LA-13095-MS

[redacted]

Knowledge Fusion: An Approach to Time Series Model Selection Followed by Pattern Recognition

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

Edited by Paul W. Henriksen, Group CIC-1
Prepared by Sharon Hurdle, Group NIS-7

This work was supported by the U.S. Department of Energy, Office of Nonproliferation and National Security.

An Affirmative Action/Equal Opportunity Employer

LA-13095-MS

UC-700
Issued: March 1996

Knowledge Fusion: An Approach to Time Series Model Selection Followed by Pattern Recognition

Shirley A. Bleasdale
Thomas L. Burr
James C. Scovel
Richard B. Strittmatter

Los Alamos
NATIONAL LABORATORY
Los Alamos, New Mexico 87545

CONTENTS

Abstract

1. Introduction and Summary 1

2. Background 2

3. Multivariate Modeling of ARIMA Series 10

4. Modeling of Gamma Data 11

5. Summary 12

References 12

Appendix A. Example of Multivariate ARIMA Modeling 13

Appendix B. Summary of Gamma Modeling 16

FIGURES

1 Residual standard errors. Data sets 1-644 are class=true and data sets 645-730 are cals. 3

2 Plots of fitted coefficients. Data sets 1-644 are class=true and data sets 645-730 are cals. 4

3 Background data: gammas and flld. 7

4 Comparison of residuals before and after transformations. 8

TABLES

1 Comparison of Residual Standard Error for Six Models 5

2 Detection Probabilities on the Testing Set for Models A and C 9

Knowledge Fusion: An Approach to Time Series Model Selection Followed by Pattern Recognition

by

Shirley A. Bleasdale, Thomas L. Burr, James C. Scovel, and Richard B. Strittmatter

ABSTRACT

This report describes work done during FY 95 that was sponsored by the Department of Energy, Office of Nonproliferation and National Security, Knowledge Fusion Project. The project team selected satellite sensor data to use as the one main example for the application of its analysis algorithms. The specific sensor-fusion problem has many generic features, which make it a worthwhile problem to attempt to solve in a general way. The generic problem is to recognize events of interest from multiple time series that define a possibly noisy background. By implementing a suite of time series modeling and forecasting methods and using well-chosen alarm criteria, we reduce the number of false alarms. We then further reduce the number of false alarms by analyzing all suspicious sections of data, as judged by the alarm criteria, with pattern recognition methods. An accompanying report (Ref. 1) describes the implementation and application of this 2-step process for separating events from unusual background and applies a suite of forecasting methods followed by a suite of pattern recognition methods. This report goes into more detail about one of the forecasting methods and one of the pattern recognition methods and is applied to the same kind of satellite-sensor data that is described in Ref. 1.

---

1. Introduction and Summary

In a companion report (Ref. 1), the following idea was presented: in many situations we monitor a facility, an operation, or an environment and watch for unusual behavior. Such monitoring inevitably produces data recorded sequentially, which is known as time series data. Sections of the time series that are judged to be unusual must somehow be scrutinized. Reference 1 covered many possible ways to model the usual behavior so that reasonable rules could be constructed for locating unusual behavior. Then, all data that is labeled unusual would be further scrutinized by a suite of pattern recognition methods. No particular method of modeling the background nor of doing

the pattern recognition was discussed in great detail in Ref. 1. This report provides more detail on two of the ways of modeling the background and on one of the ways of doing pattern recognition. The two ways of modeling the background are regression-type model fitting and vector autoregressive moving average (ARMA) modeling. The one pattern recognition method is a version of Fisher's Discriminant Analysis. This report is organized as follows. Section 2 gives the specific time series that we used and the models that were applied to forecast that time series. Results for the regression-type model fitting are given. Results for the vector ARIMA modeling were not as good as for the regression-type model fitting. In section 3 we give further details about vector ARIMA modeling. We present results of the vector ARIMA modeling on simulated data in Appendix A. In section 4 we describe our approach to regression-type modeling in more detail, and Appendix B documents some of the results of modeling the time series described in section 2, with an emphasis on the computational environment and model-rejection criteria. Section 5 is a summary.

2. Background

In Ref. 2 we documented that a pulse can be estimated by subtracting out pulses of varying widths and heights at different locations from the series that has the event, fitting an autoregressive integrated moving average (ARIMA) model, and computing the sum of the squared residuals. The subtracted pulse that minimizes the sum of the squared residuals is a good estimate of the event pulse. Given a "catalog" of possible pulses, the pulse can then be judged to be either from the catalog or not from the catalog based on the magnitude of the sum of squared residuals. ARIMA modeling of the data was tried, but we could not fit coefficients in about one-third of the data sets. These poor results could be due to the scarcity of data points in each of the individual 730 data sets or, more likely, due to a lack of a strong functional relation such as is assumed in ARIMA modeling. The original data file contained over 1300 individual data sets, each with 25 concurrent gamma counts and aggregated flld counts. The flld counts (acronym not defined here) are proportional to the sum of the charged particle (electron and proton) counts. The flld counts were aggregated to match the time stamp of the gammas. The data file contained 3 different types of data with "trigger events" supposedly occurring near point 20. The three types of data were true, calibration (cal), and operator error (OPV). For this study, the OPV data was not used because its special characteristics reduced the data file to 730 individual data sets.

Because there was such variability between the individual data sets and ARIMA models could not be used effectively, it was decided to do one-step-ahead predictions of the gamma counts and minimize the squared residuals from these predictions. To predict the gammas, they were regressed on other components of the data using least squares. The residuals are defined as the difference between the observed values and the fitted values, denoted by "e" in the following models. Different combinations of the flld counts and the gamma counts were used as predictors of the observed present gammas. The six models tried were as follows:

(1) gammas = B0 + B1*(past flld) + B2*(present flld) + B3*(prior gammas) + e

(2) gammas = B0 + B1*(past flld) + B2*(present flld) + e

(3) gammas = B0 + B1*(prior gammas) + e

(4) gammas = B0 + B1*(present flld) + e

(5) gammas = B0 + B1*(past flld) + B2*(prior gammas) + e

(6) gammas = B0 + B1*(present flld) + B2*(prior gammas) + e.

We modeled the first 20 data points, so that "trigger events," which were supposed to occur at point 20, would not effect the models. The goal was to find one model that fit all of the background data accurately. However, we discovered that this could not be done due to the variability among the different data sets and the variations caused by the orientation of the satellite. Therefore, all 730 data sets were modeled individually with residuals computed for these 20 values.

The results of the modeling showed a tremendous difference between the "true" data and the "calibration" data in that the calibration data showed much more variability than the true data. In the search for true background data, the calibration data was later eliminated from the analysis. The variability was seen through the increased standard errors of the residuals in the calibration data for all the models. Figure 1 shows plots of the individual residual standard errors for all 730 data sets, which are indexed on the x-axis for four of the models.

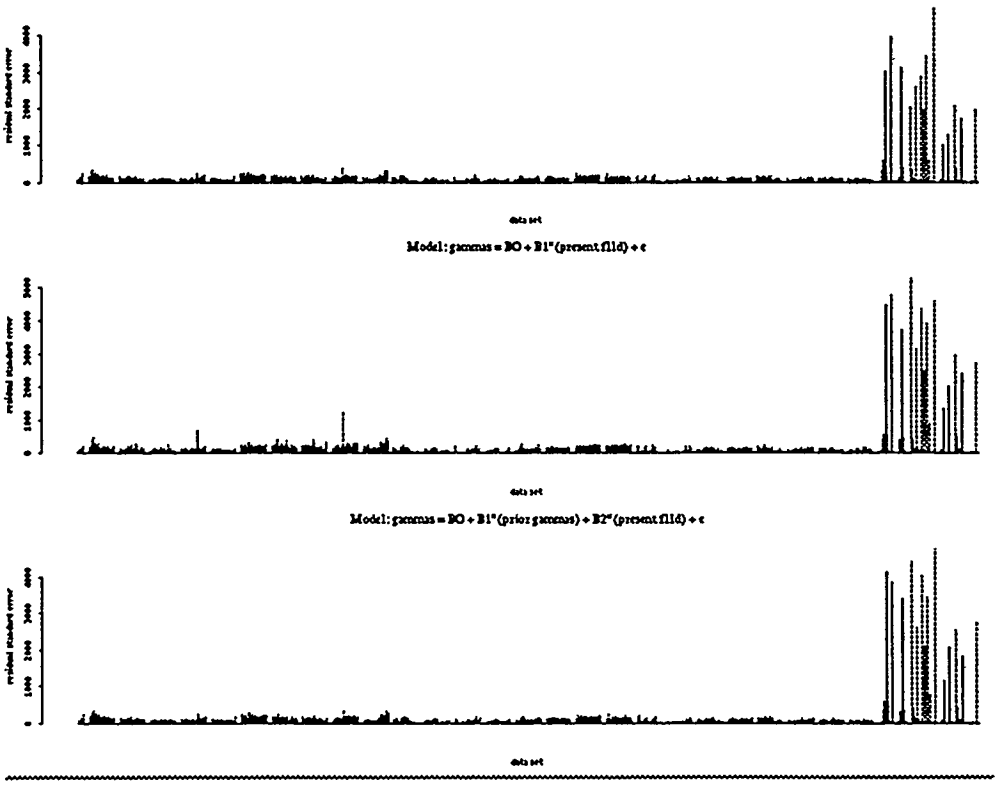

FIGURE 1. Residual standard errors. Data sets 1-644 are class=true and data sets 645-730 are cals.

