

CHARACTERIZING A MICROWAVE RADIOMETER FOR SOLAR PLASMA OBSERVATIONS WITH A LOCAL NOON-TIME POLAR REGION AS A DUMMY LOAD

JAMIE RIGGS

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

Email: jamiedses@spannedolutions.com

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. However, before meaningful measurements can be accomplished, it is necessary to characterize the radio telescope itself. 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. In the present effort, those variations are quantified at two different frequencies, and two orthogonal polarizations, by using the radiometer to observe, for two hours, a radio source of assumed constant noise temperature (one hour before and after noon in the polaris region of the sky). The resulting data are statistically analyzed to determine system mean output and variability over time for all four observation channels, as well as cross-correlations among those channels. These parameters will validate 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, 12.2 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 in the Polaris region of the noon-time sky. This is presumed a good surrogate for a 20 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. Therefore, the target sky region must satisfy (1) the target is in the antenna's electromagnetic far field (determined as a function of

wavelength, dish diameter, and solid angle subtended by the Polaris), as fluctuations in the receiver output will not be linear with noise temperature, and (2) the target completely fills the dish's aperture (Shuch, 2008). The target temperature is assumed truly uniform as the dish does not "see" the ground, which compromises the assumption that the radiometer is terminated in a 20 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 (integration). The analysis which follows is concerned primarily with the stationarity, autocovariance, and cross-correlation functions describing the Polaris 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

The descriptive analysis begins with plots of the four-channel radiometer data provided by Rodney Howe (2008). These plots (Illustrations 1 and 2) show that the Polaris data time series contain a quadratic 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), as we shall see below. Notice in Illustration 2 that, when the series are overlaid one upon the other, we see that channels 2 and 3 may be statistically equal unlike channels 0 and 1. We examine the equality of channel output next. Table 1 holds the series summary statistics.

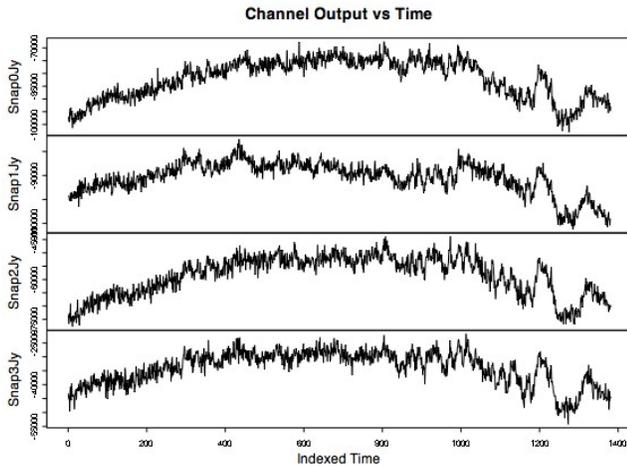


Illustration 1: Channels 0 - 3 output. See Table 1 for these series statistics.

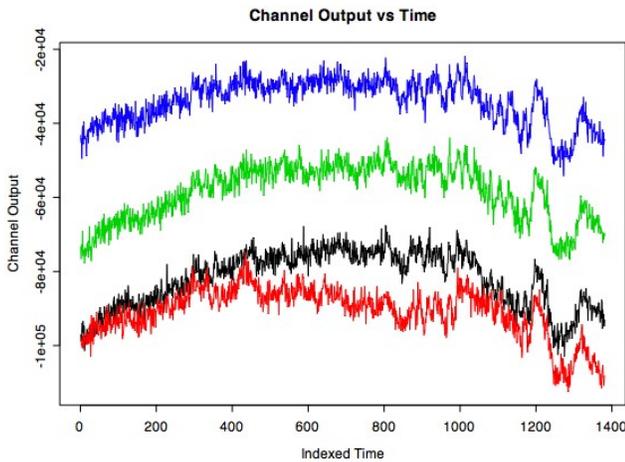


Illustration 2: Channels 0 - 3 overlaid upon each other. Note channels 2 (black) and 3 (red) may be statistically the same. The question of whether a statistically significant difference exists among all the channels is determined below.

A natural question to answer is whether these channel output values are statistically equal, and we employ box plots (Illustrations 3 and 4) for the answer. Illustration 3 shows that the medians of the channels each are statistically significantly different as the notch size (vertical) nowhere overlap. Illustration 4 shows that the standard deviations of each channel are statistically the same as the box heights all

overlap. The the median signal levels are different but the variances are the same.

We now address the issue of the nonstationarity of the series. The ACF plot in Illustration 5, which is representative of each channel, slowly decreases indicating a nonstationary series. We take the first differences of each series to produce a 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. Illustration 6 is graph of the fit of a normal distribution to the channel 0 series histogram. The Shapiro-Wilk test statistics for normally distributed data are given in the caption of this Illustration, and indicates 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.

Table 1: Channel output summary statistics.

Statistic	Channel			
	0	1	2	3
Mean	-81,993	-90,856	-58490	-34056
Std Dev	7564	6632	7380	6009
Min	-102,910	-112,471	-77,781	-54,207
Max	-67,556	-74,833	-43,861	-21,802

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 Illustration 5, is indicative of a non-stationary time series.

