Estimating uncertainties in helioseismic far-side phase shifts

Frank Hill

Dec 11, 2015

Abstract

The technique of measuring the phase shifts of solar oscillation modes that reflect from the far-side of the Sun oriented away from terrestrial observing platforms is increasingly being used for space weather forecasting since the phase shits are sensitive to the presence of large active regions. In order to provide useful forecasts with probability limits, it is essential to have uncertainty estimates, both random and systematic, for the measurements. This paper describes a method to carry out the estimates.

1 Introduction

Basics and history of far-side

2 Data and analysis

The essential approach is to compare different sets of measurements of the phase shifts obtained at the same time and over the same solar areas, but with different input data and different analysis techniques. The measurements used in this study are based on two data sets and two analysis techniques, providing four sets for comparison. While both classes of uncertainties are undoubtedly present in all cases, here we assume that comparing measurements from the same analysis methods but different data sets are dominated by random uncertainties, while a comparison of results using a common data set but different analysis methods primarily provides estimates of systematic uncertainties.

The two data sets are from SDO/HMI and GONG, while the two analysis methods are the holography method of Lindsey and Braun, but with different preprocessing choices. We designate the sets by the letters H and G, and the analysis methods by subscripts 1 and 2, so the four comparison sets are H_1, H_2, G_1 and G_2. Both H and G cover the date range of 1 Oct 2014 to 25 Jan 2015, providing a total of 233 far-side phase shifts maps. Figure 1 shows a sample pair of H_1 and G_1, emphasizing the different dimensions ($H_1 : 361 \times 181$, $G_1 : 500 \times 200$) and the different coordinate systems in longitude L and latitude

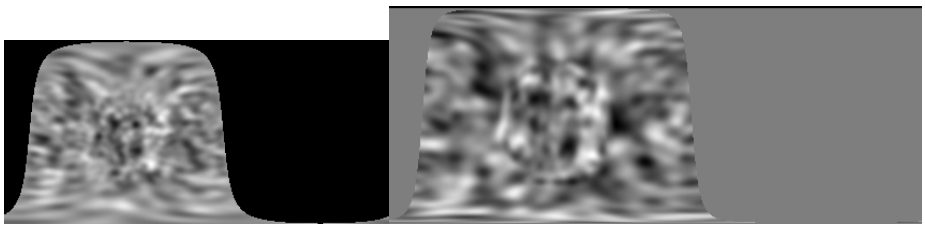

Figure 1: Example far-side phase-shift maps from SDO/HMI (left, H_1) and GONG (right, G_1) for 0 UT on 20-Oct-2014. Note the different dimensions ($H_1 : 361 \times 181$, $G_1 : 500 \times 200$), different coordinate systems ($H_1 : (L, B)$, $G_1 : (L, \sin B)$), and different near-side values ($H_1 : -0.262144$, $G_1 : 0.0$).

B ($H_1 : (L, B)$, $G_1 : (L, \sin B)$) of the two maps. An additional difference between H_1 and G_1 is the value assigned to those spatial locations in the maps which are not on the far side, but instead are on the near side. For H_1 this value H_{min} is -0.262144, while for G_1 the value G_{min} is 0.0.

The analysis is conceptually simple. First, the G maps are linearly interpolated onto the H grid of 361 points in L and 181 points in B to reconcile the raw maps. Second, the points in the interpolated G maps that correspond to the near-side points in the H maps as denoted by H_{min} are identified and, along with the H near-side points, are ignored in the subsequent processing. The resulting two maps, H' and G', are the objects that are then compared with a number of statistical comparisons including the Pearson cross-correlation coefficient P, a least-squares linear fit between H' and G' (but currently one that incorrectly assumes that the independent variable is error-free), the histograms and histogram moments of H' and G', and a scatter plot of H' vs G'.

3 Results

Figure 2 shows the scatter plot and histograms for the maps shown in Figure 1. For this example, P is 0.605, the linear fit is $G = 0.544H + 0.0226$, and the histogram moments are $(-0.0332, 0.00229, -0.138, 0.399)$ for H' and $(0.00453, 0.00186, 0.0374, 0.0850)$ for G'. The presence of the active region substantialy changes the third (skewness) and fourth (kurtosis) moments of the histograms, suggesting that these quantities may be useful for automatically detecting the presence of an active region in the maps.

