

An Introduction to PROC LOESS for Local Regression

Robert A. Cohen
SAS Institute Inc.
Cary, North Carolina, USA

Abstract

This paper describes the LOESS procedure which is a new procedure in SAS/STAT[®] software for performing local regression. Features of this procedure are outlined and a brief description of the fitting method is given. Examples are given illustrating the use of this procedure in obtaining fitted surfaces as well as prediction confidence limits for both univariate and multivariate regressor data. An automatic method for selecting the smoothing parameter based on a bias corrected Akaike information criterion is described and used in these examples.

Introduction

The LOESS procedure is a new procedure in SAS/STAT software for performing local regression. It is an experimental procedure with Release 7 and is a production procedure with Release 8. It forms one of the steps toward comprehensive support for modern nonparametric data analysis methods within SAS[®] software (Rodriguez and Stokes, 1998).

PROC LOESS implements a nonparametric method for estimating local regression surfaces pioneered by Cleveland (1979); also refer to Cleveland *et al.* (1988) and Cleveland and Grosse (1991). This method is commonly referred to as *loess*, which is short for *local regression*.

PROC LOESS allows greater flexibility than traditional modeling tools because you can use it for situations in which you do not know a suitable parametric form of the regression surface. Furthermore, PROC LOESS is suitable when there are outliers in the data and a robust fitting method is necessary.

The main features of PROC LOESS are as follows:

- fits nonparametric models
- supports the use of multidimensional predictors
- supports multiple dependent variables
- supports both direct and interpolated fitting using kd trees
- computes confidence limits for predictions

- performs iterative reweighting to provide robust fitting when there are outliers in the data
- supports scoring for multiple data sets

Local Regression and the Loess Method

Assume that for $i = 1$ to n , the i th measurement y_i of the response y and the corresponding measurement x_i of the vector x of p predictors are related by

$$y_i = g(x_i) + \epsilon_i$$

where g is the regression function and ϵ_i is a random error. The idea of local regression is that near $x = x_0$, the regression function $g(x)$ can be locally approximated by the value of a function in some specified parametric class. Such a local approximation is obtained by fitting a regression surface to the data points within a chosen neighborhood of the point x_0 .

In the loess method, weighted least squares is used to fit linear or quadratic functions of the predictors at the centers of neighborhoods. The radius of each neighborhood is chosen so that the neighborhood contains a specified percentage of the data points. The fraction of the data, called the *smoothing parameter*, in each local neighborhood controls the smoothness of the estimated surface. Data points in a given local neighborhood are weighted by a smooth decreasing function of their distance from the center of the neighborhood.

In a direct implementation, such fitting is done at each point at which the regression surface is to be estimated. A much faster computational procedure is to perform such local fitting at a selected sample of points in the predictor space and then to blend these local polynomials to obtain a regression surface.

You can use PROC LOESS to compute confidence limits for predictions under the standard statistical assumptions of independently and identically distributed normal errors with mean 0. By using iterative reweighting, PROC LOESS can also provide confidence limits for predictions when the error distribution is symmetric but not necessarily normal. Furthermore, by doing iterative reweighting, you can use

PROC LOESS to perform robust fitting in the presence of outliers in the data.

Exploratory Scatterplot Data Analysis

The following example shows how you can use PROC LOESS to perform exploratory scatterplot data analysis of complicated scatterplot data.

The data are monthly averaged atmospheric pressure differences between Easter Island and Darwin, Australia for a period of 168 months (NIST, 1998).

The following SAS statements create a data set named ENSO which contains this data.

```
data ENSO;
  input Pressure @@;
  Month=_N_;
datalines;
12.9 11.3 10.6 11.2 10.9 7.5 7.7 11.7
12.9 14.3 10.9 13.7 17.1 14.0 15.3 8.5
5.7 5.5 7.6 8.6 7.3 7.6 12.7 11.0
12.7 12.9 13.0 10.9 10.4 10.2 8.0 10.9
13.6 10.5 9.2 12.4 12.7 13.3 10.1 7.8
4.8 3.0 2.5 6.3 9.7 11.6 8.6 12.4
10.5 13.3 10.4 8.1 3.7 10.7 5.1 10.4
10.9 11.7 11.4 13.7 14.1 14.0 12.5 6.3
9.6 11.7 5.0 10.8 12.7 10.8 11.8 12.6
15.7 12.6 14.8 7.8 7.1 11.2 8.1 6.4
5.2 12.0 10.2 12.7 10.2 14.7 12.2 7.1
5.7 6.7 3.9 8.5 8.3 10.8 16.7 12.6
12.5 12.5 9.8 7.2 4.1 10.6 10.1 10.1
11.9 13.6 16.3 17.6 15.5 16.0 15.2 11.2
14.3 14.5 8.5 12.0 12.7 11.3 14.5 15.1
10.4 11.5 13.4 7.5 0.6 0.3 5.5 5.0
4.6 8.2 9.9 9.2 12.5 10.9 9.9 8.9
7.6 9.5 8.4 10.7 13.6 13.7 13.7 16.5
16.8 17.1 15.4 9.5 6.1 10.1 9.3 5.3
11.2 16.6 15.6 12.0 11.5 8.6 13.8 8.7
8.6 8.6 8.7 12.8 13.2 14.0 13.4 14.8
;
```

