Introduction

These programs and this manual were written with beginners in mind. We have a lot of sympathy for researchers who have an idea of what they want to learn from their data and who hope for easily intelligible output. They would like a set of flexible, powerful programs that they can grow into as they learn about time-series analysis. They would be likely to avoid using programs that require a great deal of technical knowledge even to punch the control cards. For example, consider researchers who think their data may be cyclic. They skim a few chapters in a time-series book and learn that there is a thing called a spectral density function, and if that has a statistically significant peak, their data has a cycle at that particular frequency. Then they read the computer manual and find that they need to answer the question, "Which spectral window do you want to use? Punch 1 for Tukey-Hanning, 2 for Bartlett, 3 for Parzen, etc." They are likely to be discouraged from ever using the program if they have to answer perplexing questions such as this.

What we have done is to make these decisions for the user by limiting the number of options. We have done this only when it makes very little practical difference which option is chosen. We specify the equations we use so that an experienced user can modify the programs to suit special needs.

We think that these are straightforward programs to use. We also think that the output is useful for making statistical decisions.

Overview of time-series analysis

For a more detailed, but still nonmathematical, overview of time-series analysis see Chapters 1 to 7 of Gottman's book *Time-Series Analysis.* There are two kinds of time-series analyses, *time-domain* and *frequency domain.* Time-domain analysis is useful for such tasks as forecasting and frequency domain analysis is useful for such tasks as detecting cyclicity. However, they really ought be go hand in hand. For example, in building a good time-domain model it is useful to know if there are cyclic, or "seasonal" components and this is accomplished in part by a frequency-domain analysis.

One of the things researchers may wish to do is to fit a model to one time-series. The model breaks the data into understandable components such as linear trend, cycles, and a stochastic stationary time-series. There are a variety of applications of model fitting. These include forecasting and testing the effectiveness of interventions. Another kind of time-series analysis is multivariate. This involves assessing the relationships among two or more time-series. Once again, this can be done in either the time domain or the frequency domain.

Overview of the programs

There are three packages of programs described in this manual. The first package is for *univariate model fitting*. It is described in the box below.

PACKAGE 1: UNIVARIATE MODEL FITTING

A. TIME DOMAIN

1. *Removing Trend.* Program DETRND removes linear trend. It tests the significance of the trend line parameters and outputs the residual time-series, which is the deviations of the original series around the trend line.
2. *Linear Filter.* Program LINFIL is useful for transforming a time-series in a variety of ways, including differencing, smoothing, and removal of a curve (see the Porges example in Chapter 9 of Gottman's book).
3. *Removing Cycles.* Program DESINE will remove a cycle of any period specified by the user. It performs a least-squares fit that finds the best amplitude and phase parameters. The output is a test of the significance of the cycle removed and a residual time-series.
4. *Autoregressive Fit.* Program ARFIT performs an autoregressive model fit up to the order specified by the user. It outputs the variance of the series, its autocovariances, autocorrelations, partial autocorrelation coefficients, the autoregressive model coefficients, tests of their significance, the residual from the model fit and a Box-Pierce test of whether the residual series is white noise.

B. FREQUENCY DOMAIN

Spectral Density. Program SPEC computes the autocovariances, autocorrelations, spectral density estimates (using a Tukey-Hanning window), confidence intervals for each frequency, and one test of significance that compares the spectral density to the theoretical density of white noise. The user must specify only the maximum lag for the calculations. Larger lags divide the frequency range into a finer grid. This program also can compute the periodogram. The density functions and confidence intervals are plotted as well as printed.

The second package is for two applications of univariate model fitting, *forecasting* and *assessing the effects of an intervention*. The second package is described in the box below.

PACKAGE 2: APPLICATIONS OF UNIVARIATE MODEL FITTING

1. *Forecasting.* Program FORCST computes the best one-step ahead prediction using the autoregressive model fitting process, and the one-step ahead prediction variance.

