

Techniques used in the analysis of data collected by the Birmingham Solar-Oscillations Network (BiSON)

II. Frequency domain analysis & data merging

W.J. Chaplin,¹ Y. Elsworth,¹ R. Howe², G.R. Isaak,¹ C.P. McLeod,¹ B.A. Miller,¹ and R. New³

¹ School of Physics and Space Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

² Now at: Queen Mary & Westfield College, University of London, Mile End Road, London, London E1 4NS, UK

³ School of Science & Mathematics, Sheffield Hallam University, Sheffield S1 1WB, UK

Received December 13, 1996; accepted January 2, 1997

Abstract. The Birmingham Solar-Oscillations Network (BiSON) consists of 6 solar observatories dedicated to the collection of full-disc solar Doppler velocity data – these data are sensitive to the lowest spatial degrees of oscillation ($0 \leq \ell \leq 4$). In Paper I of this series, we described the calibration of raw data from the BiSON instruments to produce solar velocity residuals representing the oscillations of the Sun’s surface. In this paper, we discuss the combination of velocity residuals into time series – including a thorough discussion of the treatment of data overlaps between sites – and the analysis of the power spectra computed from such time series to derive the characteristics of the acoustic modes.

Key words: Sun: oscillations — methods: data analysis

1. Introduction

The Birmingham Solar-Oscillations Network (BiSON) consists of 6 semi- and fully-automatic stations dedicated to the collection of full-disc helioseismological data (Chaplin et al. 1996a). Full details of BiSON are given in Table 1. There are two stations in each 120-degree longitude band. Moving successively east in longitude: Sutherland and Carnarvon are in band 1; Narrabri and Mount Wilson in band 2 (Mount Wilson sits on the band 2-band 3 boundary); and Las Campanas and Izaña are in band 3. All the sites lie at moderate latitudes.

The network spectrometers rely on the technique of resonance scattering to isolate two narrow passbands in the blue and red wings of a solar photospheric absorption line, formed by potassium atoms in the near infrared

(770 nm). The normalized difference between the measured blue and red intensities serves as a sensitive, near-linear measure of the line-of-sight velocity of the source. The BiSON instruments do not image the solar disc – the data collected are therefore sensitive to those solar modes possessing the largest spatial scales, i.e. for $0 \leq \ell \leq 4$. The radial wave functions of some of these modes penetrate to the core of the Sun.

The preparation of calibrated velocity residuals from Doppler velocity data collected by each of the network spectrometers, as described in Paper I of this series, is only the first step in the analysis of solar-oscillation data. In order to extract scientific results from the data, we need to combine the residuals into an extended time series, transform the time series into the frequency domain, and measure the mode characteristics in the spectra so obtained. This paper describes our approach to these tasks. In addition, we address the treatment of multi-station data overlaps between stations. Our discussion here concentrates on the signal-to-noise over the 5-minute region. Fossat (1992) considered a deconvolution treatment aimed at optimizing the signal-to-noise characteristics at low frequencies.

2. Combining residuals into time series

In order to investigate the detailed structure of the five-minute spectrum, it is necessary to combine residuals from many days and from more than one station. The first stage in this process is to produce time series of equally spaced data points. Because we changed our data-point interval from 42 to 40 seconds at different dates for different stations, and also because, even where the blocks are all of the same length, the data from different stations are not all synchronous, interpolation is sometimes necessary. Interpolation uses a simple function, of very low order compared with the number of data points, to estimate the value of a function between existing sample

Send offprint requests to: W.J. Chaplin

Table 1. The 6-station Birmingham solar-oscillations network (BiSON)

| <i>Site name</i> | <i>Latitude</i> (deg N) | <i>Longitude</i> (deg E) | <i>Altitude</i> (m) | <i>Commissioned</i> (year) | <i>Station type</i> [†] |
|-------------------------|----------------------------|-----------------------------|------------------------|-------------------------------|----------------------------------|
| Tenerife | +28.30 | −16.50 | 2368 | 1975 | A |
| Carnarvon | −24.85 | +113.75 | 10 | 1985 | B |
| Mt. Wilson [‡] | +34.13 | −118.07 | 1742 | 1992 | A |
| Sutherland | −32.38 | +20.82 | 1771 | 1990 | B |
| Las Campanas | −29.02 | −70.70 | 2282 | 1991 | B |
| Narrabri | −30.32 | +149.57 | 217 | 1992 | B |

[†] **A**: mirror-fed, semi-automatic system; **B**: equatorially-mounted, dome-based, fully-automatic system.

[‡] System moved from Haleakala, Hawaiian Islands, 1991.

points. Interpolation, by definition, involves some smoothing, which can be expected to multiply the computed spectrum by factors which change both the amplitude and phase of the Fourier components. We have tested three functions: a simple parabola, a cubic spline, and a sinc function truncated to six points on either side of the central maximum. We find that only the truncated sinc function adequately preserves the higher frequencies in the spectrum. Details are given by Elsworth (1991). The high-precision timing synchronisation needs for the network sites are handled by GPS receivers at Tenerife, Carnarvon, Sutherland, Las Campanas and Narrabri, and a WWV receiver at Mount Wilson. (The timing signals at Tenerife, Sutherland and Mount Wilson are provided by the host establishments.)

In choosing which data to include in the time series, it is convenient to quantify the quality of each day. A useful measure is the ratio of the power in the five-minute band (1.5 to 5.5 mHz) to the power above 5.5 mHz. Typical values of this “figure of merit” are: around 1.5 to 3 for Mt. Wilson, and for Carnarvon prior to a recent refurbishment; 5 for good data from Tenerife; 6 to 8 for Carnarvon after refurbishment; and between 30 and 50 for the best data from the new stations. This reflects the much improved high-frequency noise performance of the new instruments. This measure of data quality can be used both to exclude very poor data from the spectrum and to choose between stations when data from more than one are present at the same time. In general, data with a figure-of-merit less than 1 are worthless and are excluded at the daily fitting stage.