A scatterplot of these data is shown in Figure 1.

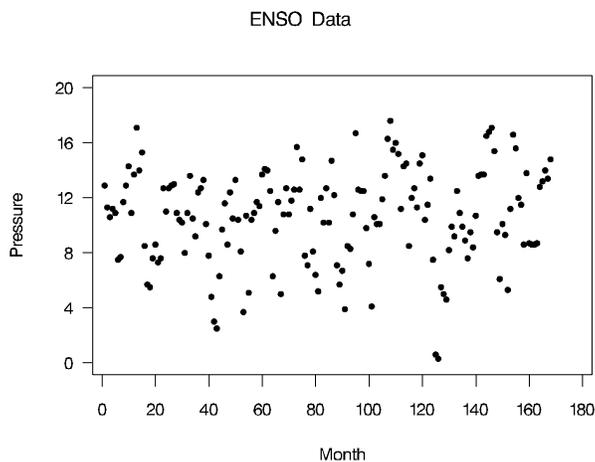


Figure 1. Average Monthly Pressure Differences between Darwin and Easter Island

The following statements compute a loess fit for these data for a range of smoothing parameters.

```
ods output OutputStatistics=ENSOstats
          FitSummary=ENSOSummary;

proc loess data=ENSO;
  model Pressure=Month /
    smooth = 0.02 to 0.2 by 0.01
    dfmethod=exact;
run;
```

The MODEL statement names the dependent variable and the regressor variables, which are separated by an equal sign. A linear function (the default) is to be fit locally, and the SMOOTH= option specifies the list of smoothing parameters. The DFMETHOD= option specifies that degree of freedom information used for statistical inference should be computed.

While all output of PROC LOESS can be optionally displayed, most often the LOESS procedure is used to produce output data sets that are viewed and manipulated by other SAS procedures. PROC LOESS uses the Output Delivery System (ODS) to place results in output data sets. This is a departure from older SAS procedures that provide OUTPUT options and statements to create SAS data sets from analysis results.

The ODS statement in the preceding SAS code requests that the “Output Statistics” table and the “Fit Summary” table be placed in SAS data sets named ENSOstats and ENSOSummary respectively. The “Fit Summary” table contains information about the fitting parameters you specify in the MODEL statement and summary statistics for the fit you obtain. The “Output Statistics” table contains the fitting points and predicted values as well as optional columns such as confidence limits that you request with MODEL statement options. The “Fit Summary” table is displayed by default. The “Output Statistics” table is not displayed by default but you can use the DETAILS option in the MODEL statement to request that this table and other optionally displayed tables appear in the output. Note that the tables produced by PROC LOESS can be placed in output data sets using the ODS OUTPUT statement, even if they are not included in the displayed output.

You can produce scatterplots of the data including the fitted loess curve for each smoothing parameter using the following statements:

```

symbol11 c=black i=join value=dot;
symbol12 c=black i=join value=none;

proc gplot data=ENSOstats;
  by SmoothingParameter;
  plot (DepVar Pred)*Month/overlay;
run;

```

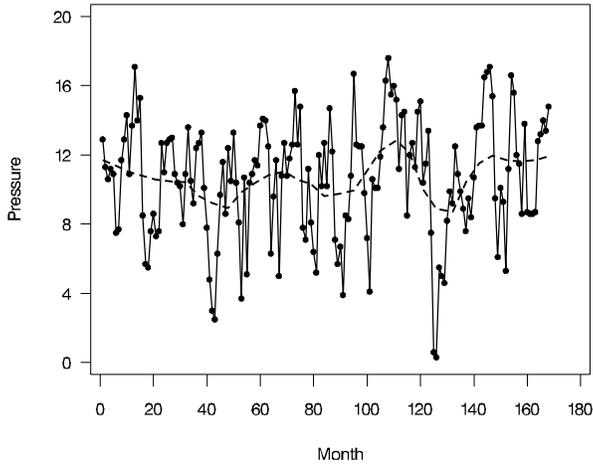


Figure 2. Over and Under Smoothed Loess Fits

In Figure 2 the loess fits for smoothing parameters 0.02 and 0.2 are plotted on the same scatterplot. You can see that the loess fits obtained depend strongly on the smoothing parameter used. When the smoothing parameter is 0.02 there is overfitting and the fit interpolates the data. When the smoothing parameter is 0.2 the loess fit suffers from excessive smoothing. A good smoothing parameter lies somewhere between these two extremes.

