Space Sci Rev (2009) 147: 55–87
DOI 10.1007/s11214-009-9550-5

Interplanetary Coronal Mass Ejections Observed in the Heliosphere: 2. Model and Data Comparison

S. James Tappin · Timothy A. Howard

Received: 6 March 2009 / Accepted: 8 June 2009 / Published online: 27 June 2009

Abstract With the recent advancements in interplanetary coronal mass ejection (ICME) imaging it is necessary to understand how heliospheric images may be interpreted, particularly at large elongation angles. Of crucial importance is how the current methods used for coronal mass ejection measurement in coronagraph images must be changed to account for the large elongations involved in the heliosphere. We present results comparing a new model of interplanetary disturbances with heliospheric image data, from the Solar Mass Ejection Imager. A database containing a range of ICMEs simulated with varying parameters describing its topology, orientation, location and speed was produced and compared with two ICMEs observed in February and December 2004. We identify the simulated ICME that best matches the data, and use the parameters required to identify their three-dimensional leading-edge structure, orientation and kinematics. By constant comparison with the data we are able to keep track of small changes to the ICME topology and kinematic properties, thus for the first time are able to monitor how the dynamic interaction between the ICME and the interplanetary medium affects ICME evolution. This is the second part of a series of three papers, where the theory behind the model is presented in an accompanying paper and the physical implications are discussed in the third part. The first part considers the effects of Thomson scattering across the entire span of the disturbance and includes its apparent geometry at large elongations. We find that the model converges reliably to a solution for both events, although we identify four separate structures during the December period. Comparing the 3-D trajectory and source location with known associated features identified with other spacecraft, we find a remarkable agreement between the model and data. We conclude with a brief discussion of the physical implications of the model.

Keywords Coronal mass ejections · Solar-terrestrial relations · Interplanetary medium

S.J. Tappin (✉) · T.A. Howard
Air Force Research Laboratory, National Solar Observatory, Sunspot, NM 88349, USA
e-mail: jtappin@nso.edu

Present address:
T.A. Howard
Department of Space Studies, Southwest Research Institute, Boulder, CO 80302, USA

1 Introduction

With the launch of recent missions including heliospheric imaging instruments, interest among the community in heliospheric disturbances, particularly interplanetary coronal mass ejections has been renewed. Interplanetary coronal mass ejections, or ICMEs (Zhao 1992; Dryer 1994), are the heliospheric counterpart of coronal mass ejections (CMEs), which may be described as large clouds of magnetised plasma originating from the Sun. Analysis of heliospheric images requires a proper treatment of not only the physics behind how ICMEs are observed, but also the effects of the relative geometry of the ICME at large elongation angles. In an accompanying paper, Howard and Tappin (2009a) (hereafter referred to as Paper 1) showed that the intensity of Thomson scattered light decreases at a lower rate for larger elongations as it moved further from the so-called Thomson surface. This indicated that not only were CMEs detectable at larger than expected distances from the Thomson surface, but also that estimations of ICME mass based on Thomson surface locality remained largely accurate even if the location of the mass was unknown. The cost was a higher uncertainty in identifying the location of the observed part of the ICME. The Thomson surface, or "half-AU" sphere as it is otherwise known, is the locus of all points where the observer-P vector is orthogonal to the Sun-P vector, where P is the point at which the intensity measurement is made. This forms a sphere with a diameter of the Sun-observer line (e.g., Vourlidas and Howard 2006).

In Paper 1 we also showed that the location of the leading edge of the ICME changes as its elongation angle increases. It was shown that the location of the leading edge, from which distance and geometry measurements are typically made, moves closer to the observer as the elongation increased. This demonstrated that measurements based on the assumption of same location of the leading edge will be larger than the true leading edge location, so that much larger apparent speeds result when elongation measurements are converted to distance. Based on a simple assumption of an expanding spherical bubble and then as an expanding arc, Paper 1 developed equations showing the difference between the observed and true leading edge location, which was dependent on the elongation, the distance of the leading edge from the Sun and the angle of its central axis. It was suggested that these equations may be used as a first-order correction for large-elongation measurements of ICME distance from the Sun.

In the conclusion, Paper 1 stated that the best way to accommodate for all of the parameters was to simulate ICMEs under its theory and then compare the simulations with the data. The present paper is the second part of this study, where we produce the simulation and perform the data comparison. We have generated a database containing a range of ICMEs simulated with varying parameters describing the topology, orientation, location and speed. These are compared with two ICMEs that were observed in all-sky heliospheric image data in February and December 2004. We identify the simulated ICME that best matches the data, and use the parameters required to simulate the ICME to identify their three-dimensional structure, orientation and kinematics. By constant comparison with the data we are able to use this technique to adjust the kinematic properties, thus for the first time are able to monitor how the dynamic interaction between the ICME and the interplanetary medium affects ICME evolution. In the following sections we discuss the data, data processing and selected events, and describe the simulated ICME model and the methods employed to compare the model with the data. The results from these comparisons are shown, including the likely characteristics of each ICME, along with estimates of the uncertainties and constraints of the parameters. The associated discussion considers the physical implications of the monitored changes to the evolution of each of the two ICMEs.

2 Data and Events

2.1 Data

Heliospheric image data were provided by the Solar Mass Ejection Imager (SMEI) (Eyles et al. 2003; Jackson et al. 2004) on the Coriolis spacecraft. SMEI observes the entire sky in white light beyond 20° elongation and builds up each image through its ∼100 minute polar orbit, so each image has a cadence of ∼100 minutes. SMEI observations and measurements of ICMEs have been discussed in previous papers (e.g. Tappin et al. 2004; Howard et al. 2006; Webb et al. 2006; Tappin 2006). The manner with which the image sky is scanned is such that each section of the sky represents a different part of the spacecraft orbit, and as such measurements at different parts of the image are at different times. Hence two measurements at two different parts of a single SMEI image may be separated in time by up to ∼100 minutes. For increased accuracy, all measurements in the present paper are the true time measurements at their respective point in the image. SMEI images shown here are Sun-centred Hammer/Aitoff projections (e.g. Leighly 1955) where the extreme left and right points of each image represents the same point directly behind the observer when looking at the Sun. Also shown are zenithal equidistant (often referred to as "fisheye") projections which map the sky such that elongation contours are perfect circles (e.g. http://www.astron.nl/aips++/docs/memos/107/node10.html).

For the purpose of this study, we have used an improved SMEI processing pipeline which incorporates significantly better methods of (*inter alia*) dark charge removal, star subtraction and of particle hit detection, than has been available for earlier work. While a detailed description of this is beyond the scope of this paper, the major improvements can be summarized as:

- The CCD dark charge is subtracted as a pattern derived from the weekly calibrations (rather than a simple gradient across the CCD). In addition we identify and attempt to correct those pixels which have changed significantly since the last calibration.
- Star subtraction is carried out in the CCD frames, rather than in the final sky maps.
- Particle hits are primarily detected in the time domain (i.e. by changes from frame to frame) rather than in the spatial domain.
- The average time of the data from which each point in the sky map was generated is recorded, thus allowing us to handle the movement of the ICME over the course of an orbit.
- The scaling is to "true" ADU (Analogue to Digital Units: the internal units of SMEI images) values rather than "scaled" ADU values, thus in instrument units these images appear to be 4/3 times brighter than older versions.

In order to gain a better perspective of the solar source for the ICMEs, we have also investigated solar coronal data. Information on the coronal mass ejection (CME) counterpart of each ICME was obtained using the Large Angle Spectroscopic Coronagraph (LASCO), (Brueckner et al. 1995). LASCO has two functioning coronagraphs at present, C2 which has a field of view (FOV) of 1.5–6 $R_\odot$ and a cadence of around 30 minutes, and C3 with a FOV of 3.7–30 $R_\odot$ and cadence of around 50 minutes. Both coronagraphs observe the white light corona. To give a better idea of the direction of propagation of each CME/ICME, we identified associated solar "surface" eruptions using EUV data. These were provided by the Extreme-ultraviolet Imaging Telescope (EIT) (Delaboudinière et al. 1995) also aboard SOHO. The wavelength most frequently used by this instrument is the 195 Å Fe XII line, which is the dominant EUV emission line from plasma at temperatures of around 1.5 ×

10^6 K. EIT observes the low-corona and produces full solar disk images in the 195 Å line roughly every 12 minutes.

2.2 Events

Two events have been chosen to test our modelling system specifically for their differences in relative direction.

The first (hereafter referred to as Event 1) was a CME that erupted near the solar limb. It first appeared in LASCO on 15 February 2004 at 03:54 UT with a sky-plane speed of $\sim 700\ \mathrm{km\,s^{-1}}$ and later in SMEI on 16 February 2004 at about 07:00 UT. While it was not associated with an X-ray or Hα flare, it was associated with an erupting prominence and post-eruptive arcade in EUV near the SE limb. This event has been discussed in detail by Howard et al. (2007), in which a first attempt at geometry and kinematic measurements are shown (their Event 2).

The second event (hereafter referred to as Event 2) was an Earth-directed ICME first observed by LASCO on 3 December 2004 at 00:26 UT with a sky-plane speed of $\sim 1100\ \mathrm{km\,s^{-1}}$, and then later by SMEI on 4 December 2004 at about 06:00 UT. It was associated with a GOES M1 class X-ray flare at 00:00 UT on 3 December, and a post-eruptive arcade and coronal dimming followed in EUV. These were located just north of solar disk centre. Later, the ACE SWEPAM data (http://www.srl.caltech.edu/ACE/ASC) show a shock at L1 on 5 December 2004 at 06:56 UT, there was also a small geomagnetic storm ($\mathrm{Dst_{min}} = -58$, $\mathrm{K}_{p\ \max} = 4+$) at the Earth on 6 December (http://swdcwww.kugi.kyoto-u.ac.jp/dstdir/index.html, http://sgd/ngdc.noaa.gov/sgd/jsp/solarindex/jsp). (We later show that one component of Event 2 is more likely to be associated with a separate LASCO CME on 30 November).

At a first glance, we may make the following revelations about the SMEI image sequences for each event. Event 1 is relatively bright in the innermost region of SMEI's FOV and then it abruptly vanishes at an elongation of around 40°. Event 2, on the other hand, is observed across a large elongation range, from ∼35° to ∼115°. Event 1 has a well-defined leading edge while Event 2 is more hazy, particularly at large elongation angles. Finally, Event 1 is localised across a fixed position angle (PA) range, while Event 2 is observed at different times in several different parts of the sky. These observations are consistent with predictions made using the theory described in Paper 1. For example, we showed that a limb CME/ICME would maintain a roughly constant observed intensity out to elongations of around 35° from the Thomson surface, which for a limb CME is the plane of the sky near the Sun. At elongations beyond 35° the observed intensity drops off rapidly, exactly as observed in SMEI images for Event 1. For events with a large component directed toward the Earth this is not the case, not only because ICMEs are observed at larger distances from the Thomson surface at large elongation angles, but also because the Thomson surface intersects the Earth as well as the Sun. Hence one would expect Earth-directed ICMEs to be observed not only across a larger range of elongation, but also close to and beyond 90°, which for an Earth-directed ICME corresponds to the location of the observer, or at the Earth in the case of SMEI. This was observed in the SMEI observations of Event 2. Sample SMEI images for both events are shown in Fig. 1, accompanied by tracings of the leading edges of the ICME structures.

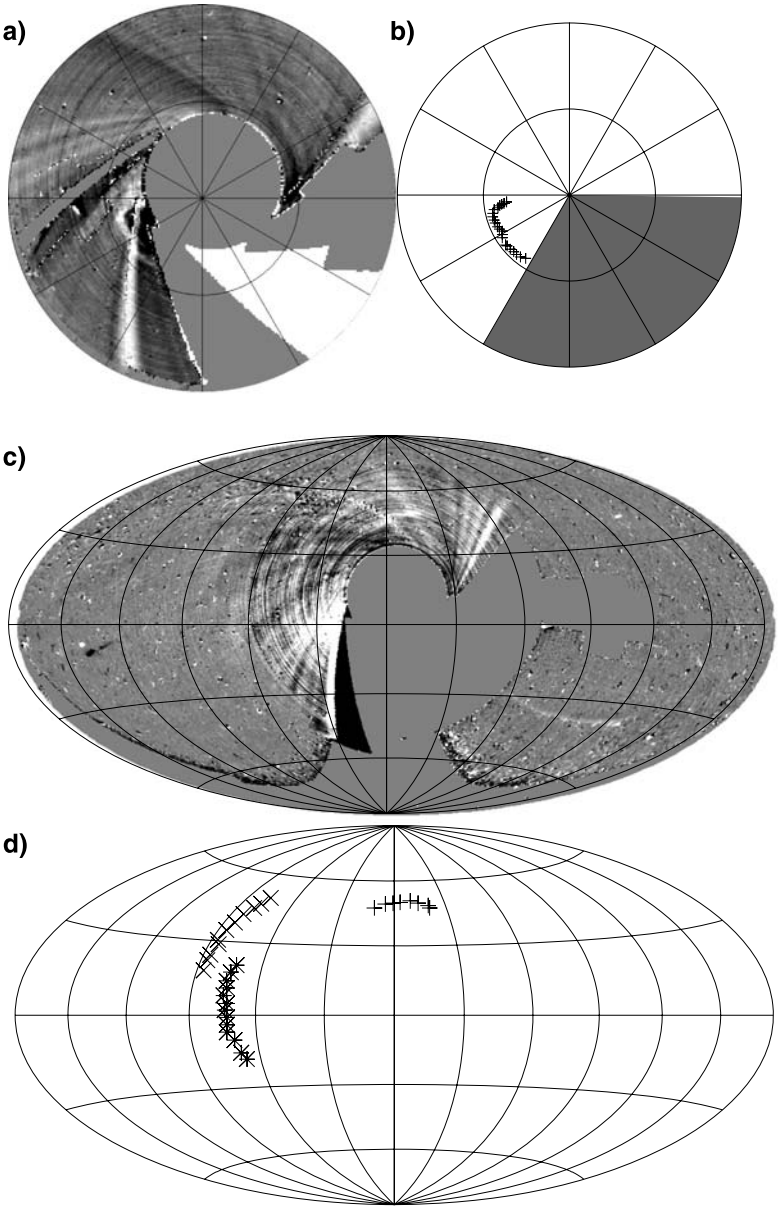

Fig. 1 SMEI images compared with the leading edge traces, for (**a**) and (**b**) Event 1 in a fisheye projection extending to 60° elongation for the orbit starting at 17:03 UT on 16 Feb. (**a**) shows the actual sky map with a range of ±3 ADU, while (**b**) shows the marked positions of the leading edge, the greyed out area indicates regions where we could not have identified the feature if it were present. (**c**) and (**d**) show Event 2 in a Hammer-Aitoff projection for the orbit starting at 15:57 on 4 Dec; the observations in (**c**) are scaled to ±2 ADU. In the interpretation in (**d**) Feature 1 is denoted by the + signs, Feature 2 by the × signs and Feature 3 by the ∗ signs (Feature 4 had not become visible by this time). For Event 2 the regions of undetectability are not included as they are different for the different features

3 Data Analysis

For ease of analysis and automation it is helpful to reduce the problem from three dimensions to two by determining the locus of the leading edge of the ICME in terms of two coordinates, elongation (ε) and position angle (PA), and using that in the fitting.