The results shown in Figure 2 are for a time when there were no significant active regions on the far side. Figure 4 shows the scatter plot and histograms for the maps with a large active region shown in Figure 3. For this example, P is 0.677, the linear fit is $G = 0.593H + 0.00919$, and the histogram moments are $(-0.0386, 0.00315, -0.908, 3.66)$ for H' and $(-0.0137, 0.00242, -1.36, 7.78)$ for G'.

Figure 5 shows the complete sequence of P for the data set. The vertical line

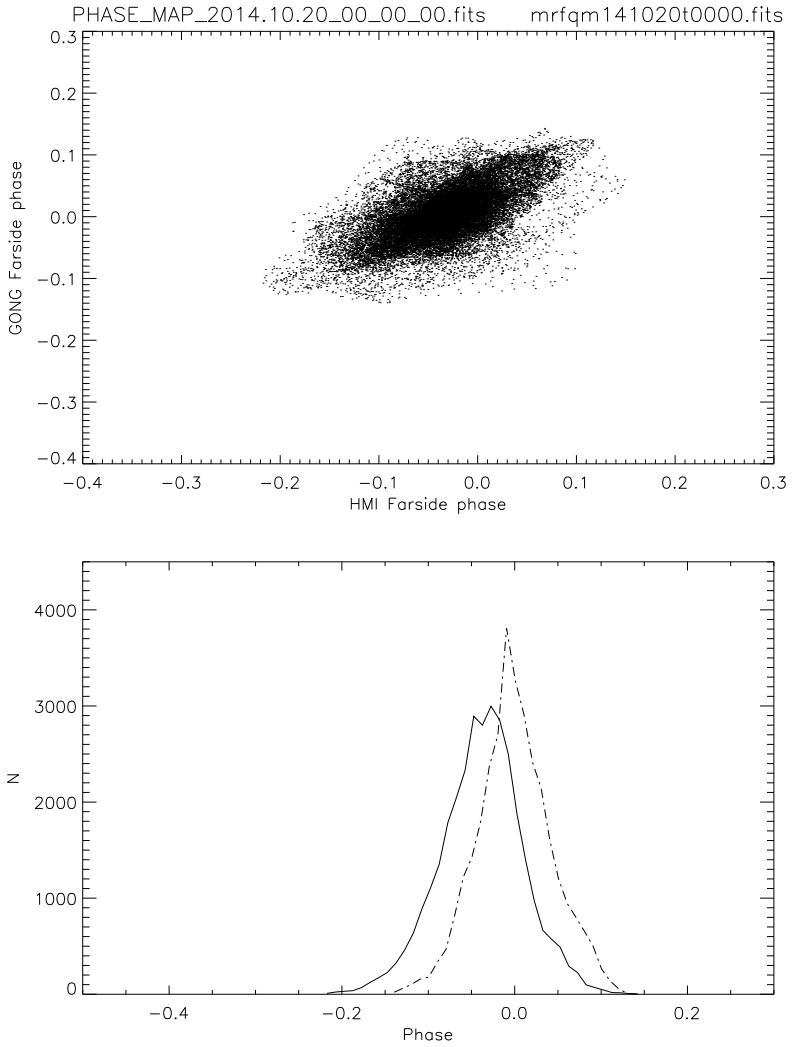

Figure 2: Scatter plot and histograms for the maps shown in Figure 1. In the histogram plot, the solid line is for H' and the dash-dot line for G'.

Figure 3: Example far-side phase-shift maps from SDO/HMI (left, H_1) and GONG (right, G_1) for 12 UT on 05-Nov-2014. A large active region was present in these maps.

indicates the time of the maps in Figure 3. There is a local maximum when the strong active region is present, but it is not particularly noticeable. Figures 6 and 7 show the coefficients a and b of the least squares fit between H' and G', $G' = bH' + a$. Again, there is no noticeable feature at the time when the large active region is present.

Figures 8 and 9 show the histogram moments for H' and G', respectively. Here, the kurtosis shows a noticeable maximum and the skewness shows a minimum when the active region is present. The absolute value of the product of these two moments are shown in Figures 10 and 11, suggesting that this quantity may be an effective way to automatically identify far-side maps containing large active regions.