The use of data where there are substantial overlaps between stations has become important now that we have six stations operating. The total fractional time for which we had overlap coverage during the calendar year 1995 was 0.45. Overlaps of several hours can occur between stations; shorter three-station and (more rarely) four-station overlaps also occur (see later (Fig. 2)). The existence of

such overlapping data allows us to investigate the solar noise by cross-correlation techniques, as described by Elsworth et al. (1994a). It also means that we can check, and where necessary correct, the time keeping at the stations by looking for obvious lags between the residuals. This is possible only when the data from each site are of reasonable quality. With the highest-quality data, we can check the timing to within a few seconds by interpolating to find the maximum cross-correlation: where the data quality is not quite so high the timing can be checked only to the nearest sample. Historically, we simply chose the best set of data (that is, the one with the highest figure of merit) at each overlap for inclusion in the time series: we now go on to discuss a straightforward method of merging residuals from different sites, that takes advantage of the extra statistical precision afforded, over the 5-minute region, by the overlap data.

3. The merging of multi-site velocity residuals

The need for quasi-continuous coverage in the study of any periodic astronomical signal has long been recognized – the introduction of artefacts characteristic of the window function of the data is a well-known phenomenon.

Quasi-periodic gaps are usually treated in time-series analysis by the introduction of zero-valued residuals. In a time series whose window-function characteristics are dominated by a day-night cycle, power is then redistributed into a series of sidebands, which straddle signal peaks in the frequency domain at multiples of $\pm 11.57 \mu\text{Hz}$. This has two important consequences.

First, power leakage from the main signal peaks leads to a reduction in the signal to noise in the modes. This is clearly of much more importance near the low-frequency end of the p -mode spectrum, where the observed signal-to-noise ratios in full-disc helioseismology data are very

small, and therefore the impact on mode detectabilities – and attainable frequency uncertainties – is high.

Secondly, since the $\ell = 0, 2$ spacing in solar p -mode spectra is of the order of $\sim 11.5 \mu\text{Hz}$ near $n \sim 14$ (at about 2.1 mHz) the presence of sidebands can “pull” the fitted frequencies of modes, giving not-insubstantial systematic frequency errors.

Various techniques can be applied in an effort to remove – or at least reduce – the effects of the window function. Perhaps the best-known of these is the CLEAN algorithm (e.g., Roberts et al. 1987). In addition, given the assumed stochastic nature of the p -mode signal, a direct deconvolution (e.g., Lazrek & Hill 1993) can in principle also be performed on the data. Lazrek & Hill note that some of the assumptions used to justify this procedure might invalidate its use with real data. These authors present a series of simulations that have been applied to a single, isolated p mode with a high signal-to-noise ratio – the real solar p -mode spectrum is of course quite complicated, with a wide variety of signal-to-noise ratios observed over the p -mode domain. We are currently investigating the application of such techniques to realistic oscillation spectra, in particular, for modes with low signal to noise, and for p -mode fine structure spacings near $\sim 11.57 \mu\text{Hz}$.

Owing to the high- Q nature of the solar oscillatory modes, the filling of gaps with reliable estimates of the true data should be possible (Brown & Christensen-Dalsgaard 1990). However, due to the rich nature of the low-degree p -mode spectrum, i.e., many modes, at fairly close intervals in frequency space, Brown & Christensen-Dalsgaard note that the techniques they have investigated might not be wholly reliable. In addition, a reasonably high signal-to-noise ratio is required in the modes to give a good reconstruction of the missing signal. As the fill in a time series of velocity residuals decreases, so the requirements on the signal to noise needed for signal reconstruction become more severe. From the typical gap distributions observed in BiSON data, we find that by filling all gaps of 1 hr or less, we would improve our overall fill by ~ 10 per cent. It is clear from the observed signal-to-noise ratios in long BiSON spectra that low (say, below 1.5 mHz) and high-frequency (rather lower Q) modes might be precluded from any gap-filling analysis. This may also be true for the $\ell = 4$ modes we have observed in full-disc BiSON data (Chaplin et al. 1995b; 1996b). Nevertheless, we are investigating possible solutions to the gap-filling problem.

In an effort to provide reliable, round-the-clock coverage of the Sun, some degree of redundancy is desirable in a network. With several network sites fairly evenly distributed in longitude, there will be many instances where data are collected simultaneously at more than one site. These regions of “overlap” give rise to a potential gain in signal-to-noise. It is this point we wish to address in detail below.

3.1. The weighted combination of residuals

The principal aims of a multi-station network are to reduce the number, and length, of data gaps in a 1-dimensional time series constructed from the velocity residuals collected at all the sites, and to provide as much signal overlap between stations as possible. The presence of multi-station overlaps, i.e., stretches where data are collected simultaneously from more than one site, provides the opportunity to increase the signal-to-noise ratio in the modes by appropriately combining the overlap data from each site for inclusion in the final time series. The residuals in the overlap region must be properly weighted in order to take full statistical advantage of the extra data (see below).

The presence of low-frequency offsets between stations demands the application of techniques to remove discontinuities in the merged time series – this we do by high-pass filtering the velocity residuals (see below). Here, we concern ourselves with the statistical advantages of merging such data.

Let the appropriate weight (see below for discussion) for the overlap data from each site be W_j , and let there be n sites in the overlap (such that $1 \leq j \leq n$). With no low-frequency offsets between sites, the data can be combined in the time domain – over the region of the overlap – to give a combined velocity residual $v_{\text{com}}(t)$, according to

$$v_{\text{com}}(t) = \frac{\sum_{j=1}^n v_j(t) \times W_j}{\sum_{j=1}^n W_j}.$$

The data can also be combined in the frequency domain. Here, the frequency spectrum for each site – over the duration of the overlap – must be calculated. The weighted Fourier components at a given frequency are combined from each site, and the combined residuals are then recovered by taking the inverse Fourier transform of the combined Fourier spectrum. This is the approach we have adopted with the solar data, since the inherent noise in the solar velocity residuals – and therefore the appropriate weighting factor – can be determined by assessing power levels in the frequency domain. We have performed a variety of rigorous tests in order to quantify any systematic errors introduced by applying Fourier techniques on short data strings to recover the merged signal – we find that these are of the order of 10^{-7}ms^{-1} , i.e., at a level that introduces no significant artefacts into the data.