Each point on a SMEI image can be measured using two coordinates: Elongation (ε) and Position Angle (PA). The elongation is the angle between the point and the Sun. For ease of comparison with default measurements in the LASCO data the convention adopted for PA using SMEI images is such that directly northwards from the Sun represent a PA of 0°, with increasing PA in the anticlockwise direction. Hence, as with LASCO, due east is at PA = 90° and due west is at 270°. Because of the way in which an image is built up over an orbit, each (ε, PA) point has a different time, which must also be recorded.

It is nonetheless important to note here that SMEI and LASCO use different coordinate systems. LASCO uses heliographic coordinates which are centred on the solar equator, while SMEI uses heliocentric-ecliptic coordinates which are centred in the plane of the ecliptic. Because the Earth orbits the Sun at an inclination to the solar equator, there is a difference between the location of solar north and ecliptic north that changes throughout the year. For a non-leap year this difference is 0° on 4 March and 3 September, and a maximum of +7.25° on 3 June and minimum of −7.25° on 3 December. The direction (+ or −) represent the tilt angle of any LASCO image when observed in heliocentric-ecliptic coordinates. So for example on 3 June a LASCO image appears rotated to the right by 7.25° when aligned with a SMEI image. To correct for this deviation, we subtract 7.25° from any PA measurement on any LASCO image obtained on this day (e.g. Allen 1973 (Sect. 86)). In the present study, all measurements are made in heliocentric-ecliptic coordinates, so the PA measurements from the LASCO data have been altered accordingly.

Using background-subtracted images it is possible to measure the elongation of the leading edge to an accuracy of about 1° at any position angle, currently this is done manually (and it seems unlikely that it will be possible to automate this for SMEI data). The background subtraction used here is to subtract the average of the three previous and 3 following orbits. This method does a better job of removing the zodiacal light contribution than a simple orbit-to-orbit running difference, and also does not suppress slow-moving features as badly.

When traced along the leading edge of a CME observed in LASCO or an ICME observed in SMEI, a two-dimensional geometry profile of the leading edge can be produced. Each CME/ICME image hence becomes a single trace in ε, PA, time coordinates. The traces for each event in both fisheye and Aitoff projections is shown in Fig. 1. A leading edge is hence a convolution of the three-dimensional structure of the CME and its location relative to the observer. Coupled with a series of measurements across a time range, it also contains information on the kinematics of the ICME. It does, however, contain little information on the density of the ICME.

4 Model

4.1 The Modelling System

Most CME and ICME modelling in recent times has attempted to make a direct comparison between observations and physics-based models (generally either MHD simulations or kinematic models). This has lead to obtaining a fit by *ad hoc* tweaking of input parameters

near the Sun to match the data under consideration (frequently with little observational evidence for those tweaks). In this paper we attempt to remove some of this arbitrariness by deriving the CME structure and propagation directly from the observations and returning to a simple phenomenological model similar to that used by Tappin (1987). However unlike that work where observations and models were compared by eye we here numerically fit the observations and models.

This phenomenological modelling system combines the scattering and geometrical effects described in Paper 1 for specified structures. We are thus able to generate synthetic sky maps which can then be parameterised and compared with the data. In this way we are able to determine which properties of the ICME are well-constrained by the observations and which are not.

The modelling system integrates the Thomson-scattering integral along lines of sight from the observer to infinity:

$$I_{\rm rec} = \int_0^\infty f(r,\theta,\phi) N_e(r) z^2 I_{tot}(z)\,{\rm d}z, \tag{1}$$

where I_{tot} is the power scattered by an electron towards the observer (equation (29) of Paper 1), the coordinates (r,θ,ϕ) are the location of the scattering centre relative to the Sun, z is the distance of the scattering point from the observer, $N_e(r)$ is the quiet-solar-wind electron density at a distance r from the Sun and f is the factor by which the density at the scattering centre differs from the quiet value. For the electron number-density we use the same fit to the tabulations of Allen (1973) that was used by Tappin (2006), namely $N_e = 5\times10^6 r^{-2.45}$ m^{-2} inside 1 AU and an inverse square law beyond.

The integration is carried out through a conceptual grid in spherical coordinates with a resolution of (0.05 AU $\times$ 3° $\times$ 3°) with the Sun at the origin, and the z-axis perpendicular to the ecliptic plane. For each step we compute the distance to the end of the "cell", and compute the integrand at the mid-point of the passage of the line of sight through the cell, this is illustrated conceptually in Fig. 2(a). We terminate the integration when $z \geq z_{\rm max} = 2(1+\sin(\varepsilon/2))$ where ε is the elongation of the line of sight.

This procedure may be described as follows:

1. Starting at the observer, compute the distance to the next cell boundary in each coordinate.
2. The distance through the cell is the shortest of these.
3. Compute the integrand at the mid-point of the path though the cell, and multiply it by the f-factor for the cell if that is not equal to 1.
4. Add the contribution to the total.
5. Subtract the distance used from the other two, and recompute the one used from the new position.
6. Repeat from step 2 until $z \geq z_{\rm max}$.

The determination of the steps through the grid were presented by Tappin (1987), however as that paper contains a number of egregious typographical errors we quote those relations again here with corrections. The radial step (Fig. 2(b)) is given by:

$$\begin{aligned}
{\rm d}z_r &= [(r+dr)^2 - p^2]^{1/2} - z_s \quad &\text{when } z_s > 0,\\
{\rm d}z_r &= [(r-dr)^2 - p^2]^{1/2} - z_s \quad &\text{when } z_s < 0,\\
{\rm d}z_r &= 2(r^2-p^2)^{1/2} &\text{when } p > r - {\rm d}r.
\end{aligned}$$

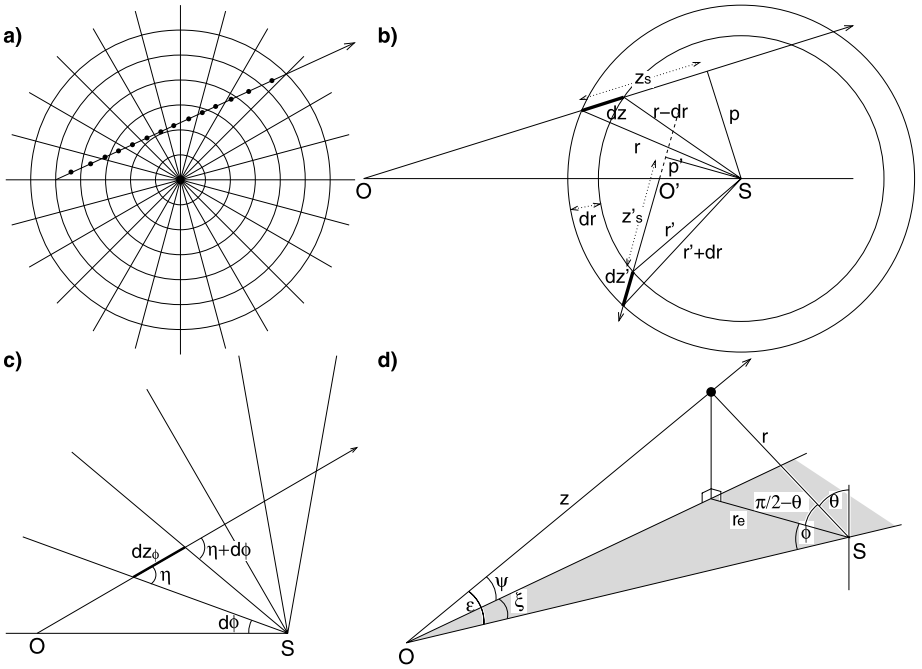

Fig. 2 Diagrams showing the geometry of the integration process in the models. (**a**) the basic concept of the integration (in the ecliptic plane); (**b**) the determination of the radial step; (**c**) the determination of the longitude step and (**d**) the determination of the latitude step

The azimuthal step (Fig. 2(c)) is given by:

$$\mathrm{d}z_\phi = \frac{r_e \sin \mathrm{d}\phi}{\sin(\eta + \mathrm{d}\phi)} \frac{1}{\cos \psi}$$

where r_e is the distance from the Sun projected into the ecliptic plane such that $r_e = z \tan\theta \sin\psi$. The latitudinal step (Fig. 2(d)) is determined using $\mathrm{d}z_\theta = z(\theta_2) - z(\theta_1)$ where:

$$z(\theta) = \frac{\cos\varepsilon \pm [\cos^2\varepsilon + (\sin^2\psi \tan^2\theta - \cos^2\psi]^{1/2}}{\cos^2\psi - \sin^2\psi \tan^2\theta}.$$

The appropriate root has the positive or negative sign according to whether $|\phi|$ is greater than or less than $\pi/2$.

4.2 Synthetic Sky Maps

As demonstrated in Paper 1, the three-dimensional structure and kinematic information can be extracted from a series of ICME leading edges by comparing leading edges extracted from data with those produced from a range of simulated ICMEs. Based on this theory, we have developed an ICME model that produces a map of ICME appearance relative to a fixed observer, and compiles the range of leading edges into a database, for comparison with the data.

We assume a range of simulated ICMEs and compute their appearances at different distances from the Sun and directions relative to the Sun-observer line. Two kinds of ICME

are considered: one assuming a distorted spherical structure (bubble) and another with a structure of a simple spherical shell.

For the bubble case a distortion factor D was introduced, which governs the distortion of the spherical structure of the bubble. Figure 3(a) shows the geometry of the bubble structure. The surface is determined by the relation:

$$\psi = \cos^{-1}\left(\frac{r'}{r}\right)D, \tag{2}$$

where r is the distance from the Sun to the leading edge of the structure along the central axis of the ICME and r' is the distance from the central axis to the surface of the bubble at some angle ψ to the axis. Figure 3(b) shows cross-sections of the bubble model with distortion factors of 0.5, 1.0 and 1.5. As shown here, values of $D < 1$ produce prolate bubbles, $D > 1$ are oblate, while $D = 1$ produces a perfect sphere.

Figure 3(c) shows a cross-section of the geometry of the spherical shell, which is otherwise known as the cone model (Zhao et al. 2002; Zhao 2008). The common parameters λ (the angle from the axis to the Sun-observer line) and r are shown. The three-dimensional cone structure is governed by the cone angle Ψ. We vary the cone angle independently in the latitudinal and longitudinal planes to generate a set of elliptical cones. In both cases the fixed thickness of the ICME, δr is shown.

For any given set of ICME parameters, the value of $f(r, \theta, \phi)$ in (1) was determined at the centre of each cell in the grid, and this was then used in the integration.

The ranges of parameters used are summarised below:

- The density enhancement factor was fixed at 2.0 for all ICMEs. For bubbles the enhancement factor inside was set to 0.5. We do not allow this parameter to vary is it has only a small effect on the location of the leading edge.
- All ICMEs had a thickness of 0.05 AU at the leading edge. Note that since we are comparing just the leading edges of the models and observations, the results are largely insensitive to the thickness. However, in comparisons by eye the thickness of 0.05 AU appeared more similar to the observed sky maps than did thicker models.
- The radius of the leading edge was varied from 0.3 to 2 AU in steps of 0.05 AU.
- The longitude of the central axis relative to the observer was varied from $-120°$ to $0°$ in steps of $6°$.
- The central latitude was varied from $0°$ to $42°$ in steps of $6°$.
- To fill in the other quadrants of the sky we simply reflect the synthetic maps in the N-S and E-W axes.
- For shell models, we use latitude and longitude cone angles of: $12°$, $18°$, $24°$, $30°$, $36°$, $48°$ and $60°$, but with a maximum aspect ratio of 2.
- For bubbles, the distortion factor was varied from 0.5 to 1.5 in steps of 0.125.

The range of the variable parameters was chosen to cover the range over which we would expect to be able to observe an ICME, while the spacing of the tabulation points was chosen to balance being close enough to interpolate reasonably, while maintaining a manageable size for the database.

Synthetic skymaps can be generated in a Hammer-Aitoff projection, a fisheye projection or as a simple rectangular grid in PA and ε coordinates (i.e. the x-axis is position angle and the y-axis is elongation). To generate a map of the excess brightness we subtract the computed brightness without an ICME present.

Note that in the following analysis the shell and bubble models are treated separately and we choose which is the better fit to a particular set of observations by comparing fitting residuals.

