

## LOWESS SMOOTH

### PURPOSE

Carries out (robust) locally-weighted time series and scatter plot smoothing for both equispaced and non-equispaced data. LOWESS stands for “locally weighted least squares.”

### DESCRIPTION

LOWESS is a data analysis technique for producing a “smooth” set of values from a time series which has been contaminated with noise, or from a scatter plot with a “noisy” relationship between the 2 variables. In a time series context, the technique is an improvement over least squares smoothing when the data is not equally spaced (as least squares smoothing assumes). For LOWESS smoothing, the analyst can vary the size of the smoothing window. This size is given as the fraction (0 to 1) of the data that the window should cover. The default window size is 0.1 (which states that the smoothing window has a total width of 10% of the horizontal axis variable). To specify the width of the LOWESS smooth, use the LOWESS FRACTION command prior to using the LOWESS SMOOTH command. This width must be a decimal value between 0 and 1. By default, the LOWESS SMOOTH command performs a weighted linear least squares fit of the points in the current data window. The LOWESS DEGREE command can be used to specify a weighted quadratic least squares fit.

The LOWESS SMOOTH command saves the predicted values in the variable PRED and the residual values in the variable RES. These variables can be plotted and manipulated just like any user created variable. The residual standard deviation, the residual degrees of freedom, the replication standard deviation, and the replication degrees of freedom are stored in the internal parameters RESSD, RESDF, REPSD, and REPDF, respectively. These parameters are described in the documentation for the FIT command.

### SYNTAX 1

```
LOWESS SMOOTH <y> <SUBSET/EXCEPT/FOR qualification>
```

where <y> is the vertical axis variable under analysis;  
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax can be used for equispaced data.

### SYNTAX 2

```
LOWESS SMOOTH <y> <x> <SUBSET/EXCEPT/FOR qualification>
```

where <y> is the vertical axis variable under analysis;  
<x> is the horizontal axis variable;  
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is required for non-equispaced data.

### EXAMPLES

```
LOWESS SMOOTH Y X
LOWESS SMOOTH Y
LOWESS SMOOTH CONC DAY
LOWESS SMOOTH CONC
```

```
LOWESS FRACTION .3
LOWESS SMOOTH Y X
```

### NOTE 1

The LOWESS fraction controls the smoothness of the curve. For example, if it is 1.0, then the LOWESS curve is a single straight line. In general, the smaller the fraction, the more that LOWESS curve follows individual data points. To obtain a smoother LOWESS curve, increase the value of the LOWESS FRACTION.

### NOTE 2

The LOWESS DEGREE command is used to specify whether local linear fits or locally quadratic fits are applied to the points in the current data window. Quadratic fits are normally recommended if there are local minimum and maximums in the data. Otherwise, linear fitting should be adequate. Linear fitting is the default.

### NOTE 3

The following is a brief sketch of the LOWESS algorithm.

1. Compute the number of points in the neighborhood. This is the LOWESS fraction times the number of data points rounded to the

nearest integer. Call this number  $q$ .

2. Use a tricube function to generate a weighted least squares fit. Let  $d_i$  be the distance from  $x_i$  to its  $q$ th nearest neighbor. The weight given to point  $(x_k, y_k)$  is:

$$t_i(x_k) = T\left(\frac{(x_i - x_k)}{d_i}\right) \quad (\text{EQ 3-54})$$

where  $T$  is the tri-cube weight function.

3. Compute the residuals of this fit.
4. Compute a bi-square weight function of the residuals.
5. Use the horizontal weights (from the tri-cube function) multiplied by the vertical weights (from the bi-square of the residuals) in a weighted least squares fit.

See the documentation for TRICUBE and BIWEIGHT for the details on these weighting functions.

#### DEFAULT

For lowess smoothing, the default width is 0.1 (that is, 10% of the range of the horizontal axis variable) and local linear fitting is used.

#### SYNONYMS

None

#### RELATED COMMANDS

LOWESS FRACTION	=	Sets the width for lowess.
LOWESS DEGREE	=	Specify linear or quadratic local fitting.
SMOOTH	=	Carries out a least squares smoothing.
MEDIAN SMOOTH	=	Carries out a median smoothing.
ROBUST SMOOTH	=	Carries out a robust smoothing.
FIT	=	Carries out a least squares fit.

#### REFERENCE

“Graphical Methods for Data Analysis,” Chambers, Cleveland, Kleiner, and Tukey, Wadsworth, 1983.

“Visualizing Data,” Cleveland, Hobart Press, 1993.