The fitted coefficients for the models showcased the difference between the true and calibration data. Figure 2 displays the B1 coefficients for Model 3 and Model 4 for all 730 data sets. Notice that the B1 coefficients of the prior gammas were especially affected. The average B1 coefficient of the prior gammas in Model 3 was .69 and fairly stable for the true data, but the calibration data had several very large coefficients changing the mean of that data to 4.48.

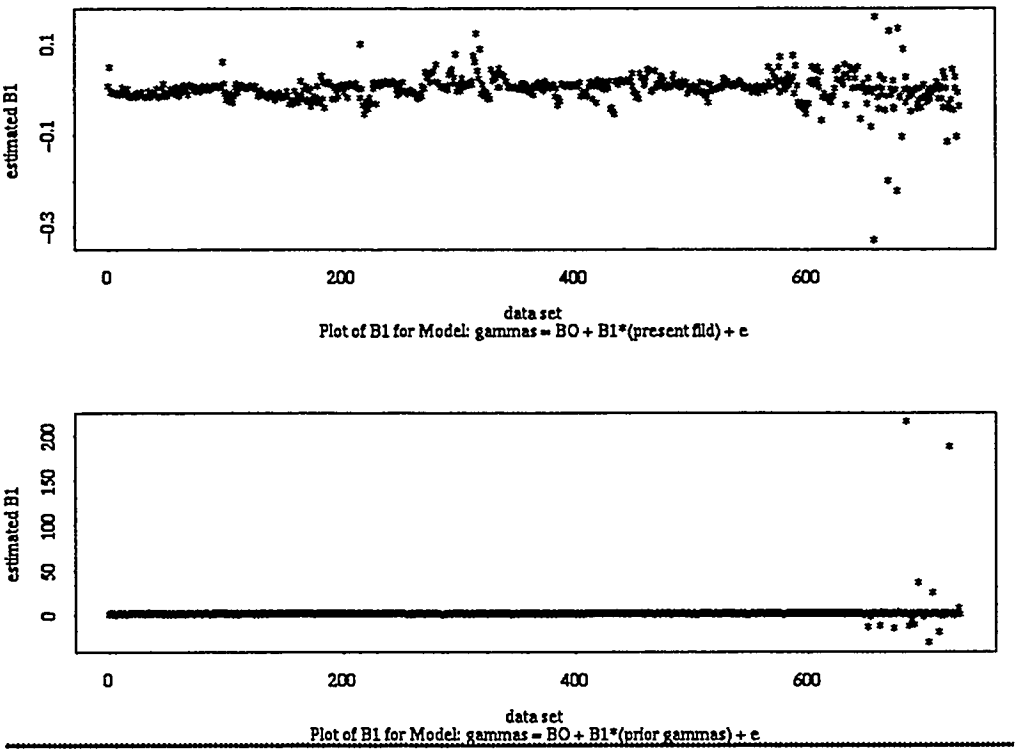

FIGURE 2. Plots of fitted coefficients. Data sets 1-644 are class=true, and data sets 645-730 are cals.

To choose the best model, we compared residual standard errors. With this method, the average magnitudes of the residual standard errors are assessed as the number of variables in the regression increases, indicating the best number of variables in the regression. The residual standard errors of all sets of "*p*" predictors, where *p* is the number of predictors in the model including B0, are compared. The residual standard errors from all 730 data sets were averaged to use in the comparison. The true and calibration data were separated to determine if the results were affected by the variability noted in the calibration data. Table 1 summarizes the results. In Table 1, *p* designates the number of predictors in the model. Thus, for the first line in "All data," 2 is the number of predictors including B0 in the model; 119.3498 and 224.4401 are the average residual standard errors for Models 3 and 4, respectively; and 211.895 is the average of these two standard errors.

The average residual standard errors are high for all the models, once again indicating that more data points per set would have been beneficial in establishing good models of the background data. Model 1 has the lowest average residual standard errors for the combined data (true and cal). Model 1 contains both past and present flld counts and prior gamma counts. The next best model is model 6, which uses the present flld counts and the prior gamma counts. In the nudet detection procedure, models 1 and 6 will be used along with model 3, which uses only prior gamma counts, and model 4, which uses present and past flld counts.

TABLE 1. Comparison of Residual Standard Error for Six Models

All Data:

Model	p	Residual Standard Error	Avg. RSE for models with given p
1	4	163.2	163.2
2	3	203.4	187.9
3	2	199.4	211.9
4	2	224.4	211.9
5	3	183.3	187.9
6	3	177.1	187.9

True data:

Model	p	Residual Standard Error	Avg. RSE for models with given p
1	4	163.2	103.2
2	3	132.9	115.9
3	2	113.2	126.9
4	2	140.5	126.9
5	3	108.3	115.9
6	3	105.4	115.9

Cal data:

Model	p	Residual Standard Error	Avg. RSE for models with given p
1	4	612.4	612.4
2	3	731.4	730.3
3	2	844.6	848.7
4	2	852.9	848.7
5	3	745.1	730.3
6	3	714.5	730.3

Detection of the Pulse

To test our methods in detecting a pulse, a library of one pulse was used because a generalization to a larger library could be made by a multi-dimensional extension of the Neyman-Pearson Lemma or other multivariate discriminant procedures.

The four models selected previously were

Model A: gammas = B0 + B1*(past flld) + B2*(present flld) + B3*(prior gammas) + e,

Model B: gammas = B0 + B1*(present flld) + B2*(prior gammas) + e,

Model C: gammas = B0 + B1*(past flld) + B2*(present flld) + e, and

Model D: gammas = B0 + B1*(prior gammas) + e.

Points 16-20 were predicted using one-step-ahead forecasting. Models A, B, and D were better predictors than Model C, as noted by the magnitude of the residuals and the standard errors of the fits. The standard error of the fits takes into account the variation in the estimated coefficients and the variation because the predicted value will not equal its expectation. Model C, which uses only flld counts to predict the gamma counts, has the largest residuals and standard errors of all the fits.

A pulse was then added in counts 16-20 in one-third of the data sets. The addition of the pulse did not change the individual estimation of the coefficients significantly. In the model, which used only flld counts to predict the gammas, the amount of change in the average value of the coefficient was less than .0001 and in the model that used gamma counts to predict gammas, the average change in the coefficient was about .006. Also, initial examination of the data and the residuals with the pulse in the data did not show a substantial difference from the data without the pulse. The models did show an increase in the second residual of the five points that constituted the pulse.

Because none of the models performed very well, a search for "pure" background data was initiated. All of the calibration data was eliminated because the variability in this data was so great. Then, the remaining true data was tested using the "trigger criteria" for points 10-25. The trigger criteria consisted of averaging the ten previous points, subtracting this average from the point being tested, which gives a residual value, and comparing this value to ten times the square root of the average of the ten points. If the residual value is larger than the trigger criteria, the point being tested is considered to be a trigger point. All data sets containing trigger points in counts 10-23 were eliminated from consideration and a subset of this data was taken for further analysis.

We then fitted Model A, which uses both flld and gamma counts, to fit the gammas; Model C, which uses only flld counts to fit the gammas; and Model D, which uses just gamma counts to fit the gammas. The residuals still showed a highly linear trend for all three models. Residuals from a least squares linear regression are assumed to be independent with zero mean and constant variance and follow a normal distribution. The linearity in the residuals was thought to be from the nonstationarity of the gammas and the periodicity of the flld counts as seen in Fig. 3. Figure 3 shows the averaged gamma counts for points 1-20 and the averaged flld counts for the same period.

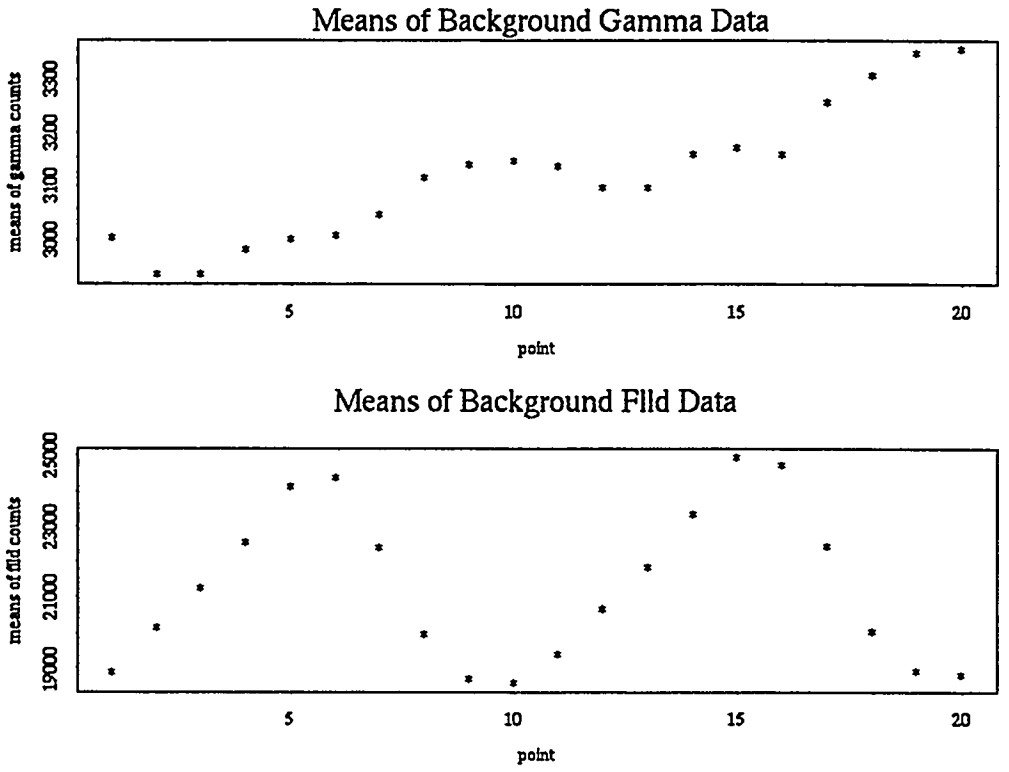

FIGURE 3. Background data: gammas and flld.

