

# CHARACTERIZING A MICROWAVE RADIOMETER FOR SOLAR PLASMA OBSERVATIONS

Jamie Riggs

Deep Space Exploration Society, 5921 Niwot Road, Longmont, CO 80503, USA.

Email: [jamedses@spannedolutions.com](mailto:jamedses@spannedolutions.com)

## Abstract

A research project is contemplated that seeks to observe solar plasma motions, by monitoring amplitude changes over time in the Sun's microwave emissions. A small KU-band radio telescope is employed as a total-power radiometer. Before meaningful measurements can be made, the radio telescope must be characterized. Circuit instabilities, local oscillator drift, and thermal changes can cause the radiometer's gain to fluctuate over time, and its output amplitude to vary, independent of the phenomenon being observed. The present effort quantifies these variations at two different frequencies, and two orthogonal polarizations, with the radiometer observing for two hours a radio source of assumed constant noise temperature (a group of trees subtending the antenna's beamwidth). The resulting data are statistically analyzed to determine system mean and variance over time for all four channels, as well as cross-correlations among these channels. This analysis validates the system for subsequent solar observations.

## 1. Introduction

The phenomenon of interest is the Extremely Low Frequency (ELF) wave motions found in the transition zone of the Sun. These motions, which are thought to cause temperature increases from the photosphere (~15,000 Kelvin), through the transition zone (~17,000 Kelvin), to the solar corona (~1,000,000 Kelvin), can be measured indirectly, by observing with a total power radiometer the slow amplitude variations in the Sun's microwave signature. If appropriate frequencies are chosen, one close to the solar surface, say, 11.7 GHz and one above the 2 km breakpoint (transition zone), say, 12.7 GHz, it may be possible to test for two possible phenomena: (1) a low frequency (ELF), cyclic Doppler shift in the solar microwave emission, indicative of the phase of plasma motions, and (2) polarized recordings of the microwave emission, allowing one to discern differences in the E and H fields of the electromagnetic waves.

The suggestion was made that a reasonable target for characterizing the radio telescope as a total power radiometer, is to aim it at trees. These are presumed a good surrogate for a 300 Kelvin dummy load. The variability in receiver output over time, in terms of means, variances, and correlations, characterize the cyclic variations in receiver gain over time (caused by circuit drift, ambient temperature fluctuation, etc.) and may well influence the solar phenomenon of interest, and thus must be quantified.

The radiometer radiation collection dish is a 90-centimeter dish with offset feed. The trees with leaves were about 50 meters from the antenna and there may have been other obstacles, like the side of the house in the field of view. These conditions require answers to the questions raised by Paul Shuch (2008), viz.: (1) whether the trees are in the antenna's

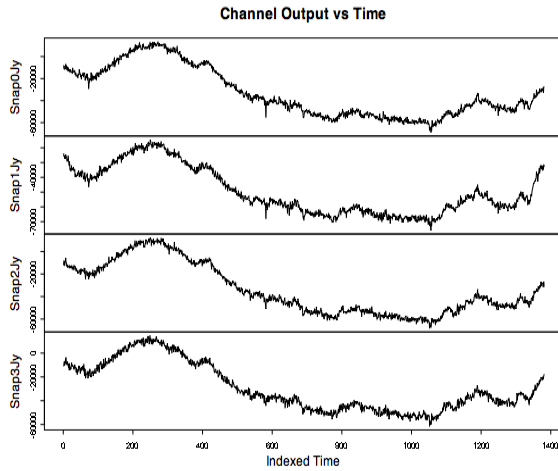
electromagnetic far field (determined as a function of wavelength, dish diameter, and solid angle subtended by the trees), as fluctuations in the receiver output will not be linear with noise temperature; (2) do the trees completely fill the dish's aperture? Shuch further comments that it is hard to find trees that are both far enough away to give far-field readings, and large enough to exceed the dish's beamwidth as viewed from the telescope. Therefore the target temperature may not be truly uniform as the dish also "sees" ground and sky, thus compromising the assumption that the radiometer is terminated in a 300 Kelvin load.

The nature of the data collection mechanism for the four-channel radiometer dictates that the statistical method known as time series analysis be used in this characterization study. Discrete-time series, the type of series from the four-channel radiometer, results because the observations are made at fixed time intervals over a specified time interval. The analysis which follows is concerned primarily with the stationarity, autocovariance, and cross-correlation functions describing the trees data. The analysis begins with a descriptive analysis (Section 2) of the radiometer data, Section 3 continues with the fitting of times series models, Section 4 describes the cross-correlation of the four channel data, which is followed by a summary of the analysis results and their attendant conclusions.

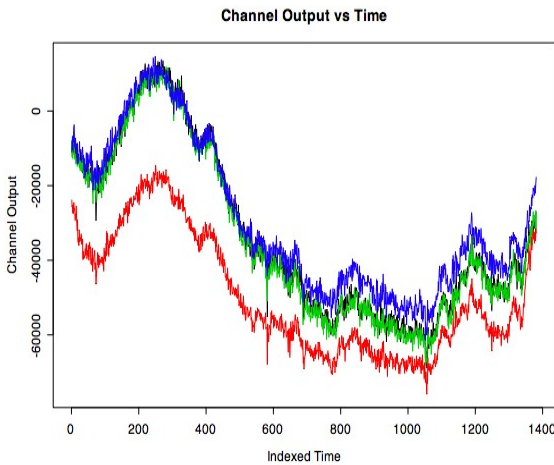
## 2. Descriptive Analysis

This portion of the analysis begins with plots of the four-channel radiometer data provided by Rodney Howe (2007). These plots (Figures 1 and 2) show that the trees data time series contain at the very least a downward trend, though possibly they are portions of longer period cycles. The most obvious feature of each series is their apparent lack of stationarity (for