There are several strategies that you can use to select the smoothing parameter. One strategy is to examine plots of the fit residuals versus the predictor variable and to choose the largest smoothing parameter that yields no clearly discernible information in the fit residuals.

Alternatively, a variety of automatic methods for choosing the smoothing parameter exist. Many of these methods choose a smoothing parameter which minimizes a criterion that incorporates both the tightness of the fit and model complexity, of the form

$$\log(\hat{\sigma}^2) + \psi(L)$$

where $\hat{\sigma}^2$ is an average residual sum of squares and $\psi(\cdot)$ is a penalty function designed to decrease with increasing smoothness of the fit. Here L is the smoothing matrix of the method. This matrix satisfies

$$\hat{y} = Ly$$

where y is the vector of observed values and \hat{y} is the corresponding vector of predicted values of the dependent variable. Examples of specific criteria obtained with this methodology are generalized cross-validation (Craven and Wahba, 1979), the Akaike information criterion (Akaike, 1973), and the bias corrected Akaike information criteria (Hurvich and Simonoff, 1998).

Hurvich and Simonoff (1998) show that the bias corrected Akaike information criteria avoid the tendency to undersmooth that often occurs when using the classical Akaike information criterion or generalized cross-validation. In this paper one of these bias corrected Akaike information criteria is used. This criterion is given by

$$AIC_{C_1} = n \log(\hat{\sigma}^2) + n \frac{\delta_1 / \delta_2 (n + \nu_1)}{\delta_1^2 / \delta_2 - 2}$$

where

$n \equiv$ Number of observations

$\delta_1 \equiv \text{Trace}(I - L)^T (I - L)$

$\delta_2 \equiv \text{Trace}((I - L)^T (I - L))^2$

$\nu_1 \equiv \text{Trace}(L)$

The statistics that you use to compute AIC_{C_1} are also used to perform statistical inference in loess models. You can find these numbers in the “Fit Summary” table of PROC LOESS provided that you specify at least one of the options ALL, CLM, DFMETHOD=EXACT, STD, and T in the MODEL statement. As with all tables of output, you can put these numbers in a SAS data set by using an ODS output statement to output the “Fit Summary” table. The SAS macro SMOOTH-SELECT provided in Appendix 1, takes such an output data set as an argument and finds the smoothing parameter that yields the smallest AIC_{C_1} statistic for all the smoothing parameters that you specified in the MODEL statement.

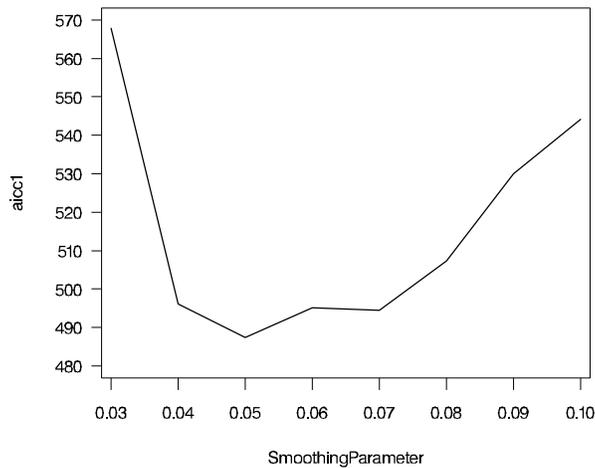


Figure 3. AIC_{C_1} versus SmoothingParameter

Figure 3 is a plot of the AIC_{C_1} versus SmoothingParameter for the ENSO data of our example. Recall that the “Fit Summary” tables are in the ODS generated output data set named ENSOsummary. The statement

```
%SmoothSelect(ENSOsummary)
```

selects 0.05 as the optimal smoothing parameter.

Figure 4 shows the loess fit with this smoothing parameter.

