SAS/ETS® 14.1 User’s Guide
The SPECTRA Procedure
## Overview: SPECTRA Procedure

The SPECTRA procedure performs spectral and cross-spectral analysis of time series. You can use spectral analysis techniques to look for periodicities or cyclical patterns in data.

The SPECTRA procedure produces estimates of the spectral and cross-spectral densities of a multivariate time series. Estimates of the spectral and cross-spectral densities of a multivariate time series are produced using a finite Fourier transform to obtain periodograms and cross-periodograms. The periodogram ordinates are smoothed by a moving average to produce estimated spectral and cross-spectral densities. PROC SPECTRA can also test whether or not the data are white noise.

### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview: SPECTRA Procedure</td>
<td>2329</td>
</tr>
<tr>
<td>Getting Started: SPECTRA Procedure</td>
<td>2331</td>
</tr>
<tr>
<td>Syntax: SPECTRA Procedure</td>
<td>2332</td>
</tr>
<tr>
<td>Functional Summary</td>
<td>2332</td>
</tr>
<tr>
<td>PROC SPECTRA Statement</td>
<td>2333</td>
</tr>
<tr>
<td>BY Statement</td>
<td>2334</td>
</tr>
<tr>
<td>VAR Statement</td>
<td>2334</td>
</tr>
<tr>
<td>WEIGHTS Statement</td>
<td>2335</td>
</tr>
<tr>
<td>Details: SPECTRA Procedure</td>
<td>2336</td>
</tr>
<tr>
<td>Input Data</td>
<td>2336</td>
</tr>
<tr>
<td>Missing Values</td>
<td>2336</td>
</tr>
<tr>
<td>Computational Method</td>
<td>2336</td>
</tr>
<tr>
<td>Kernels</td>
<td>2336</td>
</tr>
<tr>
<td>White Noise Test</td>
<td>2339</td>
</tr>
<tr>
<td>Transforming Frequencies</td>
<td>2339</td>
</tr>
<tr>
<td>OUT= Data Set</td>
<td>2339</td>
</tr>
<tr>
<td>Printed Output</td>
<td>2341</td>
</tr>
<tr>
<td>ODS Table Names: SPECTRA procedure</td>
<td>2342</td>
</tr>
<tr>
<td>Examples: SPECTRA Procedure</td>
<td>2342</td>
</tr>
<tr>
<td>Example 33.1: Spectral Analysis of Sunspot Activity</td>
<td>2342</td>
</tr>
<tr>
<td>Example 33.2: Cross-Spectral Analysis</td>
<td>2349</td>
</tr>
<tr>
<td>References</td>
<td>2352</td>
</tr>
</tbody>
</table>
PROC SPECTRA uses the finite Fourier transform to decompose data series into a sum of sine and cosine waves of different amplitudes and wavelengths. The finite Fourier transform decomposition of the series $x_t$ is

$$x_t = \frac{a_0}{2} + \sum_{k=1}^{m-1} f_k (a_k \cos \omega_k t + b_k \sin \omega_k t)$$

$$f_k = \begin{cases} 
1/2 & \text{if } n \text{ is even and } k = m - 1 \\
1 & \text{otherwise}
\end{cases}$$

where

- $t$ is the time subscript, $t = 0, 1, 2, \ldots, n - 1$
- $x_t$ are the equally spaced time series data
- $n$ is the number of observations in the time series
- $m$ is the number of frequencies in the Fourier decomposition: $m = \frac{n+2}{2}$ if $n$ is even, $m = \frac{n+1}{2}$ if $n$ is odd
- $k$ is the frequency subscript, $k = 0, 1, 2, \ldots, m - 1$
- $a_0$ is the mean term: $a_0 = \frac{2}{n} \sum x_t$
- $a_k$ are the cosine coefficients
- $b_k$ are the sine coefficients
- $\omega_k$ are the Fourier frequencies: $\omega_k = \frac{2\pi k}{n}$

Functions of the Fourier coefficients $a_k$ and $b_k$ can be plotted against frequency or against wavelength to form *periodograms*. The amplitude periodogram $J_k$ is defined as follows:

$$J_k = \frac{n}{2} (a_k^2 + b_k^2)$$

Several definitions of the term periodogram are used in the spectral analysis literature. The following discussion refers to the $J_k$ sequence as the periodogram.

The periodogram can be interpreted as the contribution of the $k$th harmonic $\omega_k$ to the total sum of squares (in an analysis of variance sense) in the decomposition of the process into two-degree-of-freedom components for each of the $m$ frequencies. When $n$ is even, $\sin(\omega_n)$ is zero, and thus the last periodogram value is a one-degree-of-freedom component.

The periodogram is a volatile and inconsistent estimator of the spectrum. The spectral density estimate is produced by smoothing the periodogram. Smoothing reduces the variance of the estimator but introduces a bias. The weight function used for the smoothing process, $W()$, often called the kernel or spectral window, is specified with the WEIGHTS statement. It is related to another weight function, $w()$, the lag window, that is used in other methods to taper the correlogram rather than to smooth the periodogram. Many specific weighting functions have been suggested in the literature (Fuller 1976; Jenkins and Watts 1968; Priestley 1981). Table 33.3 later in this chapter gives the relevant formulas when the WEIGHTS statement is used.

