Paper P256-26

Fitting Generalized Additive Models with the GAM Procedure

Dong Xiang, SAS Institute Inc., Cary, NC

Abstract

This paper describes the use of the GAM procedure for fitting generalized additive models (Hastie and Tibshirani, 1990). PROC GAM, production in Release 8.2, provides an array of powerful tools for data analysis, incorporating nonparametric regression and smoothing techniques as well as generalized distributional modeling. Compared with other regression procedures such as PROC REG and PROC GLM, the methodology behind PROC GAM relaxes the usual assumption of linearity, enabling you to uncover hidden structure in the relationship between the independent variables and the dependent variable. Moreover, you can use PROC GAM to model not only Gaussian data, but also data from binary, Poisson, and other non-Gaussian distributions.

Introduction

The GAM procedure is the most versatile of several new procedures for nonparametric regression in SAS® software. The methodology behind the GAM procedure has greater flexibility than traditional parametric modeling tools such as linear or nonlinear regression. It relaxes the usual parametric assumption, and enables you to uncover structure in the relationship between the independent variables and the dependent variable that you might otherwise miss.

Many nonparametric methods, such as the thin-plate smoothing spline used in PROC TPSPLINE and the local regression method used in PROC LOESS, do not perform well when there is a large number of independent variables in the model. The sparseness of data in this setting inflates the variance of the estimates. The problem of rapidly increasing variance for increasing dimensionality is sometimes referred to as the “curse of dimensionality.” Interpretability is another problem with nonparametric regression based on kernel and smoothing spline estimates (Hastie and Tibshirani 1990).

To overcome these difficulties, Stone (1985) proposed additive models. These models estimate an additive approximation to the multivariate regression function. The benefits of an additive approximation are at least twofold. First, since each of the individual additive terms is estimated using a univariate smoother, the curse of dimensionality is avoided, at the cost of not being able to approximate universally. Second, estimates of the individual terms explain how the dependent variable changes with the corresponding independent variables.

To extend the additive model to a wide range of distribution families, Hastie and Tibshirani (1990) proposed generalized additive models. These models assume that the mean of the dependent variable depends on an additive predictor through a nonlinear link function. Generalized additive models permit the response probability distribution to be any member of the exponential family of distributions. Many widely used statistical models belong to this general class, including additive models for Gaussian data, nonparametric logistic models for binary data, and nonparametric log-linear models for Poisson data.

PROC GAM implements the generalized additive model proposed by Hastie and Tibshirani (1990). The GAM procedure

- fits nonparametric or semiparametric additive models
- supports the use of multidimensional data
- supports multiple SCORE statements
- enables you to specify the model degrees of freedom or smoothing parameter

PROC GAM can fit Gaussian, binomial, Poisson, and Gamma distributions. For each distribution, although theoretically more than one link can exist, PROC GAM always uses the canonical link. This is because the difference between link alternatives will be less pronounced for nonparametric models, in light of the flexibility of nonparametric model forms.

In the next section, we will introduce a data set appropriate for GAM analysis. Then, we will briefly discuss the GAM methodology. Finally, we will show a couple of examples using PROC GAM, comparing it with other procedures in the SAS/STAT system.

Kyphosis Study

The data introduced in this section are based on a study by Bell et al. (1989). Bell and his associates studied the results of multiple-level thoracic and lumbar laminectomy, a corrective spinal surgery commonly performed on children. The data in the study consist of retrospective measurements on 83 patients. The specific outcome of interest is the presence (1) or absence (0) of kyphosis (a severe forward flexion of the spine). The available predictor variables are Age in months at the time of the operation, the first vertebrae level involved in the operation (StartVert), and the number of levels involved (NumVert). The goal of this analysis is to identify risk factors for kyphosis.

Figure 1 is a plot of the incidence of Kyphosis against StartVert, Age, and NumVert. It is difficult to see any trends in the data. This is often true of binary data, especially with multiple predictors. You can try to use linear logistic regression (PROC GENMOD or PROC LOGISTIC) to summarize the relationships, but without prior knowledge it is difficult to do so. PROC GAM, however, does not require any prior assumptions on the underlying relationship. It can find a valid form to represent the data in a smooth format, making it useful as a tool to visualize the pattern among the variables before using one of the parametric modeling procedures.

