A NON-LINEAR STATISTICAL MODEL FOR PREDICTING SHORT-RANGE TEMPERATURE

A THESIS

Presented to

The Faculty of the Graduate Division

by

Ponnattu Kurian George

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in the School of Industrial Engineering

Georgia Institute of Technology

December, 1967

A NON-LINEAR STATISTICAL MODEL FOR

PREDICTING SHORT RANGE TEMPERATURE

Approved:

Chairman

Date approved by Chairman: Nov. 2, 1967

ACKNOWLEDGMENTS

The author wishes to express his deep sense of gratitude to Dr. Albert F. Hanken for suggesting the present research problem and his patient guidance through almost the completion of this work, in his capacity as thesis advisor.

Thanks are also due to Dr. David E. Fyffe, who served as chairman of the Thesis Reading Committee in the absence of Dr. Hanken, and Dr. Lynwood A. Johnson and Dr. William W. Hines, also of the reading committee.

TABLE OF CONTENTS

LIST OF TABLES

SUMMARY

An attempt is made to make use of a non-linear approach for predicting 24-hour ground temperature. The model used does not make use of any specific non-linear formula. Instead, the technique of Piecewise Linear Regression is used. In other words, the space G of all sample points is divided into groups G_i and the regression function in each group is approximated by a collection of linear relationships.

First, the original piecewise linear regression program was translated from Scatran to Algol. To investigate the efficiency of this technique as a tool for non-linear regression, the program was first run on a set of artificially selected data samples.

In selecting predictors for predicting temperature, two entirely different approaches were taken. These are here called "local" and "global" variables. In the former, the predictors are values of a selected set of weather elements at the station under consideration, and in the latter, the predictors are the temperatures at a few surrounding stations. To make predictions using the above model and compare with the linear predictions, a program was developed, which also calculates the error mean square.

The results of the analysis of the weather data do not show a significant reduction in residual variance. It is concluded that temperature does not exhibit enough non-linearity to warrant the use of a complicated non-linear model. However, a comparison of the results of the local and global variables shows that the latter

easily leads to better predictions. Also, it will always be advantageous to use a minimum number of predictors to check the reduction in degrees of freedom.

CHAPTER I

INTRODUCTION

Statistical forecasting, usually classified under the general field known as objective forecasting, is more or less equivalent to forecasting by the systematic use of empirical knowledge. An objective forecasting system may be defined as any method of deriving a forecast which does not depend for its accuracy upon the forecasting experience or the subjective judgment of the meteorologist using it. Strictly speaking, an objective system is one which can produce one and only one forecast from a specific set of data.

The goal of objective forecasting is simply to eliminate, as many as possible, the subjective elements which enter into the forecasting procedure. Objective forecasting is not so much concerned with the source of hypothetical relationships as it is with the practical value of the ideas and the extent to which they contribute to the accuracy of the forecasts. Hence, objective forecasting studies are characterized by the use of historical data to demonstrate the reliability of forecasting relationships, and the expression of the forecast itself in quantitative terms.

From the standpoint of discovering and understanding the relationships which hold in the atmosphere, forecasting investigations have been relatively ineffective because of their stress on lag relationships, and it seems clear that only a complete physical explanation of the atmosphere, together with complete observational data will make it possible to produce perfect weather forecasts (1). Practically, however, uncertainties exist which make the maximum attainable accuracy something less than perfection. The forecasting problem is thus, in essence, one of estimating what is likely to occur with any given state of the atmosphere and its environment. More precisely, the problem is to state the probability that any specified event will occur within any specified time interval.

Historical Background

One of the earliest forecast techniques, which did not involve the prognosis of the pressure pattern, was that developed by Russell (8) and described as an objective method of predicting precipitation. Schuster (8) in 1898 introduced harmonic analysis of the predictions within weather elements. His periodogram illustrated the systematic recurrence of a parameter, after fixed periods of time, and formed the basis of extrapolating into the future.

Besson (8) in 1905 introduced statistical graphical techniques to meteorology for the first time. One of the first attempts to break away from purely synoptic forecasting was by Rolf (8) in 1917, wherein he expressed the probability of rain as a function of selected meteorological parameters.

Statistical methods applied to meteorology lay practically dormant until the mid 1940's. Schuman (10) in 1944 presented a detailed explanation of the regression equation applied to meteorological forecasting. Since then, statistical forecasters have devoted a major portion of their efforts to the realm of determining mathematical formulas giving predicted numerical values of weather elements as functions of observable values of chosen predictors. Most of the formulas have been linear, and have been established by standard linear regression methods.

Statement of the Problem

Linear forecasting has been popular, both due to its simplicity and availability of a highly developed mathematical theory. However, there is no reason to assume that we have reached the ultimate plateau of predictability and that there is very little to be done. The problem under study, therefore, can be generally stated as that of investigating the use of non-linear methods for purposes of predicting values of weather elements or combinations of weather elements.