We can decrease the uncorrelated background and increase any persistent correlated signal in the far-side maps by summing all of the 233 maps used in this study. The result of this is shown in Figure 12 for HMI (H_{sum}) and GONG (G_{sum}). These sums show active region bands and a number of long-lived regions. After interpolating G_{sum}) onto the HMI grid, we can produce the scatter plot comparing the phase shifts as seen in Figure 13. This plots shows a number of clusters of points, with the active regions clearly visible as the long tail towards the lower left of the plot. Another cluster is probably the background noise components, but there are three "'mystery region"' clumps whose origin remains to be determined. The characteristics of these clusters, such as their widths, can be used to estimate the uncertainties in the phase maps.

4 Discussion

There are a number of fairly obvious next steps in this analysis.

1. Inspect the maps that produce large values in Figures 10 and 11. What is the percentage of these maps that actually contain active regions?

2. Map the points in the clusters in Figure 13 back to the images in Figure 12 and see what features they correspond to.

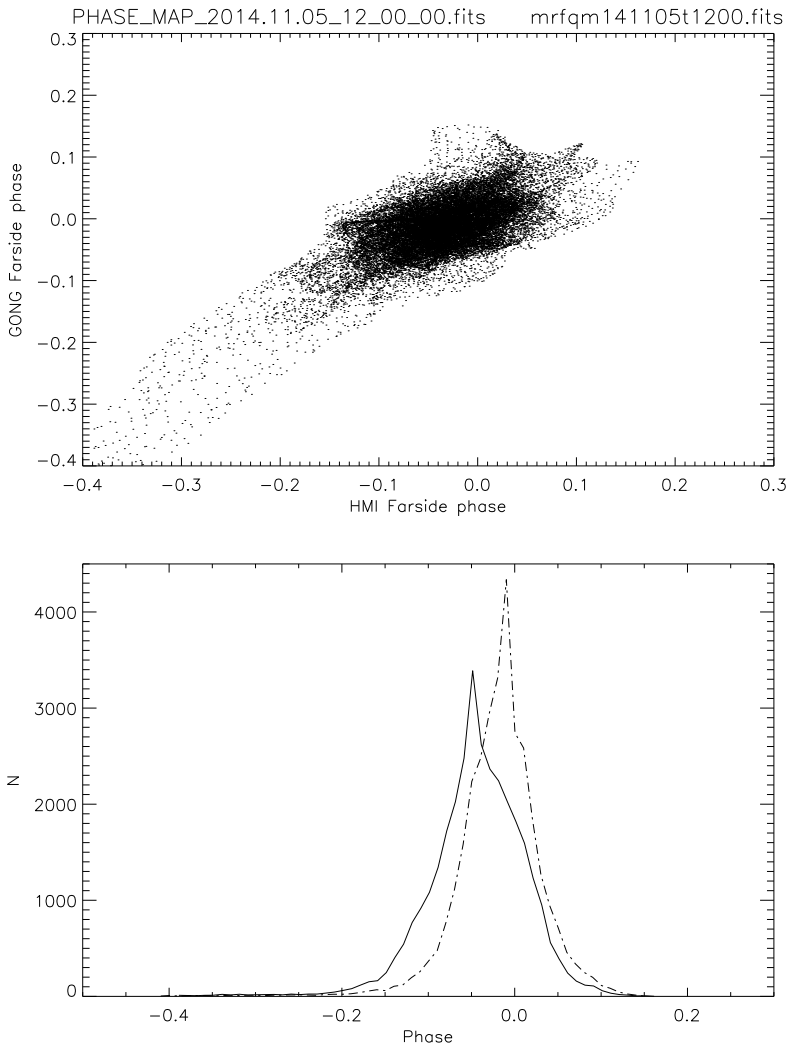

Figure 4: Scatter plot and histograms for maps containing an active region. A tail-like structure extending to large negative phase shifts is visible in the scatter plot. The histograms display substantially larger skewness and kurtosis when compared to those in Figure 2.