Before using PROC GAM to analyze this data set, we will briefly introduce the generalized additive model in the next section.

Generalized Additive Model

Let Y be a response random variable and $X_1, \cdots, X_p$ be a set of predictor variables. A regression procedure can be viewed as a method for estimating how the value of Y depends on the values of $X_1, \cdots, X_p$. The standard linear regression model assumes the expected value of Y has a linear form.

$$E(Y) = f(X_1, \cdots, X_p) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

Given a sample of values for Y and X, estimates of $\beta_0, \beta_1, \cdots, \beta_p$ are often obtained by the least squares method.

The additive model generalizes the linear model by modeling the expected value of Y as

$$E(Y) = f(X_1, \cdots, X_p) = s_0 + s_1(X_1) + \cdots + s_p(X_p)$$

where $s_i(X), i = 1, \ldots, p$ are smooth functions. These functions are not given a parametric form but instead are estimated in a nonparametric fashion.

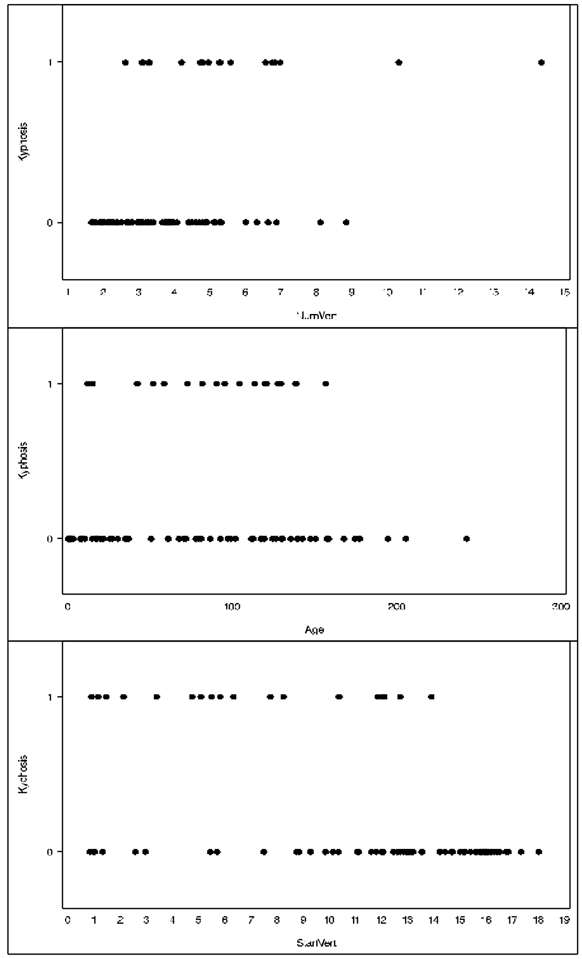

Figure 1. Plot of Kyphosis Data

Generalized additive models extend traditional linear models in another way, namely by allowing for a link between $f(X_1, \cdots, X_p)$ and the expected value of Y. This amounts to allowing for a alternative distribution for the underlying random variation besides just the normal distribution. While Gaussian models can be used in many statistical applications, there are types of problems for which they are not appropriate. For example, the normal distribution may not be adequate for modeling discrete responses such as counts, or bounded responses such as proportions. Thus, generalized additive models can be applied to a much wider range of data analysis problems.

Generalized additive models consist of a random component, an additive component, and a link function relating these two components. The response Y, the random component, is assumed to have a density in the exponential family

$$f_Y(y;\theta;\phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\}$$

where θ is called the natural parameter and ϕ is the scale parameter. The normal, binomial, and Poisson distributions are all in this family, along with many others. The quantity

$$\eta = s_0 + \sum_{i=1}^{p} s_i(X_i)$$

where $s_1(\cdot), \cdots, s_p(\cdot)$ are smooth functions defines the additive component. Finally, the relationship between the mean μ of the response variable and η is defined by a link function $g(\mu) = \eta$. The most commonly used link function is the canonical link, for which $\eta = \theta$.