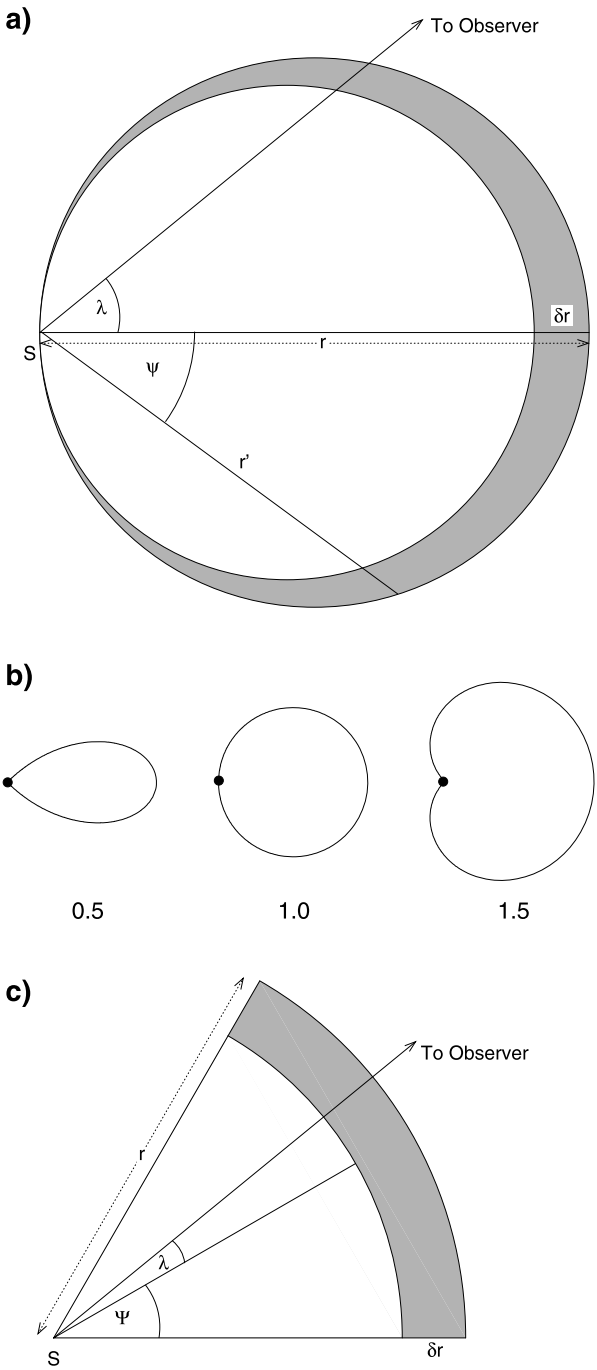

Fig. 3 Diagrams describing the chosen structures of the ICMEs for the model. (**a**) The geometry showing the necessary parameters for the bubble structure, including angle ψ and distance r'. (**b**) A cross-section of the bubble structure for three distortion factors: $D = 0.5$, 1.0 and 1.5. (**c**) A cross-section of the model for the shell structure, showing the cone angle Ψ in units of latitude (latitude width) or longitude (longitude width). In both the bubble and shell, λ is the angle between the central axis of the ICME and the Sun-observer line, r is the distance from the Sun to the leading edge of the ICME along the central axis, and δr is the thickness of the ICME, fixed at 0.05 AU for this model. In all three panels, the location of the Sun is denoted 'S'

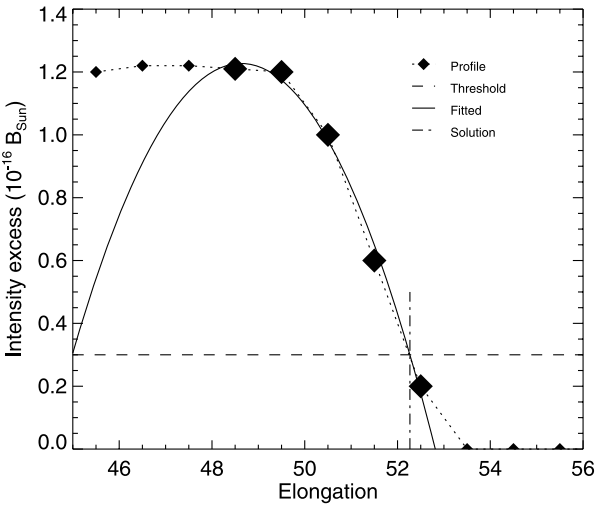

Fig. 4 Conceptual illustration of the determination of the lading edge of a synthetic sky map. The larger symbols on the synthetic skymap profile points indicate the ones used in the second-order fit

4.3 Leading Edge Generation

To determine the leading edge of the synthetic ICME, we use the simple rectangular projection as this removes the need for interpolations. We look along lines of position angle at 1° intervals and determine the largest elongation at which the brightness excess exceeds a specified threshold. That elongation, the three closer to the Sun and the one further away were then taken (or in the case of elongations close to the edge of the valid data, the innermost or outermost five valid points), a second-order polynomial was then fitted to these five points, and the elongation at which this polynomial crossed the threshold was used as the elongation of the leading edge. This is illustrated in Fig. 4. Two thresholds were selected, at 3×10^{-17} and 2×10^{-17} $B_{\odot}$, or around 0.15 and 0.1 ADU, since 1 $B_{\odot} = 8.66 \times 10^{15}$ S10 and 1 S10 = 0.55 ADU in the SMEI images (Buffington et al. 2007, D. Mizuno, private communication. Note that Buffington et al. use "scaled" ADUs which are 4/3 raw ADU). The higher threshold provides the leading edge. If there was a greater than 5° elongation separation between the two thresholds, then we considered that the leading edge was too diffuse to be detected in SMEI observations. In such a case, the model assumed no leading edge was present.

The leading edges were tabulated in 2° intervals of PA for all the synthetic skymaps generated to produce a five or six-dimensional database of leading edges. Hence, the database from which the model was compared with the data consists not of 3-D images, but 2-D ε-PA contours of the leading edges deemed sharp enough for detection by SMEI. The data cube contains a total of 311 640 ICME leading edges, 52 920 derived from the bubbles and 258 720 from the shells.

5 Comparison Between Data and Model

The measured leading-edge points are tabulated as a list of (*inter alia*) position angle, elongation and mean observation time of the point. The aim of the fitting process is to minimise the discrepancy between the observations and the model.

This is a multi-dimensional inversion problem, and as such is fraught with pitfalls. The obvious choice of using a simplex method carries a major risk of converging to a local minimum. The alternative of using a genetic algorithm (GA) (e.g. Poli et al. 2008) to maximise

Table 1 The scaling and shifts for mapping the model parameters into the genome

Quantity	Min	Max	Scale	Shift
Distance (AU)	0	2	32 000	0
Speed ($km\,s^{-1}$)	0	4000	16	0
Colatitude (°)	0	180	300	0
Longitude (°)	−180	180	150	180
Longitude size (°)	0	60	1000	0
Latitude size (°)	0	60	1000	0
Distortion	0	2	32 000	0

the goodness of fit ("fitness") of the model to the data can still converge to a local minimum, the risk is much less as it makes a much more thorough search of parameter space on the large scale. However with a GA there is a possibility that it will not find the bottom of the minimum unless a very large population and number of generations is used. For these reasons we have adopted a "hybrid" fitting procedure in which we first use a GA to find an approximate solution and then apply the simplex method to refine that solution on a small scale. In other words, the GA finds an approximate solution near to the global minimum and the additional simplex method then finds the bottom of that minimum.

5.1 The Genetic Algorithm

As discussed above the models used here are described by a small number of parameters (five or six according to the model type). The "genotype" for the models is therefore created by mapping these parameters onto 16-bit integers with suitable scalings according to the meaningful range of the parameters. The scalings and shifts which are listed in Table 1 were chosen such that meaningful values would use all bits of, but not overflow, an unsigned 16-bit integer. This method is chosen rather than simply using the floating-point bit patterns for two reasons. Firstly, with floating-point values many of the mutations and crosses would occur in parts of the number where there would be no appreciable effect on the resulting edge. Secondly, crossing floating-point bit patterns in the exponent can cause bizarre results.

While the ICME speed and sizes are always allowed to vary over the entire range of the available parameter space, the latitude and longitude can be limited according to whether the ICME appears to be in the north-east, north etc. The range of initial heights is also limited by whether the ICME is seen in Camera 3 of SMEI or not. These ranges are listed in Table 2.

Within the range of acceptable parameters an initial random population of 400 ICMEs is created, and for each of these a fitness is determined as 1/(mismatch). Figure 5 illustrates how this mismatch is determined. There are three cases to consider in the error determination:

1. When the model is defined at the position angle of an observation, then the error contribution is the difference in elongation between the observation and the model. This is the measurement $\delta\varepsilon_1$ in Fig. 5.
2. When the model is not defined at the position angle of an observation, then the error contribution is given by the lesser of a specified threshold (generally 15°) and $\varepsilon - 20°$. This is $\delta\varepsilon_2$ in Fig. 5.
3. When the model is defined, but the observations do not indicate an edge defining a sensible error contribution is somewhat more difficult. To handle this we take a set of position angles at 5° intervals around the Sun. If the model is not defined at any of these then

Table 2 The ranges of parameters for different apparent ICME directions

Direction	Latitude (°)		Longitude (°)	
	Min	Max	Min	Max
NE	−10	40	−120	0
N	−10	40	−60	60
NW	−10	40	0	120
E	−30	30	−120	0
Halo	−30	30	−60	60
W	−30	30	0	120
SE	−40	10	−120	0
S	−40	10	−60	60
SW	−40	10	0	120
Unknown	−40	40	−120	120

Camera	Starting height (AU)	
	Min	Max
Unknown	0.3	1.5
3	0.3	1.0
1 or 2	0.7	1.5

there is no contribution, likewise if the model is defined and there is a measurement or a region of bad data within 5° of that position angle then there is no contribution. If however there is neither observation nor data gap within 5°, the contribution is the smaller of the position angle distance to the nearest measurement ($\delta\Pi_1$ in Fig. 5) and the distance to the nearest data gap ($\delta\Pi_2$ in Fig. 5). For the purposes of this part of the calculation, the model is computed at the mid-point time of the orbit.

The mismatch is then simply defined as the mean absolute value of all the error contributions over the data interval under consideration (this corresponds to using the L1 norm in regular curve fitting, and is chosen as it is less sensitive to occasional wild points than the L2 norm). If the model attempted lies outside the range of the models that have been created, then an infinite value is returned for the mismatch.

Since a mismatch greater than 15° indicates a failure to fit, any model with a mismatch greater than 15° is assigned a fitness of 0.

After the fitnesses are computed for each model, they are assigned a probability of breeding as:

$$P(\text{breed}) = \frac{\text{Fitness}}{\sum \text{Fitness}}. \tag{3}$$

Using these probabilities, pairs of models are selected. For each pair, there is a probability of 0.8 of a "crossover" occurring, in which case after a random point in the genome the two models are switched (this process is illustrated in Fig. 6). One of the two resultant genomes is then selected at random. There is also a probability of 0.02 that a random mutation will occur causing 1 bit in the genome to be flipped. In addition to the new genomes, the best of the previous generation is retained.

This process is repeated for 50 generations to find a solution.

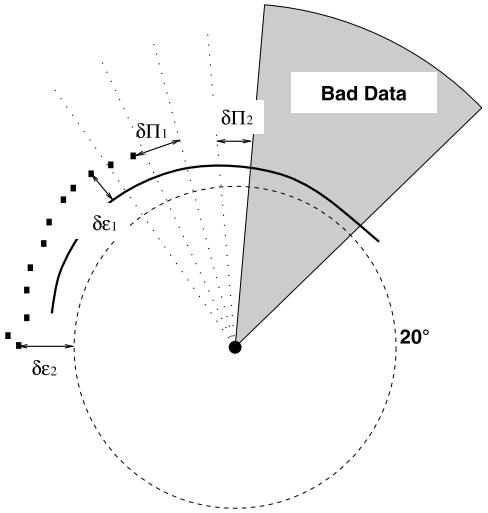

Fig. 5 Schematic illustrating the determination of the fitting error between the model CME and the observations. The points represent the observations, and the solid arc the model being tested. The *shaded sector* represents a region of bad data. The *dashed circle* indicates an elongation of 20°, and the *dotted radial lines* are representative of a set of regularly-spaced position angles. The illustrated mismatch contributions are: $\delta\varepsilon_1$ the difference in elongation where model and observations are both present, $\delta\varepsilon_2$ the contribution where data are present but the model is not, $\delta\Pi_1$ the difference of position angle from where a model edge is present to the nearest data point ($> 5°$), and $\delta\Pi_2$ the difference in position angle from where a model edge is present to the nearest place that there could be an edge in the data. Note that this is a deliberately bad match between model and data to make the error contributions clearer

	Height	Speed	Co-latitude	Longitude	Long size	Lat Size
Input values	0.357000	402.500	96.4000	-46.3000	21.4000	37.8000
	0.860000	320.400	58.3000	-22.1000	32.6000	41.0000
Input ints	11424	6440	28920	20055	21400	37800
	27520	5126	17490	23685	32600	41000
Input bits	0000010100110100	0001010010011000	0001111100001110	1110101001110010	0001100111001010	0001010111001001
	0000000111010110	0110000000101000	0100101000100010	1010000100111010	0001101011111110	0001010000000101
Output bits	0000010100110100	0001010010011000	0001111100001010	1010000100111010	0001101011111110	0001010000000101
	0000000111010110	0110000000101000	0100101000100110	1110101001110010	0001100111001010	0001010111001001
Output ints	11424	6440	20728	23685	32600	41000
	27520	5126	25682	20055	21400	37800
Output values	0.357000	402.500	69.0933	-22.1000	32.6000	41.0000
	0.860000	320.375	85.6067	-46.3000	21.4000	37.8000

Fig. 6 Schematic of the process of crossing two sets of model parameters. Note that the bit patterns are little-endian even if the fitting is run on a big-endian machine. The crossing in the example is marked by the vertical bar at bit 45

5.2 The Fitting Process

The actual fitting is performed as a two-stage process. In the first stage we use the GA described above to estimate ICME parameters. Once that is complete, we then use that set of parameters as the starting point for a simplex method to refine the solution by minimising the mismatch parameter. The scale of the simplex is chosen to be smaller than the typical sizes of the minima which we observe so that it is unlikely to jump into a different minimum from that chosen by the GA.

This first stage fitting yields a direction, size, starting height and average speed for the simulated ICME which best matches the measured leading edge.

