Automatic recognition and characterisation of supergranular cells from photospheric velocity fields

H.E. POTTS, D.A. DIVER
Dept. Physics & Astronomy, University of Glasgow, Glasgow, G12 8QQ, U.K.

Received ; accepted

Abstract. We have developed an exceptionally noise resistant method for accurate and automatic identification of supergranular cell boundaries from velocity measurements. In this paper we describe the method, and test it against simulated data. We then apply it to the analysis of velocity fields derived from high resolution continuum data from the SOHO Michelson Doppler Imager. From this, we can identify certain basic properties of supergranulation cells, such as their characteristic sizes, the flow speeds within cells and their dependence on cell areas at exceptionally high resolution. The effect of the noise and smoothing on the derived cell boundaries is investigated and quantified using simulated data. We show in detail the evolution of supergranular cells over their lifetime, including observations of emerging, splitting and coalescing cells. A key result of our analysis of cell internal velocities is that supergranules appear to be scale-independent in this respect.

Keywords: Sun Photosphere, Supergranulation, Magnetic elements, Granules, Photospheric Flow

1. Introduction

The convection processes in the Sun have been studied for many years, with the first observation of the solar granulation made by William Herschel in 1801. Granulation is a small scale rapid convectional process ($\sim 1\,\mathrm{Mm}$, few minutes lifetime, $1\,\mathrm{km\,s^{-1}}$ typical flow speed) which has been well described and modelled. More recently, larger scale patterns with weaker flows have been observed superimposed on the basic granulation flow. The clearest of these is the supergranular flow, first discovered by Leighton, Noyes & Simon (1961) by analysing dopplergrams taken at the Mount Wilson Observatory. Supergranulation is a flow with a cellular form, probably convectional in nature, with characteristic size around 30 Mm, 20 hr lifetime and typical flow speeds around $300\,\mathrm{m\,s^{-1}}$. The supergranular pattern is most dramatically seen as patterns in the line-of-sight velocity of the solar surface as seen in full disk MDI dopplergrams from the SOHO satellite. The velocities in supergranulation, as seen in the photosphere, are predominantly horizontal. A consequence of this is that the supergranulation pattern in dopplergrams is only visible towards the limb of the sun, and disappears at disk centre. In order to observe the supergranulation at higher resolutions and near disk centre, analysis of the horizontal photospheric flow fields is required. Although supergranulation has been observed for nearly 50 years remark-

ably little is known about it – indeed it is not even certain if it driven by convection, as no temperature gradient has been measured across the cells. The best attempts to date at measuring any temperature variation have concluded that it must be less than 3K across a cell (Lin & Kuhn, 1992). Part of the difficulty in analysing such flows is that they are masked by the much more rapidly varying granulation flows which have large temperature variations, and the higher concentration of magnetic fields at cell boundaries.

With the advent of long time series, high resolution data from satellite missions such as SOHO and TRACE we can now track the motion within these larger structures. When looking at the large scale but weak motions found in supergranulation, it can be hard to interpret the data to identify coherent structures. In this paper show how supergranular boundaries can be constructed with confidence from flow fields, and once constructed how such structures can be exploited to derive further essential diagnostic information. The article is constructed as follows: Section 2 describes the algorithm for finding and displaying the cell boundaries, Section 3 tests the algorithm against simulated data, and compares results with a conventional divergence algorithm. Section 4 shows how the algorithm works on real solar data. Section 5 contains notes about the potential problems when using the method, and how to avoid them.

2. Cell analysis method

This Section describes a noise tolerant way of deriving the cell structure from a flow field by following the motion of tracker particles in the time-reversed flow field.

2.1. Data preparation

In order to show the supergranulation pattern and its time variation in detail, large amounts of high resolution photospheric velocity data are required. This may be obtained by tracking the motion of the granules as they are advected by the supergranular flow fields. Long, continuous time series of data are required in order to see the slow evolution of the supergranular flows, and high cadence, high resolution images are required to resolve the granulation patterns. As a result the only possible data source at the present time is the high resolution (0.6 arcsec/pixel, 1 min cadence) continuum data sets from the MDI instrument on SOHO. The images were processed to get the velocity field by tracking the barely resolved granulation signal using the *Balltrack* method (Potts, Barrett and Diver, 2003).