Letting $i$ represent the imaginary unit $\sqrt{-1}$, the cross-periodogram is defined as follows:

$$J_k^{xy} = \frac{n}{2} (a_k^x a_k^y + b_k^x b_k^y) + i \frac{n}{2} (a_k^x b_k^y - b_k^x a_k^y)$$
The cross-spectral density estimate is produced by smoothing the cross-periodogram in the same way as the periodograms are smoothed using the spectral window specified by the WEIGHTS statement.

The SPECTRA procedure creates an output SAS data set whose variables contain values of the periodograms, cross-periodograms, estimates of spectral densities, and estimates of cross-spectral densities. The form of the output data set is described in the section “OUT= Data Set” on page 2339.

---

### Getting Started: SPECTRA Procedure

To use the SPECTRA procedure, specify the input and output data sets and options for the analysis you want in the PROC SPECTRA statement, and list the variables to analyze in the VAR statement. The procedure produces no printed output unless the WHITETEST option is specified in the PROC SPECTRA statement. The periodogram, spectral density, and other results are written to the OUT= data set, depending on the options used.

For example, to compute the Fourier transform of a variable X in a data set A, use the following statements:

```plaintext
proc spectra data=a out=b coef;
  var x;
run;
```

This PROC SPECTRA step writes the Fourier coefficients $a_k$ and $b_k$ to the variables COS_01 and SIN_01 in the output data set B.

When a WEIGHTS statement is specified, the periodogram is smoothed by a weighted moving average to produce an estimate of the spectral density of the series. The following statements write a spectral density estimate for X to the variable S_01 in the output data set B.

```plaintext
proc spectra data=a out=b s;
  var x;
  weights 1 2 3 4 3 2 1;
run;
```

When the VAR statement specifies more than one variable, you can perform cross-spectral analysis by specifying the CROSS option in the PROC SPECTRA statement. The CROSS option by itself produces the cross-periodograms for all two-way combinations of the variables listed in the VAR statement. For example, the following statements write the real and imaginary parts of the cross-periodogram of X and Y to the variables RP_01_02 and IP_01_02 in the output data set B.

```plaintext
proc spectra data=a out=b cross;
  var x y;
run;
```

To produce cross-spectral density estimates, specify both the CROSS option and the S option. The cross-periodogram is smoothed using the weights specified by the WEIGHTS statement in the same way as the spectral density. The squared coherency and phase estimates of the cross-spectrum are computed when the K and PH options are used.
The following example computes cross-spectral density estimates for the variables X and Y.

```
proc spectra data=a out=b cross s;
  var x y;
  weights 1 2 3 4 3 2 1;
run;
```

The real part and imaginary part of the cross-spectral density estimates are written to the variables CS_01_02 and QS_01_02, respectively.

**Syntax: SPECTRA Procedure**

The following statements are used with the SPECTRA procedure:

```
PROC SPECTRA options ;
  BY variables ;
  VAR variables ;
  WEIGHTS < weights > < kernel > ;
```

**Functional Summary**

Table 33.1 summarizes the statements and options that control the SPECTRA procedure.

<table>
<thead>
<tr>
<th>Description</th>
<th>Statement</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statements</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>specify BY-group processing</td>
<td>BY</td>
<td></td>
</tr>
<tr>
<td>specify the variables to be analyzed</td>
<td>VAR</td>
<td></td>
</tr>
<tr>
<td>specify weights for spectral density estimates</td>
<td>WEIGHTS</td>
<td></td>
</tr>
<tr>
<td><strong>Data Set Options</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>specify the input data set</td>
<td>PROC SPECTRA</td>
<td>DATA=</td>
</tr>
<tr>
<td>specify the output data set</td>
<td>PROC SPECTRA</td>
<td>OUT=</td>
</tr>
<tr>
<td><strong>Output Control Options</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>output the amplitudes of the cross-spectrum</td>
<td>PROC SPECTRA</td>
<td>A</td>
</tr>
<tr>
<td>output the Fourier coefficients</td>
<td>PROC SPECTRA</td>
<td>COEF</td>
</tr>
<tr>
<td>output the periodogram</td>
<td>PROC SPECTRA</td>
<td>P</td>
</tr>
<tr>
<td>output the spectral density estimates</td>
<td>PROC SPECTRA</td>
<td>S</td>
</tr>
<tr>
<td>output cross-spectral analysis results</td>
<td>PROC SPECTRA</td>
<td>CROSS</td>
</tr>
<tr>
<td>output squared coherency of the cross-spectrum</td>
<td>PROC SPECTRA</td>
<td>K</td>
</tr>
<tr>
<td>output the phase of the cross-spectrum</td>
<td>PROC SPECTRA</td>
<td>PH</td>
</tr>
</tbody>
</table>
Table 33.1  continued

<table>
<thead>
<tr>
<th>Description</th>
<th>Statement</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothing Options</td>
<td>WEIGHTS</td>
<td>BART</td>
</tr>
<tr>
<td>specify the Bartlett kernel</td>
<td>WEIGHTS</td>
<td>PARZEN</td>
</tr>
<tr>
<td>specify the Parzen kernel</td>
<td>WEIGHTS</td>
<td>QS</td>
</tr>
<tr>
<td>specify the quadratic spectral kernel</td>
<td>WEIGHTS</td>
<td>QS</td>
</tr>
<tr>
<td>specify the Tukey-Hanning kernel</td>
<td>WEIGHTS</td>
<td>TUKEY</td>
</tr>
<tr>
<td>specify the truncated kernel</td>
<td>WEIGHTS</td>
<td>TRUNCAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other Options</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>subtract the series mean</td>
<td>PROC SPECTRA</td>
<td>ADJMEAN</td>
</tr>
<tr>
<td>specify an alternate quadrature spectrum estimate</td>
<td>PROC SPECTRA</td>
<td>ALTW</td>
</tr>
<tr>
<td>request tests for white noise</td>
<td>PROC SPECTRA</td>
<td>WHITETEST</td>
</tr>
</tbody>
</table>

PROC SPECTRA Statement

PROC SPECTRA options ;

The following options can be used in the PROC SPECTRA statement:

A

outputs the amplitude variables (A_{nn \_mm}) of the cross-spectrum.