3.1.1. Signal content

In order to quantify the potential gains from a weighted merge, we must consider some straightforward statistics, and make some assumptions concerning the content of the velocity residuals.

The p -mode signal “seen” by each instrument is assumed to be identical – this assumption is not strictly true for the following reasons.

Consider two stations separated by several tens of degrees in longitude – any overlap between these sites will contain data collected during the early part of the day at one of the sites, and the latter part of the day at the other. The line-of-sight velocity of the Sun with respect to each observatory (the solar topocentric velocity) will therefore be different for each set of data in the overlap. Since the passbands of a basic resonance-scattering device are narrow compared to the width of the solar reference line, different parts of the photospheric line – and by implication, different depths in the solar atmosphere – will be sampled by each instrument. Since the amplitude of the p modes changes with height in the solar atmosphere, the measured signal amplitudes for each station in the overlap will differ.

In order to estimate the magnitude of this effect, we have used the calculated heights of formation of Underhill & Speake (1996). These authors determined the atmospheric heights of formation for the blue and red passbands of a BiSON resonance-scattering spectrometer, as a function of the line-of-sight velocity of the Sun. We assume that the effective height sampled by the spectrometer is the average of that sampled by the blue and red passbands. Furthermore, we assume that the exponential variation of the velocity amplitude of the modes with height can be characterized by a velocity-amplitude scale height H_v . Underhill & Speake found $H_v \sim 580$ km from an analysis of BiSON data. Taking this value, the estimated fractional difference in the mode velocity amplitudes for the extreme line-of-sight velocities 0 and ~ 1500 m s^{-1} is then approximately -0.07 . Since the difference in the line-of-sight velocities of “overlapping” sites is somewhat less than ~ 1500 m s^{-1} , we estimate that the typical fractional amplitude difference in overlaps should be of the order of only 2 to 3 per cent.

The separation of the spectrometer passbands is fixed by the Zeeman splitting induced by a permanent longitudinal magnetic field imposed on the potassium reference vapour in the instrument – different field strengths between network instruments would therefore give rise to a similar (small) effect.

There is also an additional line-of-sight velocity effect now known as Doppler imaging – here, the use of narrow spectral passbands when viewing the rotating Sun results in the “mapping” of different parts of the solar Fraunhofer line onto different parts of the visible solar disc (Brookes et al. 1978; van der Raay 1991).

We also assume that the predominant velocity noise source over the p -mode region has the characteristics of a normal distribution, and that the noise contributions are completely incoherent between stations. This, of course, does not quite fully apply, when, for example, the solar

background velocity noise is a substantial fraction of the total noise power (see later).

The non p -mode noise will be made up of contributions from photon shot noise, instrumental noise, atmospheric noise, and the solar velocity background noise continuum. The first three contributions serve to produce a combined noise background that is approximately Gaussian. However, the final contribution is common signal between each overlap station (allowing for some loss of coherence due to Doppler imaging). In addition, near the centre of the p -mode regime, there will be a large contribution due to the slowly-decaying Lorentzian wings of the modes, due to the presence of sidebands, and from a small, diffuse high- ℓ background. For the basic equations given below, we shall consider a single, isolated p mode – in a later section, the full effects of this additional power will be discussed.

Finally, we assume that there are no large, low-frequency discontinuities between residuals from different sites. As previously noted, for the purposes of the 5-minute analysis, the BiSON residuals are high-pass filtered with a moving mean that cuts in just below ~ 1 mHz. If differences between the observed signal from each station are assumed to be purely due to contributions from Gaussian-like noise, there is no need to employ techniques to “smooth” the joins between single-station and overlap parts of the time series.

3.1.2. Quantitative signal-to-noise gains

We now develop straightforward signal-to-noise expressions for an n -station overlap consisting of N data points. These will be expressed in terms of the straight ratio of the fitted mode-peak power to the mean background noise power level.

First we derive an expression for the noise power from single-site data. Let the sample standard deviation of the “white” noise source, in the time domain, be σ . Its magnitude is determined by the distribution of noise-source velocities v_i : for N points in the time series

$$\sigma^2 = \sum_{i=1}^N (v_i - \bar{v})^2 / (N - 1).$$

If the time series of white noise possesses a zero mean level, then for $N \gg 1$, the right-hand side of the above is simply the sum of the powers P_i in each real bin of the frequency domain of the Fourier transform of the data. Therefore, the average power per bin in the frequency domain – due to the Gaussian noise source – is given by

$$\sum_{i=1}^{N/2} P_i / (N/2) = \sigma^2 / (N/2). \quad (1)$$

If the time series also contains some periodic signal of interest, and if this signal gives rise to a peak in the

frequency spectrum of height H_ν , then we may express the signal-to-noise ratio, s/n , in the mode for the duration of the overlap according to

$$s/n \sim (H_\nu N/2) 1/\sigma^2. \quad (2)$$

Now consider the weighted merging of residuals. The weights must be set appropriately, in order to take full statistical advantage of the extra data available in the overlap. Here, the correct weights are proportional to the inverse of the square of the sample standard deviations of the noise sources for each of the overlapping time series.

If there are n stations in the overlap, each with a white noise source characterized by σ_j ($1 \leq j \leq n$), and if the signal is common, then Eq. (2) will be modified to

$$s/n \sim (H_\nu N/2) \sum_{j=1}^n 1/\sigma_j^2, \quad (3)$$

assuming that the overlap persists for the full duration of the time series.