Table 3 The fitting results for Event 1. The bold row was the "best" fit and the italicised rows are the other fits used in the Stage 2 fitting. The struck out data were rejected from consideration as being aberrant. The starting heights are determined at 03:31 UT on 16 February (the start of the first orbit in which the structure was visible)

Run #	Start Height AU	Average Speed km s^{-1}	Central Latitude °	Central Longitude °	Longitude Size °	Latitude Size °	Mismatch °
0	**0.32**	**562.26**	**−30.96**	**−107.67**	**22.87**	**30**	**1.24**
1	0.34	696.51	−28.19	−114	18	24	1.36
2	0.32	454.12	−36	−103.16	32.69	36	1.19
3	~~0.33~~	~~403.99~~	~~−36~~	~~−27.87~~	~~30.27~~	~~36~~	~~1.13~~
4	0.32	526.46	−31.89	−103.24	22.73	30	1.23
5	*0.32*	*409.22*	*−37.18*	*−99.81*	*48.82*	*34.81*	*1.18*
6	*0.35*	*850.9*	*−26.44*	*−119.15*	*13.77*	*23.37*	*1.42*
7	0.32	435.29	−33.07	−104.46	41.93	29.85	1.19
8	0.32	598.11	−30.63	−107.88	19.81	29.06	1.25
9	0.32	526.32	−32.56	−100.59	20.34	30	1.23
Avg	0.33	562.13	−31.88	−106.66	26.77	29.68	
σ	0.01	139.83	3.4	6.37	11.82	4.17	

In order to determine whether the ICME is accelerated or decelerated, we carry out a second-stage fitting. In this process we take the orbits in groups of five (i.e. orbits 1–5, 2–6 etc.), and using the first-stage fit as a starting point we run a new fit in which only the speed and height are allowed to vary. This is done in the same way as the first stage, i.e. a genetic algorithm, followed by a refinement using the simplex method.

6 Results

For each feature (see Sect. 2.2), we ran ten first stage fits so as to verify the consistency of the fitting procedure. In cases where the initial test runs showed a marked difference between the quality of fit for shell models and bubble models, only the type giving the better fit was used, but in those cases where the qualities were comparable we ran ten of each, and then made the selection as to which better matched the observations.

From the ten fits, we then selected a "best" and two "extremal" fits to use for the second stage. In the tabulated fits (Tables 3 to 7) these are denoted by bold and italic fonts respectively. To verify consistency we ran the second stage twice for the "best" fits, but in no case did we find any noticeable difference between the resulting fits. In addition, we did not find any significant differences in acceleration profiles between the "best" fits and the "extremal" cases, even for those which had significantly different average speeds.

In order to see which parameters were well-constrained by the observations and which were not, and whether certain parameters were dependent on one another, it was necessary to examine the shape of the minimum in the mismatch parameter. Since it is clearly not possible to display the multi-dimensional space of the full parameter set, we selected suitable pairs of parameters to examine and produced 2-D contour plots of those. These are discussed for the individual structures in the following sections.

6.1 Event 1: 16–17 February 2004

The first stage of fitting converged very well to a shell type ICME somewhat behind the east limb, and located well to the south. The ten fits are tabulated in Table 3, and as can be seen the mismatch parameter is typically about 1.2°. This indicates a very good match between the modelled and measured leading edges. Both the range of fitting parameters and the error contours in Fig. 7 show that the central longitude is not particularly well-constrained, but is strongly correlated with the longitude size (Fig. 7(d)). This is an indication that for this, and indeed any limb event, the location of the far flank does not have a major influence on the appearance of the leading edge. The latitudinal structure, on the other hand, is much better-constrained, as might intuitively be expected for a disturbance in this location. We also note that the speed is correlated with the central longitude in the sense that the smaller the magnitude of the longitude the smaller the speed (again as we would expect from considering the geometry). We should here note that the discontinuities seen in the error surfaces for some parameters in the shell models (e.g. Fig. 7(b)) are a consequence of the fact that we use a multi-dimensional linear interpolation between the points in our database of leading edges, combined with the close match of the shell-type model with the structure of the integration grid (it is a consequence of this latter factor that means that such edges do not appear in the bubble models). The lack of contours on the left hand side of Fig. 7(a) are due to the fact that our database does not extend below 0.3 AU, since the smallest elongations we are able to observe with SMEI are about 20° (which corresponds to a height of at least 0.34 AU).

Figure 8(a) shows the resulting second stage height-time (h-t) plot. The average speed determined from the groups of five points is shown on this plot as the lines through each point. Their values are shown in Fig. 8(b). These two plots show what appears to be a two-stage kinematic evolution, with a deceleration occurring up to around 18:00 on 16 February, where it then begins to accelerate, followed by a possible deceleration. The largest difference in speed is an increase of some 400 $\mathrm{km\,s^{-1}}$ between 14:00 on 16 February and 00:30 on 17 February. The overall change from start to finish is an increase of around 50 $\mathrm{km\,s^{-1}}$. Figure 8(c) extends the view back to the origin, and includes the projected h-t plot of the likely associated CME. The appearance of the CME in LASCO matches well with that of the ICME in SMEI (Howard et al. 2007), and the linear extraction back to the solar origin matches the CME data exceptionally well. This validates the accuracy of the model in approximating the structure and kinematics of this ICME. It also implies that for this event, there is little variation in the kinematics as it moved through the undetectable region between LASCO and SMEI.

Further analysis of this event has been conducted by Howard et al. (2007) (their Event 2), who associated this ICME with the same CME. Using a separate analysis procedure based on 3-D recreation from solar surface association, they also found that this event experienced a net acceleration from LASCO to SMEI. The associated surface event was an eruptive prominence followed by a post-eruptive arcade (PEA) which, upon close inspection, does appear to be just behind the solar limb (evidenced by the visibility of the limb with the prominence and PEA occurring behind it). This is consistent with the model-determined location of the ICME, at 108°E. The central latitude, at 31°S, is very close to the central position angle of the LASCO CME, which when rotated to ecliptic coordinates is around 110°, or 30°S. This difference could easily be accounted for by the slight projection effects of the event originating around 18° behind the limb.

Given the accuracy of the model in identifying both an accurate leading edge in the SMEI data and associated events in independent, spatially and temporally separate data, we may now recreate the 3-D structure of this ICME using the parameters derived in the model.

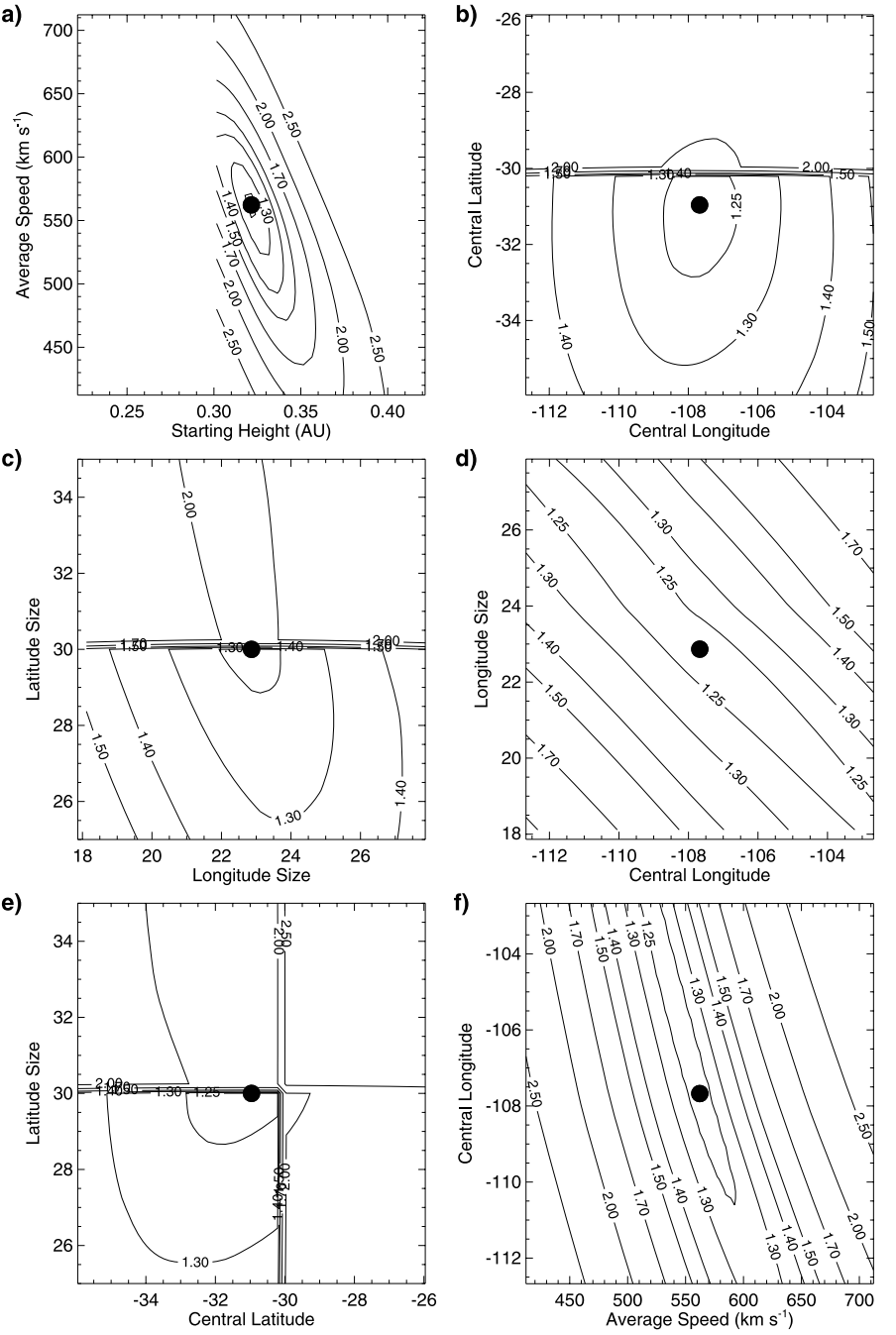

Fig. 7 Error surfaces for selected pairs of fitting parameters for Event 1. (**a**) The relationship between starting height and speed, (**b**) the latitude and longitude of the ICME central axis, (**c**) the size of the ICME in latitude and longitude, (**d**) the central longitude and the longitude size of the ICME, (**e**) the central latitude and the latitude size of the ICME, and (**f**) the average speed and the central longitude of the ICME. The central point represents the "best" fit, and the contours are of the mismatch parameters in degrees

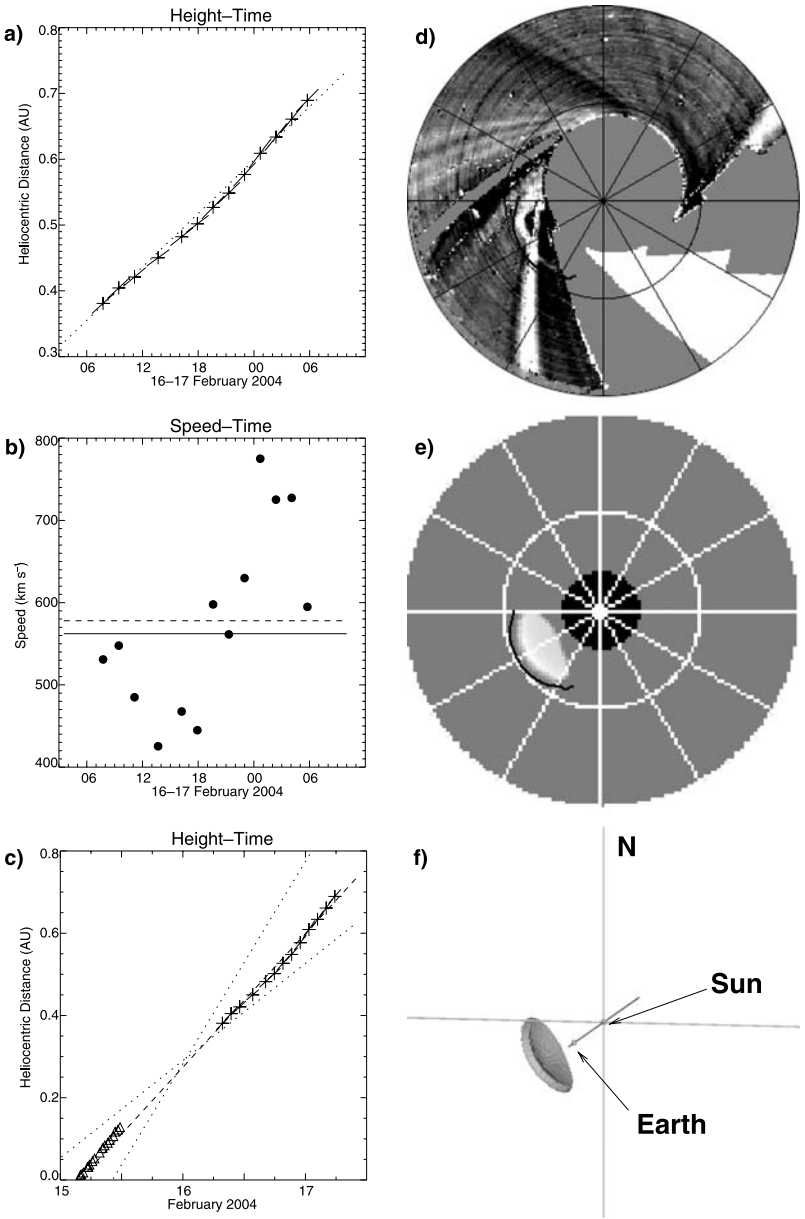

Fig. 8 A summary of the fitting results for Event 1. (**a**) The fitted height-time relationship. The *dotted line* shows the fit to all the data, the + signs and the *short lines* indicate the heights and speeds of the 5-orbit groupings in the stage 2 fitting. (**b**) The stage 2 fitted speeds as a function of time. (**c**) Similar to (**a**) but now extended back to the Sun, with the most likely LASCO association (*triangles*) added (note these are sky-plane projected heights). The *dashed line* is the "best" stage 1 fit as shown in panel (**a**), while the *dotted lines* are for the "extremal" stage 1 fits. (**d**) The SMEI image for the orbit covering 17:03 to 18:45 on 16 Feb. in a fisheye projection extending to an elongation of 60°. The calculated leading edge for the mid-point of the orbit is superposed. (**e**) The synthetic sky map for the modelled disturbance when it was at a height of 0.50 AU. (**f**) A view of the 3-D structure of the disturbance at the same time as it would be seen from the viewpoint of Ulysses located at 7.4°N, 10.4°W relative to Earth

