Helios `corefit` data product user guide

David Stansby
david.stansby14@imperial.ac.uk

1 Introduction

This document describes the Helios `corefit` data product. This is a new data set which contains estimates of number density, velocity, and temperatures of the proton core population created from systematic fitting of all the original Helios 3D distribution functions.

2 Data availability

The data set is freely available at PUT DOI LINK HERE. If you are using the data in a publication, please cite this doi. Data are split by probe, year, and day of year, with data for each day contained in a single .csv file with a header row. Table 1 summarises the variables present in the data product.

The original 3D ion distribution function data are available at `ftp://apollo.ssl.berkeley.edu/pub/helios-data/E1_experiment/helios_raw/`. The source code used to read in and fit the distribution functions is available at PUT SOURCE CODE URL HERE.

3 Fitting method

The parameters in the `corefit` product are generated from fitting bi-Maxwellian functions to the original 3D ion distribution functions. Each distribution is fitted using the following process:

1. If magnetic field data is available, an average magnetic field is calculated from values taken whilst the distribution function was measured. The distribution function is then rotated into the field aligned frame.

2. A 3D bi-Maxwellian fit is done of the following form (fit parameters in bold):

$$f_{fit}\left(v_{\parallel}, v_{\perp 1}, v_{\perp 2}\right) = \mathbf{A} \cdot \exp - \left\{ \left(\frac{v_{\parallel} - \mathbf{u}_{\parallel}}{\mathbf{w}_{\parallel}} \right)^2 + \left(\frac{v_{\perp 1} - \mathbf{u}_{\perp 1}}{\mathbf{w}_{\perp}} \right)^2 + \left(\frac{v_{\perp 2} - \mathbf{u}_{\perp 2}}{\mathbf{w}_{\perp}} \right)^2 \right\} \quad (1)$$

   The 6 fit parameters are amplitude (A), 3 velocities ($u_{\parallel}$, $u_{\perp 1}$, $u_{\perp 2}$), and 2 thermal speeds ($w_{\perp}$, $w_{\parallel}$). The fit is done using least squares minimisation of the residuals calculated from ($f_{data} - f_{fit}$).

3. The number density is calculated from

$$n = A \cdot \pi^{3/2} w_{\perp} w_{\perp} w_{\parallel} \quad (2)$$

   and two temperatures from

$$T_{\perp/\parallel} = \frac{m_p w_{\perp/\parallel}^2}{2k_B} \quad (3)$$

Parameter	Parameter label	Units	Note
Time	Time		
Fitting status	status		1
Instrument	Instrument		2
Magnetic field instrument	B instrument		3
Proton number density	n_p	cm^{-3}	
Proton x velocity	vp_x	$km \cdot s^{-1}$	
Proton y velocity	vp_y	$km \cdot s^{-1}$	
Proton z velocity	vp_z	$km \cdot s^{-1}$	
Proton perpendicular thermal speed	vth_p_perp	$km \cdot s^{-1}$	
Proton parallel thermal speed	vth_p_par	$km \cdot s^{-1}$	
Proton perpendicular temperature	Tp_perp	Kelvin	
Proton parallel temperature	Tp_par	Kelvin	
B_x	Bx	nT	
B_y	By	nT	
B_z	Bz	nT	
Magnetic field standard deviation	sigma B	nT	4
Sun-spacecraft distance	r_sun	AU	
Carrington longitude	clong	Degrees	
Carrington latitude	clat	Degrees	
Carrington rotation number	carrot		
Spacecraft-Earth angle	earth_he_angle	Degrees	

Table 1: Description of variables present in the `core_fit` data product.

1. Successful fit flags take the following values: 1: successful fit; 2: successful fit, no magnetic field data available; 3: failed fit
2. Ion instrument flags take the following values: 1: I1a ion instrument data; 2: I3 ion instrument data
3. Magnetic field instrument flags take the following values: 1: E3 magnetometer data used; 2: E2 magnetometer data used
4. Magnetic field standard deviation is calculated as $\sqrt{\sigma_{Bx}^2 + \sigma_{By}^2 + \sigma_{Bz}^2}$, with individual standard deviations taken over the interval that the ion distribution function was measured

	B available	**B** not available	Failed fits
`B_instrument` variable value	1 or 2	-1	
Number of points	1,880,138	387,852	132,000
% points	78.3%	16.1%	5.6%
Number density	x		
Velocity	x	x	
Temperatures	x		
Magnetic field	x		

Table 2: Summary of which fitted parameters are available depending on magnetic field availability

If no magnetic field is available in step 1, steps 2 and 3 still take place, but the thermal speeds and subsequently the number density are ignored. Table 2 summarises which variables are available in each case.

Figure 1 shows an example of the original distribution function along with the fit.

4 Differences between `corefit` and `merged` data sets

The `merged` data set[1] contains estimates of proton bulk parameters calculated from numerical moments of reduced 1D distribution functions. This means that the `merged` data does not discriminate between the proton core and proton beam population. Differences between each parameter are as follows:

4.1 Number density

The number density in the `merged` data set contains contributions from the proton beam, so is systematically higher than the corefit number density.

4.2 Velocity

The radial component of velocity is systematically higher in the `merged` data set compare to the corefit values, due to the presence of the proton beam. The azimuthal and polar components are not affected by this and the two data sets contain similar values.