Several transformations were then applied to the data to obtain better fits in the modeling process and a normal structure in the residuals. To remove the nonstationarity in the gammas, a first difference, as in time series modeling, was done. Also, to remove the periodicity in the flld counts, the mean flld counts were subtracted off. These two transformations together provided the best structure in the average residual values. Figure 4 shows a comparison of the residuals for Model C before and after the transformation.

Points 16-20 were predicted in various ways to establish a viable means of prediction. The following methods were used: (1) one-step-ahead predictions using data previous to that point to fit the model; (2) one-step-ahead predictions using only counts 1-15 to fit the model; and (3) predicting points 16-20 using the model obtained from counts 1-20. More reasonable predicted values were obtained from methods (1) and (3). Due to the scarcity of data points in each data set, method (2) performed very poorly.

Once the needed transformations of the data and methods to predict points 16-20 were established, procedures to classify the data with and without the pulse were explored. The first method examined used the sum of the inner products between the residuals of the predicted points and the counts constituting the known pulse as a means to discriminate between series that contained pulses and series that did not. This is an application of the Neyman-Pearson Lemma. The level of the discriminant was determined to contain at least 95% of the population without the nudet. However, the detection probabilities were dismally low at 6%.

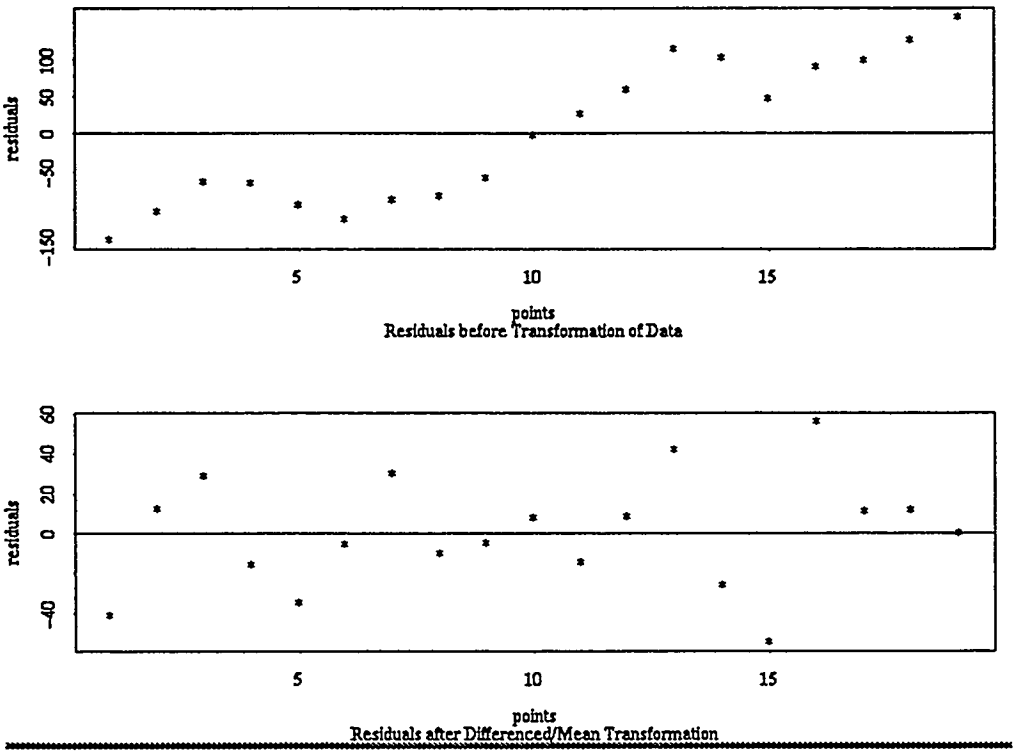

FIGURE 4. Comparison of residuals before and after transformations.

Even with the transformations to obtain Gaussian white-noise residuals, the residuals of the predicted points (the last five residuals) were not meeting these criteria. Also, it was found that the models without a pulse and with a pulse had different covariance structures. Thus, the theoretical distribution was not being followed. Hence, a quadratic discriminant function based on the Neyman-Pearson Lemma was developed and implemented empirically.

The data was randomly divided into 4 subsets of 45 cases each. Two of the subsets were used as training sets and the other two subsets were reserved to be used as testing sets. We only used 180 cases for this stage because we put further restrictions on the candidate cases to ensure that those considered in this stage were more representative of ordinary background. For example, we eliminated all cases whose gamma counts over points 16-20 had variance in the upper 10%. One of the training sets contained a pulse in points 16-20 and the other set was background data. Model A, which uses both fild and gamma counts to fit the gammas, and Model C, which uses only fild counts to fit the gammas, were then applied to both training sets. Points 16-20 were predicted using one-step-ahead forecasting as described earlier. For each model, a mean vector consisting of the residuals from the predictions of points 16-20 and a covariance matrix for these residuals were extracted from the empirical fits. These mean vectors and covariance structures were used to compute the quadratic discriminant with a discriminant being developed for each model. The formula for the quadratic discriminant D is (Ref. 3):

$$D = -\frac{1}{2}\tilde{x}'\left(\Sigma_1^{-1} - \Sigma_2^{-1}\right)\tilde{x} + \left(\tilde{\mu}'_1\Sigma_1^{-1} - \tilde{\mu}'_2\Sigma_2^{-1}\right)\tilde{x} \qquad (1)$$

where Σ_1 and $\tilde{\mu}_1$ are the covariance matrix and mean vectors, respectively, from the training set without the pulse, and Σ_2 and $\tilde{\mu}_2$ are the respective covariance matrix and mean vector from the training set with the pulse added. The vector $\tilde{x}$ is the 5-component vector of residual values from the forecasts of points 16-20. This process was done separately for Model A and Model C because each model would give different predicted values. We empirically selected the threshold k_1 for D such that the training set's false alarm rate was 3/45 under Model A or C. It was assumed that Model C would best discriminate between the distributions with and without the pulse for two reasons: the sum of the entries of the training-set covariance matrix (which is the variance of the sum of the five forecast errors) for Model A was 1529, while that for Model C was 799. Because the variance of the sum of the five forecast errors is a measure of the "noise" during the nudet pulse, the noise using Model A was larger than the noise for Model C, and any model that uses prior gammas will be penalized in terms of signal detection because the gamma pulse lasted for more than one time period. However, that turned out not to be the case, as will be shown. The detection probabilities for the testing set are given in Table 2. Note that the detection probability in the "without pulse" case is the false alarm probability. Recall that the false alarm probability was set at approximately 5% (3/45) on the training data.

TABLE 2. Detection Probabilities on the Testing Set for Models A and C

Model	Detection Probabilities	
	Without Pulse	With Pulse
A	17%	33%
C	35%	20%

Here is our explanation for the inferior performance of Model C. The sum of the entries of the test-set covariance matrix (no nudet set) for Model A was 1302 and for Model C was 6693. It is impossible to extrapolate general conclusions about models such as C that use only flld to forecast the gammas. There is simply too much variability in the relation between the gammas and the flld counts.

The testing data had 45 without-pulse cases and 45 with-pulse cases, so the percent misclassified by Model A on the testing data was (83% + 33%)/2 = 58% and by Model C was (65% + 20%)/2 = 42.5%.

Different subsets of the data gave even lower detection probabilities, some as low as 11%, which indicates the large variability between the individual data sets.

Sometimes in the application of discriminant analysis, the two types of misclassification have the same "costs." In this case, if a false alarm has the same "cost" as a "failure-to-detect," then it makes sense to use the discriminant D in a classification rule as follows: allocate the data set associated with the residual vector $\tilde{x}$ to the distribution with no nudet if

$$D > k_2 \tag{2}$$

where

$$k_2 = \frac{1}{2}\ln\left(\Sigma_1/\Sigma_2\right) + \frac{1}{2}\left(\tilde{\mu}'_1\Sigma_1^{-1}\tilde{\mu}_1 - \tilde{\mu}'_2\Sigma_2^{-1}\tilde{\mu}_2\right). \quad (3)$$

In Eq. (3), all symbols mean the same as in Eq. (1).

We allocate $\tilde{x}$ to the distribution with the pulse otherwise. With this method, Model A classified approximately 68% of the data sets correctly, while Model C classified approximately 59% of the data sets correctly. These two percentages should be compared to the percentages 58% and 42.5%, respectively, that were achieved under the constraint that the false alarm rate would be 5% on the training data. On test sets, Model C performed consistently while there was much variation in the performance of Model A.

The poor detection probabilities could be a result of the scarcity of data points per individual data set, but they are more likely to be a result of a poor-to-nonexistent functional relation between prior gammas and present gammas or between present flld and present gammas and the weakness of the pulse compared to the noise of the ordinary background. Also, the variability between individual data sets, even for "true background" data, did not allow the fitting of one model for all the data sets. For a specified false alarm rate, given enough data points to fit adequate background models, we feel that the detection probabilities could be significantly improved. Also, see Ref. 1 for other ways to apply pattern recognition to this problem.

3. Multivariate Modeling of ARIMA Series

As mentioned in the introduction, no vector ARIMA model was better than the AR models presented in section 2. Therefore, we have not reported results of vector ARIMA modeling applied to the "event" records from section 2. However, to support the effort to maintain a suite of old and modern forecasting methods, we implemented the ability to apply vector ARIMA models to any time series. Vector ARIMA modeling is reasonably well-treated in some time series texts. See Ref. 4 for example.