ADJMEAN

CENTER

subtracts the series mean before performing the Fourier decomposition. This sets the first periodogram ordinate to 0 rather than 2n times the squared mean. This option is commonly used when the periodograms are to be plotted to prevent a large first periodogram ordinate from distorting the scale of the plot.

ALTW

specifies that the quadrature spectrum estimate is computed at the boundaries in the same way as the spectral density estimate and the cospectrum estimate are computed.

COEF

outputs the Fourier cosine and sine coefficients of each series.

CROSS

is used with the P and S options to output cross-periodograms and cross-spectral densities when more than one variable is listed in the VAR statement.
DATA=SAS-data-set
names the SAS data set that contains the input data. If the DATA= option is omitted, the most recently created SAS data set is used.

K
outputs the squared coherency variables (K_{nn \_mm}) of the cross-spectrum. The K_{nn \_mm} variables are identically 1 unless weights are given in the WEIGHTS statement and the S option is specified.

OUT=SAS-data-set
names the output data set created by PROC SPECTRA to store the results. If the OUT= option is omitted, the output data set is named by using the DATA_n convention.

P
outputs the periodogram variables. The variables are named P_{nn}, where nn is an index of the original variable with which the periodogram variable is associated. When both the P and CROSS options are specified, the cross-periodogram variables RP_{nn \_mm} and IP_{nn \_mm} are also output.

PH
outputs the phase variables (PH_{nn \_mm}) of the cross-spectrum.

S
outputs the spectral density estimates. The variables are named S_{nn}, where nn is an index of the original variable with which the estimate variable is associated. When both the S and CROSS options are specified, the cross-spectral variables CS_{nn \_mm} and QS_{nn \_mm} are also output.

WHITETEST
prints two tests of the hypothesis that the data are white noise. See the section “White Noise Test” on page 2339 for details.

Note that the CROSS, A, K, and PH options are meaningful only if more than one variable is listed in the VAR statement.

BY Statement

BY variables ;

A BY statement can be used with proc spectra to obtain separate analyses for groups of observations defined by the BY variables.

VAR Statement

VAR variables ;

The VAR statement specifies one or more numeric variables that contain the time series to analyze. The order of the variables in the VAR statement list determines the index, nn, used to name the output variables. The VAR statement is required.
WEIGHTS Statement

The WEIGHTS statement specifies the relative weights used in the moving average applied to the periodogram ordinates to form the spectral density estimates. A WEIGHTS statement must be used to produce smoothed spectral density estimates. You can specify the relative weights in two ways: you can specify them explicitly as explained in the section “Using Weight Constants Specification” on page 2335, or you can specify them implicitly by using the kernel specification as explained in the section “Using Kernel Specifications” on page 2335. If the WEIGHTS statement is not used, only the periodogram is produced.

Using Weight Constants Specification

Any number of weighting constants can be specified. The constants should be positive and symmetric about the middle weight. The middle constant (or the constant to the right of the middle if an even number of weight constants are specified) is the relative weight of the current periodogram ordinate. The constant immediately following the middle one is the relative weight of the next periodogram ordinate, and so on. The actual weights used in the smoothing process are the weights specified in the WEIGHTS statement scaled so that they sum to $\frac{1}{4\pi}$.

The moving average reflects at each end of the periodogram. The first periodogram ordinate is not used; the second periodogram ordinate is used in its place.

For example, a simple triangular weighting can be specified using the following WEIGHTS statement:

```
weights 1 2 3 2 1;
```

Using Kernel Specifications

You can specify five different kernels in the WEIGHTS statement. The syntax for the statement is

```
WEIGHTS [PARZEN][BART][TUKEY][TRUNCAT][QS] [c e];
```

where $c >= 0$ and $e >= 0$ are used to compute the bandwidth parameter as

$$l(q) = cq^e$$

and $q$ is the number of periodogram ordinates +1:

$$q = \text{floor}(n/2) + 1$$

To specify the bandwidth explicitly, set $c =$ to the desired bandwidth and $e = 0$.

For example, a Parzen kernel can be specified using the following WEIGHTS statement:

```
weights parzen 0.5 0;
```

For details, see the section “Kernels” on page 2336.
Details: SPECTRA Procedure

Input Data

Observations in the data set analyzed by the SPECTRA procedure should form ordered, equally spaced time series. No more than 99 variables can be included in the analysis.

Data are often detrended before analysis by the SPECTRA procedure. This can be done by using the residuals output by a SAS regression procedure. Optionally, the data can be centered using the ADJMEAN option in the PROC SPECTRA statement, since the zero periodogram ordinate corresponding to the mean is of little interest from the point of view of spectral analysis.

Missing Values

Missing values are excluded from the analysis by the SPECTRA procedure. If the SPECTRA procedure encounters missing values for any variable listed in the VAR statement, the procedure determines the longest contiguous span of data that has no missing values for the variables listed in the VAR statement and uses that span for the analysis.