4.3 Temperature

The `merged` data set contains only one proton temperature value. This value was calculated from the 1D distribution function which means it contains variable contributions from the true parallel and perpendicular temperatures of the protons. The perpendicular and parallel temperatures in the `corefit` data set are therefore much more accurate estimates of the temperatures of the proton core.

[1] Available at `ftp://cdaweb.gsfc.nasa.gov/pub/data/helios/helios1/merged/` and `ftp://cdaweb.gsfc.nasa.gov/pub/data/helios/helios2/merged/`

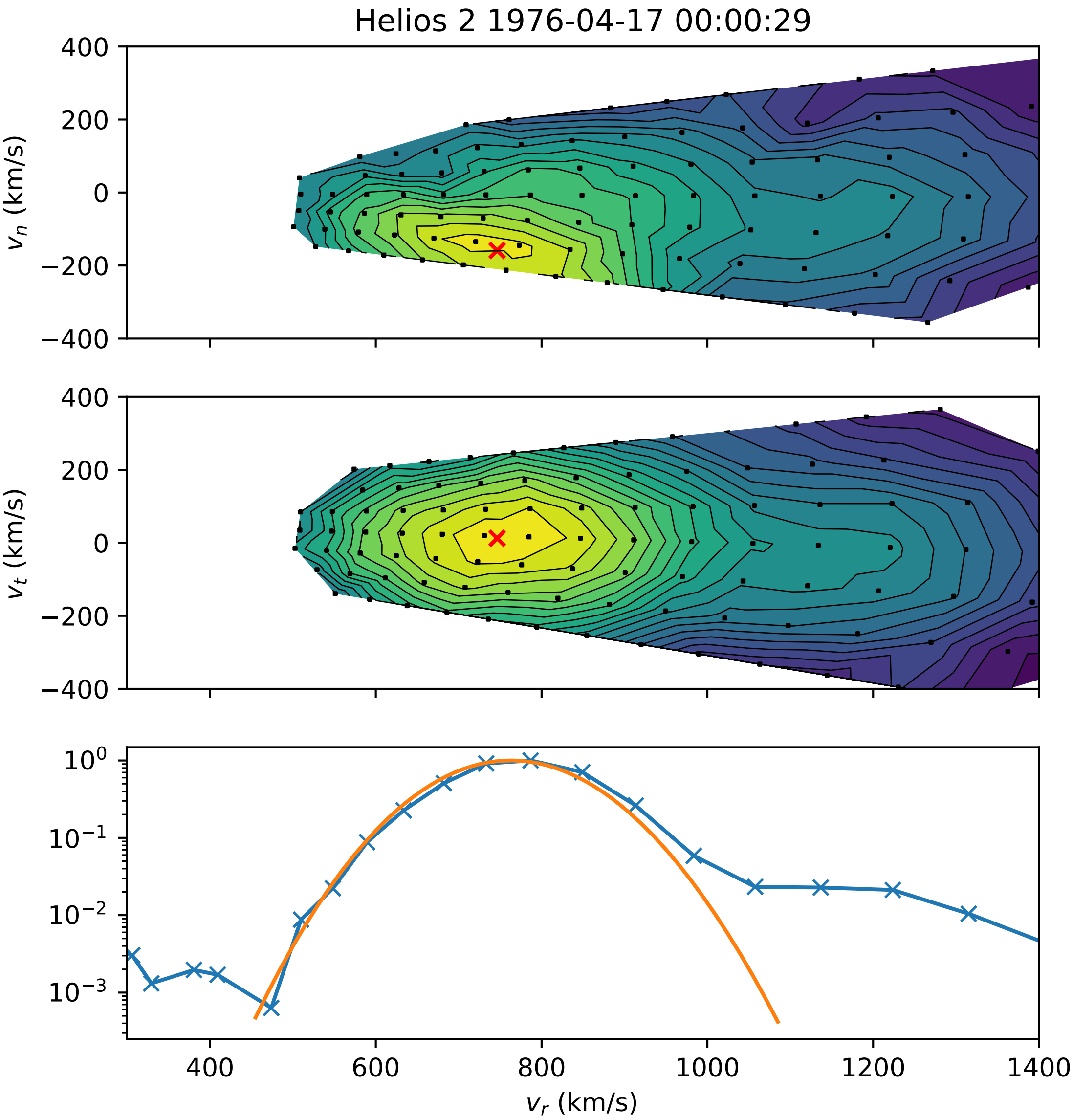

Figure 1: Example of distribution function data and corresponding fit. Top panel shows reduced 2D distribution function in the RN plane of an RTN co-ordinate system. Middle panel shows reduced distribution function in the RT plane. Black dots show positions in velocity space where the distribution function was sampled. Red crosses show the bulk velocity estimate from the fit, which agree well with the data. Bottom panel shows reduced 1D distribution function in blue, with the fit in orange. The fit agrees very well with the data, and is not sensitive to the proton beam present between 900 to 1100 km/s.

5 A note on timestamps

The timestamps in this data set are taken directly from the timestamps given in the filenames of the individual distribution function files. The distributions were recorded with a nominal cadence of 40.5 seconds, but it is believed anything after the decimal point in the timestamp has been dropped. This means that the time difference between consecutive time stamps goes like $0, 40, 81, 121, 162...$, whereas the true time differences were $0, 40.5, 81, 121.5, 162....$

A Acknowledgements

The work was primarily carried out by David Stansby, in collaboration with Chadi Salem, Lorenzo Matteini, and Tim Horbury. David Stansby is supported by STFC studentship ST/N504336/1.