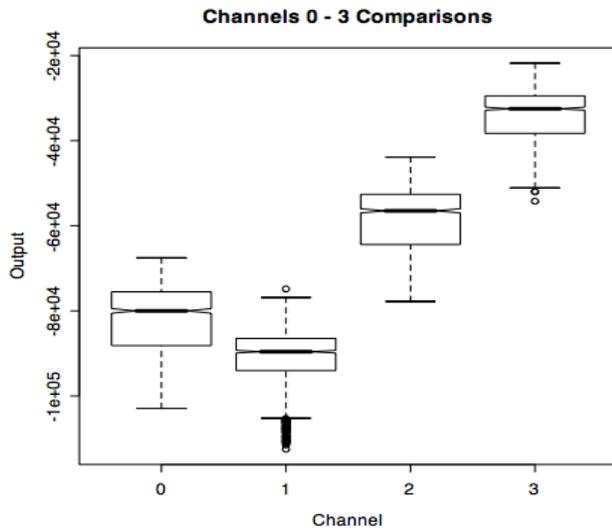


Illustration 3: The notches are the 95% confidence interval sizes about each channel median. As none of the notches overlap, the medians are statistically significantly different.

Illustrations 7 and 8 show the separated and the overlaid, normally-distributed 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 (Illustration 8) shows practically no difference among the channels.

We now turn to, perhaps, the most important questions about these series: is each channel autocorrelated; and are these series are cross-correlated?

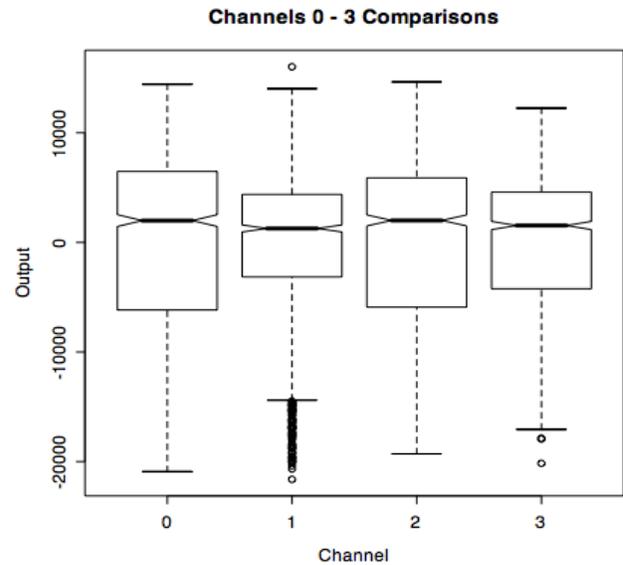


Illustration 4: The entire box height is the interquartile range (75th percentile minus the 25th percentile), which is an approximation of the variance of the channels. Each box overlaps all the others which indicates that the variances are likely statistically equal.

Table 2: First-difference channel output summary statistics.

Statistic	Channel			
	0	1	2	3
Mean	2.925	-6.873	2.670	-1.052
Std Dev	2,631	2,752	2,798	2,770
Min	-10,404	-8,887	-10,171	-8385
Max	9,971	8,905	8,285	10,252

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.

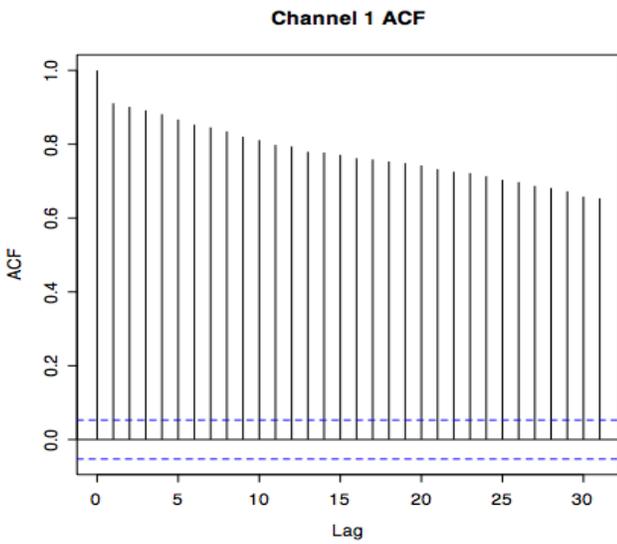


Illustration 5: ACF for channels 0 – 3 raw data are similar to this Channel 1 ACF. The slowly decreasing amplitudes over the lags signify nonstationary series.

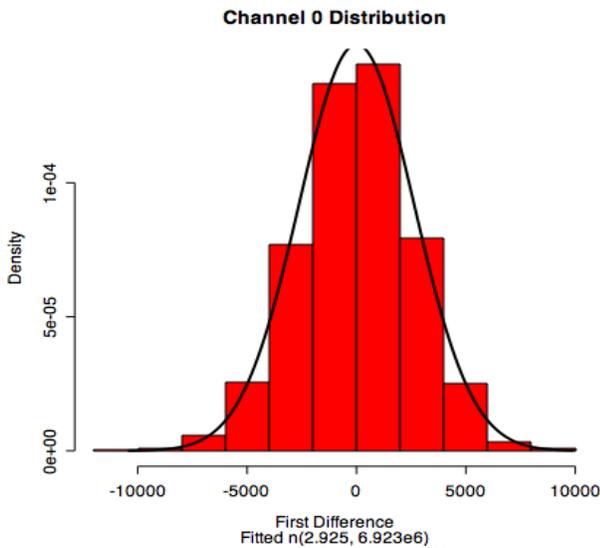


Illustration 6: The distributions of the first-differenced data are all similar to this Channel 0 histogram with overlaid normal distribution. The Shapiro-Wilk statistics are: Channel 0 $W = 0.998$, $p = 0.034$; Channel 1 $W = 0.999$, $p = 0.712$; Channel 2 $W = 0.998$, $p = 0.160$; Channel 3 $W = 0.999$, $p = 0.527$.