Computational Method

If the number of observations $n$ factors into prime integers that are less than or equal to 23, and the product of the square-free factors of $n$ is less than 210, then PROC SPECTRA uses the fast Fourier transform developed by Cooley and Tukey and implemented by Singleton (1969). If $n$ cannot be factored in this way, then PROC SPECTRA uses a Chirp-Z algorithm similar to that proposed by Monro and Branch (1977). To reduce memory requirements, when $n$ is small, the Fourier coefficients are computed directly using the defining formulas.

Kernels

Kernels are used to smooth the periodogram by using a weighted moving average of nearby points. A smoothed periodogram is defined by the following equation.

$$\hat{J}_i(1(q)) = \sum_{\tau=-1(q)}^{1(q)} w\left(\frac{\tau}{1(q)}\right) \hat{J}_{i+r}$$
where \( w(x) \) is the kernel or weight function. At the endpoints, the moving average is computed cyclically; that is,

\[
\tilde{J}_{i+\tau} = \begin{cases} 
J_{i+\tau} & 0 \leq i + \tau \leq q \\
J_{-(i+\tau)} & i + \tau < 0 \\
J_{q-(i+\tau)} & i + \tau > q 
\end{cases}
\]

The SPECTRA procedure supports the following kernels. They are listed with their default bandwidth functions.

**Bartlett:** KERNEL BART

\[
w(x) = \begin{cases} 
1 - |x| & |x| \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
l(q) = \frac{1}{2} q^{1/3}
\]

**Parzen:** KERNEL PARZEN

\[
w(x) = \begin{cases} 
1 - 6|x|^2 + 6|x|^3 & 0 \leq |x| \leq \frac{1}{2} \\
2(1 - |x|)^3 & \frac{1}{2} \leq |x| \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
l(q) = q^{1/5}
\]

**Quadratic spectral:** KERNEL QS

\[
w(x) = \frac{25}{12\pi^2 x^2} \left( \frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5) \right)
\]

\[
l(q) = \frac{1}{2} q^{1/5}
\]

**Tukey-Hanning:** KERNEL TUKEY

\[
w(x) = \begin{cases} 
(1 + \cos(\pi x))/2 & |x| \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
l(q) = \frac{2}{3} q^{1/5}
\]

**Truncated:** KERNEL TRUNCAT

\[
w(x) = \begin{cases} 
1 & |x| \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

\[
l(q) = \frac{1}{4} q^{1/5}
\]
A summary of the default values of the bandwidth parameters, \( c \) and \( e \), associated with the kernel smoothers in PROC SPECTRA are listed below in Table 33.2:

<table>
<thead>
<tr>
<th>Kernel</th>
<th>( c )</th>
<th>( e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bartlett</td>
<td>( 1/2 )</td>
<td>( 1/3 )</td>
</tr>
<tr>
<td>Parzen</td>
<td>( 1 )</td>
<td>( 1/5 )</td>
</tr>
<tr>
<td>quadratic</td>
<td>( 1/2 )</td>
<td>( 1/5 )</td>
</tr>
<tr>
<td>Tukey-Hanning</td>
<td>( 2/3 )</td>
<td>( 1/5 )</td>
</tr>
<tr>
<td>truncated</td>
<td>( 1/4 )</td>
<td>( 1/5 )</td>
</tr>
</tbody>
</table>

See Andrews (1991) for details about the properties of these kernels.
**White Noise Test**

PROC SPECTRA prints two test statistics for white noise when the WHITETEST option is specified: Fisher’s Kappa (Davis 1941; Fuller 1976) and Bartlett’s Kolmogorov-Smirnov statistic (Bartlett 1966; Fuller 1976; Durbin 1967).

If the time series is a sequence of independent random variables with mean 0 and variance $\sigma^2$, then the periodogram, $J_k$, will have the same expected value for all $k$. For a time series with nonzero autocorrelation, each ordinate of the periodogram, $J_k$, will have different expected values. The Fisher’s Kappa statistic tests whether the largest $J_k$ can be considered different from the mean of the $J_k$. Critical values for the Fisher’s Kappa test can be found in Fuller 1976.

The Kolmogorov-Smirnov statistic reported by PROC SPECTRA has the same asymptotic distribution as Bartlett’s test (Durbin 1967). The Kolmogorov-Smirnov statistic compares the normalized cumulative periodogram with the cumulative distribution function of a uniform(0,1) random variable. The normalized cumulative periodogram, $F_j$, of the series is

$$F_j = \frac{\sum_{k=1}^{j} J_k}{\sum_{k=1}^{m} J_k}, \quad j = 1, 2 \ldots, m - 1$$

where $m = \frac{n}{2}$ if $n$ is even or $m = \frac{n+1}{2}$ if $n$ is odd. The test statistic is the maximum absolute difference of the normalized cumulative periodogram and the uniform cumulative distribution function. Approximate $p$-values for Bartlett’s Kolmogorov-Smirnov test statistics are provided with the test statistics. Small $p$-values cause you to reject the null-hypothesis that the series is white noise.

**Transforming Frequencies**

The variable FREQ in the data set created by the SPECTRA procedure ranges from 0 to $\pi$. Sometimes it is preferable to express frequencies in cycles per observation period, which is equal to $\frac{1}{2\pi}$ FREQ.