Our current computational environment for time series modeling is within the statistical programming environment called S+. With the addition of the time series library written by Paul Gilbert of the Bank of Canada (Ref. 5), ARIMA modeling of time series may be implemented in S+. This addition to our library of capabilities strengthens our ability to model data accurately.

In the Gilbert library of S+ functions, if a model is not specified, then vector autoregressive (VAR) modeling is automatically done resulting in the best VAR model. Vector ARIMA modeling may be done by specifying the ARMA model to be fitted. Diagnostics may then be performed on the resulting model to determine the best fit. Also, the Gilbert library provides an automated process that compares various models using selected well-known criteria such as the Portmanteau test for goodness of fit, Akaike's information criteria (AIC), final prediction error values, and several other tests. See Appendix A for an example of vector ARIMA modeling on simulated data with a comparison of the models.

4. Modeling of Gamma Data

A series of different ARIMA models were tried on the gamma data to obtain the appropriate model that best fit the background data. No model has been completely satisfactory on all the different data sets. There is a periodicity due to the spin of the satellite that is hard to determine. Several periodic models have been tried with no success.

Also, there is much variation among the different sets of gamma data along with variations in the data obtained from different systems. In particular, System 2 gives much higher counts with a larger variance than System 1.

Two models seemed to give the best fits from preliminary modeling attempts. These models were a moving average of order one fitted to the first difference of the gammas, i.e., an ARIMA(0,1,1), and an autoregressive moving average of order (1,2) fitted to the first difference of the gammas, i.e., an ARIMA(1,1,2). These two models were compared on approximate one-hour segments of the gamma data.

The ARIMA(1,1,2) model consistently gave lower AIC values and a lower estimated variance of the innovation process "e." AIC is a penalized version of the log-likelihood function, which takes into account the total number of parameters estimated. The best model is supposed to be given by the model with the lowest AIC value. However, both the AIC value and the estimated variance of the errors were very high for both models and the diagnostics often showed very little difference between the models. Because neither model consistently fit well, it may be advisable to use the ARIMA(0,1,1) for its simplicity.

However, the best method to fit the gamma data may well be to use the electrons as an exogenous variable. This can be done using the Gilbert library for times series, which can be implemented in S+. The Gilbert library (Ref. 5) was developed by Paul Gilbert of the Bank of Canada and may also be used to fit multivariate time series. We have heard, but not confirmed, that the commercial software known as SCA is capable of fitting vector ARMA models. To date, using the Gilbert library is our only implemented way of fitting vector ARMA models.

Using the electrons to fit the gammas gives lower variances for the innovations and lower log-likelihood values than the previous models. However, the log-likelihood values and, thus, the AIC values should not be strictly compared with the previous modeling attempts because the gammas must be aggregated to fit the same time sequence as the electrons, giving different series than previously used in the modeling process.

Fitting the gammas using the electrons seems to be a viable means of modeling the data although it has not been extensively tested. This leads to further research in using our Kalman filtering approach to disaggregate data. If the electrons could be disaggregated appropriately giving more information instead of the current loss of information from aggregating the gamma counts, we feel a better model for the gammas could be established.

See Appendix B for a partial summary of modeling efforts for the gamma data. Each of the 10 data files used is specified followed by the range and variance of the series with a brief description of any trends or unusual patterns noted. The information about each model is then given with a

description of the diagnostic plots. Many of the series also contain information on other models not previously mentioned, which were also used to fit to the data. The information from modeling the gammas using the electrons is presented under the last data series.

5. Summary

This report documents two main accomplishments that complement the work in Ref. 1: (1) developing a connection between the pulse-subtraction method from Ref. 2 and the Neyman-Pearson Lemma from classical statistics and (2) implementing the ability to model vector ARIMA models. Prior to this, we were restricted to vector AR models. One next step is to explore the possibility of nonlinear ARMA-type models.

Reference 1 includes a special case of the Neyman-Pearson Lemma as one of seven candidate discrimination methods, but the setting is slightly different and particular details differ. At the time of writing Ref. 1, we were not able to apply vector ARIMA models. However, for the problem presented, we do not think that vector ARIMA models were needed. Therefore, this report includes the performance of vector ARIMA modeling only on simulated data, in Appendix A.

References

1. Tom Burr, Justin Doak, JoAnn Howell, Dave Martinez, and Richard Strittmatter, "Knowledge Fusion: Time Series Modeling Followed by Pattern Recognition Applied to Unusual Sections of Background Data," Los Alamos National Laboratory report LA-13075-MS (February 1996).

2. S. Bleasdale, T. Burr, A. Coulter, J. Doak, B. Hoffbauer, D. Martinez, J. Prommel, C. Scovel, R. Strittmatter, T. Thomas, and A. Zardecki, "Knowledge Fusion: Analysis of Vector-Based Time Series with an Example from the SABRS Project," Los Alamos National Laboratory report LA-12931-MS (April 1995).

3. R. Johnson and D. Wichern, *Applied Multivariate Statistical Analysis*, Second Edition (Prentice-Hall, 1988).

4. P. Brockwell and R. Davis, *Time Series: Theory and Methods*, Second Edition (Springer-Verlag, 1991).

5. P. Gilbert, "State Space and ARMA Models: An Overview of the Equivalence," Bank of Canada Working Paper 93-4.

Appendix A

Example of Multivariate ARIMA Modeling

In Appendix A we illustrate some of the multivariate ARIMA modeling capability in the new (1995) Gilbert Library (Ref. 5 describes 1993 version). We simulate data from a bivariate ARIMA model, then compare 2 attempts to estimate model parameters. The first attempt specifies the correct model, and the second attempt specifies an incorrect model.

The data was simulated in S+ from model 1:

$$\begin{bmatrix} 1 & -0.7 \\ 1 & -0.6L \end{bmatrix} \tilde{x}_t = \begin{bmatrix} 1 & 0.5 + 0.6L \\ 1 & -0.7 + 0.8L \end{bmatrix} \tilde{e}_t$$

where $\tilde{x}_t$ is a two-dimensional time series, $(x_a, x_2)_t$, $(e_1, e_2)_t$ is a two-dimensional vector of the innovation process, and L denotes the backshift operator (Ref. 4). The first equation from model 1 is therefore:

$$x_{1t} - 0.7x_{2t} = e_{1t} + 0.5e_{2t} + 0.6e_{2(t-1)} \ .$$

The second equation from model 1 is similar. Model 1 is fairly complicated, at least among linear models, primarily because of the presence of the unobserved innovation process. Estimating the coefficients of a specified ARIMA model is difficult, but one way to do it is to use the goodness of fit or likelihood value (from Gilbert's library of S+ functions) for a range of candidate coefficient values (for a given model). By specifying the model, we mean to specify the ARMA order (AR lag and MA lag). Model 1 is an example of specifying both the model and the coefficients, as must be done for the Gilbert library. So, if the Gilbert library were to be used to estimate the coefficients, each of a range of trial values of the coefficients must be tried, and the coefficients that maximized the negative log likelihood could be chosen as the estimated coefficients.

The simulated data was evaluated with the correct model and another model with the same dimensions. The results are summarized below along with a comparison of the two models.

Evaluation with the model used to simulate data:

> multi.testfit <- l(multi.testmodel, multi.test.sim) # function to fit maximum likelihood ARMA Model with negative log-likelihood value:
neg. log likelihood= 1521.26063023462
A =
1 -0.7
1 0-0.6L1

```
B =
1     0.5+0.6L1
1    -0.7+0.8L1


Summary statistics of the model fit:
neg. log likelihood= 1521.26063023462 sample length= 500
        [,1]     [,2]
RMSE 1.860476 1.096139

ARMA model:
    input dimension =                  # If there is no input variable specified for the
   output dimension = 2                # model, the input dimension is denoted by a
    order A = 1                        # blank space in the output summary.
    order B = 1
    order C = NA
     6 parameters
     4 non-zero constants
```

Evaluation with second model of same dimensions:

```
                                       # Note: This is known to be the wrong
                                       # model.

> multi.testfit3 <- l(multi.testmodel3, multi.test.sim)
Model with negative log-likelihood value:
neg. log likelihood= 2178.78854539421
A =
1     0.7-0.2L1
1    -0.1+0.6L1

B =
1    -0.5-0.6L1
1     0.7+0.2L1


Summary statistics from model fit:
neg. log likelihood= 2178.78854539421  sample length= 500
        [,1]     [,2]
RMSE 1.486102 3.209154

ARMA model:
    input dimension =
    output dimension = 2
    order A = 1
    order B = 1
```

order C = NA
8 parameters
4 non-zero constants

Comparison of Models:

The Gilbert library also has an excellent tabled format to compare the performance of the different models. The Portmanteau test is a goodness of fit test which looks at the sum of the squared values of the autocorrelations of the residuals. If the fitted model is appropriate, the statistic should be approximately distributed as a chi-square with $(M - p - q)$ degrees of freedom, where M is the number of autocorrelations used in the calculation, and p and q are the number of AR and MA terms, respectively, in the model.

The Akaike information criteria (AIC), Bayes information criteria (BIC), generalized cross validation (GVC), Rice criteria (Rice), and final prediction error (FPE) statistics are all similar and based on the negative log-likelihood but with different penalties for the introduction of additional parameter values. For instance, the BIC attempts to correct for overfitting with too many parameters of the AIC by putting a higher penalty on the addition of parameters. It should be noted that the FPE test was designed for pure autoregressive models, so it should not be considered when selecting the best ARMA models. In all these tests, the best fit is determined by the model that minimizes the statistic.