The data set used throughout this paper is from a 64 hour continuous run by MDI from 15-18th January 1997. The run consists of high resolution

continuum, magnetogram and dopplergram images taken with a 1 minute cadence. We use a 480x1024 pixel (4.8x10.2 arcmin), 33 hour (2001 frame) subset of this. This set was used as it gives the longest time series of a large area co-rotating with the Sun that is available at the current time. The continuum data was rigidly derotated at a rate of 0.243 pixels/minute and filtered for p-modes using a fourier filter, and the velocity then derived using *Balltrack*. The velocity samples obtained were smoothed in time by binning in 3 hour segments and spatially smoothed by convolving with a 2D gaussian with FWHM of 5.4 arcsec. Absolute calibration of the data was achieved by adding a small offset velocity to the raw granulation data during the derotation operation, and measuring the effect of this in the derived velocity. A discussion of this method and the intrinsic noise within the any measurement of the flow field may be found in Potts, Barrett and Diver (2003).

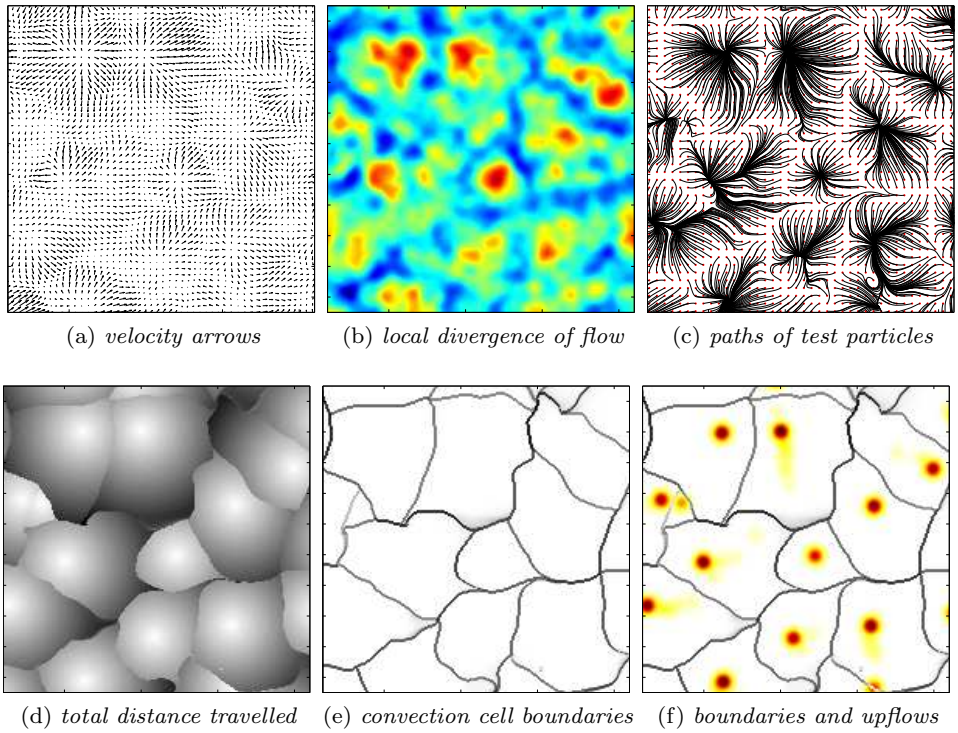
2.2. Overview

The images in Figure 1 show the steps in generating images of the supergranular pattern. Figure 1(a) shows the raw velocity field derived from a *Balltrack* analysis. Some cell structures, particularly the strong ones, are visible, but the cell boundaries are indistinct. Producing a divergence map (outflow regions in red) of the data as shown in Figure 1(b) helps to clarify the situation somewhat, although the cellular structure is still not clear

Part of the problem is that the data are intrinsically noisy: aside from observational errors, the motion itself has a stochastic element due to the turbulent nature of the small-scale flow. Any local method to find the inflows and outflows that requires taking the spatial derivative of the data is dominated by the small-scale features, at the scale used to smooth the data. This can be overcome by analysing the integral effect of the flow, as is usually done by using 'cork' like tracking particles. We use the fact that the flow patterns are asymmetric: mostly point-like sources going to line-like sinks. To exploit this asymmetry we take a regular array of starting points and send tracking particles flowing in the opposite direction to the streamlines. All trackers that end up at the same final area (corresponding to the cell upflow region) must be part of the same convection cell.

2.3. Method in detail

1. Take the initial velocity field, (see Figure 1(a)) $\mathbf{v}(x, y)$, reverse the flow direction, and normalise to the maximum speed of the flow field $\mathbf{v_n} = -\mathbf{v}/v_{max}$. It is important that the mean velocity of the data set (normally dominated by the rotation and differential rotation of the sun) is very much less than the flows due to the supergranules: we depend on the convergence of the tracers at the upflow regions of the cells, so careful



(a) *velocity arrows* (b) *local divergence of flow* (c) *paths of test particles*

(d) *total distance travelled* (e) *convection cell boundaries* (f) *boundaries and upflows*

Figure 1. Stages in the identification of the cell boundaries. The data shown is a 2x2 arcmin region near disk centre under quiet Sun conditions

derotation of the dataset is important. It may be useful to subtract the mean velocity, or a fitted differential rotation correction of the whole flow field from the data at this point to avoid artifacts. See Section 5 for more details.