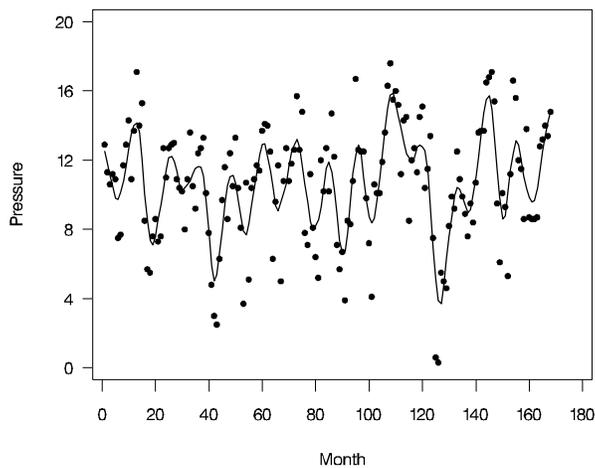


Figure 4. Loess Fit with Smoothing Parameter 0.05

You can clearly see a 12 month cycle in the loess fit shown in Figure 4. Are there other less obvious cycles in the data? You can determine this by filtering the 12 month cycle from the data. One method for doing this is to use the residuals you obtain after fitting a

parametric model of the form

$$\text{Pressure} = \beta_0 + \beta_1 \cos\left(\frac{\text{Month}}{12}\right) + \beta_2 \sin\left(\frac{\text{Month}}{12}\right)$$

The following statements use the REG procedure to produce these residuals in the output data set ENSO1.

```
data ENSO(drop=pi);
  set ENSO;
  pi = 4 * atan(1);
  cos1 = cos(2*pi*Month/12);
  sin1 = sin(2*pi*Month/12);

proc reg data=ENSO;
  model Pressure = cos1 sin1;
  output out=ENSO1 r=FilteredPressure;
run;
```

You fit a loess model for this filtered data with the following statements:

```
ods output OutputStatistics=ENSO1stats;

proc loess data=ENSO1;
  model FilteredPressure=Month/
    smooth = 0.12;
run;
```

The smoothing parameter 0.12 is chosen using the same procedure described earlier. A plot of this fit is shown in Figure 5. Vertical reference lines at months 45, 87, and 129 have been added to this plot.

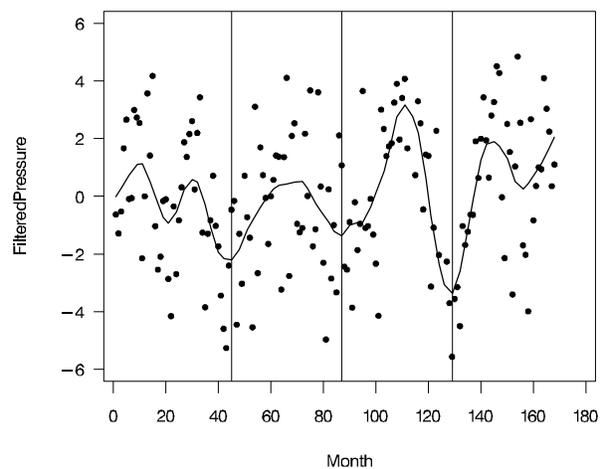


Figure 5. Loess Fit of 12 Month Filtered ENSO Data

You can see an approximately 42 month cycle in the filtered ENSO data. Are there still other significant

cycles in the data? Repeating the previous method, you first filter both the 12 and 42 month cycles from the original data. Then you fit a loess curve through the filtered data, as shown in the following statements:

```
data ENSO(drop=pi);
  set ENSO;
  pi = 4 * atan (1);
  cos2 = cos(2*pi*Month/42);
  sin2 = sin(2*pi*Month/42);

proc reg data=ENSO;
  model Pressure = cos1 sin1 cos2 sin2;
  output out=ENSO2 r=Filtered2Pressure;
run;

ods output OutputStatistics=ENSO2stats;

proc loess data=ENSO2;
  model Filtered2Pressure=Month/
    smooth = 0.12;
run;
```

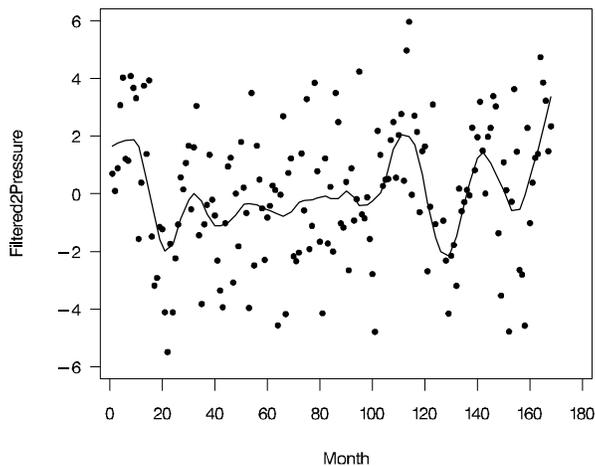


Figure 6. Loess Fit of 12 and 42 Month Filtered ENSO Data

Figure 6 shows the loess fit for the data after filtering both the 12 and 42 month cycles from the original data. You can see evidence of a cycle with a period in the middle 20's in this data, although this cycle is not as well defined as the 12 and 42 month cycles previously detected.