Objectives

The principal objective of the present research is to investigate the feasibility of using a non-linear model for predicting maximum temperature 24 hours in advance. However, the method employed does not make use of any specific non-linear formula. Instead, the technique of "Piecewise Linear Regression" developed by Hanken (5), as elaborated in Chapter III, is used. In essence, this consists of breaking down the data space into several groups and performing linear regression in each of these groups, instead of linear regression on the whole data space. The linear regression equation corresponding to the whole data space is here called linear model, and the collection of linear regression equations corresponding to all the groups is called non-linear model.

The original piecewise linear regression program written by Hanken in Scatran is translated into Algol. Initially, to test the applicability of this technique as a tool for non-linear regression, analysis is made of a set of arbitrary data generated by a non-linear surface.

The efficacy of this non-linear model in predicting temperature is tested by comparing the non-linear and linear predicted values against the actual values. In other words, the error mean squares in either cases are compared.

Hand calculation of the above values for several sets of data, especially of locating the points under consideration in its proper group, could get highly involved and tedious. In order to eliminate this difficulty, another program written by the author is used for calculating various values on the computer.

Scope and Limitations

The present research does not delve into the actual screening of predictors for predicting 24-hour maximum temperature. However, a rough screening of a number of arbitrarily selected predictors is done by using the partial correlation coefficient as a criterion. For this purpose, initially, a standard multiple linear regression program is run on the data. Obviously, this does not give an exact partial correlation coefficient, corresponding to the variable under consideration, when the actual non-linear model is taken into account. For instance, there could be a strong non-linear correlation between two variables, and yet the linear correlation coefficient could be near zero. However, this could serve as a mere guideline for cutting down the large volume of data.

The suggested model is not meant to eliminate, at least, some of the inadequacies of the linear model, namely, the use of less important and insufficient number of parameters and the inaccuracies in the actual data.

CHAPTER II

LITERATURE SURVEY

If we regard the atmosphere as a dynamic model, it soon becomes apparent that the whole process behaves as a complicated mechanism in which past and present values do not determine the future, as in most linear processes, but they themselves have an effect upon the system, which in turn produces non-linearity of a very peculiar nature.

The Forecasting Problem

Basically, the meteorological prediction problem boils down to the following: Given an initial state of the atmosphere with all its ramifications, details and complexities at some particular time, determine what its state would be at some subsequent time. In principle, the problem lends itself to solution, provided the following conditions are satisfied:

(1) The initial state can be specified precisely and in detail.

(2) The physical process by which one state is transformed into another is thoroughly understood.

(3) The process can be cast in mathematical form.

(4) The mathematical representation of the physical process can be dealt with numerically.

In practice, none of these requirements can be rigorously satisfied. However, encouraging progress has been made in recent years in dealing quantitatively with meteorological prediction. This progress has been along two converging lines, namely, the dynamical approach and the statistical approach. Oftentimes one thinks of these two methods as conflicting programs, but actually they are trying to get at one and the same thing. In the dynamical approach, the laws connecting various meteorological phenomena are investigated. These laws are considered to be precise in action, even though the data are subject to fluctuations which are random and are thus necessarily incomplete. On the other hand, in the statistical approach, the quantities are taken as they are found and the distributions examined both singly and in such combinations as one chooses for purposes of investigation. In an ideal survey all possible statistical parameters would be considered and not merely a partial selection, but such an undertaking would be impractical because of the sheer magnitude of the task.

It is clearly beyond the scope of this thesis to go into the details of the dynamical approach.

Time Series Analysis

The meteorologist is primarily interested in the behavior of his phenomenon as a function of time, and all his observations are taken either continuously in terms of this variable or are taken at discrete points in time. Each one of the series, therefore, which are considered in the problem of prediction of single meteorological elements or combinations of them can be looked at from the time-series point of view. The analysis of stationary time series, of late, has been highly aided by generalized harmonic analysis. In the case of stationary time series, the actual dynamics which motivates the series itself is assumed constant from one period of time to the next.

Considering the characteristics of these time series, we note that the sequence of observations has certain inherent dynamic properties as well as a superimposed random component. In the case of weather phenomena, the thing that we observe contains the original signal plus a random noise, and the problem is to locate the signal. In other words, if we know the past performance of some meteorological phenomenon α from some period of time in the past up to the present time t_o, we want to determine the predictable part of this sequence. The predictable part represents the dynamical component, and it is this component which permits one to extrapolate into the future to give a prediction. Naturally, the further one extends this prediction the worse the agreement is between the predicted and actual values, if there exists a random component.

Unfortunately, the basic arguments and ideas are true only if the process or processes are stationary. This, of course, cannot be true in the case of meteorological phenomena, since the basic movements of the weather systems certainly are different, at least from season to season, and this seriously hampers progress in this direction (3).

Analogue Method

A great deal of thought has been given by many meteorologists to the use of analogues as a method of forecasting. This is based on the premise that if the same weather situation should be observed twice, either it will develop the same way on both occasions, so that the analogue method will give the correct forecast, or else it will develop differently on the two occasions, and no method of forecast will give the correct forecast. The basic argument against the use of analogues is that the weather never follows exactly the same pattern. We can try to get around this difficulty by being very liberal in deciding when two situations are analogues, but then we may find that analogues are not enough alike to behave similarly.