Table 4 Stage 1 fitting results for Event 2 Feature 1. The denotations are the same as in Table 3. The starting heights are at 04:06 UT on 4 December

Run #	Start Height AU	Average Speed km s^{-1}	Central Latitude °	Central Longitude °	Longitude Size °	Latitude Size °	Mismatch °
0	0.56	666.27	36.19	58.08	48	32.33	3.66
1	**0.57**	**644.68**	**26.87**	**49.42**	**48**	**46.62**	**3.66**
2	0.57	642.85	27.28	48.04	48	36.36	3.67
3	0.56	658.88	31.77	37.3	34.01	60	3.56
4	~~0.57~~	~~653.32~~	~~37.89~~	~~12.03~~	~~12.01~~	~~15.04~~	~~3.29~~
5	0.57	639.62	27.26	53.15	47.98	36.35	3.68
6	*0.56*	*689.96*	*38.23*	*72.79*	*53.99*	*46.1*	*3.79*
7	~~0.57~~	~~646~~	~~36~~	~~12.04~~	~~12~~	~~17.54~~	~~3.28~~
8	*0.56*	*658.16*	*31.94*	*36.86*	*33.29*	*59.1*	*3.57*
9	~~No convergence~~						
Avg.	0.57	657.2	31.36	50.81	44.75	45.27	
σ	0	17.47	4.56	12.44	7.9	11.08	

Figure 8(d) shows a selected fisheye SMEI image of this event with the modelled projected leading edge superimposed, while Fig. 8(e) shows the same image with the leading edge and modelled ICME superimposed and mapped into the same projection. Figure 8(f) shows an arbitrary view of the 3-D structure of this event. We have chosen our viewpoint as that of the Ulysses spacecraft, which at the time of this event was at 7.4°N, 10.3°W of the Sun-Earth line, and at a distance of around 5.35 AU from the Sun. It is clear that this event does not pass the Earth or Ulysses, an observation supported by the absence of shock or ejecta signatures in both the Ulysses and ACE spacecraft.

6.2 Event 2: 4–6 December 2004

This event appeared as four separately distinguishable features in the SMEI data, and we have found by testing possible combinations that they do all appear to be distinct structures, although three of the four appear to be related.

6.2.1 Feature 1

The first feature of Event 2 was also well-fit by a shell model, although it is not as good as for Event 1, with a mismatch of about 3.6°. In this case the shell was directed to the north-west (Table 4). The error surfaces (Fig. 9) show that while the direction and speed are well-defined, the size is very poorly constrained (Fig. 9(c)). This implies that the latitudinal size can be varied significantly without altering the apparent leading-edge structure of the ICME. The fits do agree that the extent is much greater than that of the observable portion of the disturbance. This ability to at least indicate that an event has a greater extent than is actually seen is a significant result. For a shell structure, to determine whether the ICME will hit the Earth (or ACE) it is necessary to determine whether the extent of the structure in the ecliptic plane (ϕ) is greater than the magnitude of the longitude, i.e. the disturbance will

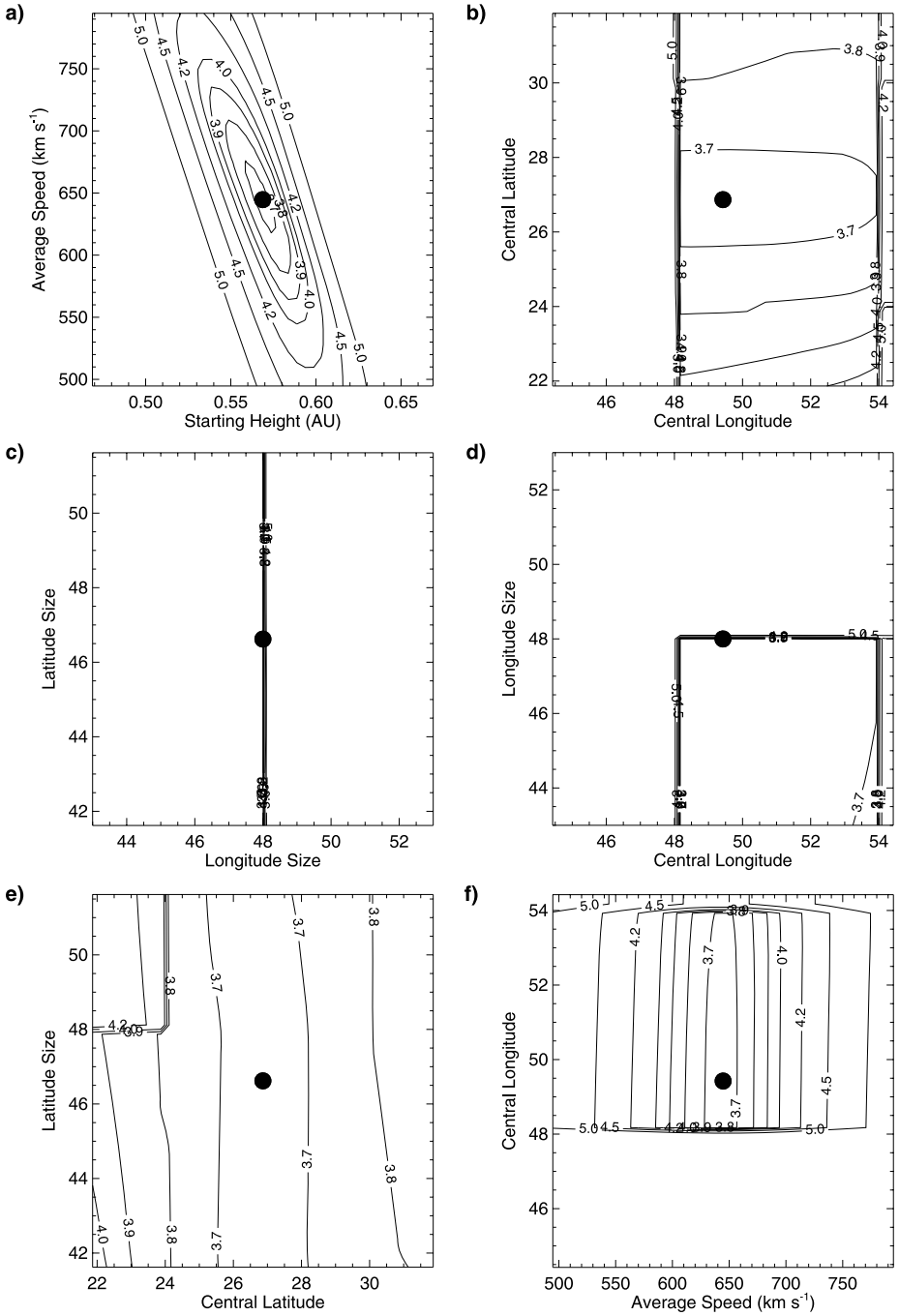

Fig. 9 Error surfaces for Event 2 Feature 1, the layout is as in Fig. 7

hit Earth if:

$$\phi = \sqrt{\Delta\lambda^2\left(1 - \left(\frac{\theta}{\Delta\theta}\right)^2\right)} \geq |\lambda|$$

where λ and θ are the longitude and latitude of the centre of the ICME and $\Delta\lambda$ and $\Delta\theta$ are the size of the ICME in those directions. In this case our fitted solutions do not hit the Earth, however they do come close. If we increase the size by just 8° in each direction, which is within the likely uncertainty of size, then the modelled ICME will hit Earth and ACE.

Figure 10 shows plots for Feature 1 of Event 2 in the same format as that shown for Event 1. The 3-D h-t plot in Fig. 10(a) shows a near-constant speed for this event, demonstrated by the speed-time plot in Fig. 10(b), where a variation of less than 100 km s^{-1} occurs throughout the field of view of SMEI. This is within the uncertainty limits of the determination of this parameter, but it is also possible the ICME is responding to variations of either its internal properties or the surrounding solar wind. The view is extended to the solar source in Fig. 10(c) as with Event 1, and we have associated this event with the Earth-directed halo CME that was first observed in LASCO on 3 December 2004 at around 00:00 UT. The linear projection of the SMEI ICME data to the Sun does not match the LASCO CME h-t data as well as for Event 1, but this could easily be accounted for by a deceleration between LASCO and SMEI. Given that the LASCO speed for this event was high ($>$1000 km s^{-1}) and the modelled SMEI speed was relatively low ($\sim$640 km s^{-1}), this is consistent with a deceleration in the region from 0.1 to 0.6 AU. Finally, given the absence of any other large CME on this day, we can state with large confidence that the LASCO halo is the most likely CME associated with this SMEI event.

The Earth-directed properties of this event is further evidenced by the appearance of a large shock observed at 0.99 AU by the ACE spacecraft on 5 December 2004 at 06:56 UT. Projecting the modelled trajectory of this event forward, continuing the constant speed assumption we predict an arrival time at ACE of 07:15 UT, only 19 minutes later than its actual arrival. This prediction, considering that the model was produced independently of the ACE and LASCO data, is remarkable, and better by far than any ICME model presently available. It also demonstrates that this ICME maintains a near constant speed from at least 0.6–1.0 AU, suggesting that once the deceleration process had completed somewhere between 0.1 and 0.6 AU, then no further change to its kinematics occurred.

The 3-D reconstruction of this event is shown in Figs. 10(d)–(f). Again, the viewpoint for the complete 3-D image is from the Ulysses spacecraft which at this time was at a location 7.15°S and 86°W of the Sun-Earth line and 5.33 AU from the Sun.

In solar coordinates, the modelled centre of the ICME was at 27°N49°W. The most apparent solar surface association with the LASCO halo CME was an X-ray flare observed by the GOES spacecraft. This was an M1.5 flare on NOAA Active Region 10708, with an onset of 23:44 on 2 December. It was located at 8°N2°W. A brightening in EIT at the same location and time was observed which was followed by a coronal dimming that extended mostly toward the northwest of the flare. Thus, the difference between the location of the flare and that of the centre of the CME structure was around 20° in latitude and 45° in longitude. This casts doubt on the validity of using solar flares as indicators of the source location of CMEs, and particularly on the too-often used claim that the "source" of the CME is the flare. It would appear more likely, as has been claimed in the past (Harrison and Sime 1989), that the flare indicates a footpoint of the CME. The extent of the coronal dimming structure would seem to provide a more reliable clue as to the origin of the CME itself, as its extent was in the direction of the modelled CME central location.

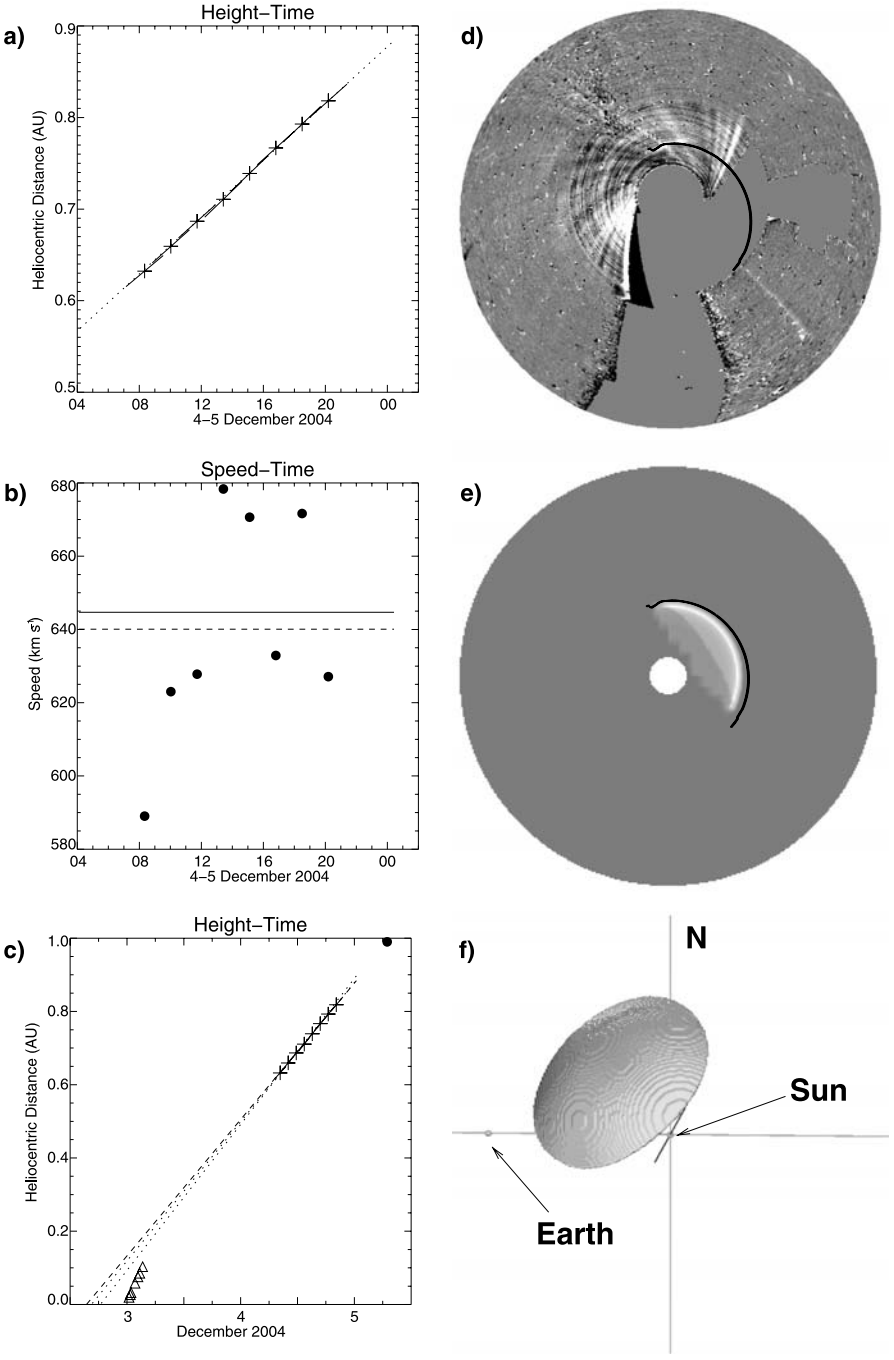

Fig. 10 Summary of the fitting for Feature 1 of Event 2. The format is similar to that of Fig. 8 with the exception that the orbit chosen for panels (**d**)–(**f**) covers the interval 15:57 to 17:39 on 4 December 2004, the fisheye projections here extend out to an elongation of 135° and in panel (**e**) the disturbance is located at 0.75 AU, while the location of Ulysses (and thus our viewpoint) in panel (**f**) is now at 7.5°S, 86°W