The exploratory data analysis of the ENSO data using PROC LOESS reveals that there are cycles in the data with periods of approximately 12 months and 42 months, and perhaps a cycle with a period of about 25 months. As it turns out, there are physical phenomena that account for the existence of these cycles. The 42 month cycle is caused by the periodic "El Nino" warming of the Pacific Ocean. The cycle with period in the mid twenties is known to climatologists as the "Southern Oscillation" and ENSO is an

acronym for the combination of these two phenomena (Battisti and Sarachik, 1995).

You can also use PROC LOESS to compute confidence limits for the predictions as follows:

```
ods output OutputStatistics=ENSOstats;

proc loess data=ENSO;
  model Pressure=Month /
    smooth = 0.05
    r clm alpha=0.01;
run;
```

The R option in the MODEL statement specifies that fit residuals are included in the "Output Statistics" table. The CLM option specifies that confidence limits are included in this table. The ALPHA=0.01 specifies that 99% limits are computed (the default is 95%).

In computing confidence limits, PROC LOESS assumes that the error distribution is normal. If the error distribution is symmetric but not normal, you can still compute asymptotically valid confidence limits using PROC LOESS by iterative reweighting. This is requested using the ITERATIONS= option in the MODEL statement.

You can check the assumption of a normal error distribution by looking at a normal probability plot of the loess fit residuals. The following statements use the CAPABILITY procedure in SAS/QC[®] software to produce this plot:

```
proc capability data=ENSOstats;
  var residual;
  qqplot/vaxis=axis1;
  axis1 label = ( r=0 a=90 );
run;
```

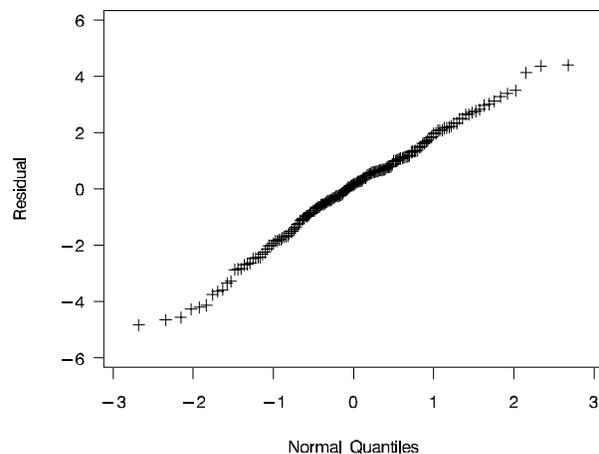


Figure 7. Normal Probability Plot of the Loess Fit Residuals

You can see from Figure 7 that the relation between the residuals and normal quantiles is linear, implying that the assumption of normal errors is reasonable. Hence the prediction confidence limits in the ENSOstats data set that you requested by specifying the CLM option in the MODEL statement of the previous SAS code are asymptotically valid. You can plot the loess fit with confidence limits with the following statements:

```

symbol1 c=black i=none value=dot;
symbol2 c=black i=join value=none;
symbol3 c=black i=join value=none;
symbol4 c=black i=join value=none;

proc gplot data=ENSOstats;
  format DepVar 3.0;
  plot (DepVar Pred UpperCl LowerCL)*Month/
  overlay
  hminor = 0
  vminor = 0
  vaxis = axis1
  frame;
  axis1 label = ( r=0 a=90 )
  order=(0 to 20 by 4);
run;

```

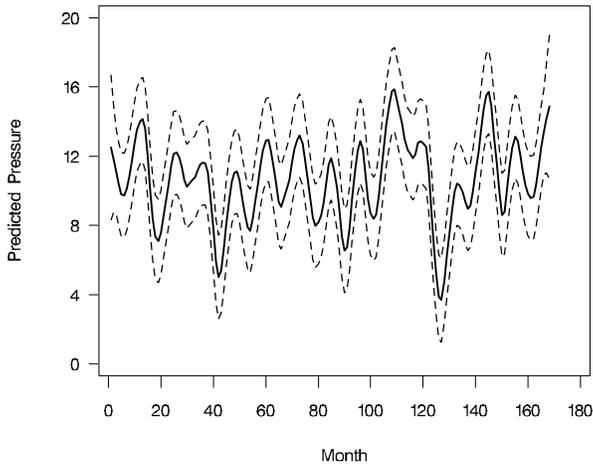


Figure 8. Loess Fit with 99% Confidence Limits

Surface Fitting

The following data set contains measurements in grams per square meter of sulfate (SO_4) deposits during 1990 at 179 sites throughout the 48 states.

```

data SO4;
  input Latitude Longitude SO4;
datalines;
32.45833 87.24222 1.403
34.28778 85.96889 2.103
...
45.07611 72.67556 2.646
;