2. Make a regular array of starting points at whatever final resolution is required for the cell structure. More points will make for a higher resolution, but take longer to calculate

3. Advect the test points with the reversed flow field. The tracks for a low resolution subset of start points are shown in Figure 1c. We use a simple second-order predictor-corrector method for efficiency. Choose a maximum step comparable to the correlation length of the velocity data for maximum efficiency. For maximum clarity enough time steps should be given for a test particle to travel from the edge of a cell to its centre. This process can be made numerically more efficient by the non-linear scaling of the velocity field, so that the particle takes less time to escape from the slow moving edges of the cell. One way to do this is to raise

the speed to a power less than unity, while maintaining the direction, for example we used $\tilde{\mathbf{v}}_\mathbf{n} = \hat{\mathbf{v}}_\mathbf{n} v_n^{0.5}$.

4. Taking the set of the final positions (x_f, y_f) of tracer particles mapped onto a grid of their initial positions (x_i, y_i), all particles that lie within the same cell will all record the same value of final position; adjacent particles that travel to different cells will record a different final position. Hence such a grid will contain regions in which the values change discontinuously. The gradient of this data grid will then reveal the cell boundaries. The quantity β expressed below quantifies this process:

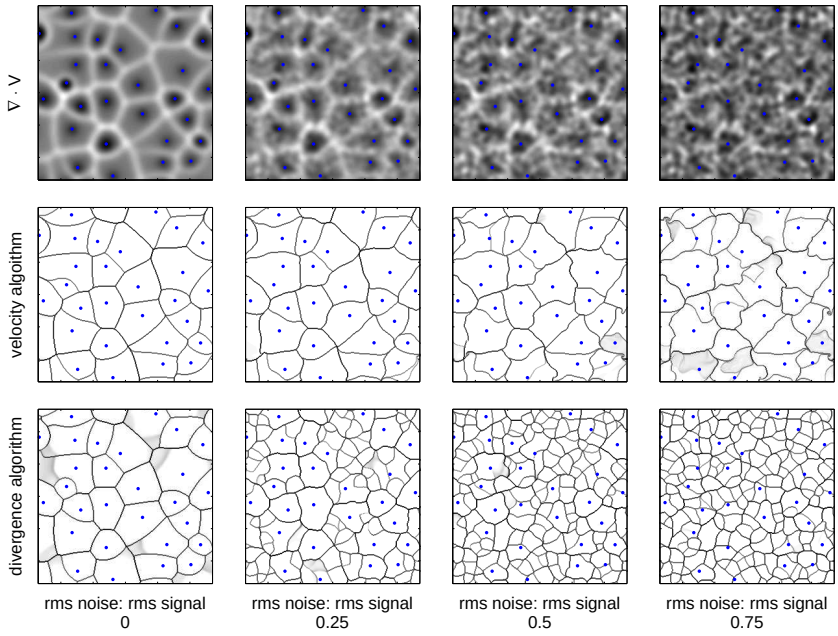
$$\beta = \left[\left(\frac{dx_f}{dx_i} \right)^2 + \left(\frac{dy_f}{dy_i} \right)^2 \right]^{\frac{1}{2}}. \quad (1)$$

A plot of a suitably normalised value of β is shown in Figure 1(e), clearly showing the cell boundaries. It is an exceptionally low noise measurement of the local divergence of the flow.