Regression Methods

Usually, the statistical approach involves the use of mathematical statistics to establish a functional relationship between a set of parameters characterizing the initial state of the atmosphere and the probability distribution of a weather element or combinations of weather elements in some subsequent state. The immediate goal of this is to determine relationships which will be valid for independent data and will minimize the uncertainty in the probability distributions. The ultimate goal is to put these functional relationships in a form amenable to physical interpretation and thereby shed some light on the nature of the physical processes which determine the successive states of the atmosphere.

Surprisingly little has been done to obtain realistic probability distributions using statistical methods. The most commonly used procedure is to apply regression methods relating a continuous predictand as a linear function of a set of continuous predictors, and to assume the errors to be normally distributed about the predicted value.

Recently, two statistical techniques which separate the information in a set of data into independent components were adopted for application to meteorological prediction problems and were programmed for use with electronic computers. The two techniques are stepwise regression, also called screening, and factor analysis, which employs empirical orthogonal functions (2). These techniques express the predictand as a linear combination of a set of predictors. It is claimed that they are superior to ordinary multiple linear regression, since they make use of a large set of predictors and express the information which is pertinent to the predictand by a smaller set of predictors.

Factor Analysis Method*

> In the factor analysis method, a matrix of normalized predictor values with σ columns corresponding to σ predictors and τ rows corresponding to τ observations in time, generates a coefficient matrix with σ columns and τ rows, where the column vectors are normalized and are mutually orthogonal. The columns of the coefficient matrix are used as new predictors where each new predictor contains only information which is independent of the other predictors and where the total set of new predictors contains all the information in the original set of predictors. Each normalized predictand is expressed as a linear combination of the new predictors, and a linear operator for each predictand is computed using the least squares principle. The contribution of each new predictor toward the prediction of the predictand is revealed by the square of the coefficient corresponding to that predictor in the linear operator. The square of the coefficient represents the per cent reduction of predictand variance attributed to the predictor. Those predictors which do not reduce the predictand variance by more than a predetermined minimum value are withdrawn from the prediction equation. The procedure usually greatly reduces the effective number of predictors appearing in the prediction equation and is the major justification for using the factor analysis method. To facilitate day-to-day use of the prediction equation, the equation obtained above is converted back in terms of the σ original predictors, where the coefficients in the final linear operator are based on the reduced number orthogonal predictors.

---

*Extracted from Aubert, p. 438. For details of the factor analysis method see Aubert, pp. 438-442.

Screening Procedure

The screening procedure, originally due to Miller (7), is an objective method of selecting from a set of those predictors which contribute most significantly to the forecast of a predictand. The forecast equation is a linear regression equation which contains only the selected predictors.

The procedure screens a large set of normalized potential predictors and chooses as the first predictor that one which is most highly correlated with the predictand. The square of the correlation coefficient represents the per cent reduction in predictand variance due to the first predictor and the correlation coefficient itself is the coefficient of the predictor in the linear-forecast operator. The remaining potential predictors are orthogonalized with respect to the first predictor and are screened again in their new form. The orthogonalized predictor which is most highly correlated with the predictand is chosen as the second predictor, and again the square of the correlation coefficient represents the per cent reduction in predictand variance due to the second predictor while the correlation coefficient appears as the second coefficient in the linear-forecast operator. The remaining predictors are then orthogonalized with respect to the second predictor which automatically makes them orthogonal to the first, and the procedure is repeated to obtain the third predictor and its coefficient. By repeatedly orthogonalizing the remaining predictors to the preceding predictor and screening them for the best correlation, a forecast equation is generated by increments, which has the properties that each predictor is independent of all the other predictors and that the contribution of each predictor to the reduction of predictand variance is given by the square of its coefficient, the predictors appearing in order of their importance. An F-test is used as a guide to stop the screening when the contribution of a predictor is judged to be not significant. To facilitate use of the forecast equation, the orthogonalized predictors are converted back in terms of the original unnormalized predictors.

Non-linear Regression

It is likely that linear forecasting procedures are popular since they are much simpler than non-linear procedures, and their mathematical theory is more highly developed. Occasionally, a joint frequency distribution indicates that the relation between the variates is definitely not linear. In such cases the linear correlation coefficient will be lower than the degree of relationship one might expect.

Panofsky and Brier (9) suggest two ways in which a valid relationship between the variates could be discovered. According to them, either a general form of the relationship can be assumed and the coefficients determined by least squares, or a stepped function could be fitted on the data. Although these methods are time consuming, they can be used with advantage when there is a theoretical reason to assume a certain form of relationship between the variates. However, it needs to be mentioned that it becomes highly difficult to visualize any existing relationship when there are more than two variates.

The establishment of non-linear formulas of specified analytical form for purposes of meteorological prediction has met with little success. Attempts to express predictands as quadratic functions of predictors have led to formulas which have equaled the linear formulas, but have not surpassed them.

CHAPTER III

THE NON-LINEAR MODEL

Piecewise Linear Regression

The theory of piecewise linear regression is based on the premise that a complicated relationship between a set of independent variables and the corresponding dependent variable can be considered as a collection of several linear relationships. In other words, the space G of all sample points is divided into groups G_i, and the regression function in each group may be approximated by a linear relationship. Thus the residual variance could be reduced by making the group as small as possible.

