Tests of a Simple Data Merging Algorithm for the GONG Project

W. E. Williams and F. Hill (NOAO/NSO)

The GONG (Global Oscillation Network Group) project proposes to reduce the impact of diurnal variations on helioseismic measurements by making long-term observations of solar images from six sites placed around the globe. The sun will be observed nearly constantly for three years, resulting in the acquisition of 1+ terabyte of image data. To use the solar network to maximum advantage, the images from the sites must be combined into a single time series to determine mode frequencies, amplitudes and line widths. Initial versions of combined, i.e., merged, time series were made using a simple weighted average of data from different sites taken simultaneously.

In order to accurately assess the impact of the data merge on the helioseismic measurements, a set of artificial solar disk images was made using a standard solar model and containing a well-known set of oscillation modes and frequencies. This undegraded data set and data products computed from it were used to judge the relative merits of various data merging schemes.

The artificial solar disk images were subjected to various instrumental and atmospheric degradations, dependent on site and time, in order to create a set of images simulating those likely to be taken at the site. The degraded artificial solar disk images for the six observing sites were combined in various ways to form merged time series of images and mode coefficients. Various forms of a weighted average were used, including an equally-weighted average, an average with weights dependent upon air mass and averages with weights dependent on various quality assurance parameters.

Both the undegraded solar disk image time series and several time series made up of various combinations of the degraded solar disk images from the six sites were subjected to standard helioseismic measurement processing. This processing consisted of coordinate remapping, detrending, spherical harmonic transformation, computation of power series for the oscillation mode coefficients and mode frequency identification.

Visual and statistical evaluation of the merged data sets themselves and differences between the merged and undegraded data set shows good agreement between the two data sets. Some slight differences in image scale and registration appear between the undegraded data set and the various merged data sets. In the set of power series made from the mode coefficients of the merged data sets, some power leakage is observed into the background and into slightly lower l-value modes, especially at higher l-value mode frequencies.

The results of the comparison of time series and mode oscillation frequencies of the undegraded data with those of the data merged using weighted averages indicate that, at least for p-mode solar oscillations, a weighted average of either the detrended remapped images or the mode coefficients gives good determination of the mode frequencies and adequate-to-good determination of the amplitudes and widths of the mode frequency lines. This conclusion is most advantageous to the analysis of the massive amounts of data to be received by the GONG network.