2. *Intervention Effects.* Program ITSE, which stands for Interrupted Time Series Analysis performs an analysis for change in level and slope parameters following an intervention. It uses the linear autoregressive model procedure described in Chapter 26 of Gottman's book. The user must supply only the order of the autoregressive model. The output includes pre- and post-intervention slope, intercept parameters, model parameters, significance tests of the parameters, an overall F-test, and t-tests for change in slope and intercept.

The third package is for *multivariate time-series analysis*. This package is described in the box below.

PACKAGE 3: MULTIVARIATE TIME-SERIES ANALYSIS

A. TIME DOMAIN

1. *Bivariate Time Domain.* Program BIVAR performs the Gottman-Ringland procedure described in Chapter 25 of Gottman's book. This analysis examines lead-lag relationships between two time-series, controlling for autocorrelation in each series. When it is used twice, once with each series as input, it can assess asymmetrical or bidirectional influence between the series. The user must specify the maximum lag values for auto and cross relationship between the two time-series. Chapter 25 describes how to obtain guesses for these values using spectral and cross-spectral analyses. The output is extensive and will not be described here. A summary table provides a chi-square test of the adequacy of the stepwise procedure the program uses as well as a test of association between the two series, with one series as

input to the second.

2. *Time-Series Regression.* Program TSREG performs a two-stage generalized least-squares regression. This predicts one criterion time-series $\mathbf{Y}_t$ from a set of other predictor time-series $\mathbf{X}_{1t}$, $\mathbf{X}_{2t}$, $\cdots$, $\mathbf{X}_{kt}$. The predictor series can be lagged, prewhitened, and so on using program LINFIL.

B. FREQUENCY DOMAIN

Cross-Spectral Analysis. Program CRSPEC performs a bivariate frequency domain time-series analysis. The user supplies only the maximum lag for the analysis. The program outputs the spectral density function for each series, the autocovariances, the cross-covariances, the coherence spectrum, with a significance test for zero coherence, and the phase spectrum, with confidence intervals. The spectra are all also plotted. The user can specify a shift parameter for the analysis (see Chapter 23, Gottman's book).

What the programs do

The chart below is an overview of the programs from a slightly different point of view.

IF YOU WANT TO	YOU WILL NEED
1. Assess the effects of an intervention, i.e., do an interrupted time-series analysis.	1. Program ITSE. It would be useful to employ the univariate model-fitting programs for examining and removing trend and seasonal components [programs DETRND, LINFIL, DESINE, ARFIT, and SPEC].
2. Examine one time-series for component cyclicities.	2. Program SPEC (also can be used to compute the periodogram).
3. Forecast data (limited to one step ahead).	3. Program FORCST.
4. Examine two series for cyclical covariation, i.e., synchronicity and cyclic lead-lag relationships.	4. Program CRSPEC.
5. Perform bivariate time-series analysis, controlling for autocorrelation in each series (Gottman-Ringland procedure).	5. Program BIVAR.
6. Do regression from a set of predictor time-series to a criterion time-series.	6. Program TSREG.

Package 1.
Univariate model fitting

Cambridge University Press
978-0-521-28059-4 - A User's Guide to the Gottman-Williams Time-Series Analysis Computer Programs for Social Scientists
Esther A. Williams and John M. Gottman
Excerpt
More information

DETRND

Program description

The program DETRND is a least-squares linear regression program. In this program, the least-squares straight line is determined for a given set of observations; hypothesis tests are performed on the slope and intercept of the regression line; and the vector of errors of the estimate is computed.

Least-squares linear regression is described in more detail in Chapter 14 of Gottman's *Time-Series Analysis.*

Program steps

In the program, the mean of the N observations (Y_i) is calculated as:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \hat{a} \ .$$

The variance is determined by:

$$s^2 = \frac{1}{N-2}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \ .$$

The value $\sum x_i^2$, which is the sum of the squared deviations from the mean of the fixed points in time of the observations is calculated as:

$$\sum x_i^2 = \frac{N(N+1)(N-1)}{12} \ .$$

This value is used in the calculation of the slope of the line:

$$\hat{b} = \frac{\sum_{i=1}^{N} iY_i - \frac{N(N+1)}{2}\bar{Y}}{\sum x_i^2} \ .$$