To express frequencies in cycles per unit time (for example, in cycles per year), multiply FREQ by $\frac{d}{\pi}$, where $d$ is the number of observations per unit of time. For example, for monthly data, if the desired time unit is years then $d$ is 12. The period of the cycle is $\frac{2\pi}{d \times \text{FREQ}}$, which ranges from $\frac{2}{d}$ to infinity.

**OUT= Data Set**

The OUT= data set contains $\frac{n}{2} + 1$ observations, if $n$ is even, or $\frac{n+1}{2}$ observations, if $n$ is odd, where $n$ is the number of observations in the time series or the span of data being analyzed if missing values are present in the data. See the section “Missing Values” on page 2336 for details.

The variables in the new data set are named according to the following conventions. Each variable to be analyzed is associated with an index. The first variable listed in the VAR statement is indexed as 01, the second variable as 02, and so on. Output variables are named by combining indexes with prefixes. The prefix always identifies the nature of the new variable, and the indices identify the original variables from which the statistics were obtained.

Variables that contain spectral analysis results have names that consist of a prefix, an underscore, and the index of the variable analyzed. For example, the variable S_01 contains spectral density estimates for the
first variable in the VAR statement. Variables that contain cross-spectral analysis results have names that consist of a prefix, an underscore, the index of the first variable, another underscore, and the index of the second variable. For example, the variable \texttt{A\_01\_02} contains the amplitude of the cross-spectral density estimate for the first and second variables in the VAR statement.

Table 33.3 shows the formulas and naming conventions used for the variables in the OUT= data set. Let \( X \) be variable number \( nn \) in the VAR statement list and let \( Y \) be variable number \( mm \) in the VAR statement list. Table 33.3 shows the output variables that contain the results of the spectral and cross-spectral analysis of \( X \) and \( Y \).

In Table 33.3 the following notation is used. Let \( W_j \) be the vector of \( 2p + 1 \) smoothing weights given by the WEIGHTS statement, normalized to sum to \( \frac{1}{4\pi} \). Note that the weights are either explicitly provided using the constant specification or are implicitly determined by the kernel specification in the WEIGHTS statement.

The subscript of \( W_j \) runs from \( W_{-p} \) to \( W_p \), so that \( W_0 \) is the middle weight in the list. Let \( \omega_k = \frac{2\pi k}{n} \), where \( k = 0, 1, \ldots, \text{floor}(\frac{n}{2}) \).

### Table 33.3 Variables Created by PROC SPECTRA

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FREQ</td>
<td>frequency in radians from 0 to ( \pi ) (Note: Cycles per observation is ( \frac{\text{FREQ}}{2\pi} ).)</td>
</tr>
<tr>
<td>PERIOD</td>
<td>period or wavelength: ( \frac{2\pi}{\text{FREQ}} ) (Note: PERIOD is missing for ( \text{FREQ}=0 ).)</td>
</tr>
<tr>
<td>COS_nn</td>
<td>cosine transform of ( X ): ( a_k^{X} = \frac{2}{n} \sum_{t=1}^{n} X_t \cos(\omega_k (t - 1)) )</td>
</tr>
<tr>
<td>SIN_nn</td>
<td>sine transform of ( X ): ( b_k^{X} = \frac{2}{n} \sum_{t=1}^{n} X_t \sin(\omega_k (t - 1)) )</td>
</tr>
<tr>
<td>P_nn</td>
<td>periodogram of ( X ): ( J_k^{X} = \frac{n}{2} \left[ (a_k^{X})^2 + (b_k^{X})^2 \right] )</td>
</tr>
<tr>
<td>S_nn</td>
<td>spectral density estimate of ( X ): ( F_k^{X} = \sum_{j=-p}^{p} W_j J_{k+j}^{X} ) (except across endpoints)</td>
</tr>
<tr>
<td>RP_nn_mm</td>
<td>real part of cross-periodogram ( X ) and ( Y ): ( \text{real}(J_k^{XY}) = \frac{n}{2} (a_k^{X}a_k^{Y} + b_k^{X}b_k^{Y}) )</td>
</tr>
<tr>
<td>IP_nn_mm</td>
<td>imaginary part of cross-periodogram of ( X ) and ( Y ): ( \text{imag}(J_k^{XY}) = \frac{n}{2} (a_k^{X}b_k^{Y} - b_k^{X}a_k^{Y}) )</td>
</tr>
<tr>
<td>CS_nn_mm</td>
<td>cospectrum estimate (real part of cross-spectrum) of ( X ) and ( Y ): ( C_k^{XY} = \sum_{j=-p}^{p} W_j \text{real}(J_{k+j}^{XY}) ) (except across endpoints)</td>
</tr>
<tr>
<td>Variable</td>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>QS nn mm</td>
<td>quadrature spectrum estimate (imaginary part of cross-spectrum) of X and Y: ( Q_{xy}^k = \sum_{j=-p}^{p} W_j \text{imag}(J_{k+j}^{xy}) ) (except across endpoints)</td>
</tr>
<tr>
<td>A nn mm</td>
<td>amplitude (modulus) of cross-spectrum of X and Y: ( A_{xy}^k = \sqrt{(C_{xy}^k)^2 + (Q_{xy}^k)^2} )</td>
</tr>
<tr>
<td>K nn mm</td>
<td>coherency squared of X and Y: ( K_{xy}^k = (A_{xy}^k)^2 / (F_x^k F_y^k) )</td>
</tr>
<tr>
<td>PH nn mm</td>
<td>phase spectrum in radians of X and Y: ( \Phi_{xy}^k = \arctan(Q_{xy}^k / C_{xy}^k) )</td>
</tr>
</tbody>
</table>

**Printed Output**

By default PROC SPECTRA produces no printed output.