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.

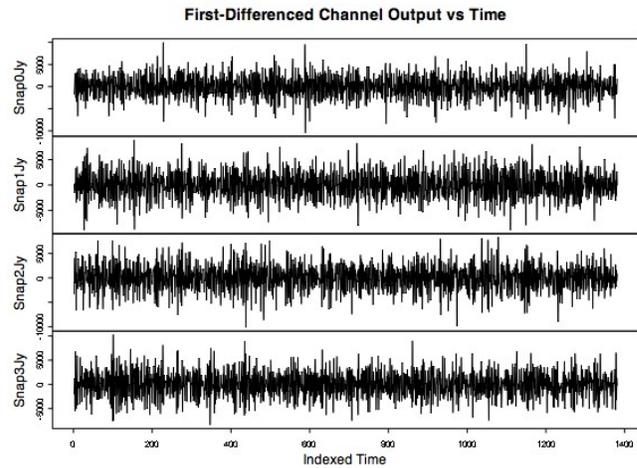


Illustration 7: Channels 0 - 3 first differences (lag 1). See Table 2 for these series statistics.

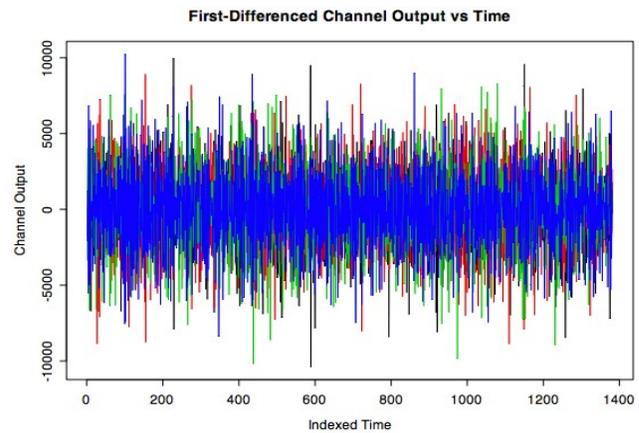


Illustration 8: Channels 0 - 3 overlaid upon each other. Note channel 1 (red) no longer clearly falls below the others.

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 periodogram of Channel 3, Illustration 9, supports the claims below that the specifiers are of low order. First, the ACF clearly shows that the time series are stationary as discussed above (Illustration 10). The specifier for the MA part (q) may be determined from the ACF plots. Examination of the PACF plots (Illustration 11) 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 random variation from the current and previous time intervals; thus, after removing the mean, we have,