any two arbitrarily chosen intervals of equal length, the mean response within the intervals is statistically the same, has finite variance, and has the same autocorrelation structure, Brockwell, (1991)), which is apparent by the slow decay across the lags in the channel AutoCorrelation Function (ACF) plot in Figure 3. Notice in Figure 2 that, when the series are overlaid one upon the other, we see that channel 1 clearly falls below channels 0, 2, and 3 (Snap0Jy, Snap1Jy, Snap2Jy, and Snap3Jy). Table 1 holds the series summary statistics.



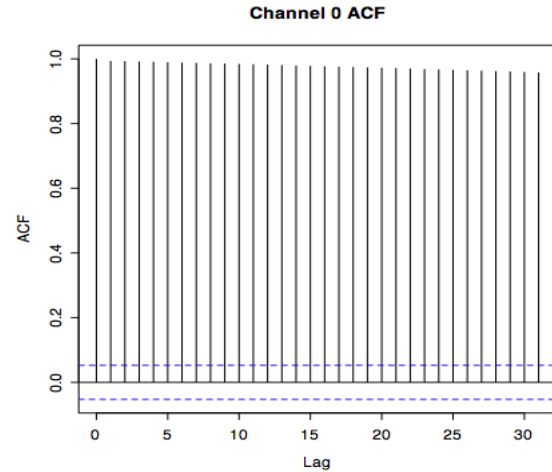
**Figure 1: Channels 0 - 3 output. See Table 1 for these series statistics.**



**Figure 2: Channels 0 - 3 overlaid upon each other. Note channel 1 (red) clearly falls below the others. The question of whether a statistically significant difference exists among all the channels is determined below.**

We now take the first differences of each series to produce stationary series. We test whether the transformed series are normally distributed as, if they are, we have Gaussian time series, which have desirable analytical properties. Figure 4 is a graph of the fit of the normal distribution to the channel 0 histogram. The fits of the normal distribution to the

channel 1 to 3 histograms are similar. The Shapiro-Wilk test statistics for normally distributed data are given in the caption of this Figure, and they indicate that each first-differenced series is distributed as a Gaussian (we do not reject the null hypothesis as  $W$  is not sufficiently small). Table 2 has the first-differenced channel output summary statistics.



**Figure 3: ACF for channels 0 – 3 raw data. The slowly decreasing amplitudes over the lags signifies nonstationary series. This ACF is representative of all four channels.**

The evidence that our time series are not stationary, apart from a visual inspection, comes from the ACF. Simplistically, the ACF describes how an observation at time  $t$  is related to the observations at times  $t-1$  though  $t-h$ ,  $h = 2, 3, \dots, 25$ . A slowly decreasing ACF, as we saw in Figure 3, is indicative of a non-stationary time series.

| Statistic | Channel |         |         |         |
|-----------|---------|---------|---------|---------|
|           | 0       | 1       | 2       | 3       |
| Mean      | -33,167 | -49,536 | -34,327 | -29,308 |
| Std Dev   | 22,495  | 15,947  | 22,301  | 20,146  |
| Min       | -68,811 | -75,876 | -68,254 | -61,537 |
| Max       | 13,537  | -14,472 | 12,333  | 14,680  |

**Table 1: Channel output summary statistics.**

And just what do these normally-distributed, first-differenced series look like? Figures 5 and 6 show the separated and the overlaid series, respectively. The salient feature is the lack of an obvious trend, and, though we need to test explicitly for cyclical behavior, there is no obvious long term periodic behavior. These plots lead us to suspect that the first-differenced transformations have given us stationary time series in addition to the normal distributions. The overlaid plot (Figure 6) shows practically no difference among the channels.

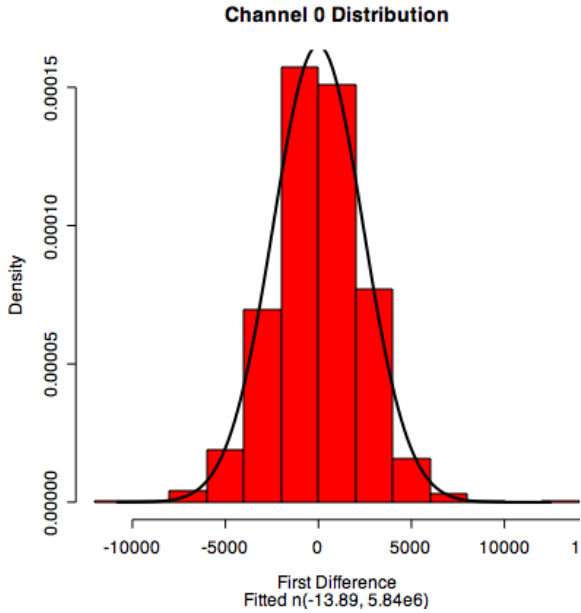


Figure 4: Channel 0 distribution with  $W = 0.993$ ,  $p = 4.2e-6$ . Channels 1 – 3 are similar with the following

respective statistics:  $W = 0.995$ ,  $p = 4.0e-6$ ;  $W = 0.999$ ,  $p = 4.0e-6$ . The next and, perhaps, the most important questions about these series is whether each channel is autocorrelated; and whether the series are cross-correlated. The next section has the autocorrelation parametric analysis.

| Statistic | Channel |        |         |        |
|-----------|---------|--------|---------|--------|
|           | 0       | 1      | 2       | 3      |
| Mean      | -13.888 | -6.445 | -13.880 | -4.814 |
| Std Dev   | 2,416   | 2,051  | 2,339   | 2,596  |
| Min       | -10,815 | -9,719 | -7,959  | -9,010 |
| Max       | 12,489  | 10,750 | 9,087   | 8,318  |

Table 2: First-difference channel output summary statistics.