When the WHITETEST option is specified, the SPECTRA procedure prints the following statistics for each variable in the VAR statement:

1. the name of the variable
2. M–1, the number of two-degree-of-freedom periodogram ordinates used in the test
3. MAX(P(*)), the maximum periodogram ordinate
4. SUM(P(*)), the sum of the periodogram ordinates
5. Fisher’s Kappa statistic
6. Bartlett’s Kolmogorov-Smirnov test statistic
7. Approximate \( p \)-value for Bartlett’s Kolmogorov-Smirnov test statistic

See the section “White Noise Test” on page 2339 for details.
**ODS Table Names: SPECTRA procedure**

PROC SPECTRA assigns a name to each table it creates. You can use these names to reference the table when you use the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in the following table.

<table>
<thead>
<tr>
<th>ODS Table Name</th>
<th>Description</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>WhiteNoiseTest</td>
<td>white noise test</td>
<td>WHITETEST</td>
</tr>
<tr>
<td>Kappa</td>
<td>Fishers Kappa statistic</td>
<td>WHITETEST</td>
</tr>
<tr>
<td>Bartlett</td>
<td>Bartletts Kolmogorov-Smirnov statistic</td>
<td>WHITETEST</td>
</tr>
</tbody>
</table>

**Examples: SPECTRA Procedure**

**Example 33.1: Spectral Analysis of Sunspot Activity**

This example analyzes Wolfer’s sunspot data (Anderson 1971). The following statements read and plot the data.

```sas
   title "Wolfer's Sunspot Data";
   data sunspot;
   input year wolfer @@;
   datalines;
   1749 809 1750 834 1751 477 1752 478 1753 307 1754 122 1755 96
   ...
   96
   run;
```

```sas
   proc sgplot data=sunspot;
   series x=year y=wolfer / markers markerattrs=(symbol=circlefilled);
   xaxis values=(1740 to 1930 by 10);
   yaxis values=(0 to 1600 by 200);
   run;
```

The plot of the sunspot series is shown in Output 33.1.1.
Example 33.1: Spectral Analysis of Sunspot Activity

Output 33.1.1 Plot of Original Sunspot Data

The spectral analysis of the sunspot series is performed by the following statements:

```sas
proc spectra data=sunspot out=b p s adjmean whitetest;
   var wolfer;
   weights 1 2 3 4 3 2 1;
run;
proc print data=b(obs=12);
run;
```

The PROC SPECTRA statement specifies the P and S options to write the periodogram and spectral density estimates to the OUT= data set B. The WEIGHTS statement specifies a triangular spectral window for smoothing the periodogram to produce the spectral density estimate. The ADJMEAN option zeros the frequency 0 value and avoids the need to exclude that observation from the plots. The WHITETEST option prints tests for white noise.

The Fisher’s Kappa test statistic of 16.070 is larger than the 5% critical value of 7.2, so the null hypothesis that the sunspot series is white noise is rejected (see the table of critical values in Fuller (1976)).

The Bartlett’s Kolmogorov-Smirnov statistic is 0.6501, and its approximate $p$-value is < 0.0001. The small $p$-value associated with this test leads to the rejection of the null hypothesis that the spectrum represents white noise.
The printed output produced by PROC SPECTRA is shown in Output 33.1.2. The output data set B created by PROC SPECTRA is shown in part in Output 33.1.3.

**Output 33.1.2** White Noise Test Results

Wolfer's Sunspot Data

The SPECTRA Procedure

<table>
<thead>
<tr>
<th>Test for White Noise for Variable <code>wolfer</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>M-1</td>
</tr>
<tr>
<td>Max(P(*))</td>
</tr>
<tr>
<td>Sum(P(*))</td>
</tr>
</tbody>
</table>

Fisher's Kappa:

\[(M-1)\times\text{Max}(P(\ast))/\text{Sum}(P(\ast))\]

Kappa | 16.70489 |

Bartlett's Kolmogorov-Smirnov Statistic:
Maximum absolute difference of the standardized partial sums of the periodogram and the CDF of a uniform(0,1) random variable.