Now consider the case where the overlap persists for some fraction, Ω , of the total time series. Let the *equivalent* sample standard deviation of the site data present in the single-station parts of the time series be σ_s . In addition, let the *equivalent* sample standard deviation for the overlap sections be σ_o – this is, of course, given by

$$1/\sigma_o^2 = \sum_{j=1}^n 1/\sigma_j^2. \quad (4)$$

The modified signal-to-noise ratio is then given by

$$s/n \sim (H_\nu N/2) \frac{1}{\Omega \sigma_o^2 + (1 - \Omega) \sigma_s^2},$$

which, substituting Eq. (4), gives

$$s/n \sim (H_\nu N/2) \frac{1}{\Omega \sum_{j=1}^n 1/\sigma_j^2 + (1 - \Omega) \sigma_s^2}. \quad (5)$$

A more realistic representation of a real network will consist of different overlap combinations and a variation in the quality of the single-station segments – under these circumstances, Eq. (5) must be expanded to a more generalized form in order to describe the resulting signal-to-noise ratio. Let there be a total of S single-station regions, each characterized by a fractional fill Ω_k , and sample standard deviation σ_k , such that $1 \leq k \leq S$; in addition, let there be a total of O regions of overlap, with fractional fills Ω_l , and sample standard deviations σ_l , such that $1 \leq l \leq O$. Here, each σ_l is derived from the sample standard deviations of the sites in the considered overlap (cf. Eq. 4). The generalized signal-to-noise expression for an “inhomogeneous” time-series composition is then:

$$s/n \sim (H_\nu N/2) \frac{1}{\sum_{l=1}^O \Omega_l \sigma_l^2 + \sum_{k=1}^S \Omega_k \sigma_k^2}. \quad (6)$$

What do these equations imply concerning potential quantitative improvements? Figure 1 shows the signal-to-noise gains for a variety of two-station overlaps, and fractional overlap fills (fills indicated next to each curve). The fractional signal-to-noise gains are calculated with respect to a time series in which data from the better site only are used, and are plotted with respect to the ratio of the sample noise standard deviations of the two stations in the simulated overlap.

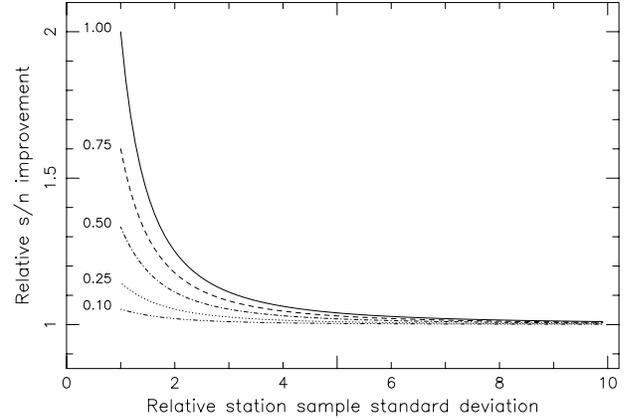


Fig. 1. Fractional signal-to-noise improvements, as derived from Eq. (6), for a variety of two-station overlaps, plotted with respect to ratio of sample noise standard deviations of the stations in the overlap – the fractional signal-to-noise gains have been calculated with respect to a time series in which data from the better station only is used. Each curve corresponds to a particular fractional overlap fill, as indicated on the plot. The sample standard deviation of the data in the single-station parts of the simulated time series is assumed to equal that of the better station

A weighted combination of overlap residuals from two sites of similar quality will lead to an improvement in signal to noise – over the region of the overlap – by a factor of 2 (Fig. 1). If the fractional time for which this type of overlap occurs is reduced from 100 to 50 per cent (assuming the single-station parts of the time series to be filled with residuals of similar quality), the signal-to-noise increase is reduced to a factor of ~ 1.3 . With a fractional overlap fill of only 10 per cent, the gain is only ~ 1.05 . The expected gains for overlaps with more stations can be readily extrapolated from these results.

What if the stations in an overlap are of different quality? This is of particular relevance here, since the quality of data differs from site to site within BiSON. With reference to Fig. 1, as the quality of the poorer site deteriorates, so the potential signal-to-noise gain in the overlap decreases. Once the ratio of the sample noise standard deviations for the sites increases beyond ~ 1.7 , the extra gain for all overlap fills shown in Fig. 1 will have halved (with respect to a two-site overlap in which the quality of the sites is the

same). We now go on to consider the effect of merging on BiSON data, i.e., on a complex time series with variable station quality and gaps.

3.2. BiSON overlap characteristics

Figure 2 indicates where multi-station overlaps were present during the calendar year 1994. Black pixels indicate no overlap; dark grey, the presence of a 2-station overlap; light grey a 3-station overlap; and white, a 4-station overlap. Most of the overlap hours are accumulated during the central UT band of each day, when the instruments at Sutherland, Izaña and Las Campanas are collecting data. The total fraction of time for which multi-station overlaps were present in the data for the full calendar year 1994 was ~ 0.45

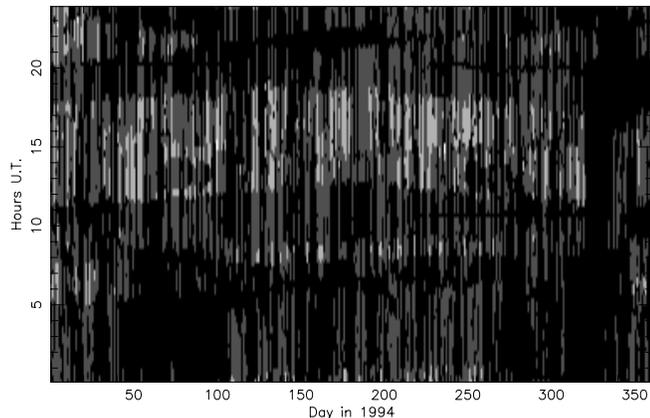


Fig. 2. Multi-station overlaps for 1994: black pixels indicate no multi-station overlap or no data; dark grey a two-station overlap; light grey a three-station overlap; and white a four-station overlap

As mentioned in the last section, the quality of the instrumentation varies between sites. The best-quality data are collected by the spectrometers at Sutherland, Las Campanas and Narrabri. Typical high-frequency noise powers for daily spectra generated from data collected at these sites are of the order of 2 to $4 \text{ m}^2 \text{ s}^{-2} \text{ Hz}^{-1}$. Data collected by the older instruments at Izaña, Carnarvon and Mount Wilson are characterized by somewhat-larger high-frequency noise powers. (This disparity is due largely to lower counting rates.) There can be as much as a difference of 10 in the high-frequency noise powers of data collected at the older and newer sites.