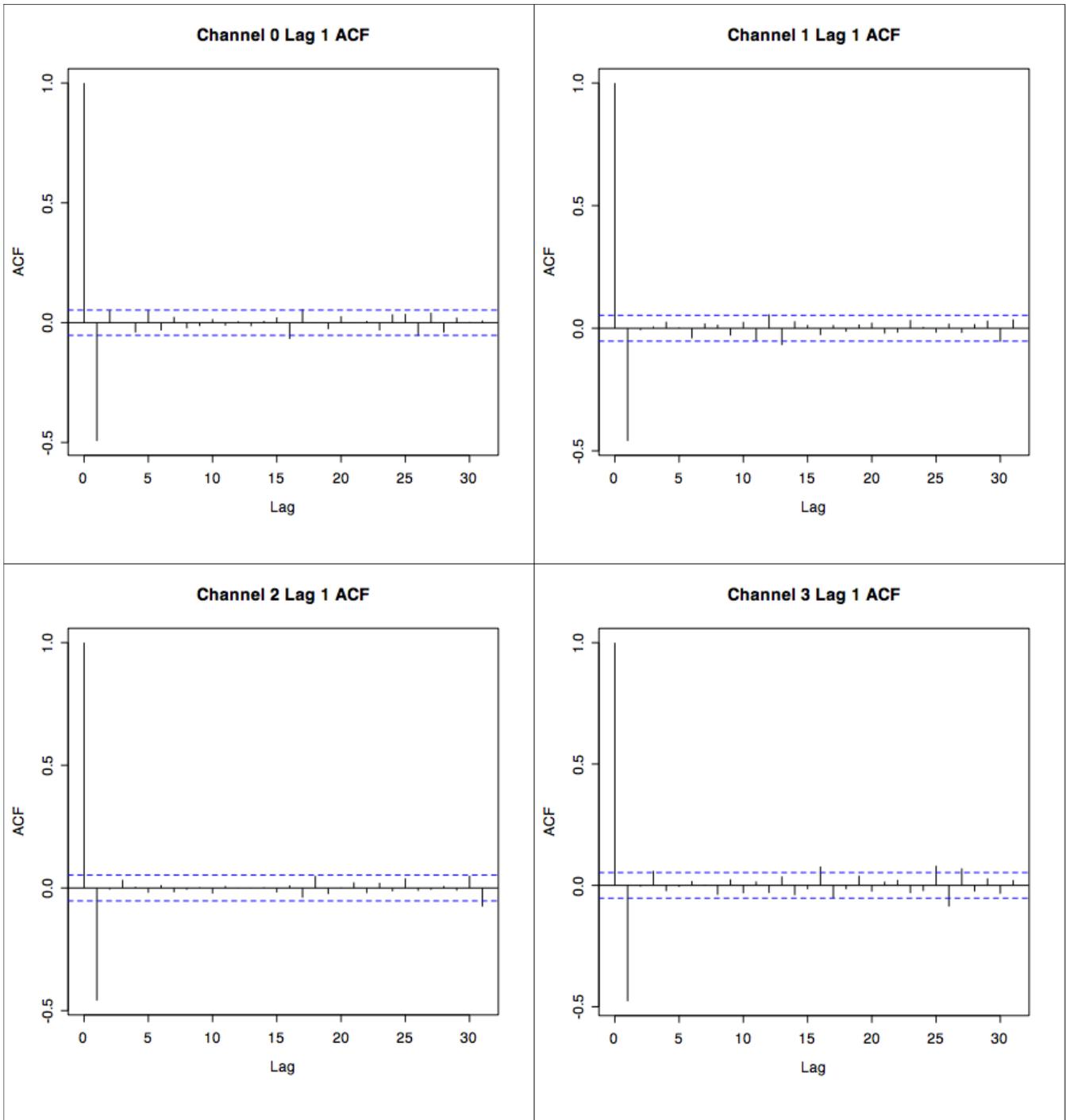


Illustration 9: ACF for channels 0 – 3 first-differenced data. The rapidly decreasing amplitudes over the lags signifies stationary series. Also, the order of the moving average component of the ARMA model (MA) can be determined.

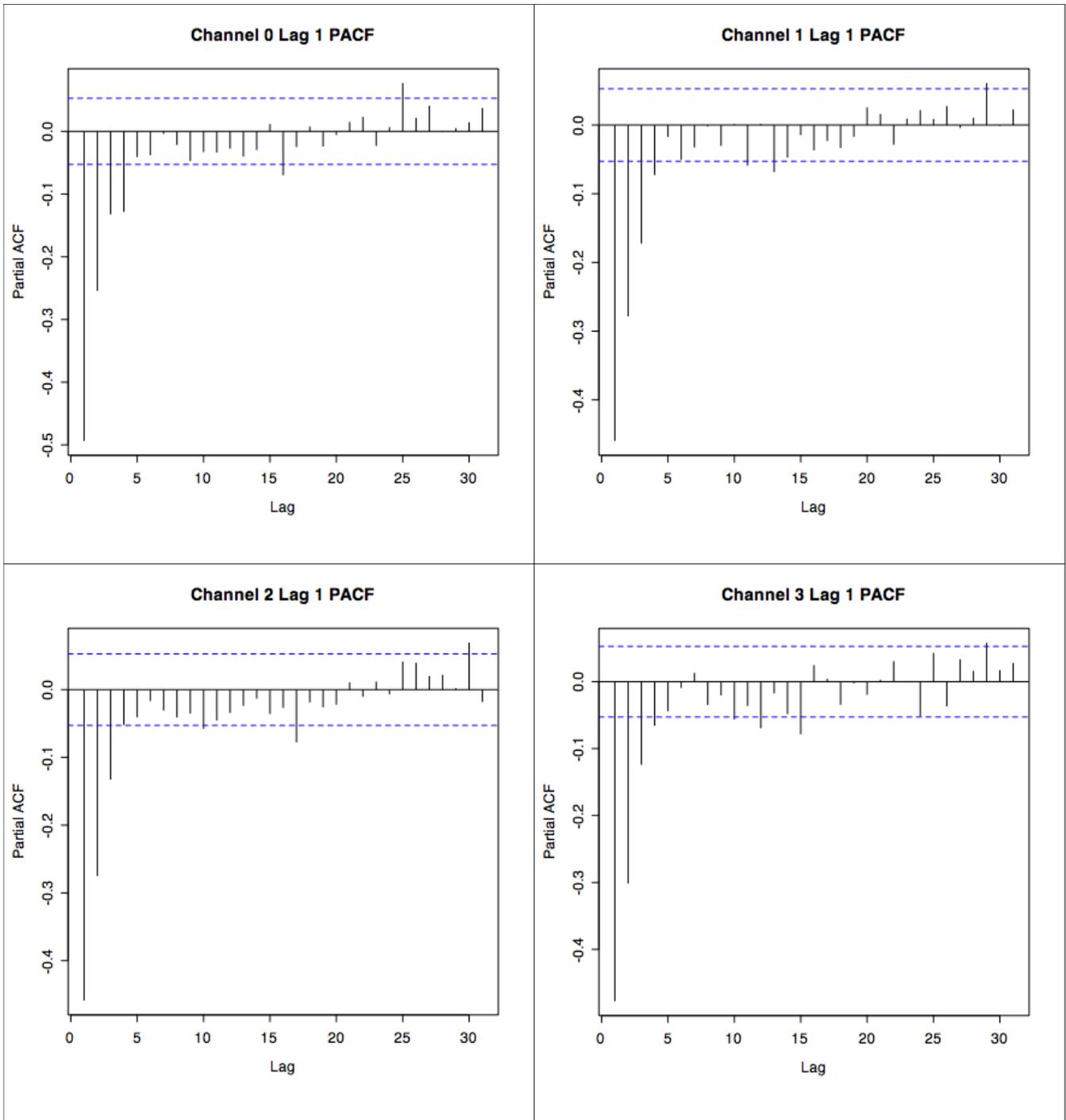


Illustration 10: PACF for channels 0 – 3 first-differenced data. The gradually changing amplitudes over the lags indicates these series likely have no autocorrelated behavior. The statistical significance of lags greater than 10 may need further investigation.