Test Statistic | 0.650055 |
Approximate P-Value | <.0001 |

**Output 33.1.3** First 12 Observations of the OUT= Data Set

Wolfer's Sunspot Data

<table>
<thead>
<tr>
<th>Obs</th>
<th>FREQ</th>
<th>PERIOD</th>
<th>P_01</th>
<th>S_01</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00000</td>
<td>.</td>
<td>0.00</td>
<td>59327.52</td>
</tr>
<tr>
<td>2</td>
<td>0.03570</td>
<td>176.000</td>
<td>3178.15</td>
<td>61757.98</td>
</tr>
<tr>
<td>3</td>
<td>0.07140</td>
<td>88.000</td>
<td>2435433.22</td>
<td>69528.68</td>
</tr>
<tr>
<td>4</td>
<td>0.10710</td>
<td>58.667</td>
<td>1077495.76</td>
<td>66087.57</td>
</tr>
<tr>
<td>5</td>
<td>0.14280</td>
<td>44.000</td>
<td>491850.36</td>
<td>53352.02</td>
</tr>
<tr>
<td>6</td>
<td>0.17850</td>
<td>35.200</td>
<td>2581.12</td>
<td>36678.14</td>
</tr>
<tr>
<td>7</td>
<td>0.21420</td>
<td>29.333</td>
<td>181163.15</td>
<td>20604.52</td>
</tr>
<tr>
<td>8</td>
<td>0.24990</td>
<td>25.143</td>
<td>283057.60</td>
<td>15132.81</td>
</tr>
<tr>
<td>9</td>
<td>0.28560</td>
<td>22.000</td>
<td>188672.97</td>
<td>13265.89</td>
</tr>
<tr>
<td>10</td>
<td>0.32130</td>
<td>19.556</td>
<td>122673.94</td>
<td>14953.32</td>
</tr>
<tr>
<td>11</td>
<td>0.35700</td>
<td>17.600</td>
<td>58532.93</td>
<td>16402.84</td>
</tr>
<tr>
<td>12</td>
<td>0.39270</td>
<td>16.000</td>
<td>213405.16</td>
<td>18562.13</td>
</tr>
</tbody>
</table>
The following statements plot the periodogram and spectral density estimate by the frequency and period.

```ods output periodogram_freq=periodogram_freq
data=b;
  series x=freq y=p_01 / markers markerattrs=(symbol=circlefilled);
run;
```

```ods output periodogram_period=periodogram_period
data=b;
  series x=period y=p_01 / markers markerattrs=(symbol=circlefilled);
run;
```

```ods output spectral_density_freq= spectral_density_freq
data=b;
  series x=freq y=s_01 / markers markerattrs=(symbol=circlefilled);
run;
```

```ods output spectral_density_period= spectral_density_period
data=b;
  series x=period y=s_01 / markers markerattrs=(symbol=circlefilled);
run;
```

The periodogram is plotted against the frequency in Output 33.1.4 and plotted against the period in Output 33.1.5. The spectral density estimate is plotted against the frequency in Output 33.1.6 and plotted against the period in Output 33.1.7.

**Output 33.1.4** Plot of Periodogram by Frequency

![Wolfer's Sunspot Data](image-url)
Output 33.1.5  Plot of Periodogram by Period

Wolfer's Sunspot Data

Periodogram of wolfer

Period
Example 33.1: Spectral Analysis of Sunspot Activity

Output 33.1.6 Plot of Spectral Density Estimate by Frequency

Wolfer's Sunspot Data

[Spectral Density of sunspots against frequency]

- Spectral Density of sunspots on the y-axis.
- Frequency from 0 to π on the x-axis.

Graph showing the spectral density estimate by frequency for Wolfer's sunspot data.
Since PERIOD is the reciprocal of frequency, the plot axis for PERIOD is stretched for low frequencies and compressed at high frequencies. One way to correct for this is to use a WHERE statement to restrict the plots and exclude the low frequency components. The following statements plot the spectral density for periods less than 50.

```
proc sgplot data=b;
  where period < 50;
  series x=period y=s_01 / markers markerattrs=(symbol=circlefilled);
  reline 11 / axis=x;
run;
title;
```

The spectral analysis of the sunspot series confirms a strong 11-year cycle of sunspot activity. The plot makes this clear by drawing a reference line at the 11 year period, which highlights the position of the main peak in the spectral density.

Output 33.1.8 shows the plot. Contrast Output 33.1.8 with Output 33.1.7.
Example 33.2: Cross-Spectral Analysis

This example uses simulated data to show cross-spectral analysis for two variables X and Y. X is generated by an AR(1) process; Y is generated as white noise plus an input from X lagged 2 periods. All output options are specified in the PROC SPECTRA statement. PROC CONTENTS shows the contents of the OUT= data set.

```sas
data a;
  x1 = 0; xll = 0;
  do i = -10 to 100;
    x = .4 * x1 + rannor(123);
    y = .5 * xll + rannor(123);
    if i > 0 then output;
    xll = x1; x1 = x;
  end;
run;

proc spectra data=a out=b cross coef a k p ph s;
  var x y;
  weights 1 1.5 2 4 8 9 8 4 2 1.5 1;
run;

proc contents data=b position;
run;
```
The PROC CONTENTS report for the output data set B is shown in Output 33.2.1.