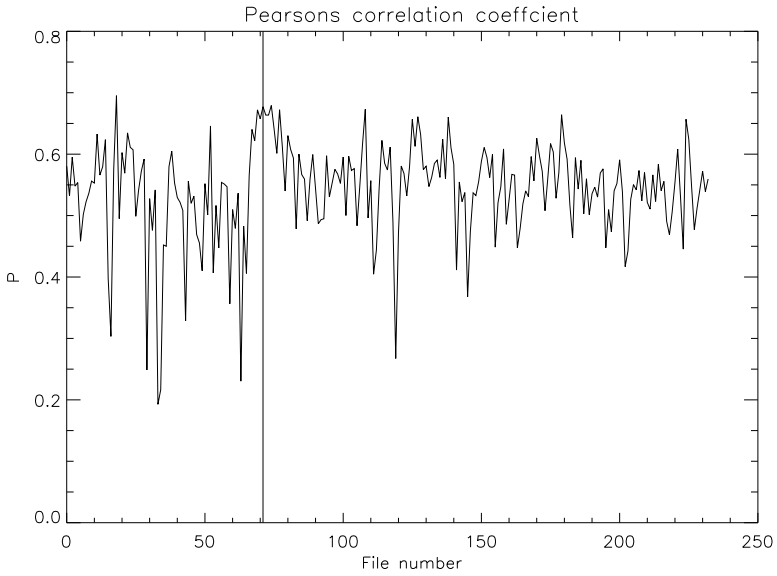

Figure 5: The time series of Pearson's correlation coefficient between H' and G'. The vertical line indicates the time of the maps in Figure 3.

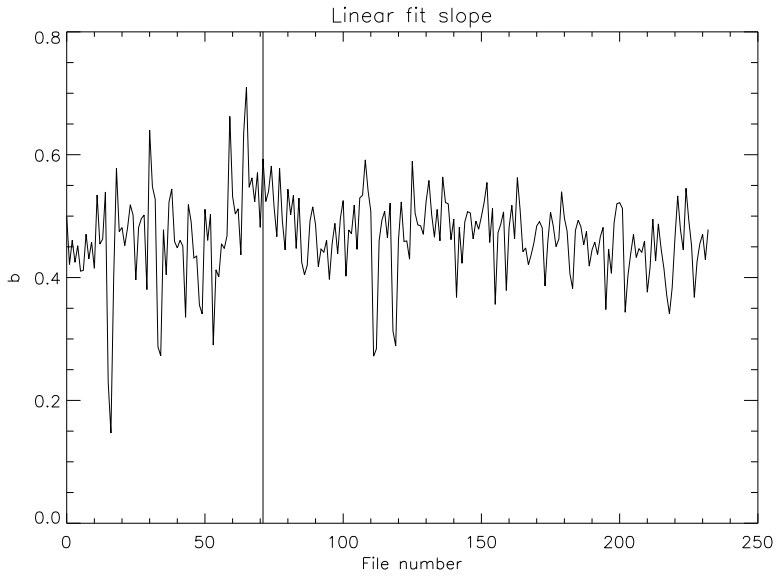

Figure 6: The time series of the linear fit slope b between H' and G'. The vertical line indicates the time of the maps in Figure 3.

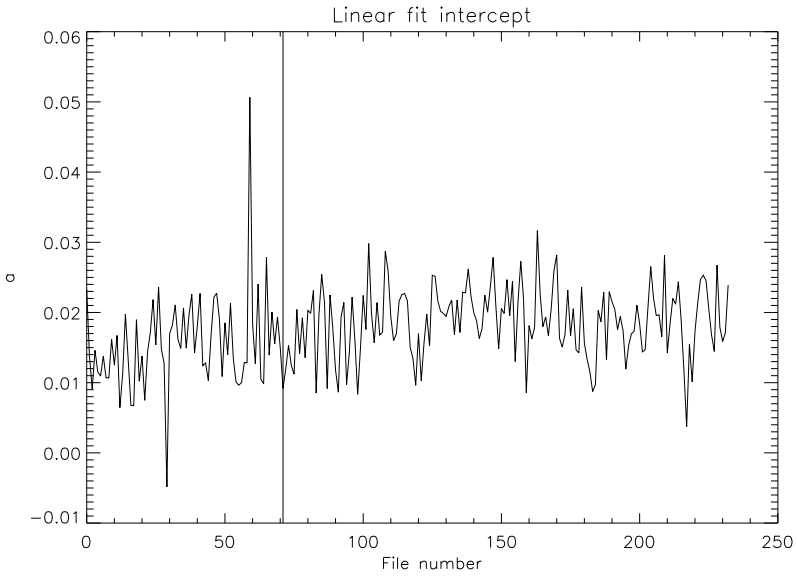

Figure 7: The time series of the linear fit y-intercept a between H' and G'. The vertical line indicates the time of the maps in Figure 3.

