePlus Survey (CGAUSS), funded by the German Aerospace Center (DLR) (grant number...).

The authors thank the Helios and OMNI PIs/teams for creating and making available the solar wind in situ data. The Helios and the OMNI data are supplied by the NASA Space Science Data Coordinated Archive (NSSDCA) and the Space Physics Data Facility (SPDF) at NASA's Goddard Space Flight Center (GSFC).

Additional thanks for maintaining and providing the international sunspot number series goes to the World Data Center – Sunspot Index and Long-term Solar Observations (WDC-SILSO) at the Solar Influences Data Analysis Center (SIDC), Royal Observatory of Belgium (ROB).

## References

- Balogh, A., Bothmer, V., Crooker, N. U., et al. 1999, *Space Sci. Rev.*, 89, 141
- Bougeret, J.-L., King, J. H., & Schwenn, R. 1984, *Sol. Phys.*, 90, 401
- Feldman, W. C., Asbridge, J. R., Bame, S. J., & Gosling, J. T. 1978, *J. Geophys. Res.*, 83, 2177
- Fox, N. J., Velli, M. C., Bale, S. D., et al. 2015, *Space Sci. Rev.*
- Kasper, J. C., Stevens, M. L., Korreck, K. E., et al. 2012, *ApJ*, 745, 162
- King, J. H. & Papitashvili, N. E. 2005, *Journal of Geophysical Research (Space Physics)*, 110, A02104
- Leblanc, Y., Dulk, G. A., & Bougeret, J.-L. 1998, *Sol. Phys.*, 183, 165
- Parker, E. N. 1958, *ApJ*, 128, 664
- Richardson, I. G. & Cane, H. V. 2012, *Journal of Space Weather and Space Climate*, 2, A2
- Rosenbauer, H., Schwenn, R., Marsch, E., et al. 1977, *Journal of Geophysics Zeitschrift Geophysik*, 42, 561
- Sanchez-Diaz, E., Rouillard, A. P., Lavraud, B., et al. 2016, *Journal of Geophysical Research (Space Physics)*, 121, 2830
- Schwenn, R. 1983, in *NASA Conference Publication*, Vol. 228, *NASA Conference Publication*
- Schwenn, R. 1990, *Large-Scale Structure of the Interplanetary Medium*, ed. R. Schwenn & E. Marsch, 99
- Sheeley, N. R., Wang, Y.-M., Hawley, S. H., et al. 1997, *ApJ*, 484, 472
- SILSO World Data Center. 1963–2016, *International Sunspot Number Monthly Bulletin and online catalogue*
- Sittler, Jr., E. C. & Guhathakurta, M. 1999, *ApJ*, 523, 812
- Vourlidas, A., Howard, R. A., Plunkett, S. P., et al. 2016, *Space Sci. Rev.*, 204, 83
- Wang, Y.-M., Sheeley, N. R., Socker, D. G., Howard, R. A., & Rich, N. B. 2000, *J. Geophys. Res.*, 105, 25133