### 3. Time Series ARIMA Models

Our first-differenced time series exhibit no apparent deviations from stationarity and, as we shall see shortly, the corresponding ACF plots decrease rapidly indicating stationarity. This leads us to consider the class of time series models known as AutoRegressive Integrated Moving Average (ARIMA) models. As we saw above, the slowly decaying ACF of the non-differenced data indicated these data were not stationary. We also saw that by taking the first differences of each series, we obtained stationary time series. Thus, we have satisfied a mandatory condition for building ARMA models (a

subclass of ARIMA models) for our data. An ARMA model is a causal model with 2 values  $p$  and  $q$  indicating the order of the autoregressive portion ( $p$ ) of the the ARMA model, and the moving average portion ( $q$ ). ARIMA models are specified by  $p$ ,  $q$ , and  $d$ , where  $p$  and  $q$  are as in the ARMA model, and  $d$  is the order of differencing where  $d = 1$  in our case, allowing us to use the ARMA class models on these first-differenced data.

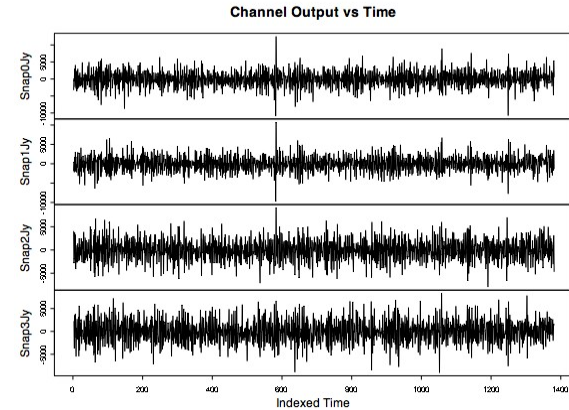


Figure 5: Channels 0 - 3 first differences (lag 1). See Table 2 for these series statistics.

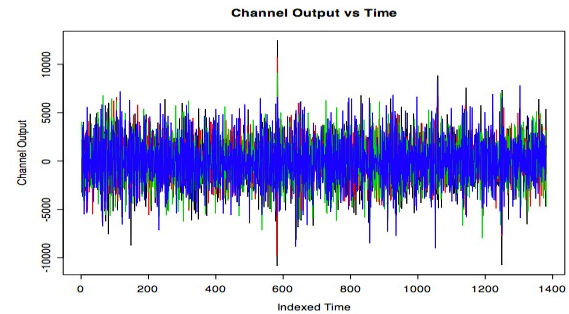


Figure 6: Channels 0 - 3 overlaid upon each other. Note channel 1 (red) no longer clearly falls below the others.

The autoregressive specifier indicates the number of time units over which an observation  $t$  is dependent on preceding observations. The moving average specifier indicates the number of white noise observations preceding  $t$  that combine to equal the combined autoregressive terms. This equality is a condition for using a causal model, and hence we can rearrange this equality such that future observations may be forecasted.

We now determine the ARMA( $p,q$ ) specifiers for each of the four channels. The combination of the ACF and the Partial Autocorrelation Function (PACF) provide the information needed to find appropriate values (Venables and Ripley, 2002). The ACF (Figures 7 and 8) clearly shows that the time series are stationary as discussed above. The specifier for the MA part ( $q$ ) may be determined from the ACF

plots. Examination of the PACF plots (Figure 9) show that weighted observation-to-observation contribution to the current observation is ambiguous. This leads us to test an ARMA(0,1) model [an ARIMA(0,1,1) model] which is abbreviated to a MA(1) (Moving Average with  $q = 1$ ) process. A moving average of order 1 averages the white noise values from the current and previous time intervals; thus, after removing the mean, we have,

$$X_t = Z_t + \theta Z_{t-1}$$

where  $X_t$  is the  $t^{\text{th}}$  observation, and  $\{Z_t\}$  is white noise with mean 0 and finite variance  $\sigma^2$  if and only if  $\{Z_t\}$  are not correlated, and is denoted  $\{Z_t\} \sim WN(0, \sigma^2)$ . The parameter  $\theta$  is to be determined from each series data. The model parameters are, by channel, in Table 3.

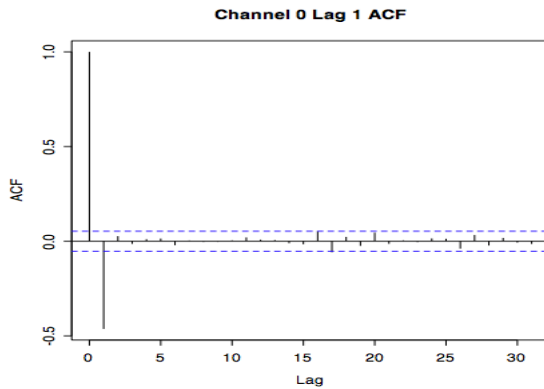


Figure 7: ACF for channels 0, 2, and 3 first-differenced data. The rapidly decreasing amplitudes over the lags signifies stationary series. Also, the order of the moving average component of the ARMA model (MA) can be determined.