3. Measure the widths of the clusters in Figure 13 to numerically estimate the uncertainties

4. Repeat the analysis to compare the far-side methods. Are there any improvements?

5. Extend the analysis to other epochs.

5 Acknowledgements

The NOAA SBIR program funded this work. This work utilizes data obtained by the Global Oscillation Network Group (GONG) program, managed by the National Solar Observatory, which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory.

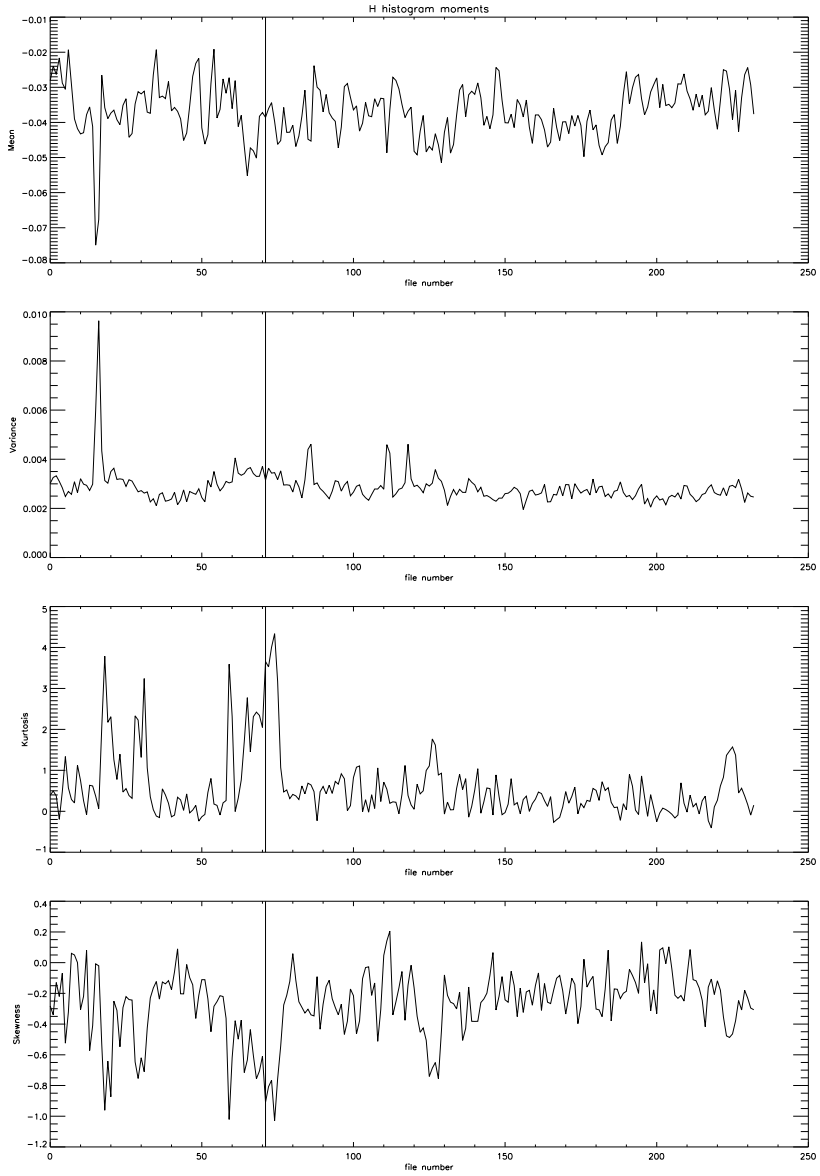

Figure 8: The time series of the histogram moments for H'. Top: mean; second from top: variance; third from top: kurtosis; bottom: skewness. The vertical line indicates the time of the maps in Figure 3.