To carry out this method an algorithm is needed which partitions the set S into subsets S_i. This should satisfy a number of practical requirements. First, the algorithm should be an automatic procedure capable of digital computer operation. Secondly, the number of points in each group should be decided by the situation under consideration. However, there should be enough points in each group to perform a linear regression with reasonable accuracy, but the total number of points would be limited to avoid problems of bad fit. Finally, the boundary between two adjacent subsets should be clearly delineated by a cutting plane, which separates two adjacent sets. The actual theory of piecewise linear regression will not be discussed here.*

The Piecewise Linear Regression Program written by Hanken (5) is based on an algorithm which will satisfy the above requirements. However, since the program is in Scatran, it cannot be used in most of the present day computers. Hence, it was first translated into Algol after removing the time lag incorporated in it, which was not needed for the present research. (See Appendix A.)

Analysis of Arbitrary Data

In order to test the applicability of the technique of piecewise linear regression as a tool for non-linear analysis, the program was first run on a set of arbitrary data generated by a non-linear surface.

To generate the arbitrary data sample the non-linear equation

$$y = x_1^2 + x_2^2 + x_3^2$$

was chosen, where x_1, x_2 and x_3 are three independent variables and y, the dependent variable. Now x_1, x_2 and x_3 were assigned random values (x_1 and x_2, between 0 and 10 and x_3, between 0 and 5) and the corresponding values of y were calculated. In order to achieve this for a data sample of 1000, an Algol program was written incorporating a standard random generator procedure (see Appendix B) and the data were generated by the computer.

*For a complete development of the theory of piecewise linear regression, see Hanken, pp. 53-94.

The translated piecewise linear regression program was then run on the data. A set of 50 data points each, at the beginning and end, was excluded from the actual analysis. The maximum number of data points in any group was limited to 49. The results of the analysis are displayed in Table 1. (See Appendix D for actual computer output.)

In Table 1, level 1 corresponds to linear regression on the whole data space. This is then cut into groups 2 and 3 using the λ_{max} criterion,* and regression performed in each of these groups. These are now cut into groups 4 and 5 and groups 6 and 7, and regression performed again. Generalizing, the mth group is cut into groups (m x 2) and (m x 2)+1. Cutting is stopped when either the number of data points in all the groups is less than 50 or the group number exceeds 128. The level number thus indicates each successive stage of cutting.

The degrees of freedom are calculated from the formula (5),

$$d.f. = N - m(k + 1),$$

where N = the total number of data points on which regression is performed,

m = the number of groups in the level under consideration, and

k = the number of independent variables or predictors.

The residual sum of squares corresponds to the total of all the groups and, hence, also the residual variance.

The regression equation corresponding to level 1, namely,

---

*See Hanken, pp. 69-76.

Table 1. Piecewise Linear Regression on Arbitrary Data

SSY = 6,203,640.00

$\bar{y}$ = 75.42

Level	Number of Groups	d.f.	SS Residuals	Residual Variance
1	1	896	107,233.00	120.00
2	2	892	35,427.20	39.80
3	4	884	26,121.08	29.60
4	8	868	20,173.96	23.40
5	15	840	13,326.78	14.70
6	16	796	9,887.66	12.40
7	27	792	9,684.72	12.20

$$y = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3),$$

where y = the dependent variable or predictand,

x_1, x_2 and x_3 = the independent variables or predictors,

$\bar{y}$ = the mean of all the y's,

$\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$ = the means of x_1's, x_2's and x_3's, respectively,

and b_1, b_2 and b_3 = the respective regression coefficients,

is called the linear model, and the collection of all the equations of the same form corresponding to all the groups in the level having the minimum residual variance, along with the equations for the required number of cutting planes, which are of the form

$$e_1x_1 + e_2x_2 + e_3x_3 - CP = 0,$$

where e_1, e_2 and e_3 are the coefficients of the respective x's and CP, the direction coefficient, constitutes the non-linear model.

All the information necessary for evaluating the models, namely, $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, b_1, b_2, b_3, CP, e_1, e_2 and e_3, are available in the computer output.* (See Appendix D.)

Analysis of Weather Data

For the purpose of the present research, maximum temperature in Atlanta, 24 hours in advance, was taken as the predictand, since data pertaining to this were available in a convenient form (11). As far as

*Evaluation of the models and calculation of the predictands for given values of predictors will be discussed later in the chapter.

the predictors were concerned, two entirely different approaches were taken.

Local Variables

In this, the maximum temperature 24 hours in advance was considered as a function of local variables, namely, maximum temperatures of the day, average cloud cover of the day, average dew point of the day, average pressure of the day, average east-west component of the resultant wind and the average difference in pressure between the day and the previous day. A total of 880 data points were considered, starting from January 1, 1965 to June 30, 1967. Out of these, 40 data points, each at the beginning and end of the time period, were set apart, without being included in the actual analysis. This was with a view to making predictions using the values of these predictors and comparing with the actual values of the predictand later.