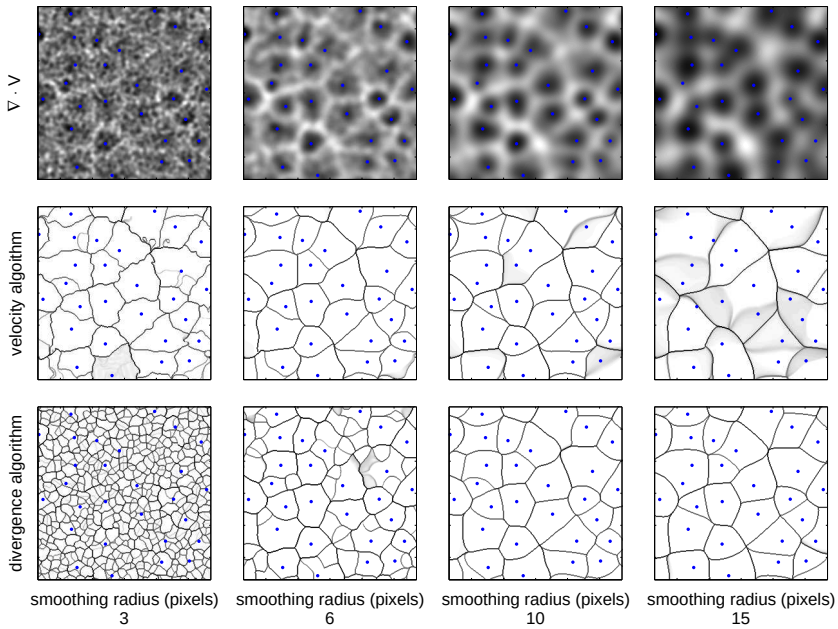
A property of a cell derived in this manner is that all the tracking particles end up in a similar area, which is the centre of the upflow for the cell. This is shown by the red blobs in Figure 1f. Notice that the distance travelled by the tracking particles is a minimum at these points, as can be seen in Figure 1d. These blobs are a smoothed image of the spatial density of the final positions of the tracking particles, so each one represents a separate upflow region. To find which cell any point on the surface belongs to, simply find out which of these regions the tracking particle ends up nearest to. The area of any cell is proportional to the number of tracking points that travel to this final location. In a movie which can be found here (*link to supergran_colour.avi*) (Potts, 2006a) each cell, as identified by the upflow regions, has been given a different random colour, and their time evolution can be clearly seen. The change in area of a few selected cells over time is shown in Figure 6, described in more detail in Section 4.3.

3. Application to test data

In order to test the accuracy of the algorithm it was run on test data with a known cell structure. The results were compared with the commonly used 'watershed basin' algorithm (DeRosa & Toomre, 2004; Hagenaar et al., 1997), which does a local minimum search in the divergence field of the flow. The test data was made to have similar properties to observations of photospheric velocity fields, at a resolution equivalent to the of the high resolution output from SOHO MDI. First, a velocity potential ϕ was made



(a) *increasing noise at smoothing radius = 4 pixels*



(b) *increasing smoothing at rms noise: rms signal ratio = 0.4*

Figure 2. Test data with increasing noise (top) and increasing smoothing(bottom). In each figure the top row is the divergence field, and the results from our velocity based segmentation algorithm are shown in the centre row, and the results from the watershed basin algorithm in the bottom row. The blue dots give the true outflow centres of the cells.

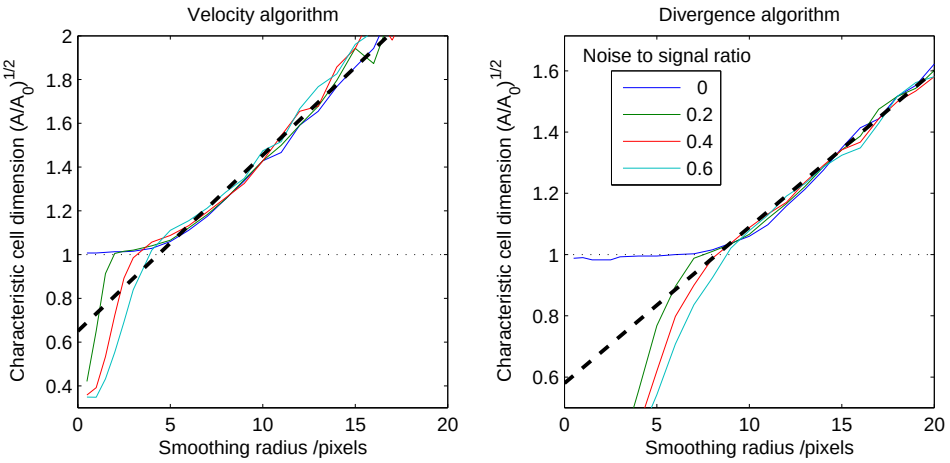


Figure 3. Recovered characteristic cell dimension as a function of smoothing radius for the velocity based method (left) and the divergence based method (right), for different noise:signal ratios.

by producing a Voronoi tessellation of cells from randomly placed generator points, with the value inside the cells a function of the distance to the cell centre. The flow pattern was then obtained by taking the x and y gradient of the potential field, and smoothed by convolving with a gaussian kernel with $\sigma = 3$ pixels to represent instrument effects. Noise was then added in variable proportion and the data smoothed again by convolution with a gaussian kernel of variable width.

The response of our supergranulation finding algorithm is summarised in Figures 2 and 3, along with the performance of the watershed basin algorithm, for comparison.