The likelihood function mentioned above considers the joint density of the random variables as a function of the parameter values rather than as a function of the observed values. The maximum likelihood estimate (mle) of the parameter vector is the vector of values that makes the observed data the "most probable." The maximum-likelihood method is both a method to choose the parameters and a measure of the goodness of fit of the covariance matrix of the parameters to the data.

A sample table is shown below for a comparison of the two models previously fitted for the simulated data. The "opt" row in the table designates the optimal model fit for each test. We see that the model from which the data was simulated gives the optimal fit for each of the tests.

```
> information.tests(multi.testfit, multi.testfit3)
              based on no.of parameters    .
      PORT  -ln(L)     AIC     BIC     GVC    RICE     FPE
Fit 1 463.4  1521.3   3054.5  3084.0  3054.6  3054.6  3054.5

Fit 2 1207   2179     4374    4413    4374    4374    4374

opt      1      1       1       1       1       1       1      #
 PORT  - Portmanteau test                    -ln(L)- neg. log likelihood
 AIC   - neg. Akaike Information Criterion   BIC   - neg. Bayes Information Criterion
 GVC   - Generalized Cross Validation        RICE  - Rice Criterion
 FPE   - Final Prediction Error
```

Appendix B

Summary of Gamma Modeling

This appendix documents some of our efforts to model the ordinary background gamma counts with a time step of 2.048 seconds. Most of the 10 data sets described here represent approximately 1 hour of data: approximately 1800 counts. The longer data sets such as data set (5) were broken into subsections of approximately 1 hour each.

The main concern in our gamma modeling effort was the dramatic nonstationarity. Our conclusion is that any "systems-study" of the nudet-detection ability should work with a "worst-case" background. Then, on a statistical-sampling basis, it should be verified that nearly any randomly selected section of background would render the nudet-detection probability at least as large as quoted for the "worst-case." The secondary concern is the general treatment of nonstationarity. In the nudet-detection case, the most reasonable approach is to work with the worst case. In other cases, a larger effort could be justified to model the nonstationarity. General techniques for modeling nonstationarity are not currently available.

Each of the 10 data files used is specified followed by the range and variance of the series with a brief description of any trends or unusual patterns noted. The information about each model is then given with a description of the diagnostic plots. Many of the series also contain information on other models not previously mentioned, which were also used to fit to the data. The information from modeling the gammas using the electrons is presented under the last data series. Again, we emphasize that the main conclusions from the analysis of these 10 randomly selected sections of data are the gamma time series is dramatically nonstationary and the relation between the gammas and the electrons or between the gammas and the protons is dramatically nonstationary.

The descriptions below are brief and have been made in an active S+ session. Comments are preceded by a # sign to improve readability.

Data Set 1

Note: Data in original form is approximately 1 hour long.

```
(a) range(gammas1.msevsmts)          # Range of the series #
  [1] 43 115

    var(gammas1.msevsmts)            # Variance of the series #
  [1] 119.8954

(b) tsplot(gammas1.msevsmts)         # Definite trend leveling out after the first 1500 counts.
```

(c) Fits:

(i) gammas1.msevsmtsft # Model #
$model:
$model$order:
[1] 0 1 1

$model$ma: # Estimated coefficient(s) of model #
[1] 0.9672681

$var.coef: # Covariance/Variance matrix for the estimates of the coefficients #
ma(1)
ma(1) 3.719956e-05

$aic: # Akaike's information criteria value #
[1] 12655

$sigma2: # Estimated variance of the innovation process #
[1] 87.37309

$n.used: # Number of observations used to compute the likelihood function. #
[1] 1731

Diagnostics: Diagnostics were very good! There were no spikes in the ACF plot of the residuals.

(ii) gammas1.msevsmtsarmaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] 0.1456266

$model$ma:
[1] 1.0943164 -0.1207158

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	1.322746	1.326049	-1.284332
ma(1)	1.326049	1.329929	-1.288094
ma(2)	-1.284332	-1.288094	1.247602

$aic:
[1] 12650.4

$sigma2:
[1] 87.28999

$n.used:
[1] 1730

Diagnostics: Diagnostics are once again excellent.

Data Set 2

Note: Data is already in approximate one hour segment.

(a) range(gammas2.msevsmts)
[1] 65 193

var(gammas2.msevsmts)
[1] 383.266

(b) tsplot(gammas2.msevsmts) # Definite trend leveling out at about count 2000.

(c) Fits:

(i) gammas2.msevsmtsft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9298777

$var.coef:

	ma(1)
ma(1)	7.817877e-05

$aic:
[1] 13298.36

$sigma2:
[1] 126.7587

$n.used:
[1] 1731

Diagnostics: Diagnostics are very good. There are no spikes in the ACF plot of the residuals.

(ii) gammas2.msevsmtsarmaft

```
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.5684401

$model$ma:
[1] 0.3612009 0.5295426

$var.coef:
             ar(1)      ma(1)      ma(2)
ar(1)   347.5116  347.3164 -323.0833
ma(1)   347.3164  347.1216 -322.9021
ma(2) -323.0833 -322.9021  300.3725

$aic:
[1] 13293.28

$sigma2:
[1] 126.689

$n.used:
[1] 1730
```

Diagnostics: Diagnostics are good, but p-values are not as good as in the ARIMA(0,1,1) model. There are no spikes in the ACF plot of the residuals.

Data Set 3

Note: Data was split into two approximate 1 hour series.

```
(a) range(gammas1.jtfsmts)
 [1] 2604 3769

 var(gammas1.jtfsmts)
[1] 38494.14

(b) tsplot(gammas1.jtfsmts)              # Definite trend .#
```

(I) First half of series: gammas1.jtfsmtsw1

```
(a) range(gammas1.jtfsmtsw1)
 [1] 2958 3769
```

var(gammas1.jtfsmtsw1)
[1] 19854.42

(b) ts.plot(gammas1.jtfsmtsw1) # Trend evident.

(c) Fits:

(i) gammas1.jtfsmtsw1ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9489382

$var.coef:
ma(1)
ma(1) 5.962626e-05

$aic:
[1] 19966.53

$sigma2:
[1] 9161.017

$n.used:
[1] 1669

Diagnostics: Diagnostics are fair. The residuals are okay with a few outliers between (-3, 3.5); the ACF plot of the residuals has two very small spikes (largest about .06); the Goodness of fit p-values fluctuate with the majority being below the desired cut-off.

(ii) gammas1.jtfsmtsw1armaft
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.4356843

$model$ma:
[1] 0.5359054 0.3898762

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.5842974	0.5908654	-0.5615307
ma(1)	0.5908654	0.5980157	-0.5682894
ma(2)	-0.5615307	-0.5682894	0.5401596

$aic:
[1] 19957.9

$sigma2:
[1] 9160.061

$n.used:
[1] 1844

Diagnostics: Diagnostics are once again fair. The residuals are good with the majority between (-2, 2.5) and a couple of outliers; the ACF plot of the residuals has a few (about 4) very small spikes (largest about .1); the Goodness of Fit p-value is good for the first third and then falls to zero.

(ii) gammas1.jtfsmtsw2armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.8500772

$model$ma:
[1] 0.09472108 0.85246845

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.01335090	0.01214253	-0.01172096
ma(1)	0.01214253	0.01119182	-0.01075532
ma(2)	-0.01172096	-0.01075532	0.01043831

$aic:
[1] 21857.45

$sigma2:
[1] 8241.077

$n.used:
[1] 1843

Diagnostics: Diagnostics are not as good as with ARIMA(0,1,1) model. The residuals are fair but have outliers which go to (-4,4); the ACF plot of the residuals has a few (about 4) very small spikes (largest about .1); the Goodness of Fit p-values are almost all zero.

Data Set 4

Note: Data was divided into two approximate 1 hour series.

```
(a) range(gammas2.jtfsmts)
 [1] 6686 8416

 var(gammas2.jtfsmts)
[1] 124255

(b) ts.plot(gammas2.jtfsmts)          # Definite trend in data --
                                        disturbance between 2000 and 3000. #
```

(I) First half of series: gammas2.jtfsmtsw1

```
(a) range(gammas2.jtsmtsw1)
  [1] 7166 8416

  var(gammas2.jtfsmtsw1)
 [1] 58222.48

(b) tsplot(gammas2.jtfsmtsw1)      # Downward trend evident with strong
                                     # disturbance (inverted V pattern)
                                     # between counts 2300 - 3000.
```

(c) Fits:

```
  (i) gammas2.jtfsmtsw1ft
  $model:
  $model$order:
  [1] 0 1 1

  $model$ma:
  [1] 0.9242624

  $var.coef:
                ma(1)
  ma(1) 8.73212e-05
```

$aic:
[1] 20316.36

$sigma2:
[1] 11299.87

$n.used:
[1] 1669

Diagnostics: Diagnostics show residuals which are very good with most contained between (-2,2) and a few outliers now appearing to be over 3; however, the ACF plot of residuals has small spikes (the largest about .2) and the Goodness of Fit p-values are zero.

(ii) gammas2.jtfsmtsw1armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.5718512

$model$ma:
[1] 0.2894483 0.5920161

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.03851194	0.03617829	-0.03297678
ma(1)	0.03617829	0.03437544	-0.03125480
ma(2)	-0.03297678	-0.03125480	0.02862657

$aic:
[1] 20295.47

$sigma2:
[1] 11216.48

$n.used:
[1] 1668

Diagnostics: Diagnostics are very like the ARIMA(0,1,1) model with the residuals appearing good -- mostly between (-2,2) with a few outliers reaching to 3: however, the ACF plot of the residuals has small spikes (largest is .2), and the Goodness of Fit p-values are zero.