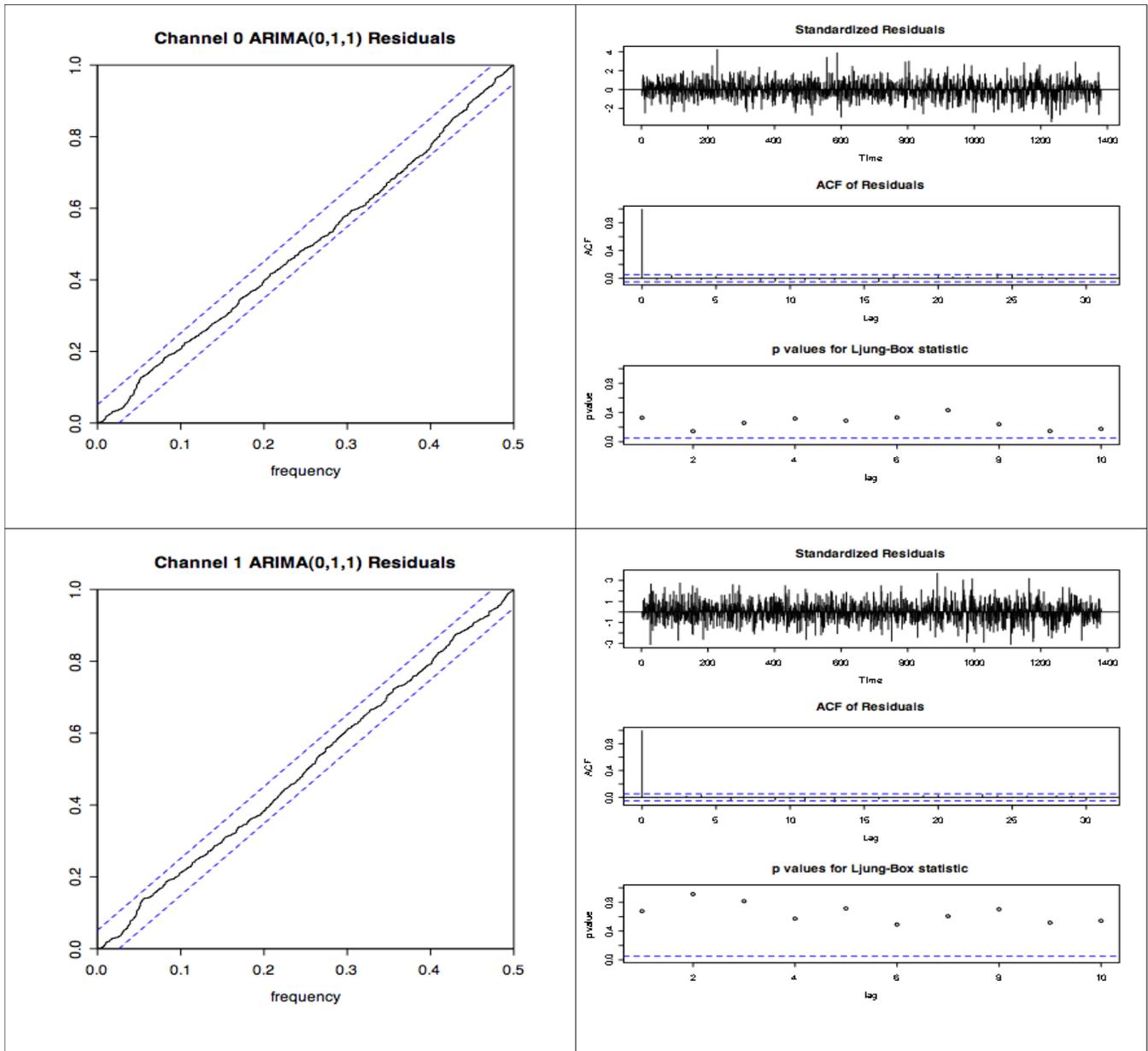


Illustration 11: ARIMA(0,1,1) model diagnostics support an ARMA(0,1) model for channels 0 and 1. The dashed lines indicate 95% confidence limits on the cumulative periodogram. As it lies completely within the limits, the residuals show the desirable characteristic of being uncorrelated. These model diagnostics 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.