The equation of the least-squares regression line is:

$$\hat{Y}_i = \hat{a} + \hat{b}\left(i - \frac{N+1}{2}\right) .$$

Student's t for the slope is determined for hypothesis testing as:

$$t(\hat{b}) = \frac{\hat{b}}{\sqrt{s^2/\Sigma x_i^2}} .$$

Degrees of freedom for testing are $N-2$. Similarly, Student's t for $\hat{a}$ is:

$$t(\hat{a}) = \frac{\hat{a}}{\sqrt{s^2/N}} .$$

Again, the degrees of freedom for testing are $N-2$.

When all of the estimates $(\hat{Y}_i)$ have been calculated with the equation of the least-squares line, the vector of the residuals (errors of estimate) (E_i) is calculated by:

$$e_i = y_i - \hat{y}_i .$$

Program input

1. First input card.

Cols. 2-5	Number of data points (N). Maximum is 1500.
Col. 20	Number of format cards to follow. Maximum is 5.
Col. 25	Set to 1 if vector of residuals (errors of estimates) is to be punched.

2. Second input card.

Cols. 1-80	The title to be printed for identification of output.

3. Third input card or cards.

Cols. 1-80	Data format is specified on this card or cards. Standard FORTRAN F-format is required, enclosed in parentheses.

4. Data cards.

Cols. 1-80	Data cards follow the last format card. The data are punched in the format specified on the format card or cards.

Program output

When the program is run, the following information is output:

1. Program name: DETRND PROGRAM.
2. The title for the printout identification.
3. Number of data points specified (N).

4. Number of format cards specified.
5. Data format is repeated back to user.
6. The first and last data points are printed for user to check format.
7. Intercept of fitted line ($\hat{a}$).
8. Slope of fitted line ($\hat{b}$).
9. Equation of fitted least-squares regression line.
10. Student's t for slope with degrees of freedom.
11. Student's t for intercept with degrees of freedom.
12. The error of estimate vector (E_i) with an error value for each data point i. The vector of errors is punched one error to a card if requested on first input card.
13. End-of-program message.

Special notes

1. If N is desired to be greater than 1500, the dimensions of the Y array and the ERROR array will have to be modified in the program as well as the statements near the beginning of the program which check for the maximum value of N.
2. In the program, output for the punch is written to unit number 7. Other input and output will be on the default system devices specified for input and output (usually the card reader and printer). A CDC computer will write unit 7 to a file called PUNCH which may have to be punched as a separate step. An IBM computer will punch directly if the card punch device is defined as unit 7.
3. None of the computation in the program is machine specific, so going from one computer to another should not create any special problems. The language used is FORTRAN IV. The user may need to remove the first card of the program (the PROGRAM card) and change double quotation marks to single quotation marks in all FORMAT statements.
4. If the user wishes to have input read from other than the default system device, the device can be redefined for the execution step and data specification and format cards must precede the data on that device. This works nicely at a time-sharing computer terminal. Or if the user wishes to read only data from another device, the program statement or statements which read data will have to be modified. In all programs, all data is read in only once just after the comment card "READ IN DATA."

Cambridge University Press
978-0-521-28059-4 - A User's Guide to the Gottman-Williams Time-Series Analysis Computer Programs for Social Scientists
Esther A. Williams and John M. Gottman
Excerpt
More information

Example

Sample input

```
   60                    1      1
SAMPLE RUN OF LINEAR REGRESSION PROGRAM
(F4.2)
1.61
1.00
1.38
2.05
2.21
2.00
1.35
2.17
2.25
2.71
3.39
2.50
3.56
4.00
3.76
3.29
3.28
2.31
2.50
2.33
2.27
2.35
3.33
1.33
2.19
1.44
1.24
1.44
4.40
4.50
3.56
3.00
2.29
1.83
2.24
1.39
2.82
3.29
3.33
3.33
2.71
2.29
1.38
2.83
1.00
1.33
3.28
3.12
3.59
3.24
```