In Figure 2(a), the performance of the algorithm is tested as the noise:signal ratio is increased. The top row shows the divergence of the velocity field, the centre row shows the cell structure recovered by our velocity based algorithm, and the bottom row shows the results from the divergence based algorithm. It is clear that out velocity based algorithm has very high immunity to noise, in comparison to the divergence based methods, returning consistent and accurate results, even when the smoothed RMS noise amplitude is similar to that of the data. Note that the left-most plots are the zero noise case, where both algorithms recover the true cell structure of the noise-free test data, and so acts as a reference. In Figure 2(b), the effect of increasing smoothing on noisy data is presented. The test data in this case had fixed amplitude noise, equivalent to a rms noise:signal ratio of 0.4 when smoothed with a radius of 4 pixels. As the smoothing level is varied, the divergence based algorithm gives much better results at higher smoothing

radii. These effects are compared in more detail in Figure 3, where the effect of derived cell-size as a function of smoothing is analysed.

The detailed effect of smoothing is shown in Figure 3 for the two methodologies. Both algorithms show an average increase in the returned cell size as the smoothing radius is increased. The true result is a value of unity on the y-axis. Our velocity-based algorithm fares much better at low smoothing radii, but becomes more than 20% inaccurate as the smoothing radius exceeds 7 pixels. The divergence algorithm, conversely, has a complementary performance, showing poor accuracy up to a smoothing radius of 7 pixels, and above 15 pixels. Both algorithms show a linear trend in derived cell size as a function of smoothing.

All real data has an element of smoothing from a variety of unavoidable sources: for example, instrumental effects, seeing and noise reduction algorithms. To mitigate these effects in the data reduction, one approach that was first used by Hagenaar et al. (1997), and was also used by DeRosa & Toomre (2004) is to smooth the data at various different smoothing scales, and then extrapolate back to infer the true result that corresponds to the zero-smoothing case. This assumes that the effect of smoothing is linear in the returned cell size. Since we have test data, we can assess the efficacy of this technique. The results of this operation are shown by the thick dashed lines in Figure 3. It is clear from this that the trend of recovered feature size being proportional to smoothing radius is only linear at large smoothing radius, where coincidentally the watershed basin algorithm works well. However, using this linear regime on this test data to extrapolate to the zero-smoothing radius case significantly underestimates the true underlying structure size for the test data, mainly because the linear behaviour is not valid at small smoothing radius. This result leads us to conclude that mean cell diameters obtained by extrapolation in this way are not necessarily secure.

4. Application to real data

Here we show how an efficient and accurate supergranular cell finder can be applied to true solar data, and exploited to reveal additional properties of the photospheric flow field. The data set used here is the same as that in Section 2.1.

4.1. Cell sizes

In order to measure the mean cell size in the real data, the same procedure was followed as described for the test data in Section 3, where the data was smoothed over a range of radii. The results are shown in figure 4, where

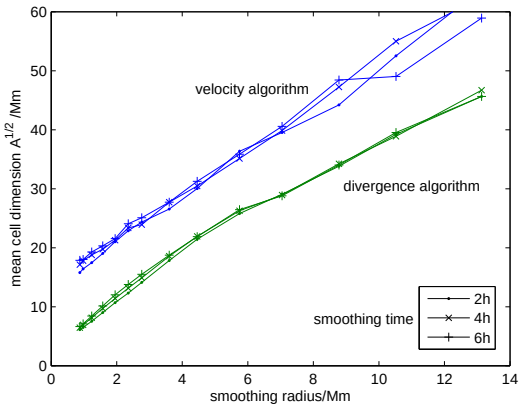


Figure 4. The effect on cell size of changing smoothing radius and time applied to high resolution velocity fields derived from MDI continuum data

the derived mean cell dimension is shown as a function of smoothing radius, for the two algorithms. There is a clear difference between these results and those of the test data: there is no sharp fall-off at small smoothing radii. This reflects the fact that there is no definite minimum scale for the features in the real data in comparison to the test data. The next clear effect is that there is almost no significant variation in the results at different smoothing timescales, showing that the supergranulation timescale must be significantly larger than our smoothing times, as observed by other authors (Del Moro et al., 2004; DeRosa & Toomre, 2004; Hagenaar et al., 1997).

For the divergence results, at smoothing radii less than 6Mm, the characteristic cell size is nearly proportional to the smoothing radius, as would be expected if it is dominated by noise. Thereafter the results show a similar linear form to that observed by DeRosa & Toomre (2004) at these smoothing radii, and when extrapolated back to zero smoothing yields cell dimension of around 8Mm, smaller than that from previous analyses of lower resolution data.

The velocity algorithm yields larger cell dimensions for a given smoothing radius, and remains linear over the whole smoothing range; unlike the divergence results, the velocity method shows evidence of real structure at low smoothing radii (right down to an unprecedented spatial resolution of 1 Mm with only 2 hours of temporal smoothing). The zero-smoothing extrapolation gives a cell dimension of approximately 15 Mm, in broad agreement with aforementioned previous studies in the literature. As the data becomes progressively noisier at smaller smoothing radii, this result gives a lower limit for the characteristic cell dimension, since noise can only decrease the derived dimension.