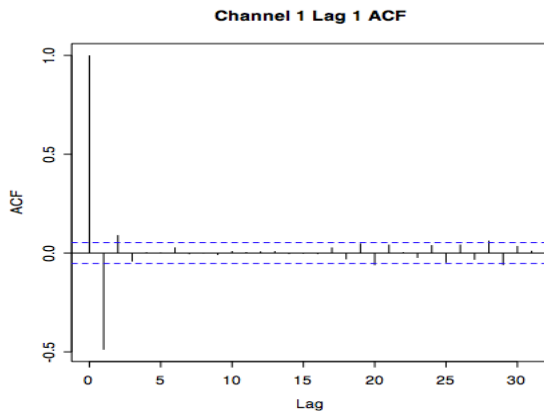


Figure 8: ACF for channel 1 first-differenced data. The rapidly decreasing amplitudes over the lags signifies stationary series. The order of the moving average component of the ARMA model (MA) can be determined.

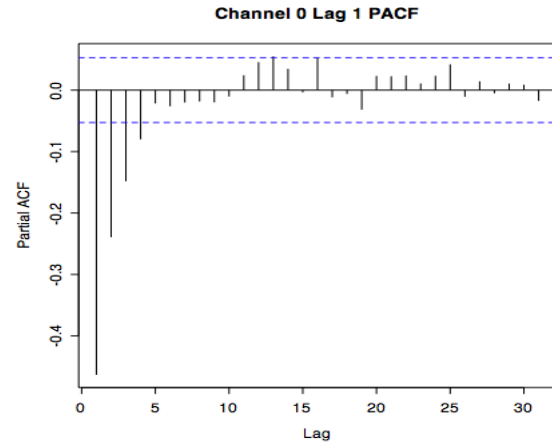


Figure 9: PACF for channels 0 – 3 first-differenced data. The gradually changing amplitudes over the lags indicates these series likely have no auto-regressive behavior. The statistical significance of lags greater than 10 may need further investigation.

|  |                |
|--|----------------|
| <pre> <code>arma(x = ch0, order = c(0, 0, 1))</code> </pre>              |                |
| Coefficients:  |                |
| ma1  | intercept      |
| -0.6071  | -13.7983       |
| s.e.   | 0.0205 21.6688 |
| sigma^2 estimated as 4188193: log likelihood = -12479.33, aic = 24964.67 |                |
| <pre> <code>arma(x = ch2], order = c(0, 0, 1))</code> </pre>             |                |
| Coefficients:  |                |
| ma1  | intercept      |
| -0.5996  | -13.7077       |
| s.e.   | 0.0209 21.6109 |
| sigma^2 estimated as 4011864: log likelihood = -12449.65, aic = 24905.29 |                |
| <pre> <code>arma(x = ch3, order = c(0, 0, 1))</code> </pre>              |                |
| Coefficients:  |                |
| ma1  | intercept      |
| -0.6371  | -7.4727        |
| s.e.   | 0.0199 21.4906 |
| sigma^2 estimated as 4826349: log likelihood = -12577.22, aic = 25160.44 |                |

Table 3: ARIMA(0,1,1) model parameters for channels 0, 2, and 3. For each channel's parameters, ma1 is the estimate of  $\theta$  and the intercept is  $Z_t$ . s.e. is the

The ARMA(0,1) model works well for channels 0, 2, and 3, as is seen in the model diagnostics in Figures 10 and 11 (Box and Pierce, 1970), and from the Akaike Information Criterion (AIC), Akaike, 1969. However it does not work well for channel 1. Closer inspection of the ACF plot (Figure 8) shows significance at lags at 1 and 2, which implies that the channel 1 data are better fitted to an ARMA(0,2)

[ARIMA(0,1,2)] model. The diagnostics (see Figures 12 and 13) confirm that channel 1 is modeled well by an ARMA(0,2). After removing the mean of the differenced series, the model is

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

where, as before,  $X_t$  is the  $t^{\text{th}}$  observation, and  $\{Z_t\}$  is white noise with mean 0 and finite variance  $\sigma^2$  if and only if  $\{Z_t\}$  are not correlated. The parameters  $\theta_1$  and  $\theta_2$  are to be determined from each series data. These parameter estimates are in Table 4. The AIC is lower for this ARMA(0,2) model than for the ARMA(0,1) model, which reinforces the diagnostics indication that the ARMA(0,2) is the appropriate model.

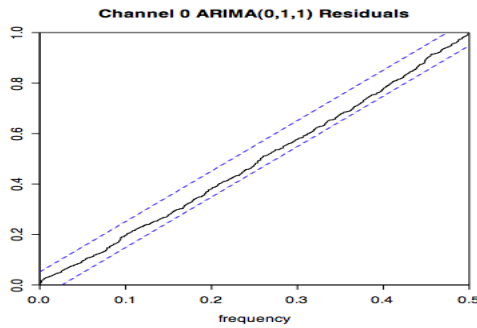


Figure 10: ARIMA(0,1,1) model diagnostics support an ARMA(0,1) model for channel 0, 2, and 3. The dashed lines indicate 95% confidence limits on the cumulative periodogram. As it lies completely within the limits, the residuals show the desirable characteristic of being uncorrelated.