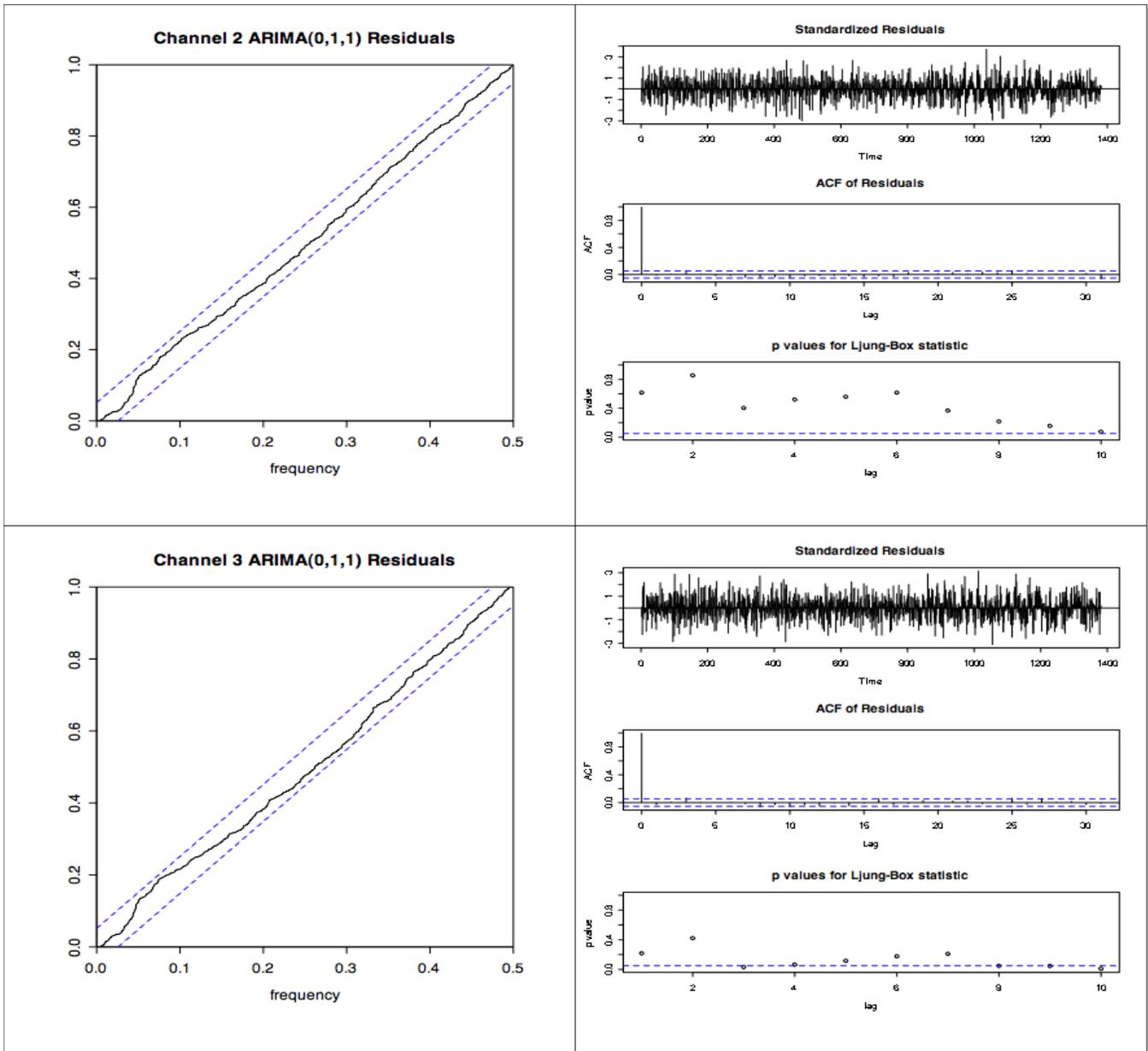


Illustration 12: ARIMA(0,1,1) model diagnostics support an ARMA(0,1) model for channels 2 and 3. The interpretation of the plots is as in Illustration 12.