Actually the variables mentioned above were selected after running the standard multiple regression program of the Burroughs Corporation on a set of data with a larger number of predictors. These variables have a partial correlation coefficient equal to or greater than 0.1. The results of the piecewise linear regression are displayed in Table 2.

The linear model is given by

$$y = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3) + b_4(x_4 - \bar{x}_4)$$

$$+ b_5(x_5 - \bar{x}_5) + b_6(x_6 - \bar{x}_6),$$

Table 2. Piecewise Linear Regression on Weather Data, Local Variables

SSY = 4,026,420.00

$\bar{y}$ = 69.35

Level	Number of Groups	d.f.	SS Residuals	Residual Variance
1	1	793	30,680.80	38.60
2	2	786	30,078.00	38.20
3	4	772	29,116.48	37.70
4	8	744	28,101.25	37.70
5	15	695	26,113.99	37.60
6	22	646	25,090.60	38.80
7	25	625	23,448.17	37.40

where y = 24-hour maximum temperature in Atlanta,

x_1 = maximum temperature of the day,

x_2 = average cloud cover of the day in tenths,

x_3 = average dew point of the day,

x_4 = average pressure of the day,

x_5 = average east-west component of the resultant wind, and

x_6 = average difference in pressure between the day and the previous day.

The non-linear model is a collection of equations of the same form corresponding to all the groups in the level having the minimum residual variance with the equations for the required number cutting planes.

Global Variables

In this, the maximum temperature in Atlanta 24 hours in advance was considered as a function of maximum temperature in Atlanta and a few stations around Atlanta within a radius of about 300 miles, namely, Raleigh, North Carolina, Mobile, Alabama and Nashville, Tennessee, and also the average east-west component of the resultant wind in Atlanta. Here also the screening of the variables was done as in the previous case and only variables of better correlation were used for the piece-wise linear regression. The results of the analysis are displayed in Table 3.

The linear model is given by

$$y = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3) + b_4(x_4 - \bar{x}_4) + b_5(x_5 - \bar{x}_5),$$

Table 3. Piecewise Linear Regression on Weather Data, Global Variations

SSY = 4,028,080.00

$\bar{y}$ = 69.35

Level	Number of Groups	d.f.	SS Residuals	Residual Variance
1	1	794	25,992.40	32.70
2	2	788	24,319.65	30.80
3	4	776	23,859.73	30.70
4	7	758	22,639.89	29.80
5	11	734	21,305.77	29.00
6	19	686	21,252.45	31.00
7	21	674	19,221.97	28.60

where y = 24-hour maximum temperature in Atlanta,

x_1 = maximum temperature of the day in Atlanta,

x_2 = average east-west component of the resultant wind in Atlanta,

x_3 = maximum temperature of the day in Raleigh, North Carolina,

x_4 = maximum temperature of the day in Mobile, Alabama, and

x_5 = maximum temperature of the day in Nashville, Tennessee,

and the collection of all the equations of the same form corresponding to all the groups in the level having the minimum residual variance, along with the equation for the required number of cutting planes, constitutes the non-linear model.

Comparison of Linear and Non-linear Predictions

As mentioned earlier, in each of the above two runs on weather data, 40 data points each were left out at the beginning and end of the time period under consideration, without being included in the actual analysis. Predictions were made using the values of these predictors with both linear and non-linear models and compared with the actual values of the corresponding maximum temperature. Also the square of the residuals for each prediction and the error mean squares were calculated. The results are displayed in Tables 4 and 5. Although the calculations were done on 80 sets of data in either case, only 20 sets are shown in the above tables. However, the mean square errors correspond to the complete set.

In order to do the above calculation an Algol program was written (see Appendix C) which could calculate the values for both the

Table 4. Comparison of Linear and Non-linear Predictions, Local Variables

Linear Error Mean Square = 58.35
Non-linear Error Mean Square = 78.75

Maximum Temperature	Cloud Cover in Tenths	Dew Point	Pressure	E-W Component of Wind	24-Hour Pressure Diff.	Actual 24-Hour Temp.	Linear Prediction	Non-linear Prediction
62.00	9.00	50.00	29.03	-5.21	-0.23	55.00	61.78	61.41
55.00	10.00	32.00	29.11	-7.20	0.08	51.00	55.47	55.73
51.00	3.00	19.00	29.29	4.79	0.18	59.00	55.83	53.49
59.00	7.00	27.00	29.21	-2.91	-0.08	63.00	58.82	57.17
63.00	9.00	41.00	29.09	-4.89	0.12	58.00	63.64	62.83
58.00	10.00	45.00	29.14	4.89	0.05	67.00	61.32	55.70
68.00	8.00	54.00	29.03	-1.32	-0.08	60.00	68.58	67.18
60.00	7.00	45.00	28.91	-8.76	-0.12	43.00	59.30	57.21
43.00	10.00	31.00	29.06	-2.95	0.15	49.00	46.83	45.16
49.00	0.00	30.00	29.11	2.91	0.05	55.00	54.70	58.95
76.00	10.00	59.00	29.01	7.14	-0.02	59.00	75.80	74.82
59.00	10.00	54.00	28.80	13.44	-0.21	56.00	59.91	63.00
56.00	10.00	49.00	28.85	11.00	0.05	75.00	58.60	62.27
75.00	0.00	50.00	28.95	2.02	0.10	84.00	76.34	74.00
84.00	0.00	54.00	28.99	-5.21	0.04	87.00	82.98	83.07
87.00	0.00	58.00	28.98	-7.27	-0.01	88.00	85.31	82.49
88.00	0.00	61.00	29.00	-6.04	0.02	89.00	87.08	87.05
89.00	5.00	63.00	28.95	-4.83	-0.05	88.00	85.82	85.44
88.00	5.00	62.00	28.81	-3.25	-0.14	85.00	83.26	84.17
85.00	2.00	62.00	28.77	-2.96	-0.04	81.00	82.29	86.28