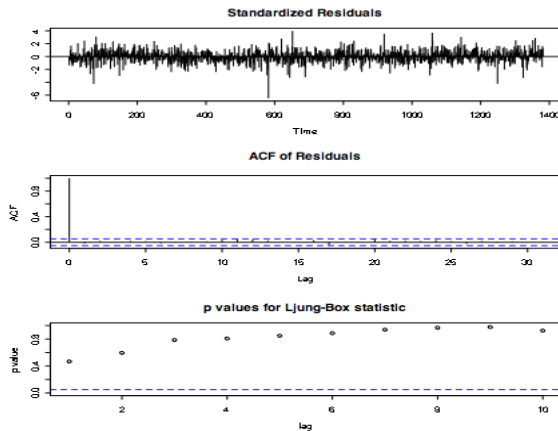


Figure 11: These model diagnostics indicate that the ARMA(0,1) model is reasonable. The standardized residuals appear, with one exception, to be white noise. This is tested in the ACF, which shows the residuals are white noise. Also, as the p values all lie above the horizontal dashed line in the Ljung-Box portmanteau test for white noise, we again conclude the ARMA(0,1) model is appropriate.

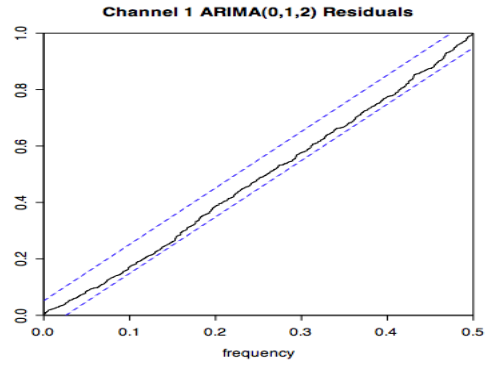


Figure 12: ARIMA(0,1,2) model diagnostics support an ARMA(0,2) model for channel 1, as the cumulative periodogram of the residuals lie within the 95% confidence limits.

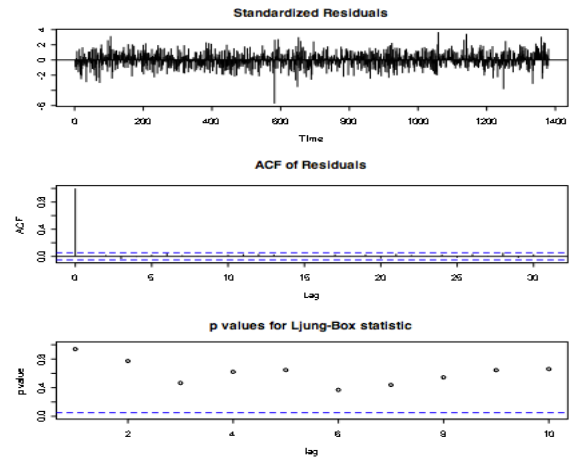


Figure 13: Similar to the reasons discussed in Figure 11, we see that these model diagnostics support an ARMA(0,2) model.

```

arma(x = tmp[, 2], order = c(0, 0, 2))

Coefficients:
      ma1      ma2  intercept
-0.6171  0.0735  -5.6181
s.e.    0.0265  0.0252   21.3657

sigma^2 estimated as 3019933: log
likelihood = -12253.66, aic = 24515.31

```

Table 4: ARIMA(0,1,2) model parameters for channel 1. For channel 1, ma1 is the estimate of  $\theta_1$ , ma2 is the estimate of  $\theta_2$ , and the intercept is  $Z_t$ . s.e. is

#### 4. Cross-Correlation

The final portion of our analysis examines the question as to whether the channel series are cross-correlated, as these series are measured over the same

time period and have the same interval size. Cross-correlation among the channels suggests one or more channels may be leading indicators or concomitant influencers. These indicators may be real, or they may be an artifact of the electronics, or of the electromagnetic physics.

Figure 14 shows the lag 1 (first-differenced) pairwise channel scatter plots with loess smoothing as the red curves (Cleveland, 1979, and Cleveland, et al., 1988). We see that channels 0 and 1 may be weakly correlated, and channels 2 and 3 appear to be more strongly correlated. To see if these relations indeed hold, we examine the multivariable ACF and PACF plots (Figures 15 and 16). It is clear that for every pair of channels, both the ACF and PACF show statistically significant leading or lagging between channel pairs. At any specific lag (x-axis), we look for vertical lag lines that exceed the dashed horizontal lines (blue). Lines exceeding the dashed lines indicate that the associated lag is statistically significant. The plots on the diagonal are the respective channel's autocorrelation plots. The off-diagonal plots are the cross-correlations.

The ACFs and PACFs for channel pairs show the channels are cross-correlated. Row 1 is the autocorrelation of channel 0 (first plot of the first row) and cross-correlation lags of channel 0 with channels 1 through 3 (graphs 2 through 4 of row 1). Row 2 begins with the cross-correlation leads of channel 1 with channel 0, followed by the autocorrelation of channel 1 (second graph of row 2), and ending with the cross-correlation lags of channel 1 with channels 2 and 3. Similarly for rows 3 and 4. Note the graphs on the diagonal are the autocorrelation plots. Any vertical line of each plot extending beyond the dashed horizontal line (blue) indicate that the lag (or lead) at that vertical line is statistically significant. Thus, lag 1 (or lead 1) is significant in all the autocorrelation plots, with channel 1 having significance at lag 2. The cross-correlation plots show similar significance at a variety of lags or leads. The reason for these cross-correlations needs to be determined.