In order to assess the impact of weighted merging, we have taken the 2-month BiSON window function for the period 1994 July to 1994 August inclusive. The fill of useful residuals for this period was quite high (~ 87 per cent), with a ~ 50 -per-cent overlap fill. A damped harmonic oscillator, excited at regular intervals by a “white” forcing term, has been used as a model for the p -mode signal.

(Full details of this model can be found in Chaplin et al. 1996c).

This signal was assumed to be common to all sites over the duration of the real 2-month window function. Appropriate quantities of noise, with the spectral characteristics of a normal distribution, were added to the simulated p -mode time series for each site, in order to match the observed relative noise levels. No allowance was made for the increased levels of noise – caused by atmospheric extinction effects – at the extremes of the daily data. We note, however, that there is only a small deterioration in quality in the real residuals, since the regular day-to-day BiSON analysis excludes data collected at high air masses (greater than ~ 4). The artificial data – modulated by the real 2-month window function for each of the stations – were then merged to give a single time series.

The weights for each of the stations in a given overlap were fixed according to the high-frequency power (from 8 to 12.5 mHz) of the Fourier spectrum of the overlap data for the site. Assuming a Gaussian-like noise source, the mean high-frequency power per bin is, of course, proportional to the square of the sample standard deviation characterizing the noise source. The simulated time series was then constructed with the weighted-merging technique – as a reference, a time series was also constructed in which overlaps were treated by taking data from the best-quality station only.

For a single, simulated p -mode, forced through the real window function, with relative station noise levels characteristic of the current status of BiSON, the signal-to-noise gain given by using the weighted-merging technique instead of the best-quality-station-only technique was about 4 per cent. This rather modest gain can be readily explained by considering the make-up of the overlaps in the 6-station time series.

About 80 per cent of all overlaps over this two-month stretch occur between sites with an old and a new spectrometer. Only 16 per cent involve sites with at least two new instruments. The distribution of overlap types is similar for other 2-month stretches. As we saw in the previous section (and with reference to the derived equations), in order to provide a reasonable overall signal-to-noise gain, the noise level of the overlapping data must be of comparable quality. The overwhelming number of overlaps between old and new sites consequently restricts any gains to a low figure. By upgrading the older spectrometers, so that all instruments would be of similar quality, the expected overall signal-to-noise gain would increase to about 30 per cent.

In deriving the above gains, we have considered the signal to be composed of a single mode, and we have neglected the common signal from the solar velocity noise continuum spectrum. At low frequencies, where this common signal is high (Elsworth et al. 1994a), the solar continuum spectrum constitutes an insurmountable noise-floor level (the noise power due to the solar velocity

continuum is comparable with the expected levels of photon shot noise). Hence, the overall gain for a homogeneous network is reduced from 30 to about 17 per cent at frequencies near ~ 1.5 mHz; for the current BiSON configuration, the expected gain at lower frequencies is about 2 per cent.

Furthermore, we also recognise that the single-mode representation constitutes an over-simplification of the true picture in the frequency domain, i.e., the close frequency spacing of modes in the low-degree spectrum gives rise to a substantial amount of power in between the modes from the slowly decaying Lorentzian peaks. This is most prominent near the centre of the p -mode envelope, where the power from the wings of the modes is comparable to, or larger than, the incoherent background noise power arising from instrumental and atmospheric contributions.

The above quantitative calculations have been based upon various assumptions concerning the content of the real data, some of which, as has been pointed out, are questionable – they do, however, provide a useful guide to the expected quantitative gains made available by using all data from regions of overlap in the construction of a multi-station time series. When the weighted-merging technique is applied for real network data over the 2-month period covering 1994 July to 1994 August inclusive, there is no discernable signal-to-noise gain compared with a spectrum generated using the best-station-only overlap technique. This validates the simulations performed above with artificial data, which indicated one might expect very small fractional gains only. We do, however, remind the reader of the other benefits to be had from the overlaps generated by BiSON, i.e., the study of the solar continuum velocity noise; confirming the solar nature of large transitory or rapid excitation phenomena; and cross checking timing and calibration between stations.

We have demonstrated that in order to produce substantial signal-to-noise gains by weighted merging (with a view to increasing the quality of long Fourier spectra generated from the coherent analysis of many-station data): the noise level of the data from sites comprising a given overlap should be comparable, i.e., an upgrade of the older BiSON spectrometers would clearly be beneficial; and there should be a large amount of overlap coverage during each observing day, which demands the presence of at least two (ideally more) observing sites in each 120-degree longitude band (i.e., observational redundancy). This would also clearly be desirable in terms of the requirement for a high data-fill, simply via the increased redundancy afforded by a greater number of network sites.

When the data from each station are not high-pass filtered, techniques must be used to smooth or reduce low-frequency steps between sites. Sharp discontinuities can produce extra Fourier components in the vicinity of the mode peaks. While the low-frequency content of the velocity residuals contains large amounts of noise due to

atmospheric and instrumental contributions, there is also signal content of interest. The fundamental solar mode is expected to possess a period of the order of an hour or more; the motion of plages and sunspots across the visible solar disc give rise to a measurable signal with a periodic time scale of several days. On-going improvements to the low-frequency behaviour of our instrumentation are being made. However, complicated methods are still required to merge residuals which still retain very low-frequency information – these will be discussed in an up-coming paper in this series.