(iii) gammas2.jtfsmtsw1ar2maft

```
$model:
$model$order:
[1] 2 1 1

$model$ar:
[1]  0.0290891 -0.1755446

$model$ma:
[1] 0.9087727

$var.coef:
                    ar(1)          ar(2)          ma(1)
ar(1)  6.800248e-04 7.481135e-05 0.0001164489
ar(2)  7.481135e-05 6.620617e-04 0.0001053123
ma(1) 1.164489e-04 1.053123e-04 0.0001374862

$aic:
[1] 20245.95

$sigma2:
[1] 10966.66

$n.used:
[1] 1667
```

Diagnostics: Diagnostics appear the same in the residuals with most of them between (-2,2) with a few outliers to 3;, the ACF plot of the residuals has spikes but they are smaller (largest is approx 1.5), but the Goodness of Fit p-values are still zero.

(II) Second half of series: gammas2.jtfsmtsw2

(a) range(gammas2.jtfsmtsw2)

```
[1] 6686 7746
```

var(gammas2.jtfsmtsw2)

```
[1] 30956.65
```

(b) tsplot(gammas2.jtfsmtsw2) # Definite downward trend that levels # out between counts of 5000 and 5500.

(c) Fits:

(i) gammas2.jtfsmtsw2ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9450301

$var.coef:
ma(1)
ma(1) 5.798165e-05

$aic:
[1] 22097.45

$sigma2:
[1] 9351.202

$n.used:
[1] 1844

Diagnostics: Diagnostics show residuals which are very good lying between (-2,2) with a few outliers to 3: however, the ACF plot of the residuals has small spikes (largest .2) and the Goodness of Fit p-values are zero.

(ii) gammas2.jtfsmtsw2armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.1815729
$model$ma:
[1] 0.6846331 0.2534347

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.07067470	0.06913138	-0.06510777
ma(1)	0.06913138	0.06812950	-0.06415164
ma(2)	-0.06510777	-0.06415164	0.06048709

$aic:
[1] 22077.13

$sigma2:
[1] 9290.148

$n.used:
[1] 1843

Diagnostics: Diagnostics are very similar to ARIMA(0,1,1) model with the residuals appearing slightly better in the area of outliers.

(iii) gammas2.jtfsmtsw2ar2maft
$model:
$model$order:
[1] 2 1 1

$model$ar:
[1] 0.07388476 -0.09843337

$model$ma:
[1] 0.9409501

$var.coef:

	ar(1)	ar(2)	ma(1)
ar(1)	6.016781e-04	2.357789e-05	7.026411e-05
ar(2)	2.357789e-05	5.933494e-04	6.553673e-05
ma(1)	7.026411e-05	6.553673e-05	7.708103e-05

$aic:
[1] 22053.91

$sigma2:
[1] 9232.596

$n.used:
[1] 1842

Diagnostics: Diagnostics are very similar to the other two models, but the ACF plot of the residuals has slightly smaller spikes (largest about .15); residuals are still good; Goodness of Fit p-values are still zero.

Data Set 5

Note: Data was divided into four approximate 1 hour series.

(a) range(gammas1.julnsmts)
[1] 596 1196

var(gammas1.julnsmts)
[1] 10597.61

(b) tsplot(gammas1.julnsmts) # Definite trend in whole data set, but may not show up in 1 hour segments. #

(I) First approximate 1 hour series: gammas1.julnsmtsw1

(a) range(gammas1.julnsmtsw1)
[1] 805 1196

var(gammas1.julnsmtsw1)
[1] 3016.162

(b) tsplot(gammas1.julnsmtsw1) # The data has a slight V-shape.

(c) Fits:

(i) gammas1.julnsmtsw1ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9708833

$var.coef:
ma(1)
ma(1) 3.450725e-05

$aic:
[1] 17924.39

$sigma2:
[1] 2800.488

$n.used:
[1] 1663

Diagnostics: Diagnostics are not good. A few of the residuals go are out of bounds (-4,4) [which may not be that bad for a data set this size]; there are spikes in the ACF plot of the residuals (largest about .2); the Goodness of Fit p-value is zero.

(ii) gammas1.julnsmtsw1arma2ft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.2970529

$model$ma:
[1] 0.7427035 0.2172329

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.09431710	0.09587969	-0.09318860
ma(1)	0.09587969	0.09804146	-0.09527644
ma(2)	-0.09318860	-0.09527644	0.09264688

$aic:
[1] 17908.2

$sigma2:
[1] 2785.428

$n.used:
[1] 1662

Diagnostics: Diagnostics are not very good. The residuals are not very bad (a few go to -4 or 4), but the ACF plot of the residuals has spikes (the largest about .2) and the Goodness of Fit p-value is zero.

(iii) gammas1.julnsmtsw1armaft
$model:
$model$order:
[1] 2 1 1

$model$ar:
[1] -0.06529752 0.04493841

$model$ma:
[1] 0.9693758

$var.coef:

	ar(1)	ar(2)	ma(1)
ar(1)	6.456263e-04	8.467541e-05	4.298284e-05
ar(2)	8.467541e-05	6.432612e-04	4.183274e-05
ma(1)	4.298284e-05	4.183274e-05	4.124380e-05

$aic:
[1] 17896.88

$sigma2:
[1] 2783.832

$n.used:
[1] 1661

Diagnostics: Diagnostics are not good. The results are almost identical to the ARIMA(0,1,1) model.

(II) Second approximate one hour series: gammas1.julnsmtsw2

(a) range(gammas1.julnsmtsw2)
[1] 761 1114

var(gammas1.julnsmtsw2)
[1] 2596.777

(b) tsplot(gammas1.julnsmtsw2) # The trend does not show up in the
graph as it did with window 1,
but ACF suggests a first difference.

(c) Fits:

(i) gammas1.julnsmtsw2ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9687148

$var.coef:
ma(1)
ma(1) 3.34737e-05

$aic:
[1] 19539.68

$sigma2:
[1] 2389.678

$n.used:
[1] 1840

Diagnostics: Diagnostics are not good. The residuals look good, but the ACF plot of the residuals has spikes (largest about .25), and the Goodness of Fit p-value is zero.

(ii) gammas1.julnsmtsw2arma2ft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.2579033

$model$ma:
[1] 0.7784477 0.1872956

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.1120445	0.1134570	-0.1106111
ma(1)	0.1134570	0.1154120	-0.1125081
ma(2)	-0.1106111	-0.1125081	0.1097206

$aic:
[1] 19520.35

$sigma2:
[1] 2373.503

$n.used:
[1] 1839

Diagnostics: Diagnostics are not bad in the plot of the residuals (there are a few near -4 or 4), but the ACF of the residuals has spikes (the largest about .25) and the Goodness of Fit p-value is zero.

(iii) gammas1.julnsmtsw2armaft
$model:
$model$order:
[1] 2 1 1

$model$ar:
[1] -0.07727684 0.02102533

$model$ma:
[1] 0.9645753

$var.coef:

	ar(1)	ar(2)	ma(1)
ar(1)	5.890838e-04	8.681054e-05	4.438434e-05
ar(2)	8.681054e-05	5.863812e-04	4.303862e-05
ma(1)	4.438434e-05	4.303862e-05	4.353074e-05

$aic:
[1] 19507.3

$sigma2:
[1] 2370.218

$n.used:
[1] 1838

Diagnostics: Diagnostics appear the same as with the ARIMA(0,1,1) model.

(III) Third approximate 1 hour series: gammas1.julnsmtsw3

(a) range(gammas1.julnsmtsw3)
[1] 661 1027

var(gammas1.julnsmtsw3)
[1] 3764.902

(b) tsplot(gammas1.julnsmtsw3) # Definite downward trend noted.

(c) Fits:

(i) gammas1.julnsmtsw3ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9587858

$var.coef:
ma(1)
ma(1) 4.385109e-05

$aic:
[1] 19395.22

$sigma2:
[1] 2196.943

$n.used:
[1] 1841

Diagnostics: Diagnostics are not good. There are several residuals that reach to (-4,4); the ACF plot of the residuals has spikes (largest about .175); and the Goodness of Fit p-value is zero.

(ii) gammas1.julnsmtsw3armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.04296537

$model$ma:
[1] 1.00488588 -0.05017581

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.06515140	0.06485310	-0.06228211
ma(1)	0.06485310	0.06509828	-0.06251568
ma(2)	-0.06228211	-0.06251568	0.06008130

$aic:
[1] 19373.92

$sigma2:
[1] 2179.433

$n.used:
[1] 1840

Diagnostics: Diagnostics are not very good -- very similar to the ARIMA(0,1,1) model.

(IV) Fourth part of series: gammas1.julnsmtsw4

(a) range(gammas1.julnsmtsw4)
[1] 596 936

var(gammas1.julnsmtsw4)
[1] 3474.853

(b) tsplot(gammas1.julnsmtsw4) # Plot shows definite downward trend.

(c) Fits:

(i) gammas1.julnsmtsw4ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.9670451

$var.coef:
ma(1)
ma(1) 3.909762e-05

$aic:
[1] 17143.03

$sigma2:
[1] 1805.963

$n.used:
[1] 1658

Diagnostics: Diagnostics are not as poor as before. There are three large residuals (which is not that bad); the ACF plot of residuals has very small spikes (largest about .07); however, the Goodness of Fit p-value is once again zero.

(ii) gammas1.julnsmtsw4armaft
$model:
$model$order:
[1] 1 1 2

```
$model$ar:
[1] -0.2769792

$model$ma:
[1] 0.7776408 0.1783456

$var.coef:
                ar(1)       ma(1)       ma(2)
ar(1)    0.06217476 0.06322449 -0.06128438
ma(1)    0.06322449 0.06487624 -0.06287207
ma(2) -0.06128438 -0.06287207  0.06099105

$aic:
[1] 17122.88

$sigma2:
[1] 1791.421

$n.used:
[1] 1657
```

Diagnostics: Diagnostics are the same as the ARIMA(0,1,1) model.