## 5. Conclusions

The time series analysis of the four-channel, total-power radiometer trees surrogate 300 Kelvin dummy load data gives the following results:

1. All four channels (0 – 3) are nonstationary time series.
2. Channels 0, 2, and 3 are described by an ARIMA(0,1,1) model, and channel 1 is described by an ARIMA(0,1,2) model.
3. All paired channel combinations are cross-correlated.

First differencing of the data from the receiver channels is necessary to obtain a stationary time series. Stationary series are required to understand autocorrelation, which is whether an observation is dependent upon previous observations. These trees data are autocorrelated which implies that some type of receiver system behavior such as cyclically varying gain, circuit drift, ambient temperature fluctuation, the inadequacy of the trees data as a 300 Kelvin dummy load, or other physical phenomena are the possible causes. The cross-correlation analysis indicates that there may be current bleed-over from one channel to another, or it may be due to the polarization physics, or a combination of the two. Engineering and physics analysis is required to ascribe actual causes to these auto- and cross-correlation behaviors, and to the greater than lag 10 auto- and cross-correlations present in Figure 16.

We see with this analysis that any observations of, say, the Sun, must account for this systematic error before useful solar activity results can be derived.

## 6. Acknowledgments

We thank Rodney Howe for providing the data used in this analysis. The insights into radiometer behavior is from Dr. H. Paul Shuch, who gave his time generously, and we thank him ipso facto.

## 7. References

- Akaike, H., (1969), "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics*, Tokyo, 21, 243-247.
- Box, G.E.P. and Pierce, D.A., (1970), "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *J. Amer. Statist. Assoc.*, 65, 1509-1526.
- Brockwell, P. J., and Davis, A.D., (1991), *Time Series: Theory and Methods*, 2<sup>nd</sup> ed., New York: Springer-Verlag New York, Inc.
- Cleveland, W.S., (1979), "Robust Locally-Weighted Regression and Smoothing Scatterplots," *J. Amer. Statist. Assoc.*, 74, 829-836.
- Cleveland, W.S., Devlin, S.J., and Grosse, E., (1988), "Regression by Local Fitting," *J. Econometrics*, 37, 87-114.
- Howe, R., (2007), November email correspondence with the author.