4. Power spectrum computation

Once the time series has been prepared, standard Fast Fourier Transform algorithms are used to obtain the power spectrum. Much of our work (Elsworth et al. 1990a,b, 1991, 1993, 1994b) has been based on two-month time series. This is a convenient length for several reasons: it gives a frequency resolution of about $0.2 \mu\text{Hz}$, which is suitable for studying modes with a natural width of the order of $1 \mu\text{Hz}$; the use of relatively short time periods also allows us to follow solar cycle trends, and to estimate the uncertainties of peak parameters from the scatter of many measurements. (For data obtained prior to the automation of the network, this was the longest suitable interval over which an adequate duty cycle could be obtained.)

We are now in a position to generate spectra over much longer periods with reasonable duty cycles, and these have enabled us to study the low-frequency end of the five-minute spectrum where the modes have small amplitudes and long lifetimes (Elsworth et al. 1995b; Chaplin et al. 1996d).

5. Power spectrum analysis

5.1. Power spectrum scaling

One must first choose a suitable scaling scheme to use for a spectrum that consists neither of incoherent noise nor of completely coherent sine waves, but of modes of finite lifetime, superimposed on a non-negligible noise background. The situation is further complicated because, in practice, a substantial fraction of the data in the time series is missing: our best 2-month spectrum so far has 94 per cent fill, but in some early years the best 2-month fills were around 25 per cent (early to mid 1980 s). For frequency measurements, of course, the scaling can be arbitrary, but in order to compare mode amplitudes obtained from different spectra the scaling needs to be carefully considered. There are two main methods of scaling power spectra: “power per unit frequency” and “equivalent sine wave”. In the former method, the total power in the spectrum is equal to the sum of the squares of the velocity residuals – or, in our case, to the sum of squares of the residuals divided by the fill to allow for the missing data. This method is appropriate when comparing incoherent noise levels.

When there are zero-valued residuals in the time series – owing to the presence of gaps – power will be re-distributed into sidebands characteristic of the Fourier transform of the resulting window function. The power in the central peak will therefore be reduced. The equivalent sine-wave scaling scheme compensates for this, by dividing again by the fill, to give an estimate of the power in the oscillatory signal that is independent of the window function. This method is appropriate for estimating the power in a given signal.

5.2. Window function

The window function can usefully be characterized by generating an artificial spectrum in which the real data in a time series are replaced by a single sinusoid of a frequency such that it appears in a single bin of the power spectrum. There is generally a dominant 24-hour period due to the regular interruption of the data by night, which gives rise to a series of sidelobes associated with each peak, spaced at 1/day or approximately $11.57 \mu\text{Hz}$. In the analysis of the low- ℓ p -mode spectrum, where the modes occur in pairs separated by about 5 to 15 μHz , this can be a serious problem, as one peak may overlap with the sidelobe of its neighbour. With a multi-station network, although some remnant of the twenty-four hour period persists, the breaks are more random and the 1/day sidelobes are greatly reduced in size. Figure 3 shows typical window function transforms for a one-station and a six-station time series. In our best six-station spectra, the sidelobes have all but disappeared into the window-function noise.

In addition to the discrete sidelobes, the window function contains background noise and some power in the wings of the central peak. The broadening of the central peak, which is particularly severe if there are long gaps in the data, can corrupt the measurement of mode strengths and must be corrected for when comparing mode amplitudes between spectra. This correction has been discussed in detail by Elsworth et al. (1993). The correction applied for 1/day sidelobes is discussed below under the heading of “sidelobe fitting”.

5.3. Power spectrum statistics and smoothing

The ragged appearance of the power spectrum of the modes arises from the stochastically excited nature of the oscillations and the fact that each Fourier spectrum is only an estimate, or realization, of the true spectrum, where the power in each bin consists of the underlying spectrum multiplied by a random number distributed as chi-squared with two degrees of freedom (Duvall & Harvey 1986; Anderson, Duvall & Jefferies 1990). If a large number of such realizations were averaged together, a better estimate of the true spectrum would be obtained, with errors tending to a Gaussian distribution. However, unlike observers studying the high-degree modes by resolv-

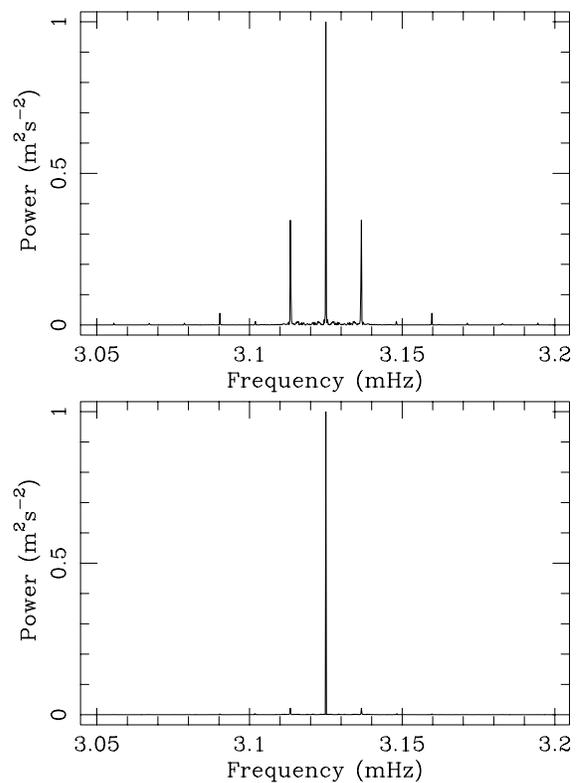


Fig. 3. Window function power spectrum for May through June 1996. Here, a sine wave with period 320 s and amplitude 1 ms^{-1} has been forced through each window function. Top: Tenerife only, fill 32 per cent. Bottom: Six stations, fill 78 per cent. The ordinate is equivalent sine-wave scaled

ing the image of the Sun, we have only a few dozen modes and two to three dozen spectra of reasonable quality to work with. In very early work the repeating pattern of the spectrum made it useful to average together corresponding portions of the spectrum at different frequencies. However, more detailed examination of well-resolved spectra reveal that the properties (including the repeat interval) of the spectrum vary considerably with frequency: the averaging of different frequencies is therefore of limited use for accurate work. Moreover, the properties of the spectrum vary with solar activity (e.g., Libbrecht & Woodard 1990, 1991; Rhodes et al. 1993; Elsworth et al. 1993, 1994b; Regulo et al. 1994): averages of many spectra from different epochs obscure this information and must be treated with caution. We therefore work mainly with individual spectra, or occasionally with the mean of two spectra close together in time.