Table 5. Comparison of Linear and Non-linear Predictions, Global Variables

Linear Error Mean Square = 48.28
Non-linear Error Mean Square = 151.09

Max. Temp. Atlanta, Ga.	East-West Component of Wind	Max. Temp. Raleigh, N. C.	Max. Temp. Mobile, Ala.	Max. Temp. Nashville, Tenn.	Actual 24-Hour Temp.	Linear Prediction	Non-linear Prediction
60.00	6.60	44.00	73.00	67.00	62.00	64.03	68.98
62.00	-5.21	63.00	75.00	63.00	55.00	64.57	65.87
55.00	-7.20	60.00	65.00	44.00	51.00	53.12	47.91
51.00	4.79	53.00	59.00	50.00	59.00	52.46	52.48
59.00	-2.91	60.00	70.00	59.00	63.00	61.02	60.67
63.00	-4.89	53.00	72.00	66.00	58.00	65.50	63.15
58.00	4.89	56.00	73.00	67.00	67.00	63.79	77.05
67.00	0.85	69.00	73.00	68.00	68.00	67.58	67.11
68.00	-1.32	72.00	75.00	65.00	60.00	67.14	65.89
60.00	-8.76	60.00	66.00	36.00	43.00	51.58	46.59
76.00	7.14	68.00	81.00	63.00	59.00	68.87	62.68
59.00	13.44	56.00	69.00	72.00	56.00	64.95	77.05
56.00	11.00	63.00	74.00	76.00	75.00	66.74	64.42
75.00	2.02	69.00	83.00	82.00	84.00	77.25	79.66
84.00	-5.21	76.00	88.00	88.00	87.00	83.97	82.56
87.00	-7.27	80.00	91.00	89.00	88.00	85.99	83.85
88.00	-6.04	86.00	93.00	89.00	89.00	86.60	85.33
89.00	-4.83	92.00	93.00	87.00	88.00	86.09	84.97
88.00	-3.25	90.00	90.00	88.00	85.00	85.62	84.54
85.00	-2.96	67.00	91.00	84.00	81.00	82.66	74.57

linear and non-linear cases. In the latter case, the program incorporates a method of locating the group to which any particular data point belongs.

This is done by substituting the values of the predictors in the equation for the cutting plane between two adjacent groups. For example, let

$$e_1x_1 + e_2x_2 + e_3x_3 + \cdot \cdot \cdot \cdot - CP = 0$$

represent any cutting plane, where e_1, e_2, e_3, etc., represent the coefficients of the independent variables x_1, x_2, x_3, etc., and CP the direction coefficient. Substituting values for x_1, x_2, x_3, etc., in the first cutting plane will give a result which is either less than zero or greater than or equal to zero. In the former case the data point under consideration will belong to the group (1 x 2) = 2, and in the latter case, to the group (1 x 2)+1 = 3. Generalizing this, in the case of the mth group, if the result is less than zero the point belongs to group (m x 2), else to group (m x 2)+1. Proceeding likewise, when a group with no further cutting is encountered the values of the predictors are substituted in the equation for the regression plane of the group and the value of the predictand is calculated.

CHAPTER IV

DISCUSSION OF RESULTS

Analysis of Arbitrary Data

The last column of Table 1 (page 17) shows the residual variance corresponding to the various levels of cutting of the arbitrary data discussed in Chapter III.

From SSY, the sum of squares of the dependent variable, $s^2(y)$, the square of the universe standard deviation can be calculated.

$$s^2(y) = \frac{SSY - N.\bar{y}^2}{N - 1} = 1209.27$$

where $\bar{y}$ represents the mean of the dependent variable and N, the total number of data points.

The residual sum of squares, and hence the residual variance, can be seen to be decreasing as the number of groups is increasing. However, the reduction in residual variance may not always stick to this pattern, as the reduction in degrees of freedom could more than offset the reduction in sum of squares.

The multiple correlation coefficient for the linear case can be calculated as

$$R_L = \sqrt{1 - \frac{120.00}{1209.20}} = 0.946$$

Considering the level with the minimum residual variance, the final level in this particular case, the non-linear correlation coefficient can be calculated as

$$R_{NL} = \sqrt{1 - \frac{12.20}{1209.20}} = 0.994$$

Analysis of Weather Data

In considering the applicability of the non-linear model for predicting temperature, the two different approaches used for selecting the predictors have got to be considered.