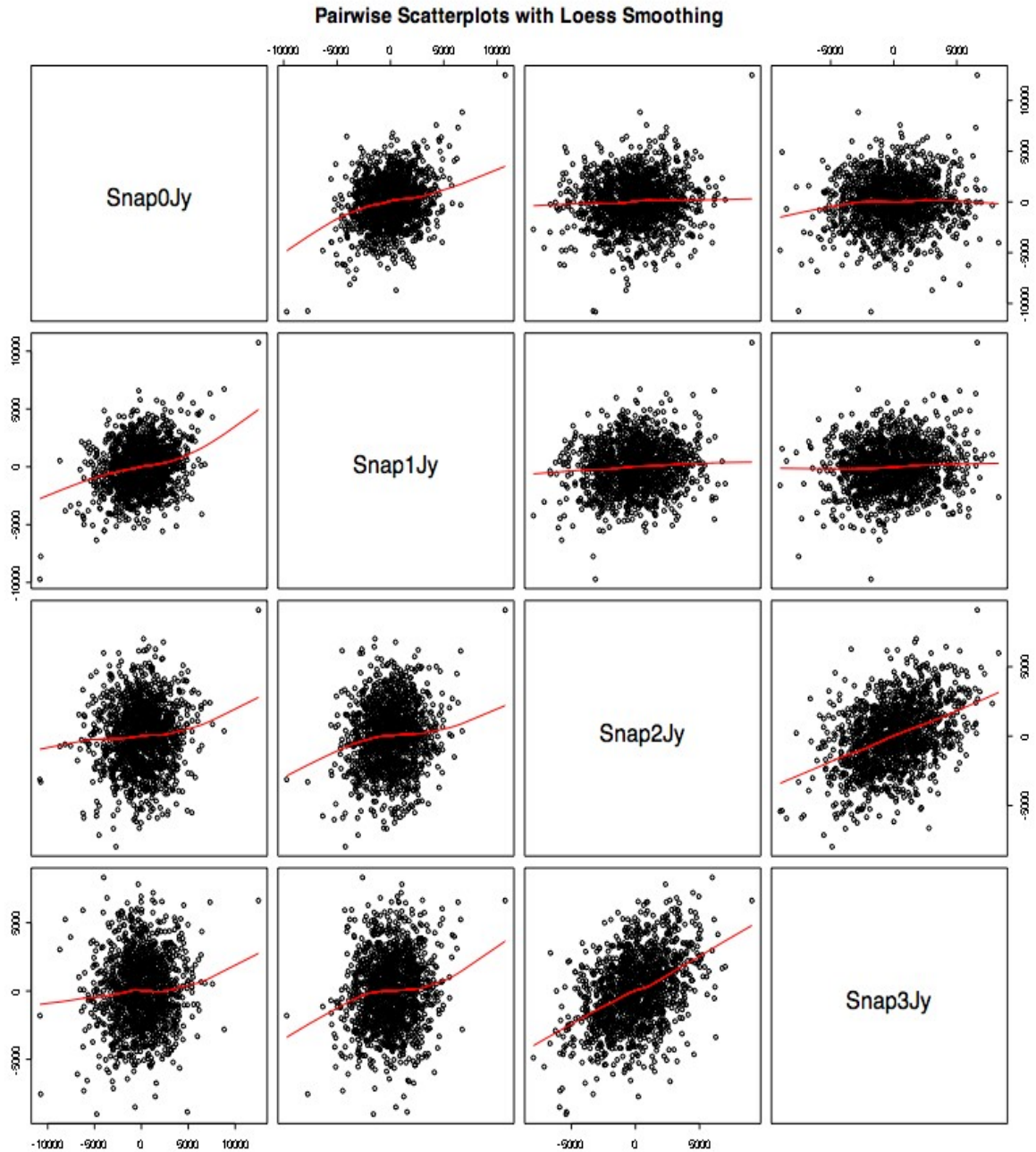
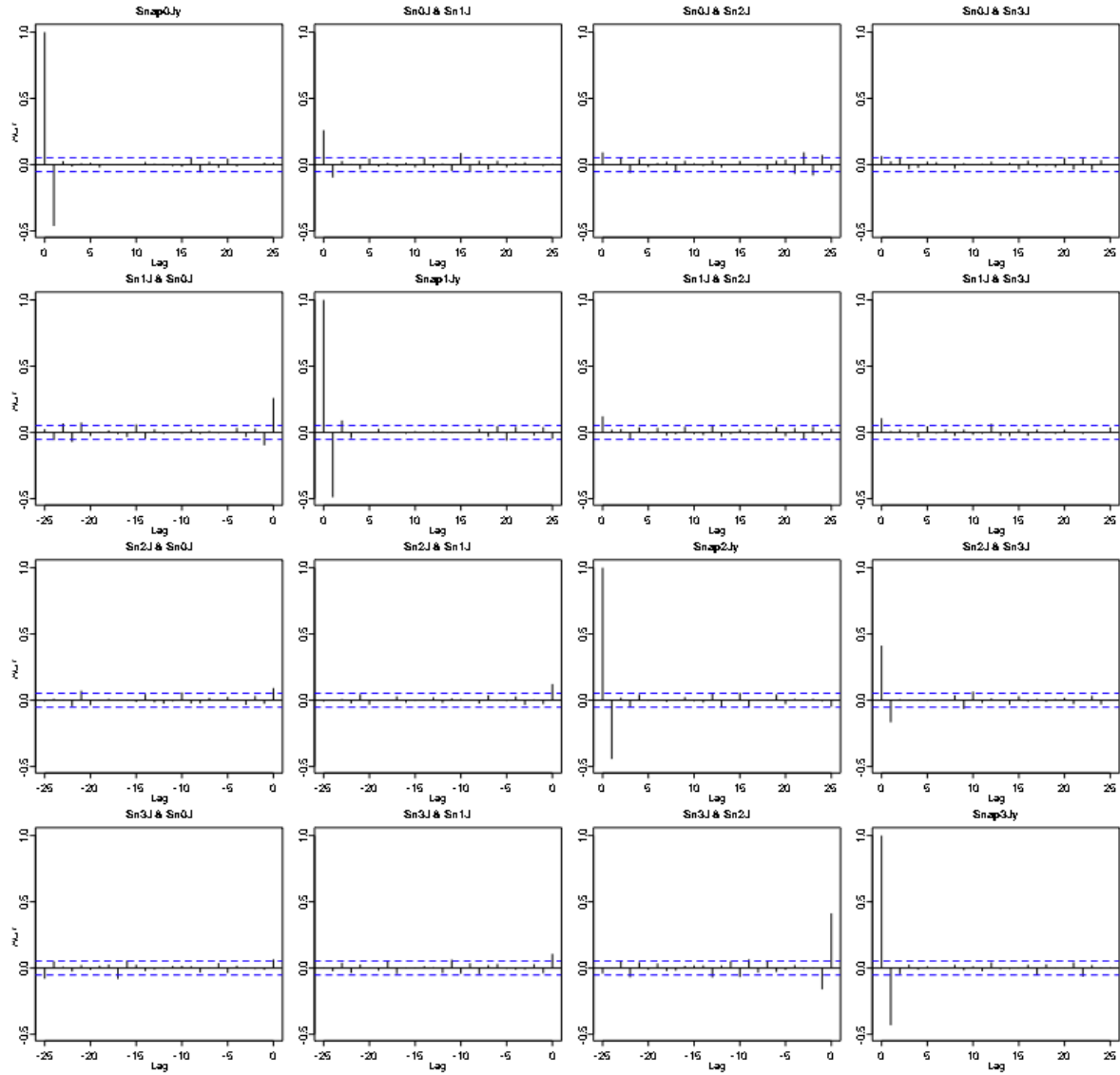


Figure 14: Pairwise scatter plots of the channel lag 1 data with Loess smoothing (red curves). Row 1 is the scatterplot of channel 0 (first plot of the first row) with channels 1 through 3 (graphs 2 through 4 of row 1). Row 2 begins with the scatterplot of channel 1 with channel 0, followed by the scatterplots of channel 1 (second graph of row 2) with channels 2 and 3. Similarly for rows 3 and 4. The scatterplots between channels 2 and 3 (rows 3 and 4) show the most amount of cross-correlation.

Shuch, H. Paul, (2008), January email correspondence with the author.

Venables, W.N., and Ripley, B.D., (2002), *Modern Applied Statistics with S*, 4<sup>th</sup> ed., New York: Springer Science+Business Media, Inc.