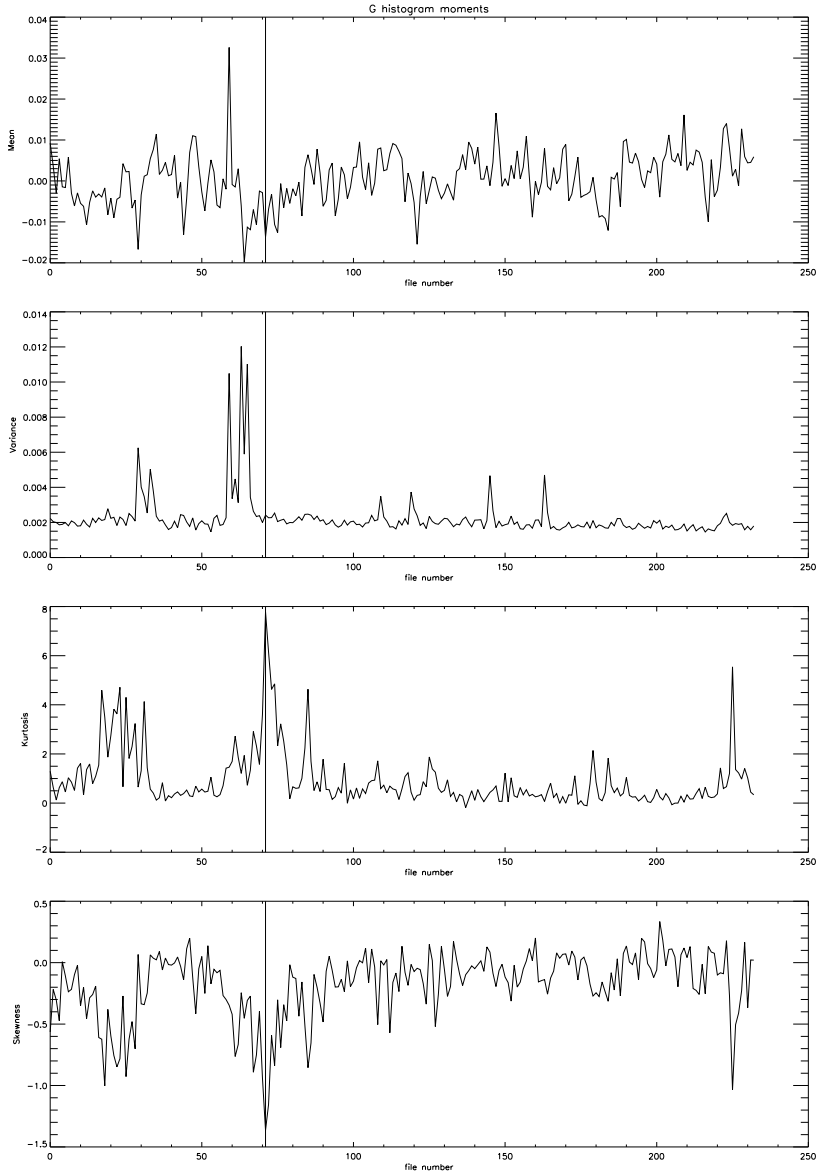

Figure 9: The time series of the histogram moments for G'. Top: mean; second from top: variance; third from top: kurtosis; bottom: skewness. The vertical line indicates the time of the maps in Figure 3.

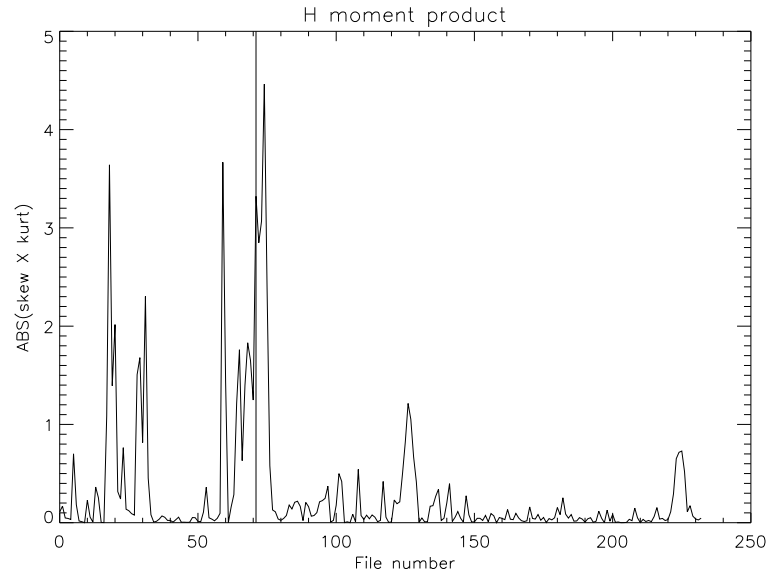

Figure 10: The time series of the absolute value of the product of the skewness and kurtosis for H'. The vertical line indicates the time of the maps in Figure 3.