As discussed by Anderson, Duvall & Jefferies, one approach to the treatment of individual spectra is to fit the model parameters to the peaks using a “maximum likelihood” method with the appropriate statistics, rather than normal least-squares fitting, which assumes a Gaussian distribution of the data points about the model. We have recently started to use this method

in studies of high-resolution spectra. In our earlier work (Elsworth et al. 1990a,b, 1991, 1993, 1994), however, we have adopted a different approach. The power spectra are first smoothed using a weighted moving mean (Pollard 1977). Pairs of peaks are then fitted to a model function using non-linear least-squares methods. A frequency range of about 50 μHz (about 250 frequency bins) is used to fit an $\ell = 0$, $\ell = 2$ pair, and a range of 70 μHz for an $\ell = 1$, $\ell = 3$ pair. As was discussed in our publication on mode strengths (Elsworth et al. 1993), the smoothing does introduce systematic errors in the height and width of the fitted peaks, but these can be corrected for. We have attempted to deconvolve the smoothing function by using the same smoothing on the model function as was used on the data, but this was found not to be effective. We have found empirically that the smoothing function chosen has no effect on a Lorentzian peak profile until the frequency range over which the smoothing is carried out exceeds the full width at half maximum of the peak. Thereafter the width of the peak is increased by the addition in quadrature of the effective width of the smoothing function, which is slightly less than one third of the full range of the function, while its height decreases.

We have carried out numerous tests in which artificial spectra, generated from known parameters, were smoothed and fitted in the same way as the real data. Although individual realizations behave in very different ways, the effect of the smoothing on the average parameters can be predicted from the behaviour of a simple Lorentzian peak under the same amount of smoothing, provided that the smoothing is not performed over too wide a range compared with the peak width. In practice, for two-month spectra, we use between 7 and 49 bins (about 1.5 to 10 μHz), using the heavier smoothing for the wider modes at the high-frequency end of the spectrum. The fitted peak parameters are corrected for the effect of smoothing as follows. The width obtained from the fit is first corrected for the effect of smoothing by subtracting in quadrature the effective width of the smoothing function. A Lorentzian peak of this corrected width and known height is then smoothed and fitted in the same way as the real data, in order to estimate the correction to be applied to the peak height. By comparing the results from the same spectrum fitted with different amounts of smoothing, we have verified that this procedure satisfactorily removes the systematic effects of smoothing.

5.4. Sidelobe fitting

The model function used to fit a pair of neighbouring peaks consists of two Lorentzian profiles, for which the adjustable parameters are the height h , the full width at half-maximum Γ , and the central frequency ν_0 , on a flat background of which the size b is also an adjustable pa-

rameter. The power, $P(\nu)$, in each single Lorentzian component, is given by

$$P(\nu) = \frac{h(\Gamma/2)^2}{(\nu - \nu_0)^2 + (\Gamma/2)^2} + b. \quad (7)$$

Associated with each peak is a pair of symmetrical sidelobes at 11.57 μHz on either side of the central peak. The sidelobes are assumed to have the same width and Lorentzian shape as the central peak, but their height relative to the main peak is left as an adjustable parameter for each peak. The relative height of the sidelobes can vary significantly from one peak to another. This is partly because of interference with neighbouring modes and noise, but we believe that a real effect is also involved: the amplitude of individual modes can vary dramatically from one week to the next, so that if a particular mode happens to have been strong for a period when only one or two stations were producing good data, it will appear to have stronger sidelobes than one that was strong when the network was running well. A more rigorous method of allowing for the sidelobes would be to convolve the model function with the known window function spectrum at each step of the minimization. This form of deconvolution would be much more expensive in computing time, and makes no allowance for the sidelobe variability discussed above.

5.5. Maximum-likelihood estimation and rotational splitting

A commonly used method for extracting mode parameters from power spectra is to use a maximum-likelihood estimator that takes into account the exponential distribution of the data about the underlying spectrum. As discussed above, we have found our own smoothing and least-squares fitting methods adequate for obtaining frequency measurements from 2-month power spectra. However, in most of this analysis we have neglected the rotational splitting of the modes of $\ell > 0$ and treated all modes as single Lorentzian peaks. As the rotational splitting is not clearly resolved – particularly in 2-month spectra – for frequencies above about 2.2 mHz, and there is neither strong empirical evidence for, nor any theoretical reason to expect (at least from a simple model) asymmetries in the amplitudes of splitting components, this seemed to us a reasonable approximation. However, when we wish to analyse higher-resolution spectra to derive the rotational splitting of lower-frequency modes, the smoothing method becomes less appropriate. As discussed in Elsworth et al. (1995a), we have therefore chosen to use maximum-likelihood estimation on power spectra covering periods up to, for example, 16 months, with model functions that explicitly include the splitting components at the theoretically expected amplitude ratios calculated by Christensen-Dalsgaard (1989). To satisfy the statistical requirements of some aspects of the fitting procedure,

we vary the logarithms of some of the parameters, rather than the parameters themselves.

The calculation of formal errors for this kind of fit, where neither the parameters nor (because of incomplete fill) the data points are completely independent, is complicated. We have in the past chosen to use the scatter of many determinations from independent spectra to estimate the errors of the peak parameters. This becomes more difficult with long spectra, of which we have only a few examples.

A thorough discussion of the errors in BiSON data can be found in Chaplin et al. (1996e), in which we outline our reservations regarding the use of formal uncertainties.