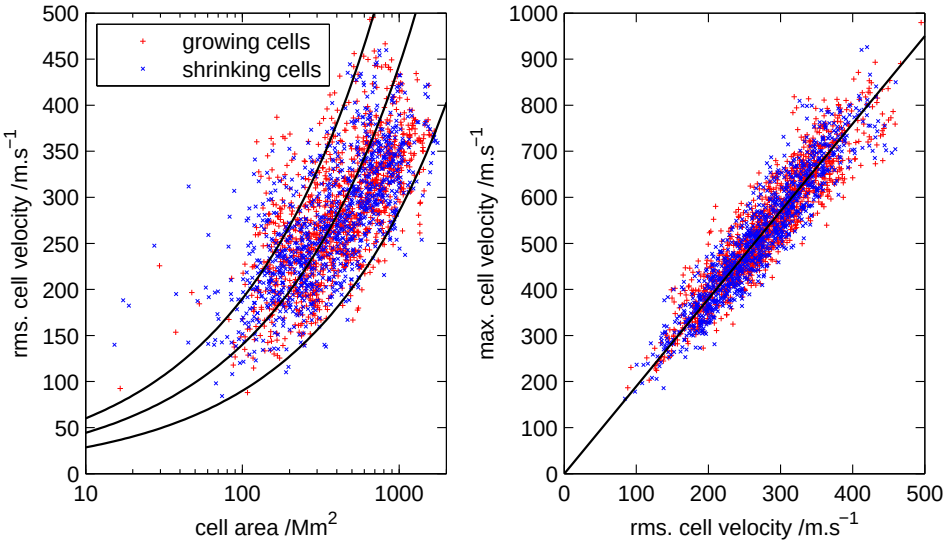


Figure 5. Relation between the area of supergranular cells and their velocities for 2500 cells. The left plot shows the rms velocity across the whole cell, and the right plot shows the maximum velocity inside the same cells. The lines represent a family of velocity relations that would be expected if the cells had identical form at all scales

4.2. Cell internal speeds

With an easy way to delimit the parts of the photosphere that belong to a particular cell other useful data can be derived simply from this. One such quantity is the velocity profile within supergranular cells. In Figure 5 around 2500 cells from the quiet Sun dataset described in Section 2.1 are analysed. The left and right plots show the internal rms speed versus cell area, and the maximum speed verses the rms speed, respectively for the same cells. Cells were chosen that could be tracked for a minimum of 3 hours, in order to determine if the they were expanding or contracting. The results are fairly clear: larger cells have larger rms speeds. A 1000 Mm^2 cell has an rms speed of around 375 $\mathrm{m\,s}^{-1}$, roughly twice that of a 100 Mm^2 cell. This is a similar trend as that seen in Paniveni et al. (2003), although we cover a substantially larger range of cell sizes and velocities due to our much improved cell recognition method. It is interesting and a little surprising to note that there is no significant difference in the distribution of cell internal speeds between cells that are growing or shrinking, as shown by the distribution of red and blue markers. Notice also that there are approximately similar numbers of growing and shrinking cells, which is discussed below.

From these data we also investigated how the velocity profile within the cells varies with cell size. If the cells of all sizes have the same velocity profile, we would expect the maximum and rms velocities in cells to be directly

proportional to each other, and the velocities to be proportional to the cell radius, or $\sqrt{A}$ where A is the cell area. These relations are shown by the families of black curves in Figure 5, and it can be seen that the data broadly conforms to them over the 100–1000 Mm^2 range. The linear relation between the maximum and mean speeds over the full area range (around 2 orders of magnitude) is significant as it suggests that the scale of supergranulation as observed here is not strongly constrained by some physical mechanism based on a critical scale size, as at around such a critical size we would expect the physical properties of the cells to be modified. The data suggests that there is in fact no particular scale length at which cells must become unstable, because at such a scale length there would be a departure from the self-similar relation: if cells split at a particular scale length then their internal rms speeds would be markedly different at that scale from those predicted by the trend.

4.3. Tracking supergranular cells over time

Identifying a cell at a particular time is interesting, but even a random velocity field, appropriately smoothed, would show some cell-like structures. Confidence that the cells obtained are real physical phenomena is gained by showing the evolution of the supergranular pattern over consecutive times using independent data sets. If the data can be processed this way it allows the growth and decay of the patterns to be monitored.

There are many problems with tracking supergranular cells over time, the most significant of which is that the cells are continuously splitting and merging. We have developed an algorithm that tracks the upflow points of the cells though time (see Figure 1(f)). This allows us to track the boundary of individual cells over long time periods (limited only by the available data), and stores all the branching and merging events from the entire data set. All other data about the cells is also stored, such as the cell area, velocities, centroid and upflow centre.