Pairwise Scatterplots with Loess Smoothing

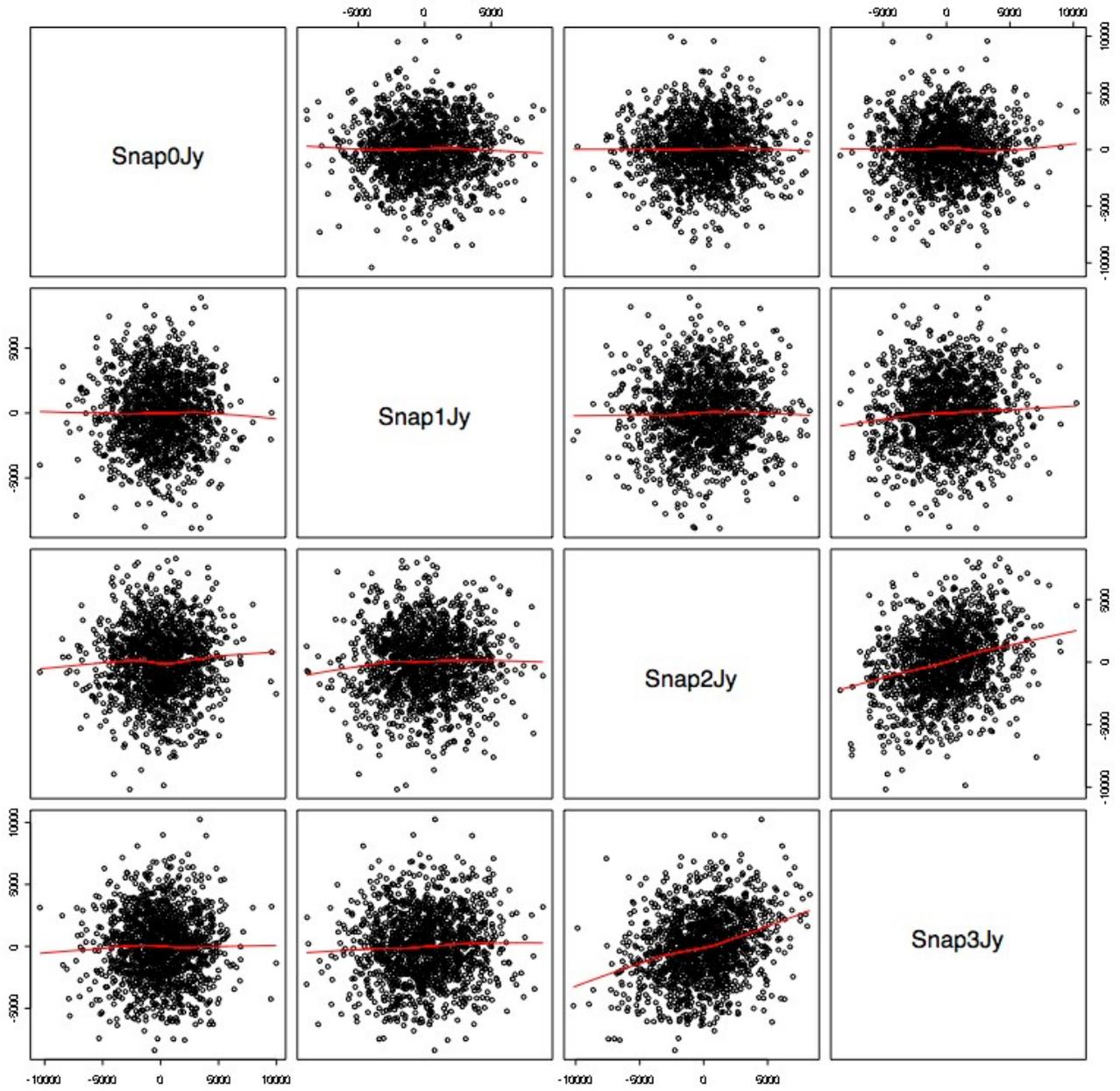


Illustration 13: Pairwise scatter plots of the channel lag 1 data with Loess smoothing.

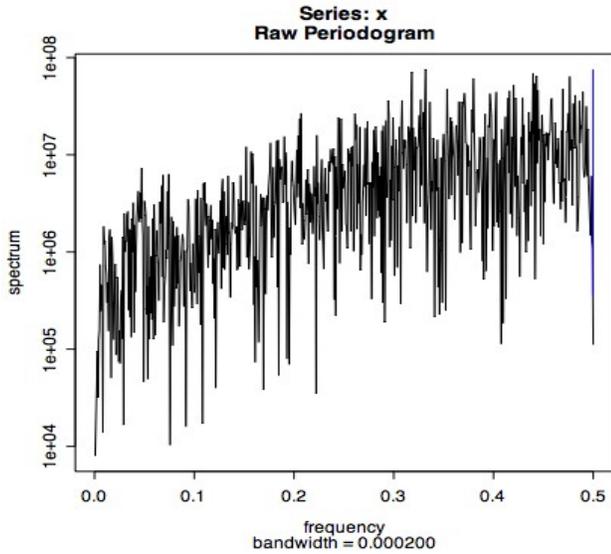


Illustration 14: Channels 0 - 3 first differences (lag 1) periodogram for Channel 3 which is representative of the other 3 channels. Note that the highest frequencies have the highest power levels, indicating low order specifiers..

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

where X_t is the t^{th} observation, and $\{Z_t\}$ is white noise with mean 0 and finite variance σ^2 if and only if $\{Z_t\}$ are not correlated, and is denoted

$\{Z_t\} \sim WN(0, \sigma^2)$. The parameter θ is to be determined from each series data. The model parameters are, by channel, in Table 3.

The ARMA(0,1) model works well (Box and Pierce, 1970) for all four channels as is seen in the model diagnostics in Illustrations 12 and 13, and from the Akaike Information Criterion (AIC), Akaike, 1969 in Table 3.

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.

Illustration 14 shows the lag 1 (first-differenced) pairwise channel scatter plots with loess smoothing (Cleveland, 1979, and Cleveland, et. al., 1988). We see that channel 0 doesn't appear correlated with channels 1, 2, and 3. Channel 1 appears not to be correlated with Channel 2, but may be correlated with Channel 3. Channels 3 and 4 appear to be more strongly correlated. To see if these relations indeed hold, we examine the cross-correlation

functions of each channel.

Each pairwise cross-correlation function is benign except that of channels 2 and 3 (Illustrations 15 and 16). The cross-correlation function in Illustration 15 depends on both the autocorrelation functions of channels 2 and 3 which confounds the test for independence of the channels 2 and 3. We circumvent this difficulty by "prewhitening" both series through their respective ARMA(0,1) filters. Thus, we cross-correlate the ARMA(0,1) residuals, for which we assume the ARMA(0,1) model is the true model. At any specific lag (the x-axis in Illustration 16), we look for lag lines that exceed the dashed horizontal lines (blue). Lines exceeding the dashed lines indicate that the associated lag is statistically significant at the 95% confidence level. We see that the ACF is 0.364 which, though statistically significant, is still marginally correlated on a ± 1 scale.

5. CONCLUSIONS

The time series analysis of the four-channel, total-power radiometer Polaris surrogate 20 Kelvin dummy load data gives the following results:

1. All four channels (0 - 3) are nonstationary time series.
2. All four channels 0 through 4 are described by an ARIMA(0,1,1) model.
3. The paired channels 2 and 3 are marginally cross-correlated, while all the other possible channel combinations are not 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 Polaris 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 Polaris data as a 20 Kelvin dummy load, or other physical phenomena is the cause. The cross-correlation analysis indicates that there may be current bleed-over from one channel to another in channels 2 and 3, 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.

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*, 2nd 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., (2008), March email correspondence with the author.

Shuch, R., (2008), March email correspondence with the author.