```

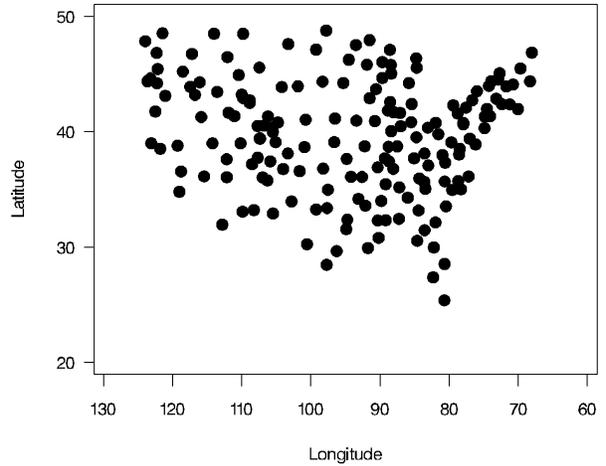


Figure 9. Locations of Sulfate Measurements

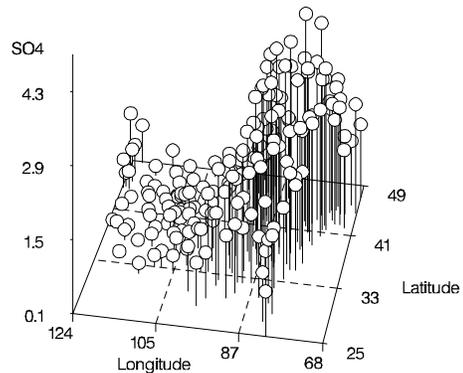


Figure 10. Scatter plot of SO_4 Data

Figure 9 shows the measurement locations and Figure 10 shows a scatterplot of the measured sulfate deposits. From these scatter plots, it is clear that the largest concentrations are in the northeastern United States. These plots also indicate that a nonparametric surface, such as a loess fit, is appropriate for these data.

The sulfate measurements are irregularly spaced. The following statements create a SAS data set containing a regular grid of points to be scored for the purposes of plotting.

```

data PredGrid;
  do Latitude = 25 to 49 by 1;
  do Longitude = 67 to 124 by 1;
  output;

```

```

end;
end;

```

You fit a loess surface to the SO4 data with the following statements:

```

ods output ScoreResults=SO4score;

proc loess data=SO4;
  model SO4=Latitude Longitude/smooth=0.12;
  score data=PredGrid;
run;

```

The preliminary steps of first fitting the surface for a range of smoothing parameters and then selecting the smoothing parameter which minimizes the bias corrected Akaike information criterion, AIC_{C_1} are not shown here. Note that you use the SCORE statement to evaluate the fit at the points in the PredGrid data set. You use an ODS OUTPUT statement to capture the “Score Results” table in a SAS data set named SO4score.

Because the data points in the PredGrid data set are spaced regularly, the scored data set, SO4score, has data in a form you can use directly with the G3D procedure and the GCONTOUR procedure in SAS/GRAPH® software.

The following statements plot the loess surface shown in Figure 11 using the G3D procedure.

```

proc g3d data=SO4score;
  format Latitude f4.0;
  format Longitude f4.0;
  format p_SO4 f4.1;
  plot Longitude*Latitude=p_SO4/
    tilt=60 rotate=80;
run;

```

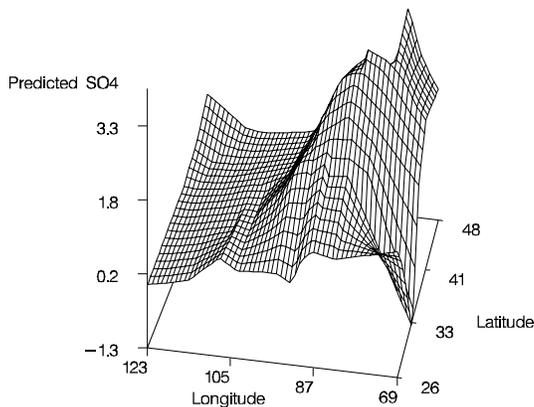


Figure 11. Loess Fit of the SO4 Data

You need to exercise some caution when interpreting the loess fit shown in Figure 11, since the scored surface is interpolated onto a rectangular grid from data that does not fill this rectangle. This can cause spurious effects near the boundaries of the surface. Figure 12 helps explain the problem.