The evolution of some representative cells found by the algorithm over a 30 hour period are shown in Figure 6. Cell A is an unusual cell that grows rapidly over a 15 hour period from the intersection of several cell boundaries. It is worth noting that this cell is unusually free of magnetic field as it grows. Cell B is a cell that stays very stable over the 30 hour period. Cell C slowly breaks up and shrinks over the 30 hour period. Cells D are a pair of fairly large cells that merge to produce a single cell.

Note that although this data set was derotated so that the mean equatorial velocity was zero, there is a clear solar westward motion of all the cells, corresponding to a rotation rate which is around 30nHz faster than the rotation of the granules themselves. This is a well known phenomenon,

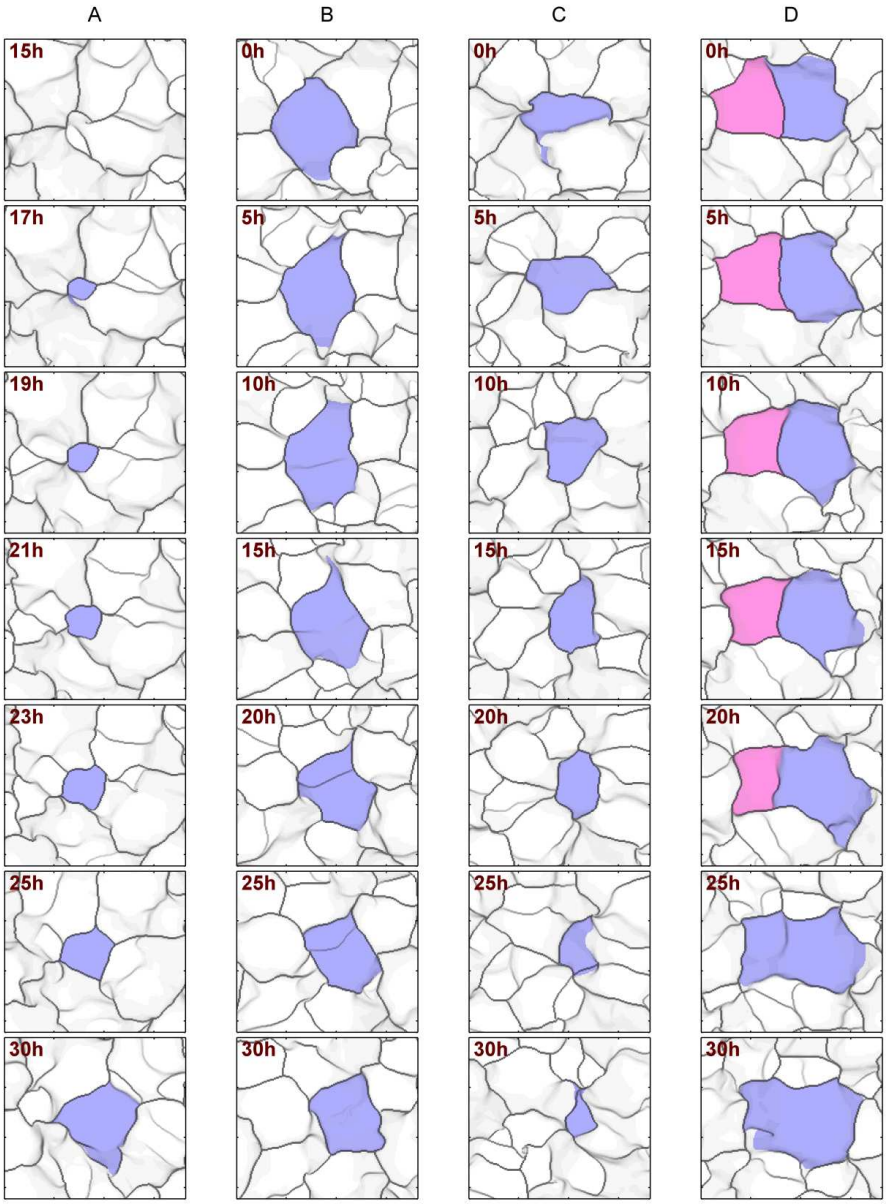


Figure 6. Evolution of a selection of supergranular cells over a 31 hour period. Each of the images is 110 × 100 arcsec. Each column shows a cell at different intervals after it was first observed, shown in hours at the top-left of each frame. This data set has been corotated with the Sun to set the mean plasma (granule) velocity to be zero, yet note the significant westwards motion of the supergranules, showing that they rotate faster than the fluid that forms them. A movie showing the time evolution of the entire data set over a 36 hour period can be found here (*link to supergran_colour.avi*). (Potts, 2006a).