Data Set6

Note: The data was already in approximate 1 hour segments.

```
(a) range(gammas1.atvsmts)
  [1] 183 425

   var(gammas1.atvsmts)
  [1] 1193.366

(b) ts.plot(gammas1.atvsmts)          # Slight trend evident, more like a
                                      # wave pattern.

(c) Fits:

   (i)  gammas1.atvsmtsft
   $model:
   $model$order:
   [1] 0 1 1

   $model$ma:
   [1] 0.9536135
```

$var.coef:

	ma(1)
ma(1)	5.43619e-05

$aic:
[1] 16317.75

$sigma2:
[1] 1041.247

$n.used:
[1] 1667

Diagnostics: Diagnostics are not very good. The residuals are fair with a few outliers which go to (-4,4); the ACF plot of the residuals has spikes throughout (largest about .22); the Goodness of Fit p-value is zero.

(ii) gammas1.atvsmtsarmaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.199984

$model$ma:
[1] 0.83863582 0.09891227

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.06593341	0.06654309	-0.06343245
ma(1)	0.06654309	0.06775278	-0.06457218
ma(2)	-0.06343245	-0.06457218	0.06162072

$aic:
[1] 16297.76

$sigma2:
[1] 1032.639

$n.used:
[1] 1666

Diagnostics: Diagnostics are not good. The residuals are fair with a several outliers to 4 (no outliers on the negative side); the ACF plot of the residuals has spikes throughout (largest about .2); the Goodness of Fit p-value is zero.

(iii) gammas1.atvsmtsmapdft

```
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.9998757

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
              ma(1)
ma(1) 1.495368e-07

$aic:
[1] 16988.28

$sigma2:
[1] 1600.653

$n.used:
[1] 1662
```

Diagnostics: The residuals are fairly good with one large outlier which goes to 4; the ACF plot of the residuals has only one spike at lag 6 with a magnitude of -.46; the Goodness of Fit p-value goes above the desired cut-off and then drops to zero.

Data Set 7

Note: Data is already in approximate 1 hour segment.

(a) range(gammas2.atvsmts)
[1] 606 1069

var(gammas2.atvsmts)
[1] 5027.686

(b) ts.plot(gammas2.atvsmts) # Definite trend with wave pattern.

(c) Fits:

(i) gammas2.atvsmtsft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.8916038

$var.coef:
ma(1)
ma(1) 0.000123001

$aic:
[1] 17700.04

$sigma2:
[1] 2387.192

$n.used:
[1] 1667

Diagnostics: Diagnostics are not good. The residuals are not bad with all being within (-3,3) range; the ACF plot of the residuals has very large spikes throughout (largest about .5); the Goodness of Fit p-value is zero.

(ii) gammas2.atvsmtsarmaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.3141006

$model$ma:
[1] 0.83228805 -0.05725149

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.004987129	0.004893088	-0.004320532
ma(1)	0.004893088	0.005399092	-0.004710031
ma(2)	-0.004320532	-0.004710031	0.004341307

$aic:
[1] 17519.97

$sigma2:
[1] 2151.94

$n.used:
[1] 1666

Diagnostics: Diagnostics are slightly better in the residuals than with the ARIMA(0,1,1) model. The residuals are mainly between (-2,2) with a few outliers at the beginning and end of the series; the ACF plot of the residuals has very large spikes throughout (largest about .4 -- which is better than previous model); and the Goodness of Fit p-value is zero.

(iii) gammas2.atvsmtsmapdft
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.9998601

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
ma(1)
ma(1) 1.683188e-07

$aic:
[1] 17555.22

$sigma2:
[1] 2251.378

$n.used:
[1] 1662

Diagnostics: The residuals are mainly between (-2,2) with a few outliers (to 4) at the beginning and end of series; the ACF plot of the residuals has one spike at lag 6 with a magnitude of -.34; the Goodness of Fit p-value is once again zero.

Data Set 8

Note: Data was divided into two approximate 1 hour series.

(a) range(gammas1.antsmts)
[1] 3281 4666

var(gammas1.antsmts)
[1] 42429.18

(b) ts.plot(gammas1.antsmts) # Definite trend with V pattern.

(I) First half of series: gammas1.antsmtsw1

(a) range(gammas1.antsmtsw1)
[1] 3498 4494

var(gammas1.antsmtsw1)
[1] 24893.11

(b) ts.plot(gammas1.antsmtsw1) # Definite trend with V-pattern

(c) Fits:

(i) gammas1.antsmtsw1ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.8798249

$var.coef:

	ma(1)
ma(1)	0.0001352744

$aic:
[1] 20723.86

$sigma2:
[1] 14321.7

$n.used:
[1] 1670

Diagnostics: Diagnostics are not very good. The residuals range from (-3,3); the ACF plot of the residuals has very large spikes throughout (largest about .3); the Goodness of Fit p-values are zero.

(ii) gammas1.antsmtsw1armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] 0.4086561

$model$ma:
[1] 1.4111664 -0.5000728

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.01569154	0.01429208	-0.01235365
ma(1)	0.01429208	0.01346675	-0.01167458
ma(2)	-0.01235365	-0.01167458	0.01017513

$aic:
[1] 20679.58

$sigma2:
[1] 14006.67

$n.used:
[1] 1669

Diagnostics: The residuals are slightly better than in the ARIMA(0,1,1) model with containment between (-2.5,3) with most between (-2,2.5); the ACF plot of residuals has very large spikes (largest about .3); the Goodness of Fit p-values are zero.

(iii) gammas1.antsmtsw1mapdft
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.9999041

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
ma(1)
ma(1) 1.152015e-07

$aic:
[1] 21130.9

$sigma2:
[1] 18916.52

$n.used:
[1] 1665

Diagnostics: The residuals have a few outliers ranging from (-3,3); the ACF plot of the residuals has one spike at lag 6 with a magnitude of -.47; the Goodness of Fit p-value falls to zero after the first lag.

(II) Second half of series: gammas1.antsmtsw2

(a) range(gammas1.antsmtsw2)
[1] 3281 4666

var(gammas1.antsmtsw2)
[1] 53425.49

(b) ts.plot(gammas1.antsmtsw2) # Definite V-pattern in data.

(c) Fits:

(i) gammas1.antsmtsw2ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.8833405

$var.coef:
ma(1)
ma(1) 0.0001239896

$aic:
[1] 21755.39

$sigma2:
[1] 12549.64

$n.used:
[1] 1772

Diagnostics: The residuals are between (-2.5,3) with one outlier to 4; the ACF plot of the residuals has very large spikes tapering off (largest about .35); the Goodness of Fit p-value is zero.

(ii) gammas1.antsmtsw2armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] 0.3991508

$model$ma:
[1] 1.3698639 -0.4542322

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.02941173	0.02771449	-0.02410930
ma(1)	0.02771449	0.02656334	-0.02314019
ma(2)	-0.02410930	-0.02314019	0.02021095

$aic:
[1] 21729.33

$sigma2:
[1] 12416.34

$n.used:
[1] 1771

Diagnostics: The residuals are mostly between (-2,3) with one large outlier to 4; the ACF plot of the residuals has spikes which taper off (largest about .3); the Goodness of Fit p-values are zero.

(iii) gammas1.antsmtsw2mapdft
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.999837

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
ma(1)
ma(1) 1.845269e-07

$aic:
[1] 22228.31

$sigma2:
[1] 16925.93

$n.used:
[1] 1767

Diagnostics: The residuals are okay with more outliers appearing in the second half of the series; The ACF plot of residuals has one spike at lag 6 which has a magnitude of -.42; the Goodness of Fit p-value is above the desired cut-off and the falls to zero about lag 10.

Data Set 9

Note: Data was divided into two approximate 1 hour series.

(a) range(gammas2.antsmts)
[1] 7968 10290

var(gammas2.antsmts)
[1] 182528.1

(b) ts.plot(gammas2.antsmts) # Definite trend and V-shaped patterns
in data.

(I) First half of series: gammas2.antsmtsw1

(a) range(gammas2.antsmtsw1)
[1] 8146 10126

var(gammas2.antsmtsw1)
[1] 93709.29

(b) ts.plot(gammas2.antsmtsw1) # Trend and V-pattern are evident.

(c) Fits:

(i) gammas2.antsmtsw1ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.7987582

$var.coef:
	ma(1)
ma(1)	0.0002167577

$aic:
[1] 21519.41

$sigma2:
[1] 23067.79

$n.used:
[1] 1670

Diagnostics: The residuals are very good; the ACF plot of the residuals has large spikes throughout (largest about .5); the Goodness of Fit p-values are zero.

(ii) gammas2.antsmtsw1armaft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] 0.5176097

$model$ma:
[1] 1.4866318 -0.6265308

$var.coef:
	ar(1)	ma(1)	ma(2)
ar(1)	0.004458140	0.003582006	-0.002694642
ma(1)	0.003582006	0.003242020	-0.002497740
ma(2)	-0.002694642	-0.002497740	0.001992693

$aic:
[1] 21428.9

$sigma2:
[1] 21943.3

$n.used:
[1] 1669

Diagnostics: Diagnostics look almost identical to the ARIMA(0,1,1) model, but the spikes in the ACF plot of the residuals only go to .4 which is an improvement.