Generalized additive models and generalized linear models can be applied in similar situations, but they serve different analytic purposes. Generalized linear models emphasize estimation and inference for the parameters of the model, while generalized additive models focus on exploring data nonparametrically. Generalized additive models are more suitable for exploring the data set and visualizing the relationship between the dependent variable and the independent variables.

Fitting Generalized Additive Models

In this section, the general method of fitting additive models will be outlined. The two important pieces are the backfitting and local scoring algorithms.

Consider the estimation of the smoothing terms $s_0, s_1(\cdot), \cdots, s_p(\cdot)$ in the additive model. Many ways are available to approach the formulation and estimation of additive models. The backfitting algorithm is a general algorithm that can fit an additive model using any regression-type smoothers.

Define the jth set of partial residuals as

$$R_j = Y - s_0 - \sum_{k \neq j} s_k(X_k)$$

The partial residuals remove the effects of all the other variables from y; therefore they can be used to model the effects against x_j.

This is the foundation for the backfitting algorithm, providing a way for estimating each smoothing function $s_j(\cdot)$ given estimates $\{\hat{s}_i(\cdot), i \neq j\}$ for all the others. The backfitting algorithm is iterative, starting with initial functions $s_0, \cdots, s_p$, and with each iteration, cycling through the partial residuals, fitting the individual smoothing components to its partial residuals. Iteration proceeds until the individual components do not change.

The algorithm so far described fits just additive models. The algorithm for generalized additive models is a little more complicated. Generalized additive models extend generalized linear models in the same manner as additive models extend linear regression models, that is, by replacing the linear form $\alpha + \sum_j X_j\beta_j$ with the additive form $\alpha + \sum_j s_j(X_j)$.

In the same way, estimation of the additive terms for generalized additive models is accomplished by replacing the weighted linear regression for the adjusted dependent variable by the weighted backfitting algorithm, essentially fitting a weighted additive model. The algorithm used in this case is called the local scoring algorithm. It is also an iterative algorithm and starts with initial estimates of $s_0, s_1, \cdots, s_p$. During each iteration, an adjusted dependent variable and a set weight are computed, and then the smoothing components are estimated using a weighted backfitting algorithm. The scoring algorithm stops when the deviance of the estimates ceases to decrease.

Overall, then, the estimating procedure for generalized additive models consists of two loops. Inside each step of the local scoring algorithm (outer loop), a weighted backfitting algorithm (inner loop) is used until convergence. Then, based on the estimates from this weighted backfitting algorithm, a new set of weights is calculated and the next iteration of the scoring algorithm starts.

Any nonparametric smoothing method can be used to obtain $s_i(x)$. The GAM procedure implements the B-spline and local regression methods for univariate smoothing components and the thin-plate smoothing spline for bivariate smoothing components. More detailed descriptions of these smoothing methods can be found in the *SAS/STAT User's Guide, Version 8*.

The smoothers used in PROC GAM have a single smoothing parameter. The generalized cross validation (GCV) function has been widely used in many nonparametric regression methods as a criterion to choose the smoothing parameters. The GCV function approximates the expected prediction error. The model selected by the GCV function is thus judged to have the best prediction ability. In addition to automatically choosing the smoothing parameter by

GCV, the GAM procedure also gives you the option of specifying the degrees of freedom for each individual smoothing component.

Analysis of Kyphosis Data

In this section, the Kyphosis data are used to illustrate the use of the GAM procedure. The syntax of the GAM procedure is similar to that of other regression procedures in the SAS® software. The following statements may be used to fit a GAM model.

```
PROC GAM data=kyphosis;
   model kyphosis = spline(NumVert,df=3)
                    spline(Age,df=3)
                    spline(StartVert,df=3)
                    /dist=logist;
    output out=estimate p uclm lclm;
 run;
```