(Thompson et al (2003), Beck (2000)), but as yet there is no clear consensus in the literature as to the underlying physics.

5. Cautionary notes

Whenever a continuous process is studied by looking at discrete time steps, the frequency of observation and the amount of smoothing in time greatly influences the results. Such influences can be highly significant in the sort of data processing addressed in this article: the flow field itself is obtained by spatially smoothing the small scale granular motions. This means that any derived supergranulation flow field obtained has been implicitly convolved with whatever temporal and spatial smoothing was used in the observation. This problem is ubiquitous in any measurement of supergranulation, whether from dopplergrams, or even from the measurements of the chromospheric network, where the lifetime of small magnetic elements imposes a natural timescale on the data.

It is also important to be very careful about how the velocity data used to derive the cells is derotated, due to the small values of the velocity field ($\sim 300\,\mathrm{m\,s^{-1}}$) in comparison to the rotation rate of the sun (equator speed $\sim 2\,\mathrm{km\,s^{-1}}$). For a weak supergranule the outflow speed near the edges is very small, so the apparent position of the lanes can change considerably due to a small derotation error. Small cells can completely disappear if the rotation offset is larger than their peak velocity. One way to help prevent this is to subtract the mean velocity with differential rotation corrections) from the velocity field.

6. Conclusions

We have developed a method for automatically identifying supergranulation cells, including an accurate measure of the position of lanes between cells and the upflow centre of the cells, without resorting to numerical differentiation of the data.

Since our method can work at exceptionally small smoothing radii, extrapolation to the zero-smoothing radius case is more secure than conventional algorithms, which tend to need high-smoothing to avoid domination by noise. This makes out method particularly well suited to the new generation of high resolution solar data.

We can track the cells over considerable times, limited only by the source data, and from this observe in detail their evolution. Over a 30 hour period we have observed cells growing, splitting, merging and shrinking.

There are many other applications for which the process outlined in this paper will be useful. For example the the problem of measuring any temperature differential across the cells will be greatly reduced by having accurate positions for the cell outflow and inter-cell lanes. Accurate measurements of the statistical properties of supergranulation, both over the supergranular evolution timescale and over the solar cycle will help expand our knowledge of this poorly understood process.

The study of small scale energetics will also benefit greatly from this method, particularly small scale magnetic interactions which are dominated by the solar surface flows. As high resolution Solar-B data comes available this will become a very interesting area to study.

Acknowledgements

We would like to thank Lyndsay Fletcher and Hugh Hudson for helpful discussions which have considerably improved this paper.

This work was funded in the UK at Glasgow University by PPARC rolling grant number PP/C000234/1.

References

Beck, J.:2000 *Solar Phys.* **191**, 47–70.

Del Moro, D., Berrilli, F., Duvall, T.L., Jr., & Kosovichev, A.G.: 2004, *Solar Phys.* **221**, 23–32.

DeRosa, M.L., Toomre, J.: 2004, *ApJ.* **616** 1242–1260.

Hagenaar, H.J., Schrijver, C.J., & Title, A.M.: 1997, *ApJ.* **481** 988–995.

Leighton, R.B., Noyes, R.W.,Simon, G.: 1961, *ApJ.* **135** 474–499.

Lin, H., Kuhn,J.R.: 1992, *Solar Phys.* **141**, 1–26.

Lui, Y., Norton, A.A.: MDI measurement errors: The magnetic perspective, *SOI Technical note 01–144*

Mehltretter, J.P.:1978, *A&A* **62**, 311–316.

Paniveni,U., Krishan, V., Singh,J., Srikanth,R: 2003, *MNRAS* **347**, 1279-1281.

Potts,H.E., Barrett,R.K., Diver,D.A.: 2003, *Solar Phys.* **197**, 69-78.

Potts,H.E.:2006 `http://www.astro.gla.ac.uk/users/hugh/supergranulation/index.html`

Potts,H.E., Khan,J.I., Diver,D.A.: 2006, *Proceedings of the Sixth Solar-B Science Meeting* eds. K.Shibata, S.Nagata, and T.Sakurai (accepted)

Rast, M.P.:1995 *ApJ.* **443**, 863–868.

Schou, J.:2003 *ApJ.* **596**, L259–L262.

Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I., MDI Engineering Team: 1995, *Solar Phys.* **162**, 129–188

Simon, G. W., Title, A.M., Topka, K.P., Tarbell, T.D., Shine, R.A., Ferguson, S.H., Zirin, H., The SOUP Team:1988 *ApJ.* **327**, 964–967.

Thompson,M.J., Christensen-Dalsgaard,J., Miesch, M.S., Toomre, J.: *Annual Review of Astronomy &Astrophysics,* **41***, 599–643*