6. Automation of fitting

Any non-linear fitting procedure requires an initial set of guessed parameters, which are then refined to find the best estimate. The initial guess often has to be supplied by the operator. We have found, however, that in the case of five-minute p -mode spectra a high degree of automation can be applied. Although the frequency, width and amplitude of each peak vary from spectrum to spectrum, with a random variation overlying the systematic solar-cycle effects, the overall pattern is consistent. The frequencies change by no more than one or two micro Hertz from spectrum to spectrum, and their average values for the stronger modes are well established: average mode widths (from less than one to several micro Hertz) vary across the spectrum in a known manner. The approximate ratio of sidelobe to peak height for each spectrum can be estimated from the window function spectrum. The only quantity for which adequate first guesses cannot be predicted from experience is the mode amplitude, which exhibits large random variations due to the stochastic nature of the excitation process as well as a 30 per cent systematic variation (Elsworth et al. 1993) over the solar cycle. We therefore proceed as follows.

For each peak pair, the frequency range to be included in the fit, and the guessed frequencies and widths, are read from a standard control file. The first-guess frequencies are suitably modified to take into account the part of the solar cycle over which the data comprising the time series to be analysed were collected. The peak heights are then estimated by analysing the power distribution in a region of a few micro Hertz around the first-guess frequencies. The program then refines these initial guesses, and the final parameters are written out to file. In this way, the spectrum can be fitted without intervention, in a few minutes of microcomputer processing time. The high degree of automation built in to the basic analysis of our data consequently makes the up-to-date maintenance of records of mode parameters very straightforward.

7. Concluding remarks

We have presented and discussed several techniques for analysing calibrated Doppler velocity residuals collected in integrated sunlight, with the principle aim of studying the phenomenology of the solar low-degree, p -mode oscillations. In addition, we have discussed a method for taking advantage of the extra statistical precision afforded, over the 5-minute region, by the simultaneous collection of data at more than one site – even with a substantial fill of such “overlaps”, the analysis demonstrates that any gains are somewhat reduced if the quality of the data from each of the sites comprising the overlap differs.

Acknowledgements. This work was funded by a PPARC grant. We would like to thank all those who are – or have been – associated with BiSON. In particular: H.B. van der Raay, H. Williams, J. Litherland & R. Lines in Birmingham and P. Fourie at SAAO; and our hosts R. Stobie (SAAO); the Carnegie Institute of Washington; the Australia Telescope National Facility (CSIRO); E.J. Rhodes (Mt. Wilson, California); and members (past and present) of the IAC, Tenerife. Some of the research presented above made use of the Birmingham node of STARLINK.

References

- Anderson E.R., Duvall T.R., Jefferies S.M., 1990, ApJ 364, 699
- Brown T.M., Christensen-Dalsgaard J., 1990, ApJ 349, 667
- Brookes J.R., Isaak G.R., van der Raay H.B., 1978a, MNRAS 185, 19
- Chaplin W.J., Elsworth Y., Howe R., et al., 1995a, in: Fourth SOHO Workshop: Helioseismology, Hoeksema J.T., Domingo V., Fleck B. & Batrick B. (eds.) ESA SP-376, Vol. 2, p. 391
- Chaplin W.J., Elsworth Y., Howe R., et al., 1995b, in: Fourth SOHO Workshop: Helioseismology, Hoeksema J.T., Domingo V., Fleck B. & Batrick B. (eds.) ESA SP-376, Vol. 2, p. 387
- Chaplin W.J., Elsworth E., Howe R., et al., 1996a, Sol. Phys. 168, 1
- Chaplin W.J., Elsworth Y., Howe R., et al., 1996b, MNRAS 280, 1162
- Chaplin W.J., Elsworth Y., Howe R., et al., 1996c, MNRAS (in press)
- Chaplin W.J., Elsworth Y., Isaak G.R., et al., 1996d, MNRAS 282, L15
- Chaplin W.J., Elsworth Y., Howe R., et al., 1996e, MNRAS 280, 849
- Christensen-Dalsgaard J., 1989, MNRAS 239, 977
- Duvall T.L. Jr., Harvey J.W., 1986, in: Seismology of the Sun and the Distant Stars, Gough D.O. (ed.). Reidel, Dordrecht, p. 105
- Elsworth Y., 1991, BiSON Technical Report Series, No. 2, BiSON, Birmingham, UK
- Elsworth Y., Howe R., Isaak G.R., McLeod C.P., New R., 1990a, Nat 345, 322
- Elsworth Y., Howe R., Isaak G.R., McLeod C.P., New R., 1990b, Nat 347, 536
- Elsworth Y., Howe R., Isaak G.R., McLeod C.P., New R., 1991, MNRAS 251, L7

- Elsworth Y., Howe R., Isaak G. R., McLeod C.P., et al., 1993, MNRAS 265, 888
- Elsworth Y., Howe R., Isaak G.R., et al., 1994a, MNRAS 269, 529
- Elsworth Y., Howe R., Isaak G.R., et al., 1994b, ApJ 434, 801
- Elsworth Y., Howe R., Isaak G.R., et al., 1995a, in: proceedings of GONG 94, Ulrich R.K., Rhodes E.J. & Däppen W. (eds.) p. 51
- Elsworth Y., Howe R., Isaak G.R., et al., 1995b, Nat 376, 669
- Fossat E., 1992, A&A 262, 443
- Lazrek M., Hill F., 1993, A&A 280, 704
- Libbrecht K.G., Woodard M.F., 1990, Nat 345, 779
- Libbrecht K.G., Woodard M.F., 1991, Sci 253, 152
- Pollard J.H., 1977, A Handbook of Numerical and Statistical Techniques. Cambridge University Press, Cambridge, p. 126
- Regulo C., Jiminez A., Palle P., Perez Hernandez F., Roca Cortes T., 1994, ApJ 434, 384
- Rhodes E.J. Jr., Cacciani A., Korzennik S., Ulrich R.K., 1993, ApJ 406, 714
- Roberts D.H., Lehar J., Dreher J.W., 1987, AJ 93, 968
- Underhill C.J., Speake C.C., 1996, Sol. Phys. 163, 231
- van der Raay H.B., 1991, A&A 252, 366