(iii) gammas2.antsmtsw1mapdft
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.9998574

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
ma(1)
ma(1) 1.713112e-07

$aic:
[1] 21437.69

$sigma2:
[1] 22744.83

$n.used:
[1] 1665

Diagnostics: The residuals look fairly good with a couple of outliers to -4; the ACf plot of residuals has no spikes after lag 10 -- there is a spike at lag 6 with magnitude of -.3 and one at lag 3 of magnitude .16; the Goodness of Fit p-values are zero.

(II) Second half of series: gammas2.antsmtsw2

(a) range(gammas2.antsmtsw2)
[1] 7968 10290

var(gammas2.antsmtsw2)
[1] 250011.7

(b) ts.plot(gammas2.antsmtsw2) # Definite trend with V-pattern.

(c) Fits:

(i) gammas2.antsmtsw2ft
$model:
$model$order:
[1] 0 1 1

$model$ma:
[1] 0.8080643

$var.coef:
ma(1)
ma(1) 0.0001958421

$aic:
[1] 22541.74

$sigma2:
[1] 19564.43

$n.used:
[1] 1772

Diagnostics: The residuals look very good with one outlier to -4; the ACF plot of the residuals has large spikes throughout (largest about .45); the Goodness of Fit p-values are zero.

(ii) gammas2.antsmtsw2armaft

```
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.8567515

$model$ma:
[1] -0.0982364  0.7452668

$var.coef:
                  ar(1)          ma(1)          ma(2)
ar(1)   0.002243993  0.0018113443 -0.0012744012
ma(1)   0.001811344  0.0017131436 -0.0009318839
ma(2) -0.001274401 -0.0009318839  0.0009747857

$aic:
[1] 22511.03

$sigma2:
[1] 19315.13

$n.used:
[1] 1771
```

Diagnostics: Diagnostics are almost identical to ARIMA(0,1,1) model (the spikes in the ACF plot of residuals are slightly smaller -- about .4).

(iii) gammas2.antsmtsw2mapdft

```
$model:
$model[[1]]:
$model[[1]]$order:
[1] 0 1 1

$model[[1]]$ndiff:
[1] 1

$model[[1]]$ma:
[1] 0.9980484

$model[[2]]:
$model[[2]]$order:
[1] 0 1 0
```

$model[[2]]$period:
[1] 5

$model[[2]]$ndiff:
[1] 1

$var.coef:
ma(1)
ma(1) 2.206788e-06

$aic:
[1] 22634.15

$sigma2:
[1] 21316.15

$n.used:
[1] 1767

Diagnostics: Residuals look fairly good with one outlier to -4; the ACF plot of the residuals has one spike of magnitude -.35 at lag 6; the Goodness of Fit p-values are zero once again.

Data Set 10

Note: Data in original form is approximately 1 hour long.

(a) range(gammas1.marscsmts)
[1] 46 187

var(gammas1.marscsmts)
[1] 615.749

(b) tsplot(gammas1.marscsmts) # First 2600 counts on graph show
no trend, but large jumps appear
after that. They almost looks like steps.

(c) Fits:

(i) gammas1.marscsmtsft
$model:
$model$order:
[1] 0 1 1

$model$ndiff:
[1] 1

$model$ma:
[1] 0.878984

$var.coef:
ma(1)
ma(1) 0.0001295653

$aic:
[1] 13534.46

$sigma2:
[1] 130.5927

$n.used:
[1] 1755

Diagnostics: The residuals are mainly between (-2,2) with a very large outlier near the end of the series; there are no spikes in the ACF plot of the residuals; the Goodness of Fit p-value is below the desired cut-off for the entire series.

(ii) gammas1.marscsmtsarma2ft
$model:
$model$order:
[1] 1 1 2

$model$ar:
[1] -0.9023966

$model$ndiff:
[1] 1

$model$ma:
[1] -0.02025184 0.78905851

$var.coef:

	ar(1)	ma(1)	ma(2)
ar(1)	0.2495446	0.2536002	-0.2238983
ma(1)	0.2536002	0.2579368	-0.2275164
ma(2)	-0.2238983	-0.2275164	0.2011029

$aic: [1] 13530.09

$sigma2:
[1] 130.5079

$n.used:
[1] 1754

Diagnostics: The residuals look much like the ARIMA(0,1,1) model as does the ACF plot of the residuals; the Goodness of Fit p-value is lower than the previous model.

(iii) gammas1.marscsmtsarmaft
$model:
$model$order:
[1] 2 1 1

$model$ar:
[1] 0.08041740 0.07409809

$model$ndiff:
[1] 1

$model$ma:
[1] 0.9124628

$var.coef:

	ar(1)	ar(2)	ma(1)
ar(1)	0.0007379196	0.0001036996	0.0001611430
ar(2)	0.0001036996	0.0007044850	0.0001444922
ma(1)	0.0001611430	0.0001444922	0.0001522093

$aic:
[1] 13509.92

$sigma2:
[1] 129.596

$n.used:
[1] 1753

Diagnostics: The plot of the residuals appears much as before as does the ACF plot of the residuals; however, the Goodness of Fit p-value is above the desired cut-off for the entire series. This appears to be the best fit of these three.

Using electrons to forecast the gammas:

(iv) Using electrons as exogenous variable to fit gammas.

Data Set 1

Note: Gamma counts were aggregated to match electron counts.

model:
A =
1-0.9L1

B =
1+0.02L1-0.5L2

C =
0.04-0.002L1

neg. log likelihood:
2286.60615004514

RMSE:
44.24452

n:
439

Diagnostics: The residuals looked very good until the last one-fourth of the plot where some were rather large, but the kernel estimate of the residual distribution appeared normal centered at zero; the auto-correlation plot of the residuals has spikes throughout which are all positive except for the first one which is approximately -.17 (largest is approximately .5 with most around .2); partial autocorrelation plot has spikes in first three positions (largest .5 others smaller).

Note: The AIC values for this model and the previous models can not be compared since the number of observations is not the same.

This model can be improved, but illustrates the possibility of obtaining better fits using the electrons as an exogenous variable in fitting the gammas. The residual mean square error is much less as is the log-likelihood.

(v) AR models with electrons as exogenous variables:

(A) Using channel one of electrons to model first channel of gammas.

(a) electgam1.marsc.armodel

A =
1-0.2684L1-0.07422L2-0.2989L3-0.2342L4-0.02245L5-0.09733L6

B =
1

C =
0.04111-0.002607L1-0.003954L2-0.02223L3+0.00488L4-0.01382L5

neg. log likelihood= 2152.12787570545 sample length= 439

RMSE 32.57037

model estimated by est.VARX.ls
inputs: ts(elect1.marsc.data, start = 61.416748, delta = 8.192)
outputs: ts(gammas1.marsc.datasm, start = 61.416748, delta = 8.192)
ARMA model:
input dimension = 1
output dimension = 1
order A = 6
order B = 0
order C = 5
12 parameters
2 non-zero constants

Diagnostics: The residual plot looks very good until the last fourth of the series where there is more fluctuation and many more outliers, but the kernel estimate of the residual distribution appears very normal centered a little to the left of zero; the ACF plot of the residuals has a few spikes at the beginning with the largest having a magnitude of .2; the Partial ACF plot of the residuals has very small spikes at lags 50 and 150.

(B) Using all three channels of electrons to model first channel of gammas.

(i) elect3gam1.marsc.armodel

A =
1-0.2344L1-0.02381L2-0.2803L3-0.1836L4-0.02597L5-0.08235L6

B =
1

C =
0.03111-0.008811L1-0.004096L2-0.02934L3+0.0009635L4-0.008111L5
0.3021+0.2568L1+0.1699L2+0.2955L3+0.1077L4-0.1398L5
-0.2368+0.5015L1-0.03324L2+0.1502L3+0.1065L4+0.2861L5

neg. log likelihood= 2142.93402995751 sample length= 439

RMSE 31.89535

model estimated by est.VARX.ls
inputs: Series 1 Series 2 Series 3
outputs: ts(gammas1.marsc.datasm, start = 61.416748, delta = 8.192)
ARMA model:
input dimension = 3
output dimension = 1
order A = 6
order B = 0
order C = 5
24 parameters
2 non-zero constants

Diagnostics: The plot of the residuals looks good for the first 3/4's of the data then starts to fluctuate and has very large outliers, but the kernel distribution of the residuals looks like a normal distribution centered at zero; the ACF plot of the residuals has a couple of spikes at the beginning of the series (with magnitude of largest .2) and then a very small spike at lag 50; the Partial ACF plot of residuals has small spikes at the beginning, lag 50, and lag 150.

Comparison of the three models:

information.tests(electgam1.marsc.mymodel7df, electgam1.marsc.armodel, elect3gam1.marsc.armodel)

based on no.of parameters

dim.	PORT	-ln(L)	AIC	BIC	GVC	RICE	FPE
Fit 1	423.1	2286.6	4583.2	4603.6	4583.3	4583.3	4583.2
Fit 2	47.76	2152.13	4328.26	4377.27	4328.59	4328.94	4328.26
Fit 3	47.01	2142.93	4333.87	4431.90	4335.23	4336.70	4333.92
opt	3	3	2	2	2	2	2

PORT - Portmanteau test
-ln(L)- neg. log likelihood
AIC - neg. Akaike Information Criterion
BIC - neg. Bayes Information Criterion
GVC - Generalized Cross Validation
RICE - Rice Criterion
FPE - Final Prediction Error

The "opt" row designates the best model for each criteria. Thus, it appears that using the first channel of the electrons does as well as using all three channels except for the first two tests. However, the values are very close in both cases. It was found that if the data was windowed, all three channels did as well fitting the data with no "disruptions", but data with disruptions was fit better by the first channel alone.

$\tilde{\mathbf{x}}$

Los Alamos
NATIONAL LABORATORY
Los Alamos, New Mexico 87545