**Output 33.2.1** Contents of PROC SPECTRA OUT= Data Set

The CONTENTS Procedure

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Type</th>
<th>Len</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>A_01_02</td>
<td>Num</td>
<td>8</td>
<td>Amplitude of x by y</td>
</tr>
<tr>
<td>3</td>
<td>COS_01</td>
<td>Num</td>
<td>8</td>
<td>Cosine Transform of x</td>
</tr>
<tr>
<td>5</td>
<td>COS_02</td>
<td>Num</td>
<td>8</td>
<td>Cosine Transform of y</td>
</tr>
<tr>
<td>13</td>
<td>CS_01_02</td>
<td>Num</td>
<td>8</td>
<td>Cospectra of x by y</td>
</tr>
<tr>
<td>1</td>
<td>FREQ</td>
<td>Num</td>
<td>8</td>
<td>Frequency from 0 to PI</td>
</tr>
<tr>
<td>12</td>
<td>IP_01_02</td>
<td>Num</td>
<td>8</td>
<td>Imag Periodogram of x by y</td>
</tr>
<tr>
<td>15</td>
<td>K_01_02</td>
<td>Num</td>
<td>8</td>
<td>Coherency**2 of x by y</td>
</tr>
<tr>
<td>2</td>
<td>PERIOD</td>
<td>Num</td>
<td>8</td>
<td>Period</td>
</tr>
<tr>
<td>17</td>
<td>PH_01_02</td>
<td>Num</td>
<td>8</td>
<td>Phase of x by y</td>
</tr>
<tr>
<td>7</td>
<td>P_01</td>
<td>Num</td>
<td>8</td>
<td>Periodogram of x</td>
</tr>
<tr>
<td>8</td>
<td>P_02</td>
<td>Num</td>
<td>8</td>
<td>Periodogram of y</td>
</tr>
<tr>
<td>14</td>
<td>QS_01_02</td>
<td>Num</td>
<td>8</td>
<td>Quadrature of x by y</td>
</tr>
<tr>
<td>11</td>
<td>RP_01_02</td>
<td>Num</td>
<td>8</td>
<td>Real Periodogram of x by y</td>
</tr>
<tr>
<td>4</td>
<td>SIN_01</td>
<td>Num</td>
<td>8</td>
<td>Sine Transform of x</td>
</tr>
<tr>
<td>6</td>
<td>SIN_02</td>
<td>Num</td>
<td>8</td>
<td>Sine Transform of y</td>
</tr>
<tr>
<td>9</td>
<td>S_01</td>
<td>Num</td>
<td>8</td>
<td>Spectral Density of x</td>
</tr>
<tr>
<td>10</td>
<td>S_02</td>
<td>Num</td>
<td>8</td>
<td>Spectral Density of y</td>
</tr>
</tbody>
</table>

The following statements plot the amplitude of the cross-spectrum estimate against frequency and against period for periods less than 25.

```
proc sgplot data=b;
  series x=freq y=a_01_02 / markers markerattrs=(symbol=circlefilled);
  xaxis values=(0 to 4 by 1);
run;
```

The plot of the amplitude of the cross-spectrum estimate against frequency is shown in Output 33.2.2.
The plot of the cross-spectrum amplitude against period for periods less than 25 observations is shown in Output 33.2.3.

```
proc sgplot data=b;
  where period < 25;
  series x=period y=a_01_02 / markers markerattrs=(symbol=circlefilled);
  xaxis values=(0 to 30 by 5);
run;
```
Output 33.2.3  Plot of Cross-Spectrum Amplitude by Period

References


Subject Index

BY groups
   SPECTRA procedure, 2334

Chirp-Z algorithm
   SPECTRA procedure, 2336
coherency
   cross-spectral analysis, 2341
   SPECTRA procedure, 2341
cospectrum estimate
   cross-spectral analysis, 2340
   SPECTRA procedure, 2340
cross-periodogram
   cross-spectral analysis, 2330, 2340
   SPECTRA procedure, 2340
cross-spectral analysis
   coherency, 2341
   cospectrum estimate, 2340
   cross-periodogram, 2330, 2340
cross-spectrum, 2341
   quadrature spectrum, 2341
   SPECTRA procedure, 2329, 2330, 2340, 2341
cross-spectrum
   cross-spectral analysis, 2341
   SPECTRA procedure, 2341
fast Fourier transform
   SPECTRA procedure, 2336
finite Fourier transform
   SPECTRA procedure, 2330
Fourier coefficients
   SPECTRA procedure, 2340
Fourier transform
   SPECTRA procedure, 2330
frequency
   SPECTRA procedure, 2339
kernels, 2336
   SPECTRA procedure, 2336
output data sets
   SPECTRA procedure, 2339
output table names
   SPECTRA procedure, 2342
periodogram
   SPECTRA procedure, 2330, 2340
quadrature spectrum

SPECTRA procedure
   BY groups, 2334
   Chirp-Z algorithm, 2336
   coherency of cross-spectrum, 2341
cospectrum estimate, 2340
cross-periodogram, 2340
cross-spectral analysis, 2329, 2330, 2340, 2341
cross-spectrum, 2341
   fast Fourier transform, 2336
   finite Fourier transform, 2330
   Fourier coefficients, 2340
   Fourier transform, 2330
   frequency, 2339
   kernels, 2336
   output data sets, 2339
   output table names, 2342
   periodogram, 2330, 2340
   quadrature spectrum, 2341
   spectral analysis, 2329
   spectral density estimate, 2329, 2340
   spectral window, 2335
   white noise test, 2339, 2341
spectral analysis
   SPECTRA procedure, 2329
spectral density estimate
   SPECTRA procedure, 2329, 2340
spectral window
   SPECTRA procedure, 2335
white noise test
   SPECTRA procedure, 2339, 2341
Syntax Index

A option
   PROC SPECTRA statement, 2333

ADJMEAN option
   PROC SPECTRA statement, 2333

ALTW option
   PROC SPECTRA statement, 2333

BY statement
   SPECTRA procedure, 2334

CENTER option
   PROC SPECTRA statement, 2333

COEF option
   PROC SPECTRA statement, 2333

CROSS option
   PROC SPECTRA statement, 2333

DATA= option
   PROC SPECTRA statement, 2334

K option
   PROC SPECTRA statement, 2334

OUT= option
   PROC SPECTRA statement, 2334, 2339

P option
   PROC SPECTRA statement, 2334

PH option
   PROC SPECTRA statement, 2334

PROC SPECTRA statement, 2333

S option
   PROC SPECTRA statement, 2334

SPECTRA procedure, 2332
   syntax, 2332

VAR statement
   SPECTRA procedure, 2334

WEIGHTS statement
   SPECTRA procedure, 2335

WHITETEST option
   PROC SPECTRA statement, 2334