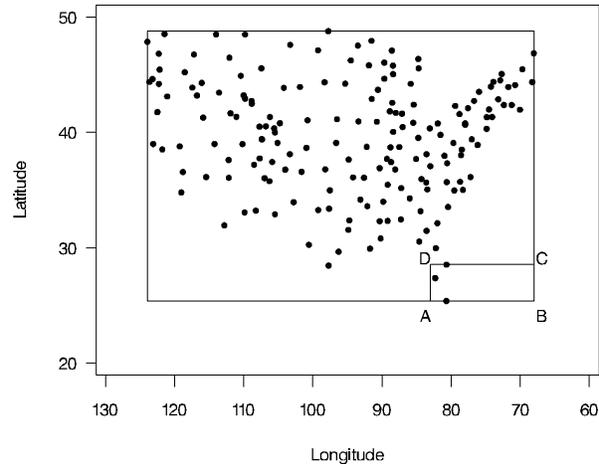


Figure 12. Measurement Locations Showing a kd Tree Cell

The rectangle with vertices “ABCD” in Figure 12 is one of the kd tree cells used in fitting the loess surface. Weighted local regression is done to obtain predicted values at these vertices. The points B and C are “far” from the data, making the predicted value at these points suspect. Since scoring at points within this rectangle is done by interpolation from the fitted values at points A,B,C, and D, the interpolated surface above this rectangle is similarly not reliable.

In such cases, a more prudent approach is to do local fitting directly at each fitting point. This can be done by specifying the DIRECT option in the MODEL statement of PROC LOESS as follows:

```

ods Output OutputStatistics=SO4Out;

proc loess data=SO4;
  model SO4=Latitude Longitude /
    smooth=0.34
    clm direct degree=2;
run;

```

As the DEGREE=2 option has been specified in the MODEL statement local quadratic polynomials are used. The smoothing parameter is selected to minimize the AIC_{C_1} criterion. Note that the “Output Statistics” table, which contains the predicted values and 95% confidence limits, has been placed in the data set named SO4Out using an ODS OUTPUT statement.

The following statements produce a scatter plot of the loess fit shown in Figure 13.

```
proc g3d data=SO4Out;
  format Latitude f4.0;
  format Longitude f4.0;
  format Pred f4.1;
  scatter Longitude*Latitude=Pred /
    shape='balloon'
    size=0.35
    rotate=80
    tilt=60;
run;
```

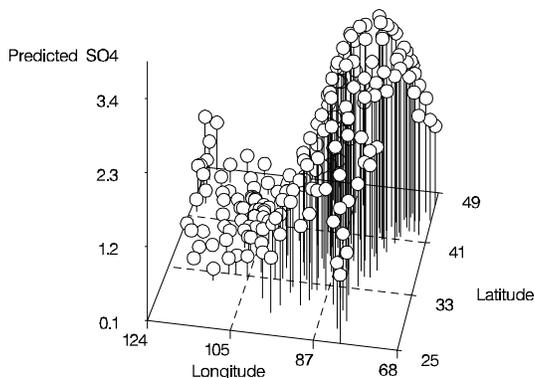


Figure 13. Scatterplot of the Loess Fit using DIRECT Fitting with Locally Quadratic Polynomials

Conclusion

This paper gives examples of how to use PROC LOESS to fit loess models to data having one or two regressors. Several interesting features of PROC LOESS are not demonstrated in this paper. You can specify that PROC LOESS use iterative reweighting to fit loess models and obtain confidence limits for predictions for data with outliers and symmetric but non-normal error distribution. See Rodriguez and Stokes (1998) for such an example. You can also specify that LOESS procedure do robust scaling of the regressor data. This is appropriate for data with two or more regressors that are a priori incompatibly scaled.

Documentation for this procedure is available in *SAS/STAT: Experimental Procedures*, which is included in SAS Online Doc, Version 7-1. Note that this version of the online documentation contains .pdf files for all chapters.