#### APPLICATIONS

Robust Smoothing, Time Series Analysis

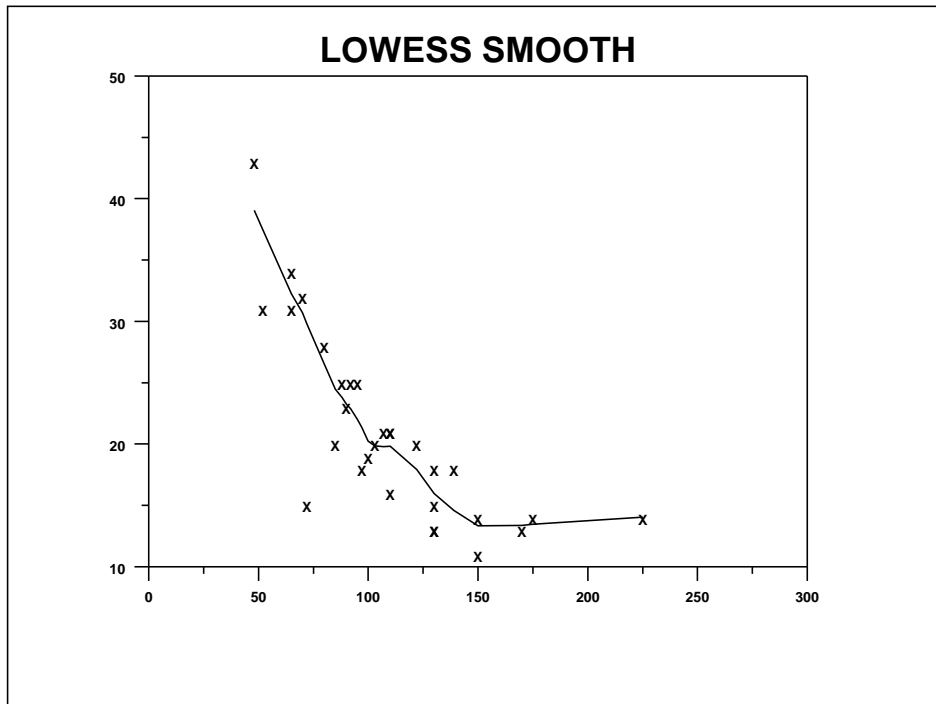
#### IMPLEMENTATION DATE

88/3 (local quadratic fitting was implemented 94/2)

PROGRAM 1

```

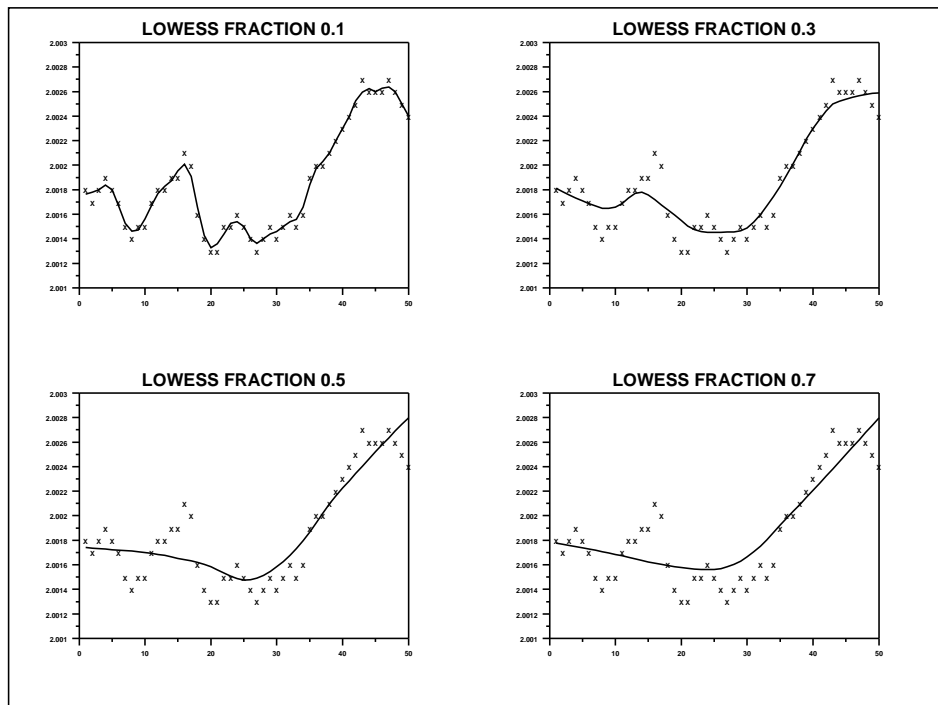
SERIAL READ Y
18 14 25 19 13 31 14 13 28 14 11 21 20 16 31
15 21 18 25 23 32 13 15 25 43 20 18 20 21 34
END OF DATA
SERIAL READ X
130 225 95 100 170 65 175 130 80 150 150 107 122 110 52
72 110 97 92 90 70 130 130 88 48 85 139 103 110 65
END OF DATA
LOWESS FRACTION 0.3
LOWESS SMOOTH Y X
CHAR X BLANK
LINES BLANK SOLID
TITLE LOWESS SMOOTH
PLOT Y PRED VS X
    
```



PROGRAM 2

```

SKIP 25
READ MAVRO.DAT Y
.
LET N = SIZE Y
LET X = SEQUENCE 1 1 N
.
MULTIPLY 2 2; MULTIPLY CORNER COORDINATES 0 0 100 100
LOWESS Y
CHARACTER X BLANK
LINE BLANK SOLID
TITLE LOWESS FRACTION 0.1
PLOT Y PRED VS X
LOWESS FRACTION 0.3
LOWESS Y
TITLE LOWESS FRACTION 0.3
PLOT Y PRED VS X
LOWESS FRACTION 0.5
LOWESS Y
TITLE LOWESS FRACTION 0.5
PLOT Y PRED VS X
LOWESS FRACTION 0.7
LOWESS Y
TITLE LOWESS FRACTION 0.7
PLOT Y PRED VS X
END OF MULTIPLY
    
```



PROGRAM 3

```

SKIP 25
READ HAYES1.DAT X ST Y
.
MULTIPLY 2 2; MULTIPLY CORNER COORDINATES 0 0 100 100
CHARACTER CIRCLE BLANK; LINE BLANK SOLID
TITLE RAW DATA
PLOT Y X
LOWESS FRACTION 0.3
LOWESS Y X
TITLE LINEAR LOWESS (0.3)
PLOT Y PRED VS X
LOWESS DEGREE 2
TITLE QUADRATIC LOWESS (0.3)
LOWESS Y X
PLOT Y PRED VS X
LOWESS FRACTION 0.15
LOWESS Y X
TITLE QUADRATIC LOWESS (0.15)
PLOT Y PRED VS X
    
```