Table 5 Stage 1 fitting parameters for Event 2 Feature 2. The starting heights are at 07:30 on 4 December

Run #	Start Height AU	Average Speed km s^{-1}	Central Latitude °	Central Longitude °	Longitude Size °	Latitude Size °	Mismatch °
0	0.97	334.39	26.41	−12	18	16.75	2.91
1	0.97	360.29	27.2	−12.85	14.88	16.47	2.94
2	~~1.05~~	~~96.94~~	~~36~~	~~−61.96~~	~~54.12~~	~~42~~	~~5.66~~
3	0.99	241.71	33.89	−12.15	13.3	23.4	2.92
4	**0.96**	**292.46**	**30.26**	**17.62**	**38.56**	**27.7**	**1.93**
5	*0.97*	*265.4*	*24*	*−23.32*	*27.79*	*24*	*3.24*
6	0.96	299.88	31.52	10.46	32.46	27.1	2.07
7	0.96	300.9	18	−19.98	23.99	18	3.15
8	0.97	375.94	38.46	−6.39	36.14	28.81	3.06
9	*0.97*	*375.19*	*38.48*	*−6.23*	*36.32*	*28.83*	*3.06*
Avg.	0.97	316.24	29.8	−7.21	26.83	23.45	
σ	0.01	48.11	6.72	13.38	9.74	5.16	

6.2.2 *Feature 2*

The second feature of Event 2 appeared at first sight to be a limb event to the North-East, however the fitting (Table 5) shows that it extends well into the region of bad data to the north of the Sun, and is actually centred near the central meridian. The error surfaces (Fig. 11) show that all of the parameters are well-constrained. As well as being a much larger event than we might at first suppose, it is also seen to have a very low speed of about 300 km s^{-1}.

Figure 12(a) shows the modelled h-t plot of the second feature of Event 2. This shows fewer data points, and, according to the model, is not detected by SMEI until it is almost at 1 AU. As with the first feature, the speed for this event is almost constant when detected by SMEI, with a variation within 60 km s^{-1}. Its LASCO association, shown in Fig. 12(c), was a high speed CME first appearing in LASCO on 30 November 2004 at 07:27 UT. It was not a halo, and occurred in the NE quadrant with a PA width of around 90°. As with Feature 1 we assume some deceleration has occurred between LASCO and SMEI, evidenced by the speed change from $>$800 km s^{-1} in LASCO to $<$300 km s^{-1} in SMEI. This explains the discrepancy between the onset time of the ICME obtained from a linear extrapolation of the SMEI h-t data and the timing of the associated LASCO event.

Given the structure and trajectory of this ICME, the model predicted that this event would not impact the Earth, a prediction supported by the absence of shocks or ejecta in the ACE spacecraft at the predicted time of arrival. The central location of this event was 30°N18°W. This is around 30° to the east of Feature 1, which erupted around 2 days 16 hours later. This is approximately the time it would take for a point on the Sun to rotate 30° to the west, which is consistent with a common magnetic structure on the Sun as a source for the CMEs responsible for Features 1 and 2. There was a C4.8 class X-ray flare observed in GOES at 9°N33°E. This is some 20° latitude and 50° longitude different from the central location of the CME. Once again it would appear there is a significant discrepancy between the location of the solar surface association and the location of the centre of CME structure.

The 3-D reconstruction of this ICME is shown in Figs. 12(d)–(f), with the complete 3-D reconstruction once again obtained from the viewpoint of Ulysses. The 3-D image shows

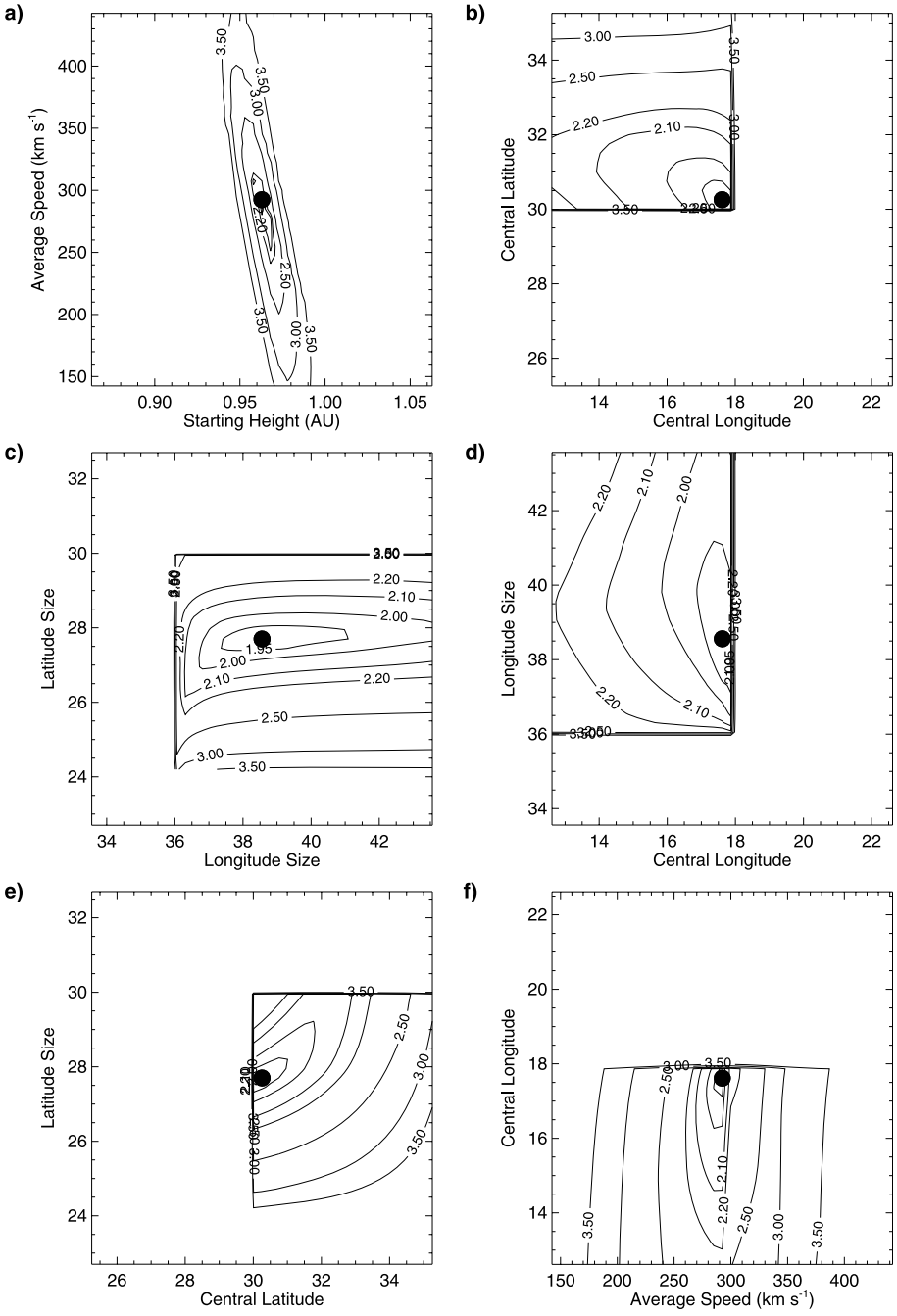

Fig. 11 Error surfaces for Event 2 Feature 2, the layout is the same as for Fig. 7

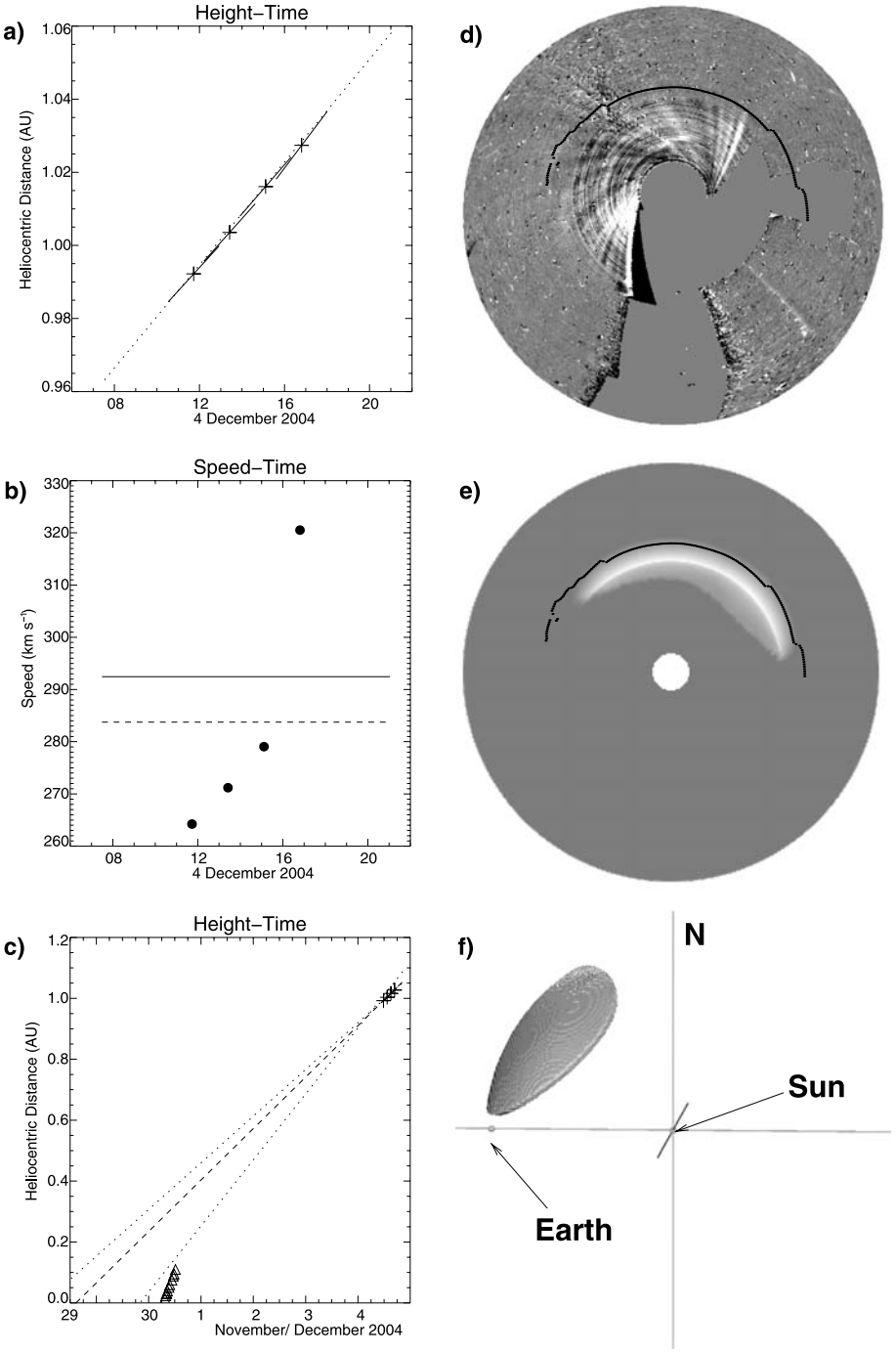

Fig. 12 Summary of the fitting for Feature 2 of Event 2. The format is identical to that of Fig. 10. The leading edge of the simulation in panel (**e**) is at 1.0 AU

Table 6 Stage 1 fitting parameters for Event 2 Feature 3. The starting heights are at 12:34 on 4 December

Run #	Start Height AU	Average Speed $km\,s^{-1}$	Central Latitude °	Central Longitude °	Distortion	Mismatch °
0	0.99	1045.31	0	−62.27	1.29	3.6
1	*0.93*	*754.2*	*0.69*	*−43.22*	*1.19*	*3.64*
2	0.98	931.7	0.38	−64.54	1.47	3.68
3	*0.99*	*1109.61*	*−0.39*	*−65.34*	*1.32*	*3.62*
4	0.99	1052.58	0.06	−62.57	1.3	3.6
5	0.94	964.34	−1.74	−45.59	1	3.62
6	0.96	979.42	0	−55.11	1.23	3.59
7	0.93	694.95	4.55	−42	1.34	3.72
8	0.96	967.9	−0.96	−52.8	1.16	3.6
9	**0.96**	**969.39**	**0**	**−54.47**	**1.23**	**3.59**
Avg.	0.96	946.94	0.26	−54.79	1.25	
σ	0.02	129.31	1.66	8.88	0.13	

that this event does not intersect the Earth or Ulysses, which is again supported by the in-situ data. We note, however, that this event does not miss the Earth by much, and that by extending its size by a small amount (as with Feature 1) an impact was predicted. In this case, we believe the 3-D structure is overestimated, probably due to the large geometrical extrapolation from the small amount of data available. A smaller structure is supported in the associated LASCO data, which shows a relatively narrow CME.