The above statements request PROC GAM to fit a logistic additive model with binary dependent variable kyphosis and ordinary independent variables NumVert, Age, and StartVert. Each term is fitted using a B-spline smoother with 3 degrees of freedom. Although this might seem to be an unduly modest amount of flexibility, it is better to be conservative with a data set of 83 data points. Also the OUTPUT statement is used to write the estimated function and confidence limits to the data set estimate. The output from PROC GAM is listed in Figure 2 and Figure 3.

```
                          The GAM Procedure
                     Dependent Variable: Kyphosis
 Smoothing Model Component(s): spline(Age) spline(StartVert) spline(NumVert)

                       Summary of Input Data Set

              Number of Observations                 83
              Number of Missing Observations          0
              Distribution                     Binomial
              Link Function                       Logit


                  Iteration Summary and Fit Statistics

     Number of local score iterations                       9
     Local score convergence criterion           2.6635758E-9
     Final Number of Backfitting Iterations                 1
     Final Backfitting Criterion                 5.2326788E-9
     The Deviance of the Final Estimate          46.610922317
```

Figure 2. Summary Statistics

The first part of the output from PROC GAM (Figure 2) summarizes the input data set and provides a summary for the backfitting and local scoring algorithms. The second part of the output (Figure 3) provides analytical information about the fitted model. The critical part of the output is the “Analysis of Deviance” table, shown in Figure 3. For each smoothing effect in the model, this table gives a χ^2-test comparing the deviance between the full model and the model without this variable. In this case, the analysis of deviance results indicate that the effects of Age and StartVert are significant, while the effect of NumVert is insignificant.

```
                          The GAM Procedure
                     Dependent Variable: Kyphosis
 Smoothing Model Component(s): spline(Age) spline(StartVert) spline(NumVert)

                       Regression Model Analysis
                          Parameter Estimates

                       Parameter      Standard
   Parameter            Estimate         Error    t Value    Pr > |t|

   Intercept            -2.01533       1.45620      -1.38      0.1706
   Linear(Age)           0.01213       0.00794       1.53      0.1308
   Linear(StartVert)    -0.18615       0.07628      -2.44      0.0171
   Linear(NumVert)       0.38347       0.19102       2.01      0.0484


                       Smoothing Model Analysis
                Fit Summary for Smoothing Components

                                                                    Num
                       Smoothing                                 Unique
 Component             Parameter            DF           GCV        Obs

 Spline(Age)            0.999996      3.000000    328.512864         66
 Spline(StartVert)      0.999551      3.000000    317.646703         16
 Spline(NumVert)        0.921758      3.000000     20.144058         10


                       Smoothing Model Analysis
                         Analysis of Deviance

                                        Sum of
  Source                     DF        Squares    Chi-Square    Pr > ChiSq

  Spline(Age)           3.00000      10.494369       16.4358        0.0009
  Spline(StartVert)     3.00000       5.494968        8.6060        0.0350
  Spline(NumVert)       3.00000       2.184518        3.4213        0.3311
```

Figure 3. Model Fit Statistics

The partial predictions for each predictor are plotted in Figure 4. Notice that the 95% confidence limits for NumVert cover the zero axis, confirming the insignificant of this term. The plot also shows that the partial predictions corresponding to both Age and StartVert have a strong quadratic pattern.

Having used the GAM procedure to discover an appropriate form of the dependence of Kyphosis the independence variables, you can use the GENMOD procedure to fit and assess the corresponding parametric model. The following codes fit a GENMOD model with quadratic terms for Age and StartVert, including tests for the joint linear and quadratic effects of each variable. The resulting contrast tests are shown in Output 5.

```
PROC GENMOD data=kyphosis (where=(NumVert ^= 14));
   model kyphosis = Age  age*age
              StartVert StartVert*StartVert
             /link = logit dist=binomial;
   contrast 'Age' age 1, age*age 1;
   contrast 'StartVert' StartVert 1,
                    StartVert*StartVert 1;
run;
```