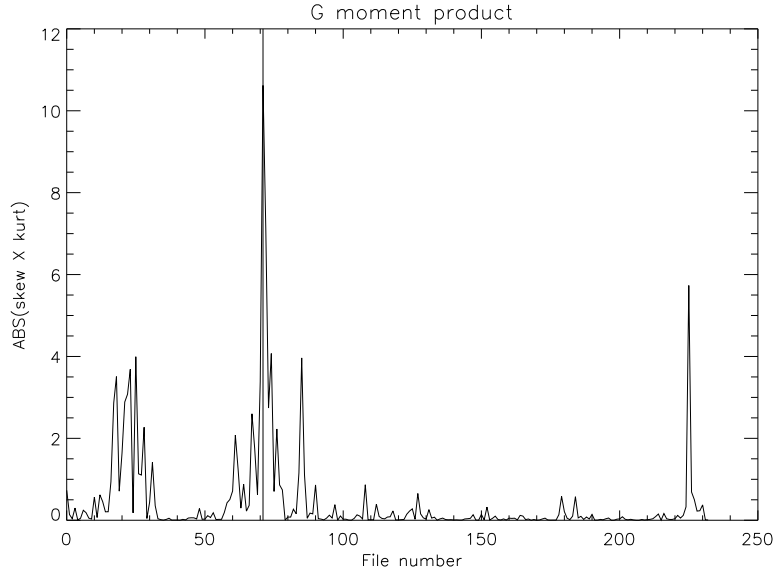

Figure 11: The time series of the absolute value of the product of the skewness and kurtosis for G'. The vertical line indicates the time of the maps in Figure 3.

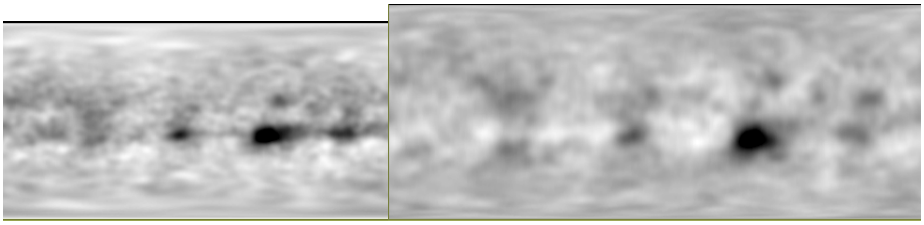

Figure 12: The sum of all 233 far-side phase-shift maps from SDO/HMI (left, H_{sum}) and GONG (right, G_{sum}).

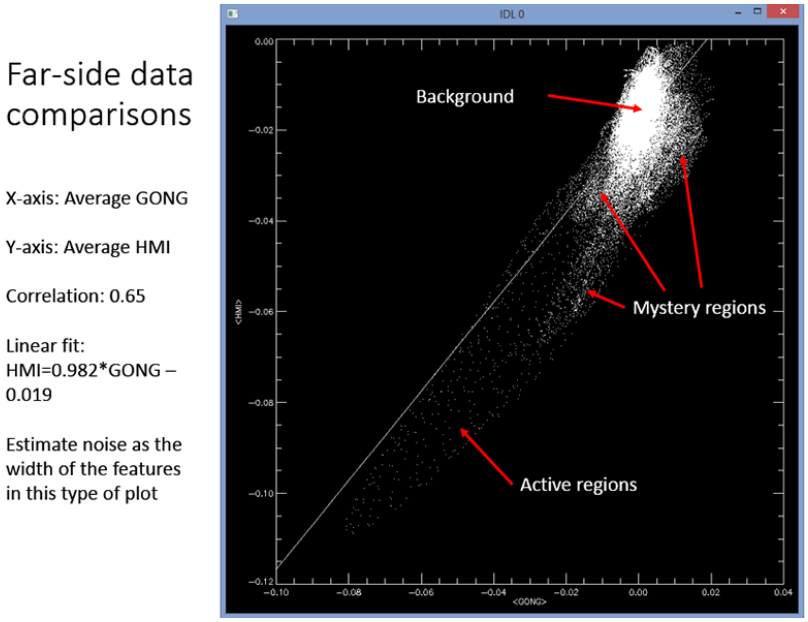

Figure 13: Scatter plot of the images in Figure 12 after interpolation onto a common grid.