Local Variables

The last column of Table 2 (page 20) shows the residual variance corresponding to the various levels of cutting in the analysis using local variables.

The maximum number of groups in the final level shows up a reduction in residual variance of 1.2 from the linear case. It can be seen that the residual variance in level 5 is less than that in level 6. This is because, in this range the reduction in degrees of freedom is faster than the reduction in sum of squares.

After calculating the square of the universe standard deviation, the linear and non-linear correlation coefficients can be calculated as follows:

$$R_L = \sqrt{1 - \frac{38.60}{223.20}} = 0.910$$

$$R_{NL} = \sqrt{1 - \frac{37.40}{223.20}} = 0.913$$

Global Variables

The last column of Table 3 (page 22) shows the residual variance corresponding to the various levels of cutting in the analysis using the global variables.

Here also the residual variance is less in level 5 than in level 6. The reason is obviously the same as the previous case. It can be seen that there is a high reduction in residual variance (4.1) from linear to non-linear, and also the actual value of the residual variance is considerably less in the case of global variables.

The linear and non-linear correlation coefficients can be calculated as in the previous case.

$$R_L = \sqrt{1 - \frac{32.70}{225.20}} = 0.924$$

$$R_{NL} = \sqrt{1 - \frac{28.60}{225.50}} = 0.934$$

Comparison of Linear and Non-Linear Predictions

It can be seen from Tables 4 and 5 that the non-linear model does not offer better predictions than the linear model. In fact, the error mean square is considerably greater in the non-linear case.

Table 6. Comparison of Linear and Non-linear Predictions, Fewer Predictors

Linear Error Mean Square = 63.08
Non-linear Error Mean Square = 60.66

Maximum Temperature	Pressure	East-West Component of Wind	Actual 24-Hour Temperature	Linear Prediction	Non-linear Prediction
60.00	29.26	6.60	62.00	64.10	64.80
62.00	29.03	-5.21	55.00	62.55	57.60
55.00	29.11	-7.20	51.00	56.68	53.83
51.00	29.29	4.79	59.00	55.98	56.08
59.00	29.21	-2.91	63.00	61.69	59.54
63.00	29.09	-4.89	58.00	64.03	60.83
58.00	29.14	4.89	67.00	61.01	60.92
67.00	29.11	0.85	68.00	68.48	66.35
68.00	29.03	-1.32	60.00	68.44	67.24
60.00	28.91	-8.76	43.00	59.26	51.96
76.00	29.01	7.14	59.00	76.47	79.28
59.00	28.80	13.44	56.00	59.80	58.57
56.00	28.85	11.00	75.00	57.25	56.63
75.00	28.95	2.02	84.00	74.45	72.93
84.00	28.99	-5.21	87.00	82.21	81.23
87.00	28.98	-7.27	88.00	84.62	84.40
88.00	29.00	-6.04	89.00	85.85	86.69
89.00	28.95	-4.83	88.00	86.44	90.42
88.00	28.81	-3.25	85.00	84.44	84.19
85.00	28.77	-2.96	81.00	81.39	80.59

However, this does not entirely rule out the desirability of using non-linear models. Table 6 (page 30) shows the results of the predictions after non-linear analysis on the same temperature. Here, only three predictors are used, namely, temperature, pressure and east-west component of the resultant wind. The error mean square in the non-linear case is seen to be slightly less than that of the linear case. This is because, in the original weather analysis, the larger number of predictors used considerably reduce the degrees of freedom, and this reduction in degrees of freedom destroys the chances of making better predictions.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

The results of the analysis of the arbitrary data shows a reduction in residual variance from 120.00 in the linear case to 12.20 in the non-linear case. The correlation coefficient is correspondingly increased from 0.946 to 0.994.

It is obvious from the above that the technique of piecewise linear regression is clearly superior to linear regression when the data exhibit a reasonable amount of non-linearity.

The results of the analysis of the weather data (both the local and the global case) do not show any significant reduction in residual variance. Also the non-linear error mean squares are greater than the linear ones in both the original weather data analysis. This implies that temperature does not show enough non-linearity to warrant the use of a non-linear model. Also, it is not at all profitable to use predictors which do not have a high degree of correlation. In other words, it will be always advantageous to use a minimum number of predictors to check the reduction in degrees of freedom.

A comparison of Tables 4 and 5 shows that global variables lead to considerably better predictions and this information can be made use of while using linear or non-linear models.

Further research on weather elements like pressure, precipitation, etc., could possibly prove the profitability of using piecewise linear regression. The time differential used in the present models, namely, one day or 24 hours, is a factor which could seriously affect the predictions. Hence, for better predictions, it is advisable to use a time differential of 6 hours.

It is recommended that the program for comparing the results of the prediction using the linear and non-linear methods be combined with the original program. This will facilitate handling of a lesser volume of data during further research in these lines.