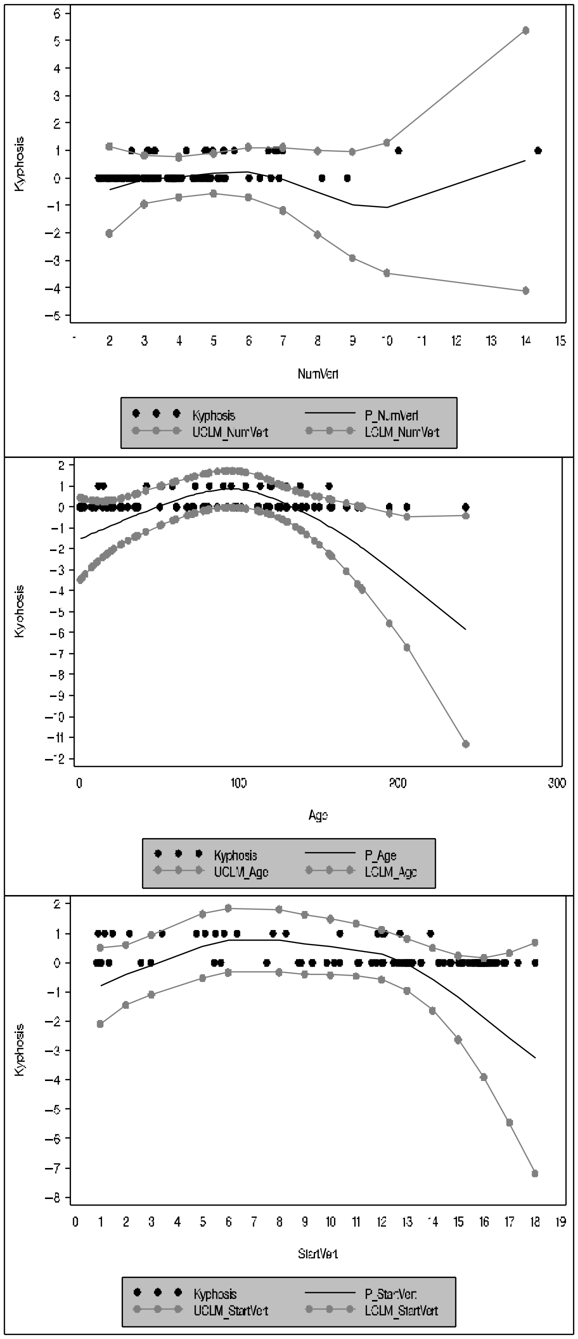

Figure 4. Partial Prediction for Each Predictor

The results for the quadratic GENMOD model are quite consistent with the GAM results.

```
                         The GENMOD Procedure

PROC GENMOD is modeling the probability that Kyphosis='1'.


                           Contrast Results

                                  Chi-
          Contrast        DF     Square    Pr > ChiSq    Type

          Age              2      11.78        0.0028    LR
          StartVert        2      21.49        <.0001    LR
```

Figure 5. The GENMOD Output

Comparing PROC GAM with PROC TP-SPLINE

In this section, we compare the difference between the model fitted by PROC GAM and model fitted by PROC TPSPLINE.

The data represent the deposition of sulfate (SO_4) at 179 sites in the 48 contiguous states of the United States in 1990. Each observation records the latitude and longitude of the site as well as the SO_4 deposition at the site measured in grams per square meter (g/m^2). The data are and plotted in Figure 6.

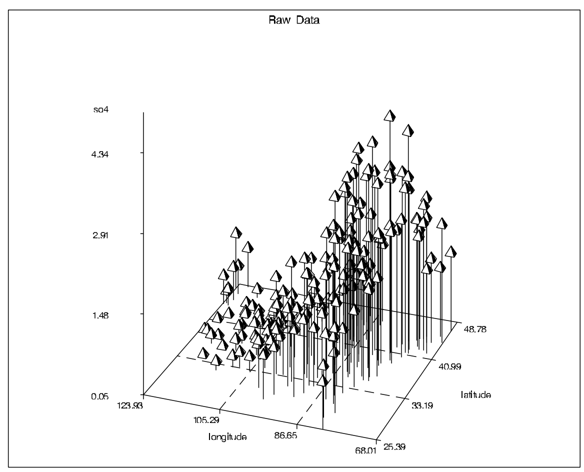

Figure 6. Raw Data Plot

GAM assumes an additive model, that is, the difference between SO_4 at two longitude points is identical across the whole range of latitudes, while the thin-plate smoothing spline method used in PROC TP-SPLINE does not make that assumption. The benefit of the assumption is that PROC GAM runs much faster than PROC TPSPLINE. The downside is that the additive assumption may not be appropriate for the spatial data like this.