Appendix 1: Macro for Smoothing Parameter Selection

The following SAS macro selects the smoothing parameter which minimizes the AIC_{C_1} statistic for the loess fits that are summarized in the data set that is the argument of this macro. You create this data set by using an ODS OUTPUT statement to output the “Fit Summary” table of PROC LOESS for a range of SMOOTH= values. The macro also produces a graph of AIC_{C_1} versus the smoothing parameter for the loess fits summarized in the argument data set.

```
%macro SmoothSelect(data);
  options nonotes;
  ods listing close;

  data temp;
    set &data(keep = Label1 nValue1 smoothingParameter
      where=(Label1 in
        ('Number of Observations',
        'Residual Sum of Squares',
        'Trace[L]',
        'Equivalent Number of Parameters',
        'Delta1',
        'Delta2',
        'Lookup Degrees of Freedom')));
  run;

  /*
  NOTE: If you are using this macro with V7 replace
  'Delta1' by 'Trace[tr(I-L)*(I-L)]' and
  'Delta2' by 'Trace[(tr(I-L)*(I-L))^2]'
  in the previous where= clause.
  */

  proc transpose data=temp(drop=Label1) out=temp;
    by smoothingParameter;
  run;

  data temp(drop=_NAME_); set temp;
    rename Col1 = n
      Col2 = rss
      Col3 = traceL
      Col4 = delta1
      Col5 = delta2
      Col6 = nul
      Col7 = lkdf;

  data SmoothCriteria(keep = SmoothingParameter aiccl);
    set temp;
    sigmaHat=rss/n;
    aiccl=n*(log(sigmaHat) + (delta1/delta2)*(n+nul)/(lkdf-2));

  proc sort data=SmoothCriteria(where=(aiccl^=.)
    out=AicclResults;
  by aiccl;
  run;

  ods listing;
  proc print data=AicclResults(obs=1);
    title2 'Smoothing Parameter Minimizing the AICCL Statistic';
    id SmoothingParameter;
  run;

  proc sort data=AicclResults;
    by SmoothingParameter;
  run;

  title1 'AICCL Criterion';
  symbol1 c=black i=join value=none width=2;
  proc gplot data=SmoothCriteria;
    plot aiccl*SmoothingParameter/
      hminor = 0
      vminor = 0
      vaxis = axis1
      frame;
    axis1 label = ( r=0 a=90 );
  run; quit;

  title1;
%mend;
```

Appendix 2: Accessing the Example Data Sets

The file, `loess.sas`, which contains the data sets of the examples and the macro listed in APPENDIX 1 has been stored on SAS Institute’s Internet gateway.

You can download this file if you have access to the Internet. To download these files, connect to

ftp.sas.com. Once you are connected, enter the following responses as prompted:

```
Name (ftp.sas.com: userid): anonymous
Password: your e-mail address
```

All SUGI 24 files are stored in the following directory:

/pub/SUGI24

There is one subdirectory for each paper that has ancillary files. For a complete index of all files in /pub/SUGI24, download the following file:

README.index

The file README.index has a description of each subdirectory. The description will contain the title of the paper and the directory name where the files are stored.

You should use a binary file transfer when transferring SAS data sets, SAS catalogs, and other binary files.

References

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle," *Proceedings of the 2nd International Symposium on Information Theory* (eds Petrov and Csaki), 267-281.

Battisti, D. S., and Sarachik, E. S. (1995) "Understanding and Predicting ENSO," *Reviews of Geophysics*, 33, 1367-1376.

Cleveland, W. S. (1979), "Robust locally-weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, 74, 829-836.

Cleveland, W. S., Devlin, S. J., and Grosse, E. (1988), "Regression By Local Fitting," *Journal of Econometrics*, 37, 87-114.

Cleveland, W. S., and Grosse, E. (1991), "Computational Methods for Local Regression," *Statistics and Computing*, 1, 47-62.

Craven, P., and Wahba, G. (1979) "Smoothing noisy data with spline functions," *Numer. Math.*, 31, 377-403.

Hurvich, C. M., and Simonoff, J. S. (1998), "Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion" *Journal of the Royal Statistical Society B* 60, 271-293.

NIST (1998), "Statistical Reference Data Sets," [<http://www.nist.gov/itl/div898/strd>], accessed 20 January 1999.

Rodriguez, R.N., and Stokes, M.E. (1998) "Recent Enhancements and New Directions in SAS/STAT

Software, Part II. Nonparametric Modeling Procedures" *Proceedings of the 23rd SAS Users Group International Conference*, 1262-1270.

SAS Institute Inc. (1999), *SAS/STAT: Experimental Procedures*, Cary, NC: SAS Institute Inc.

Contact Information

Robert A. Cohen, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513. Phone (919) 677-8000 x5849. FAX (919) 677-4444. Email sasroc@wnt.sas.com

SAS, SAS/STAT, SAS/QC, and SAS/GRAPH are registered trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.