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

Cross-Correlation of Channels 2 and 3

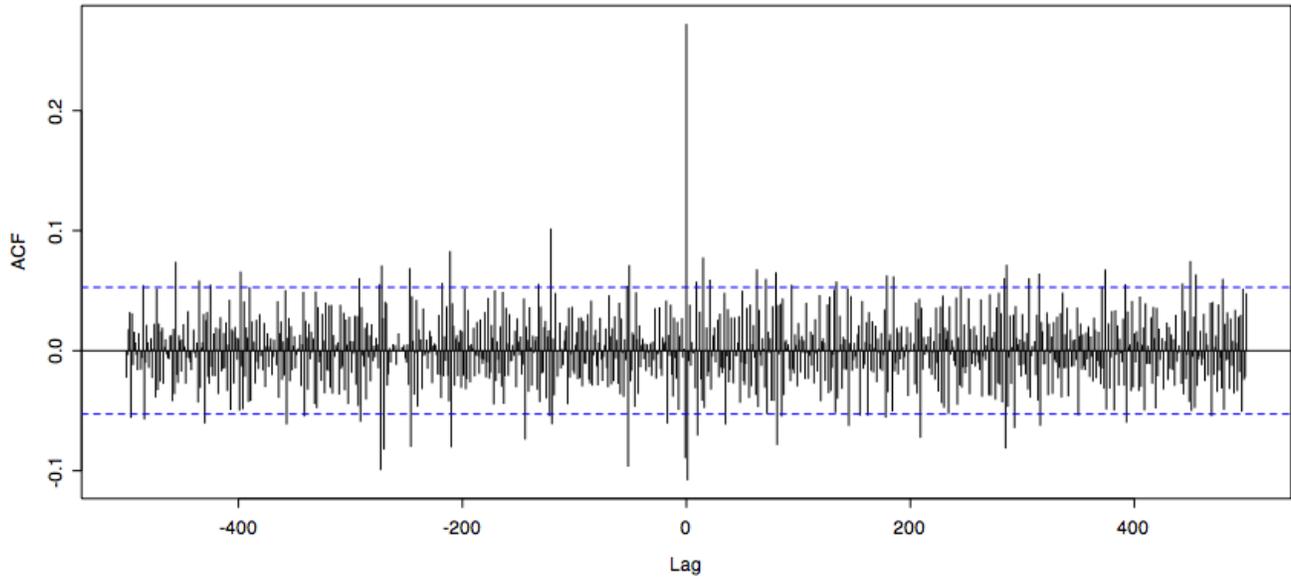


Illustration 15: Cross-correlation function for channels 2 and 3 for lags of ± 500 . The cross-correlations of the other channel combinations are benign ($ACF < 0.2 \ll 1.0$), whereas further investigation is needed for channels 2 and 3..

Cross-Correlation of Channels 2 and 3

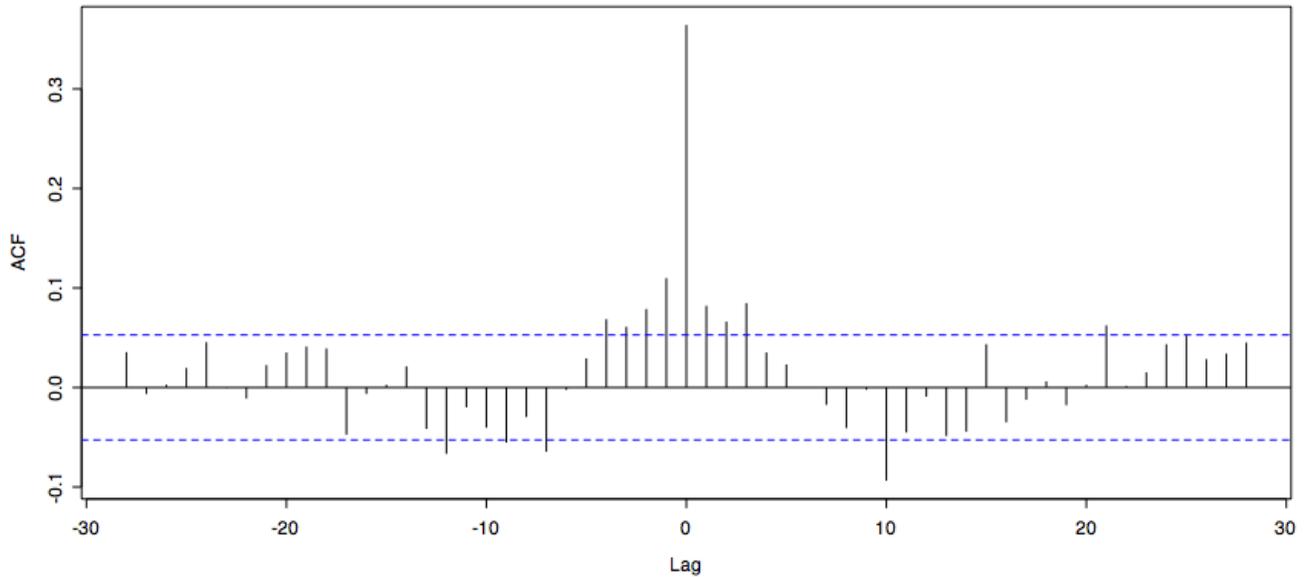


Illustration 16: Cross-correlation function of the ARMA(0,1) model residuals of channels 2 and 3. Using the residuals removes any cross-correlation ambiguity, revealing that channels 2 and 3 indeed are cross-correlated.