The following code produces the estimates plotted in Figure 7.

```
data pred;
 do latitude = 25 to 47 by 1;
  do longitude = 68 to 124 by 1;
   output;
  end;
 end;


PROC TPSPLINE data=so4;
  model so4 = (longitude, latitude)
         /lognlambda = 0.277
  score data=pred out=estimates1;
run;

PROC GAM data=so4;
  model so4=spline(longitude)
           spline(latitude)
  score data=pred out=estimate2;
run;
```

Both procedures score on a data set pred, which has grid points within the ranges of longitude and latitude. The plots of predictions for the two methods are shown in Figure 7.

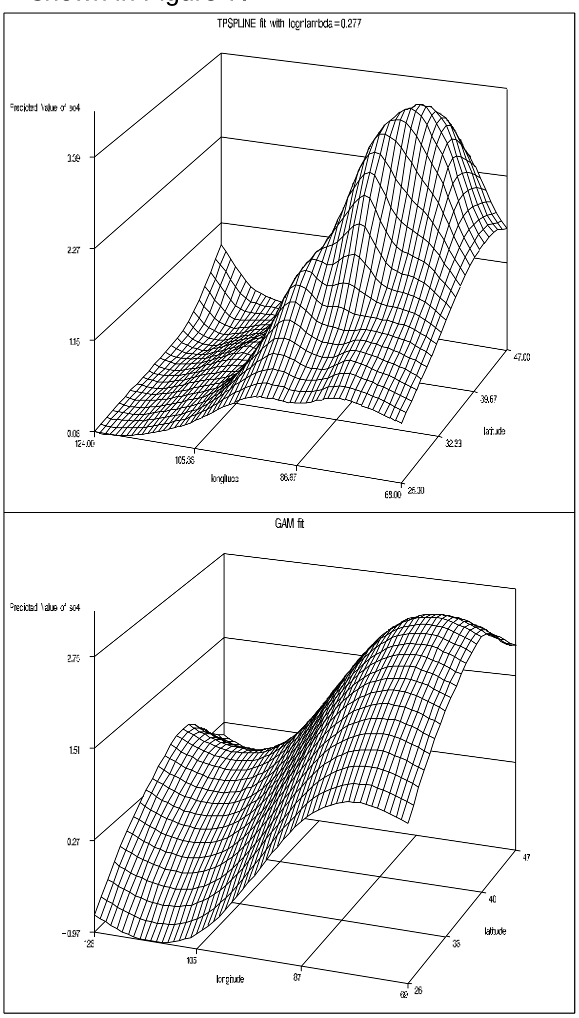

Figure 7. Comparing the GAM Fit with the TP-SPLINE Fit

From Figure 7, we can see that the TPSPLINE fit is more complicated than the GAM fit. It is flat in the western region and rises dramatically in the eastern region. Along the eastern coast, it has two maxima, which correspond to the New York and Atlanta regions. The GAM fit only shows that the eastern region has a higher SO4 concentration than the western region. However, the GAM procedure runs in about a quarter of the time required for the TPSPLINE procedure. In general, the algorithm for PROC GAM requires $O(n)$ operations while the algorithm for PROC TPSPLINE requires $O(n^3)$ operations. Thus GAM can be used on much larger datasets than TPSPLINE.

Conclusion

In this paper, we have discussed a new addition to SAS/STAT software. The GAM procedure is a very powerful tool in exploratory analysis when you have little prior information about the data or you want to find new features that parametric tools ignore. When combined with other parametric regression procedures, GAM can guide you in fitting parametric models. However, in some cases, it may produce less accurate results than other nonparametric regression procedures because of the additive assumption.

References

Bell, D., Walker, J., O'Connor, G., Orrel, J. and Tibshirani, R. (1989), "Spinal Deformation Following Multi-Level Thoracic and Lumbar Laminectomy in Children." Submitted for publication.

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

Friedman, J.H. and Stuetzle, W. (1981), "Projection Pursuit Regression," *Journal of the American Statistical Association*, 76, 817–823.

Hastie, T.J. and Tibshirani, R.J. (1990), *Generalized Additive Models*, New York: Chapman and Hall.

Stone, C.J. (1985), "Additive Regression and Other Nonparametric Models," *Annals of Statistics*, 13, 689–705.