**Figure 15: Pairwise cross-correlated channel ACF plots. Row 1 is the autocorrelation of channel 0 (first plot of the first row) and cross-correlation lags of channel 0 with channels 1 through 3 (graphs 2 through 4 of row 1). Row 2 begins with the cross-correlation leads of channel 1 with channel 0, followed by the autocorrelation of channel 1 (second graph of row 2), and ending with the cross-correlation lags of channel 1 with channels 2 and 3. Similarly for rows 3 and 4. Note the graphs on the diagonal are the autocorrelation plots. Any vertical line of each plot extending beyond the dashed horizontal line (blue) indicate that the lag (or lead) at that vertical line is statistically significant. Thus, lag 1 (or lead 1) is significant in all the autocorrelation plots, with channel 1 having significance at lag 2. The cross-correlation plots show similar significance at a variety of lags or leads.**

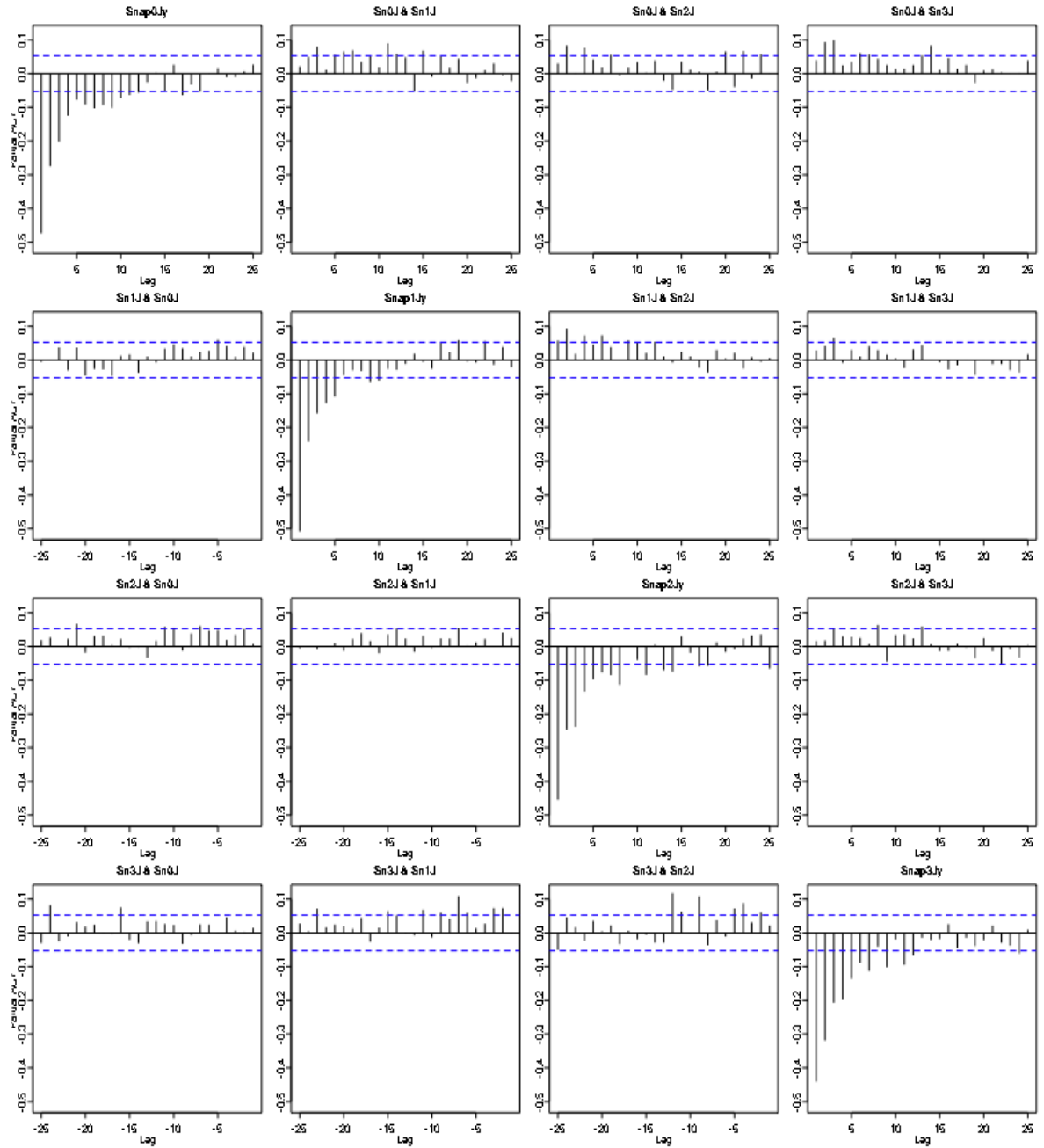


Figure16: Pairwise cross-correlated channel PACF plots. Row 1 is the partial autocorrelation of channel 0 (first plot of the first row) and cross-correlation lags of channel 0 with channels 1 through 3 (graphs 2 through 4 of row 1). Row 2 begins with the cross-correlation leads of channel 1 with channel 0, followed by the partial autocorrelation of channel 1 (second graph of row 2), and ending with the cross-correlation lags of channel 1 with channels 3 and 4. Similarly for rows 3 and 4. Note the graphs on the diagonal are the partial autocorrelation plots. Any vertical line of each plot extending beyond the dashed horizontal line (blue) indicate that the lag (or lead) at that vertical line is statistically significant. Thus, no lag (or lead) is significant in all the partial autocorrelation plots. The cross-correlation plots show significance at a variety of lags or leads.