6.2.3 Feature 3

The third feature of Event 2 was seen as an arc to the east. This was best fitted by a somewhat flattened "bubble" disturbance (Table 6). This bubble was directed around 60° to the east of Earth and close to the ecliptic plane. The error contours (Fig. 13) show that all of the parameters are well-constrained, although as might be expected for a disturbance in this location the longitude and speed are somewhat correlated.

The h-t plot and speed are shown in Figs. 14(a)–(b). As with Features 1 and 2, the speed variation for this event is limited. The speed itself, however, is very high, allowing only a slight deceleration from LASCO to SMEI. This is consistent with the LASCO comparison plot in Fig. 14(c). The best LASCO association for this event is the same halo CME that was associated with Feature 1, thus allowing for the likely possibility that we are observing two parts of the same complex structure. For bubble models, determining if and when the disturbance will hit the Earth is rather more complicated than for shells. If the bubble is defined as in (2), then if the angle between the direction of propagation and the Sun-Earth line $\Psi = \cos^{-1}(\cos\lambda\cos\theta)$ is greater than $\pi D/2$ then the bubble will not contact the Earth. If the bubble does contact the Earth then the distance of the leading edge at the time of contact (assuming that the Earth is at 1 AU) is given by:

$$r_{\mathrm{le}} = \frac{1}{\cos(\frac{\Psi}{D})}.$$

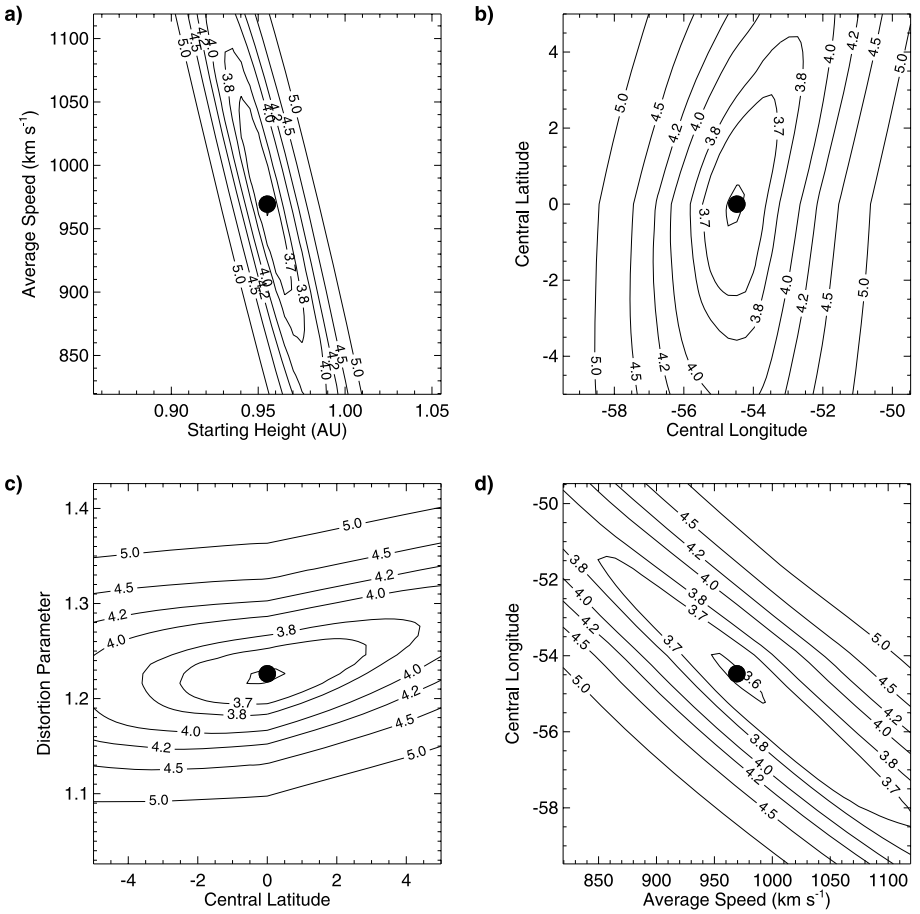

Fig. 13 Error surfaces for Event 2 Feature 3, (**a**) Starting height and average speed, (**b**) Central latitude and longitude, (**c**) Central latitude and "Distortion parameter" (Sect. 4.2), (**d**) Average speed and central longitude

The best fit model predicts that this ICME does hit the Earth, with an estimated arrival late on 8 December or early on 9 December, however since the predicted interaction with the Earth is at the extreme edge of the structure when the leading edge of the ICME is at 3.5 AU, this prediction should be treated with extreme caution.

The bubbles do not have a latitudinal or longitudinal size, and the central location of this event was 0°N54°E. This is considerably different from the location of Feature 1, suggesting that either the magnetic structure extended across a large part of the solar disk, or that there is an inconsistency in the location measurement for this event.

6.2.4 Feature 4

Feature 4 appears as an arc following Feature 3 out from the Sun and (at least from our viewpoint at Earth) appears to catch up with it. At those times where this feature could not be clearly distinguished from Feature 3, we flagged that region as missing data. We have not attempted to consider possible interactions between the features. This event was poorly

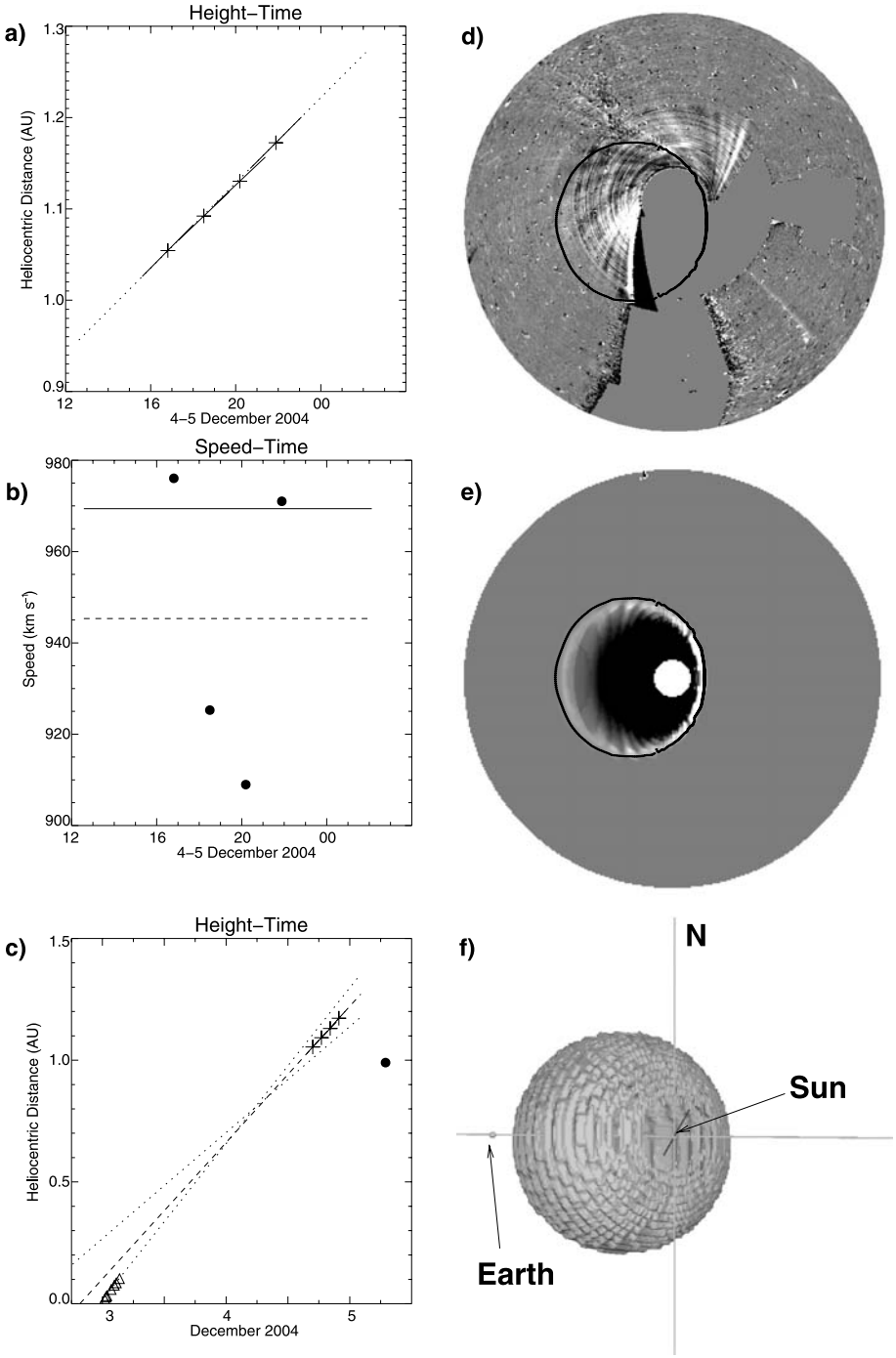

Fig. 14 Summary of the fitting for Feature 3 of Event 2. The format is identical to Fig. 10. In panel (**e**) the leading edge of the disturbance is at 1.05 AU

Table 7 Stage 1 fitting parameters for Event 2 Feature 4. The starting heights are at 02:07 UT on 5 December

Run #	Start Height AU	Average Speed $km\,s^{-1}$	Central Latitude °	Central Longitude °	Distortion	Mismatch °
0	*0.93*	*1048.45*	*−32.89*	*−32.18*	*1.48*	*7.74*
1	~~0.92~~	~~662.85~~	~~−12.19~~	~~−10.29~~	~~0.84~~	~~8.06~~
2	0.93	1053.88	−32.55	−31.98	1.47	7.74
3	0.92	900.18	−24	−21.68	1.22	7.84
4	0.9	924.54	−17.48	−12.53	0.96	7.71
5	0.94	827.92	−24.2	−30.81	1.38	8.09
6	**0.91**	**1036.17**	**−27.02**	**−22.16**	**1.24**	**7.83**
7	0.92	1017.75	−29.38	−25.43	1.32	7.8
8	*0.9*	*863.28*	*−14.28*	*−9.74*	*0.83*	*7.93*
9	0.92	1028.71	−29.61	−27.31	1.36	7.78
Avg.	0.92	966.76	−25.71	−23.76	1.25	
σ	0.01	87.83	6.45	8.16	0.22	

fit by both shell and bubble type models with the bubbles performing a little better. The mismatch parameters close to 8° indicate that neither form describes this disturbance well. The strong curvature of the leading edge would be expected to cause problems for a shell model. The fact that the fits we did obtain with the shell model showed a very elongated structure is a likely explanation of the mismatch in the bubble models since those do not allow for a non-circular section. The bubble fits are presented in Table 7, and show that the bubble was considerably flattened and centred to the south-east of the Sun-Earth line. The error surfaces (Fig. 15) show that with the exception of the speed, despite the poor overall fit, the parameters are well-constrained. Since there were only 5 orbits for which this feature could be traced, we should not be surprised at the poor constraint of the speed. The best fit model predicts that this event does not hit the Earth.

As we did not have enough SMEI measurements of this structure to complete a Stage 2 analysis for this event, we only have the constant speed solution with which to make a LASCO comparison. It is a high speed event, even at these large distances (1036 $km\,s^{-1}$), and its large apparent size implies it is likely to be associated with a halo or partial halo CME, of onset time around 12:00 on 3 December. The only LASCO halo CME around this time is the event associated with Features 1 and 3, around 00:00 on 3 December. With the modelled speed for this event, however, for this LASCO event to match the data, we would need a strong *deceleration* followed by an *acceleration* between LASCO and SMEI (Fig. 16). This seems unlikely, and it is more likely the speed for this event has been overestimated. It is difficult to identify such problems with this event as no Stage 2 processing was possible. Alternatively, it is possible that this event represents a later eruption that was not detected by LASCO, possibly because it was obscured by activity from the previous halo. The event certainly appears separately and later in the SMEI data, and is seen to catch up with the earlier Feature 3. This would imply that Feature 4 is a separate event from Feature 3, or at least a separate component of the same event.

The central location of this event was 27°S22°E. There was a minor (Class B8) flare reported by GOES with an onset time of 13:02 UT on 3 December, located at 18°S46°W.

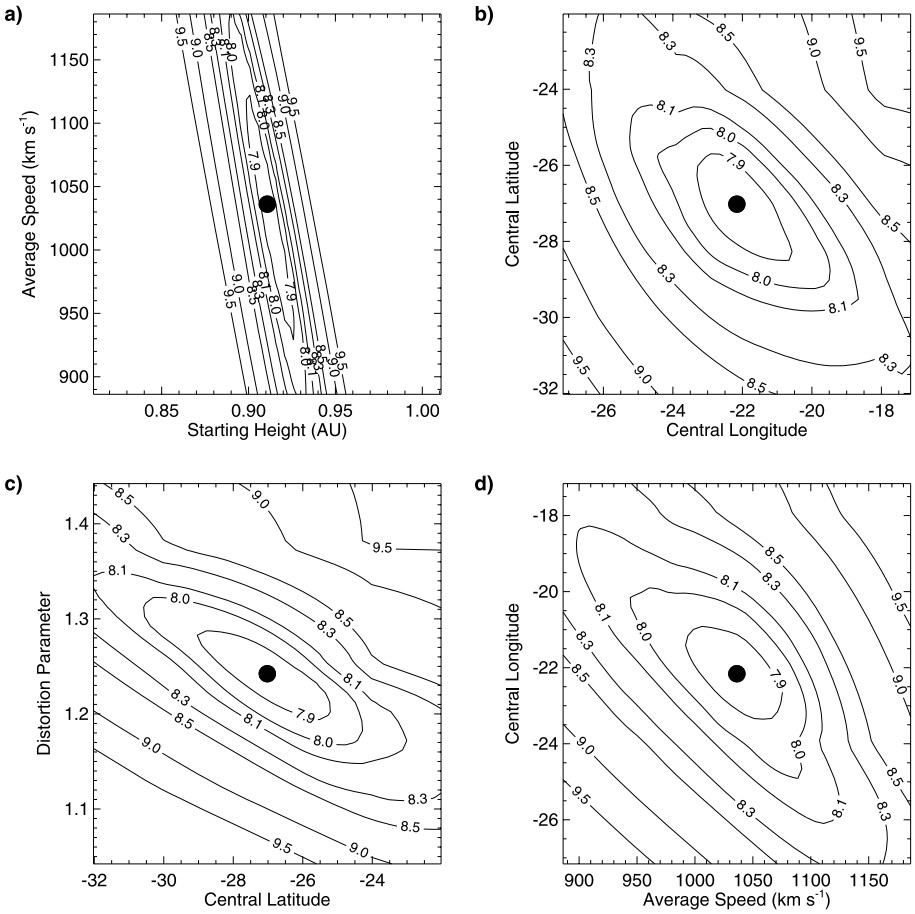

Fig. 15 Error surfaces for Event 2 Feature 4

This may be the signature of an eruption that was not detected by LASCO, but was in SMEI. Such events have been discussed by Howard and Simnett (2008).