Appendix

The two data sets used in this paper are listed in this section.

```
   data kyphosis;
      input Age StartVert NumVert Kyphosis @@;
      datalines;
   71  5  3  0     158 14 3 0    128 5   4 1
   2   1  5  0     1   15 4 0    1   16  2 0
   61  17 2  0     37  16 3 0    113 16  2 0
   59  12 6  1     82  14 5 1    148 16  3 0
   18  2  5  0     1   12 4 0    243 8   8 0
   168 18 3  0     1   16 3 0    78  15  6 0
   175 13 5  0     80  16 5 0    27  9   4 0
   22  16 2  0     105 5  6 1    96  12  3 1
   131 3  2  0     15  2  7 1    9   13  5 0
   12  2 14  1     8   6  3 0    100 14  3 0
   4   16 3  0     151 16 2 0    31  16  3 0
   125 11 2  0     130 13 5 0    112 16  3 0
   140 11 5  0     93  16 3 0    1   9   3 0
   52  6  5  1     20  9  6 0    91  12  5 1
   73  1  5  1     35  13 3 0    143 3   9 0
   61  1  4  0     97  16 3 0    139 10  3 1
   136 15 4  0     131 13 5 0    121 3   3 1
   177 14 2  0     68  10 5 0    9   17  2 0
   139 6  10 1     2   17 2 0    140 15  4 0
   72  15 5  0     2   13 3 0    120 8   5 1
   51  9  7  0     102 13 3 0    130 1   4 1
   114 8  7  1     81  1  4 0    118 16  3 0
   118 16 4  0     17  10 4 0    195 17  2 0
   159 13 4  0     18  11 4 0    15  16  5 0
   158 15 4  0     127 12 4 0    87  16  4 0
   206 10 4  0     11  15 3 0    178 15  4 0
   157 13 3  1     26  13 7 0    120 13  2 0
   42  6  7  1     36  13 4 0
   ;
```

```
  data so4;
    input latitude longitude so4 @@;
    dataline;
32.45833  87.24222 1.403 34.28778  85.96889 2.103
33.07139 109.86472 0.299 36.07167 112.15500 0.304
31.95056 112.80000 0.263 33.60500  92.09722 1.950
34.17944  93.09861 2.168 36.08389  92.58694 1.578
36.10056  94.17333 1.708 39.00472 123.08472 0.096
36.56694 118.77722 0.259 41.76583 122.47833 0.065
34.80611 119.01139 0.053 38.53528 121.77500 0.135
37.44139 105.86528 0.247 38.11778 103.31611 0.326
39.99389 105.48000 0.687 39.40306 107.34111 0.225
39.42722 107.37972 0.339 37.19806 108.49028 0.559
40.50750 107.70194 0.250 40.36417 105.58194 0.307
40.53472 106.78000 0.564 37.75139 107.68528 0.557
39.10111 105.09194 0.371 40.80639 104.75472 0.286
29.97472  82.19806 1.279 28.54278  80.64444 1.564
25.39000  80.68000 0.912 30.54806  84.60083 1.243
27.38000  82.28389 0.991 32.14111  81.97139 1.225
33.17778  84.40611 1.580 31.47306  83.53306 0.851
43.46139 113.55472 0.180 43.20528 116.74917 0.103
44.29778 116.06361 0.161 40.05333  88.37194 2.940
41.84139  88.85111 2.090 41.70111  87.99528 3.171
37.71000  89.26889 2.523 38.71000  88.74917 2.317
37.43556  88.67194 3.077 40.93333  90.72306 2.006
40.84000  85.46389 2.725 38.74083  87.48556 3.158
41.63250  87.08778 3.443 40.47528  86.99222 2.775
42.90972  91.47000 1.154 40.96306  93.39250 1.423
37.65111  94.80361 1.863 39.10222  96.60917 0.898
38.67167 100.91639 0.536 37.70472  85.04889 2.693
37.07778  82.99361 2.195 38.11833  83.54694 2.762
36.79056  88.06722 2.377 29.92972  91.71528 1.276
30.81139  90.18083 1.393 44.37389  68.26056 2.268
46.86889  68.01472 1.551 44.10750  70.72889 1.631
45.48917  69.66528 1.369 39.40889  76.99528 2.535
38.91306  76.15250 2.477 41.97583  70.02472 1.619
42.39250  72.34444 2.156 42.38389  71.21472 2.417
45.56083  84.67833 1.701 46.37417  84.74139 1.539
47.10472  88.55139 1.048 42.41028  85.39278 3.107
44.22389  85.81806 2.258 47.53111  93.46861 0.550
47.94639  91.49611 0.563 46.24944  94.49722 0.591
44.23722  95.30056 0.604 32.30667  90.31833 1.614
32.33472  89.16583 1.135 34.00250  89.80000 1.503
38.75361  92.19889 1.814 36.91083  90.31861 2.435
45.56861 107.43750 0.217 48.51028 113.99583 0.387
48.49917 109.79750 0.100 46.48500 112.06472 0.209
41.15306  96.49278 0.743 41.05917 100.74639 0.391
36.13583 115.42556 0.139 41.28528 115.85222 0.075
38.79917 119.25667 0.053 39.00500 114.21583 0.273
43.94306  71.70333 2.391 40.31500  74.85472 2.593
33.22028 108.23472 0.377 35.78167 106.26750 0.315
32.90944 105.47056 0.355 36.04083 106.97139 0.376
36.77889 103.98139 0.326 42.73389  76.65972 3.249
42.29944  79.39639 3.344 43.97306  74.22306 2.322
44.39333  73.85944 2.111 41.35083  74.04861 3.306
43.52611  75.94722 3.948 42.10639  77.53583 2.231
41.99361  74.50361 3.022 36.13250  77.17139 1.857
35.06056  83.43056 2.393 35.69694  80.62250 2.082
35.02583  78.27833 1.729 34.97083  79.52833 1.959
35.72833  78.68028 1.780 47.60139 103.26417 0.354
48.78250  97.75417 0.306 47.12556  99.23694 0.273
39.53139  84.72417 3.828 40.35528  83.06611 3.401
39.79278  81.53111 3.961 40.78222  81.92000 3.349
36.80528  98.20056 0.603 34.98000  97.52139 0.994
36.59083 101.61750 0.444 44.38694 123.62306 0.629
44.63472 123.19000 0.329 45.44917 122.15333 0.716
43.12167 121.05778 0.050 44.21333 122.25333 0.423
43.89944 117.42694 0.071 45.22444 118.51139 0.109
40.78833  77.94583 3.275 41.59778  78.76750 4.336
40.65750  77.93972 3.352 41.32750  74.82028 3.081
33.53944  80.43500 1.456 44.35500  98.29083 0.372
43.94917 101.85833 0.224 35.96139  84.28722 3.579
35.18250  87.19639 2.148 35.66444  83.59028 2.474
35.46778  89.15861 1.811 33.95778 102.77611 0.376
28.46667  97.70694 0.886 29.66139  96.25944 0.934
30.26139 100.55500 0.938 32.37861  94.71167 2.229
31.56056  94.86083 1.472 33.27333  99.21528 0.890
33.39167  97.63972 1.585 37.61861 112.17278 0.237
41.65833 111.89694 0.271 38.99833 110.16528 0.143
41.35750 111.04861 0.172 42.87611  73.16333 2.412
44.52833  72.86889 2.549 38.04056  78.54306 2.478
37.33139  80.55750 1.650 38.52250  78.43583 2.360
47.86000 123.93194 1.144 48.54056 121.44528 0.837
46.83528 122.28667 0.635 46.76056 117.18472 0.255
37.98000  80.95000 2.396 39.08972  79.66222 3.291
45.79639  88.39944 1.054 45.05333  88.37278 1.457
44.66444  89.65222 1.044 43.70194  90.56861 1.309
46.05278  89.65306 1.132 42.57917  88.50056 1.809
45.82278  91.87444 0.984 41.34028 106.19083 0.335
42.73389 108.85000 0.236 42.49472 108.82917 0.313
42.92889 109.78667 0.182 43.22278 109.99111 0.161
43.87333 104.19222 0.306 44.91722 110.42028 0.210
45.07611  72.67556 2.646
;
```

Contact Information

Dong Xiang, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513. Phone (919) 531-4854, FAX (919)

677-4444, Email dong.xiang@sas.com.