7 Discussion

In this section we consider the performance of the model described, and its advantages and disadvantages compared with other methods of interpreting heliospheric images. We also consider ways in which the model may be made more realistic.

We have seen that for the two events that we have considered, this simple phenomenological model converges to believable solutions. The modelled solutions also agree well with other data sources, such as LASCO and ACE, which were not used in the fitting process but only examined after we had produced our solutions and predictions.

The chief advantage of this model over more sophisticated physics-based models is its speed, in principle it could be run by an operational forecaster and yield results sufficiently quickly to issue a forecast.

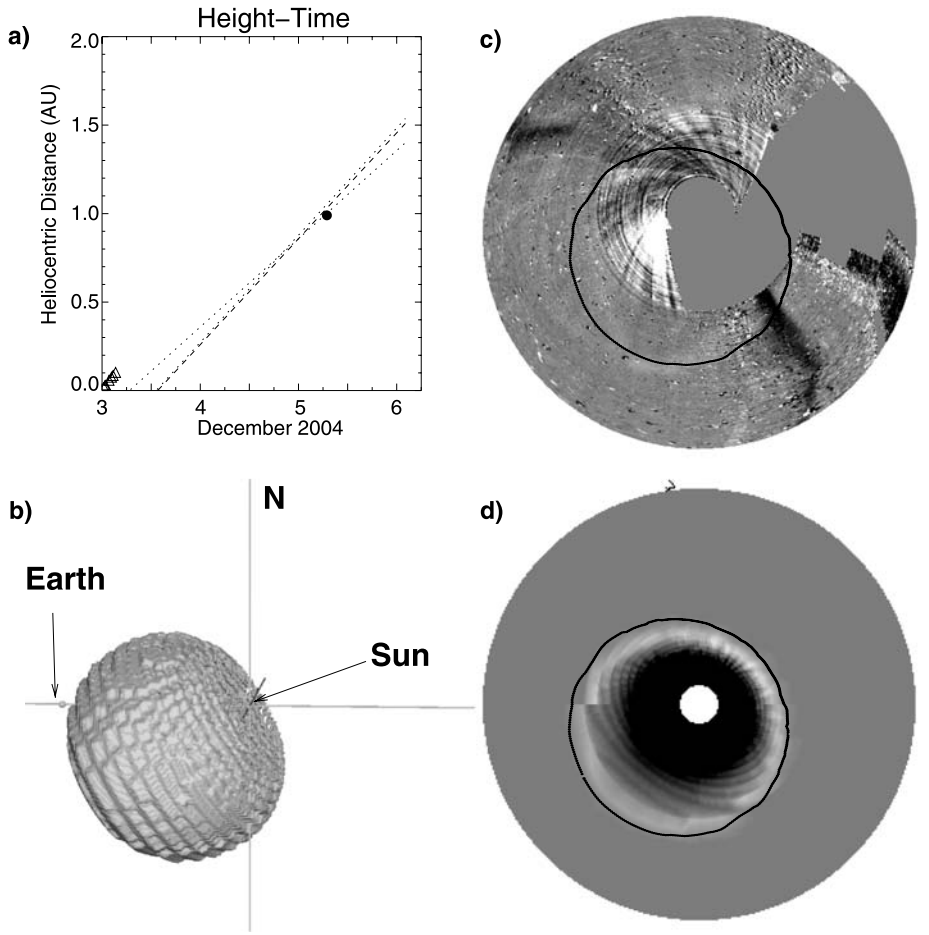

Fig. 16 Summary of the fitting for Feature 4 of Event 2. As there was no stage 2 fitting for this event the layout is somewhat different. (**a**) The height-time fit from the stage 1 fitting (*dotted line*) and the possible LASCO association (*triangles*) are shown. (**b**) a 3-D view of the modelled disturbance at 0800 on 5 December, (**c**) SMEI observations of the disturbance for the orbit covering 0711–0853 on 5 December with the calculated leading edge at 0800 superposed. (**d**) Simulated sky map for the leading edge at 1.05 AU with the 0800 leading edge superposed

This model also represents a decoupling of the physics associated with the appearance of the ICME from the physics responsible for the propagation if the ICME through the heliosphere. The coupling of these two processes is, we believe, the main reason why current models fail to accurately predict ICME evolution and arrival at 1 AU. The model presented here extracts the 3-D properties of the ICME directly from the heliospheric images alone, and as such does not consider the propagation physics or effects. This allows an unbiased study of these properties. The model therefore represents an advance in both the fundamental physics of ICME evolution and an improvement to space weather forecasting.

There are two major areas in which we might be able to improve the model. The first of these is in making a more realistic geometry for the simulated ICMEs. At present we have two distinct structure types, simple spherical shells and the bubbles (as seen in Fig. 3).

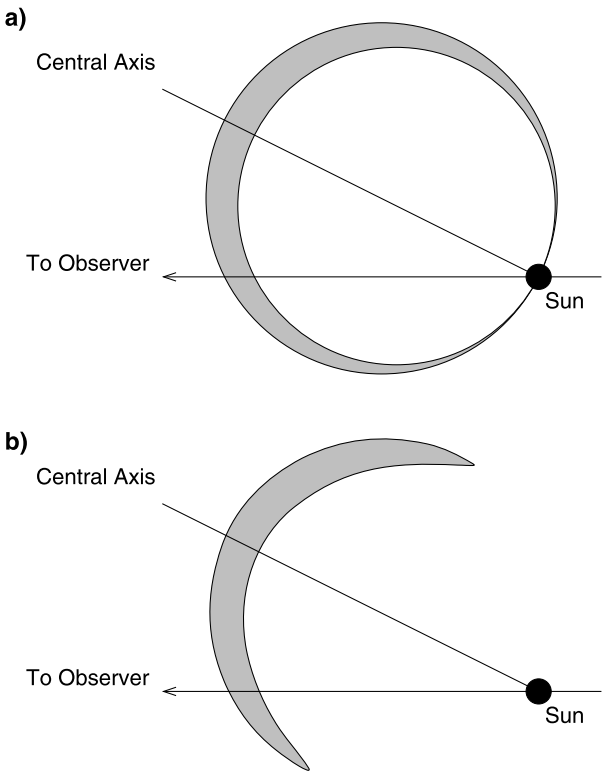

Fig. 17 A comparison of (**a**) what is modelled in a bubble ICME and (**b**) a more realistic interpretation (after Figs. 5 and 7 in Tappin 1987)

However the latter structure is not particularly realistic as it includes a density enhancement behind the widest part of the structure, and while this may occur close to the Sun it seems unlikely to persist as swept-up matter begins to dominate over the original ejecta. A more likely structure is that of a distorted shell, as illustrated in Fig. 17, where the front part of the real disturbance (Fig. 17(b)) follows the structure of a bubble, but the part behind the maximum extent is absent. We plan to modify the shell models to include an "excess curvature" parameter that will allow us to generate simulated ICMEs with a structure similar to that shown in Fig. 17(b).

A second concern with the model as it is currently implemented is evident in a number of the error surface contour plots (e.g. Figs. 11(b) and (d)), where there are discontinuities in the mismatch parameter at sizes and locations where the shell model moves across boundaries in the integration grid with consequent gradient discontinuities in the interpolation for the error contribution for the individual measurements. Because all of these occur at the same place they add up to produce the edgelike features that we see in the error contours. We note, however, that this does not occur for the bubble models where the match between the structure of the model and of the integration grid does not exist. This suggests that a possible way to improve this might to use a Cartesian grid in which none of the simulated ICME structures would have a strong match to the grid structure. This would make inserting the ICME structure into the grid more difficult, but the geometry of the integration process would be simplified.

In the final instalment of this series, Howard and Tappin (2009b) demonstrate the versatility of the model by extending its utility to include the STEREO spacecraft. This allows

a detailed study of the physical implications of the model and an accurate recreation of an event observed in late November 2007.

Acknowledgements SMEI was designed and constructed by a team of scientists and engineers from the US Air Force Research Laboratory, the University of Birmingham in the United Kingdom, the University of California at San Diego, Boston College and Boston University. SMEI data are maintained by the Air Force Research Laboratory (http://smei.nso.edu). SOHO is a project of international cooperation between ESA and NASA and LASCO data are made available courtesy of Goddard Space Flight Center and the Naval Research Laboratory. This work is supported in part by the National Research Council Fellowship Program, funded by AFOSR contract F49620-02C-0015. The National Solar Observatory is operated by AURA, Inc., under cooperative agreement with the National Science Foundation. Partial support for the NSO is provided by the USAF under a Memorandum of Agreement. We thank Max M. Hampson of North Carolina State University and the Research Education for Undergraduates program, funded by the National Science Foundation.

References

C.W. Allen, *Astrophysical Quantities*, 3rd edn. (Athlone Press, London, 1973)
G.E. Brueckner, R.A. Howard, M.J. Koomen, C.M. Korendyke, D.J. Michels, J.D. Moses, D.G. Socker, K.P. Dere, P.L. Lamy, A. Llebaria, M.V. Bout, R. Schwenn, G.M. Simnett, D.K. Bedford, C.J. Eyles, Sol. Phys. **162**, 357 (1995)
A. Buffington, J.S. Morrill, P.P. Hick, R.A. Howard, B.V. Jackson, D.F. Webb, Proc. SPIE **6689**, 1–6 (2007). 66890B
J.-P. Delaboudinière, G.E. Artzner, J. Brunaud, A.H. Gabriel, J.F. Hochedez, F. Millier, X.Y. Song, B. Au, K.P. Dere, R.A. Howard, R. Kreplin, D.J. Michels, J.D. Moses, J.M. Defise, C. Jamar, P. Rochus, J.P. Chauvineau, J.P. Marioge, R.C. Catura, J.R. Lemen, L. Shing, R.A. Stern, J.B. Gurman, W.M. Neupert, A. Maucherat, F. Clette, P. Cugnon, E.L. Van Dessel, Sol. Phys. **162**, 291 (1995)
M. Dryer, Space Sci. Rev. **67**, 363 (1994)
C.J. Eyles, G.M. Simnett, M.P. Cooke, B.V. Jackson, A. Buffington, N.R. Waltham, J.M. King, P.A. Anderson, P.E. Holladay, Sol. Phys. **217**, 319 (2003)
R.A. Harrison, D.G. Sime, J. Geophys. Res. **94**, 2333 (1989)
T.A. Howard, G.M. Simnett, J. Geophys. Res. **113**, A08102 (2008). doi:10.1029/2007JA012920
T.A. Howard, S.J. Tappin, Space Sci. Rev. (2009a), this issue
T.A. Howard, S.J. Tappin, Space Sci. Rev. (2009b), this issue
T.A. Howard, D.F. Webb, S.J. Tappin, D.R. Mizuno, J.C. Johnston, J. Geophys. Res. **111**, A04105 (2006). doi:10.1029/2005JA011349
T.A. Howard, C.D. Fry, J.C. Johnston, D.F. Webb, Astrophys. J. **667**, 610 (2007)
B.V. Jackson, A. Buffington, P.P. Hick, R.C. Altrock, S. Figueroa, P.E. Holladay, J.C. Johnston, S.W. Kahler, J.B. Mozer, S. Price, R.R. Radick, R. Sagalyn, D. Sinclair, G.M. Simnett, C.J. Eyles, M.P. Cooke, S.J. Tappin, T. Kuchar, D.R. Mizuno, D.F. Webb, P.A. Anderson, S.L. Keil, R.E. Gold, N.R. Waltham, Sol. Phys. **225**, 177 (2004)
J.B. Leighly, Geograph. Rev. **45**, 246 (1955)
R. Poli, W.B. Langdon, N.F. McPhee, *A Field Guide to Genetic Programming* (2008). ISBN 978-1-4092-0073-4, published via http://lulu.com, downloadable from http://www.gp-field-guide.org.uk
S.J. Tappin, Planet. Space Sci. **35**, 271 (1987)
S.J. Tappin, Sol. Phys. **233**, 233 (2006)
S.J. Tappin, A. Buffington, M.P. Cooke, C.J. Eyles, P.P. Hick, P.E. Holladay, B.V. Jackson, J.C. Johnston, T. Kuchar, D. Mizuno, J.B. Mozer, S. Price, R.R. Radick, G.M. Simnett, D. Sinclair, N.R. Waltham, D.F. Webb, Geophys. Res. Lett. **31**, L02802 (2004). doi:10.1029/2003GL018766
A. Vourlidas, R.A. Howard, Astrophys. J. **642**, 1216 (2006)
D.F. Webb, D.R. Mizuno, A. Buffington, M.P. Cooke, C.J. Eyles, C.D. Fry, L.C. Gentile, P.P. Hick, P.E. Holladay, T.A. Howard, J.G. Hewitt, B.V. Jackson, J.C. Johnston, T.A. Kuchar, J.B. Mozer, S. Price, R.R. Radick, G.M. Simnett, S.J. Tappin, J. Geophys. Res. **111**, A12101 (2006). doi:10.1029/2006JA011655
X.-P. Zhao, J. Geophys. Res. **97**, 15051 (1992)
X.-P. Zhao, J. Geophys. Res. **113**, A012582 (2008)
X.-P. Zhao, S.P. Plunkett, W. Lui, J. Geophys. Res. **107**, 1223 (2002)