APPENDICES

APPENDIX A

ALGOL PROGRAM FOR NON-LINEAR ANALYSIS

```
        BEGIN
COMMENT      NON LINEAR ANALYSIS PROGRAM TRANSLATED FROM SCATRAN
             P K GEORGE;
INTEGER      K,NM,NL,ML,IM,IL;
             K←3;
             NM←1000;
             IM←50;
             IL←50;
             NL←50;
             ML←128;
        BEGIN
REAL         SSY,YM,SES,P,EQ,CP,TL,MP,DET;
INTEGER      M,I,L,J,H;
REAL ARRAY   X[0:K+2,0:NM],IX[0:NM],T[0:K],XS[0:K],R[0:K,0:K],
             S[0:K,0:K],B[0:K,0:1],XM[0:K],EXS[0:K],A[0:K,0:K],
             C[0:K,0:K],E[0:K,0:1],F[0:K,0:1],G[0:K],N[0:2×ML+1];
LABEL        START,L20,P1,P2,P3,P4,L12,P5,L16,P6,L5,P7,L19;
FILE IN      PKGIN(2,10);
FILE OUT     PKGOUT 6(2,15);
FORMAT       F1(4(X4,F8.3)),
             F2(//,"SSY=",E11.5),
             F3(//,"GROUP",X2,"NO",X1,"SIGMA RES",X6,"MEANS",X3,
             "Y",X7,"X1",X7,"X2",X7,"X3",//,X41,"B1",X7,"B2",X7,"B3",
             //,X32,"CP",X7,"E1",X7,"E2",X7,"E3"),
             F4(//,X1,I3,X1,I4,X1,E11.5,X4,4F9.2,/,X34,
             3F9.2,/,X25,4F9.2),
             F5(//,X1,I3,X1,I4,X1,E11.5,X4,4F9.2,/,X34,3F9.2);
LIST         LST1(FOR H←1 STEP 1 UNTIL NM DO
             FOR I←1 STEP 1 UNTIL K+1 DO X[I,H]),
             LST2(M,N[M],SES,YM,FOR I←1 STEP 1 UNTIL K DO XM[I],
             FOR I←1 STEP 1 UNTIL K DO B[I,1],CP,FOR I←1 STEP 1
             UNTIL K DO E[I,1]),
```

```
          LST3(M,N[M],SES,YM,FOR I←1 STEP 1 UNTIL K DO XM[I],
          FOR I←1 STEP 1 UNTIL K DO B[I,1]);
$$ A A012                                                        00000000
                                                                 99999999
$$ A A014                                                        00000000
                                                                 99999999
START:    WRITE (PKGOUT [NO]);
          READ (PKGIN,F1,LST1);
          CLOSE (PKGIN,RELEASE);
          M←1;
          SSY←0.0;
          FOR L←IM+1 STEP 1 UNTIL NM-IL DO
          SSY←SSY+X[K+1,L]×X[K+1,L];
          WRITE (PKGOUT,F2,SSY);
          WRITE (PKGOUT,F3);
          N[1]←NM-(IM+IL);
          FOR L←IM+1 STEP 1 UNTIL NM-IL DO
          IX[L]←1;
L20:      FOR I←1 STEP 1 UNTIL K DO
      BEGIN
          XS[I]←0.0;
          T[I]←0.0;
          FOR J←I STEP 1 UNTIL K DO
          R[I,J]←0.0;
       END;
          YM←0.0;
          L←IM;
P1:       IF L=NM-IL THEN GO TO P2;
          L←L+1;
          IF IX[L]≠M THEN GO TO P1;
          YM←YM+X[K+1,L];
          FOR I←1 STEP 1 UNTIL K DO
      BEGIN
          XS[I]←XS[I]+X[I,L];
          T[I]←T[I]+X[I,L]×X[K+1,L];
          FOR J←I STEP 1 UNTIL K DO
          R[I,J]←R[I,J]+X[I,L]×X[J,L];
       END;
          GO TO P1;
P2:       YM←YM/N[M];
```

```
            FOR I←1 STEP 1 UNTIL K DO
     BEGIN
            B[I,1]←T[I]-XS[I]×YM;
            XM[I]←XS[I]/N[M];
            FOR J←I STEP 1 UNTIL K DO
     BEGIN
            S[I,J]←R[I,J]-(XS[I]×XS[J]/N[M]);
      END;
      END;
            FOR I←1 STEP 1 UNTIL K DO
           FOR J←1 STEP 1 UNTIL K DO
           S[J,I]←S[I,J];
           INVERT(K,S,DET,PKGOUT);
           MATPROD(K,K,1,S,B,B);
           IF N[M]<NL THEN GO TO L5;
           FOR I←1 STEP 1 UNTIL K DO
    BEGIN
           EXS[I]←0.0;
           FOR J←I STEP 1 UNTIL K DO
           A[I,J]←0.0;
     END;
           SES←0.0;
           L←IM;
P3:        IF L=NM-IL THEN GO TO P4;
           L←L+1;
           IF IX[L]≠M THEN GO TO P3;
           P←0.0;
           FOR I←1 STEP 1 UNTIL K DO
           P←P+B[I,1]×(X[I,L]-XM[I]);
           X[K+2,L]←(X[K+1,L]-YM-P)*2;
           SES←SES+X[K+2,L];
           FOR I←1 STEP 1 UNTIL K DO
    BEGIN
           EXS[I]←EXS[I]+X[K+2,L]×X[I,L];
           FOR J←I STEP 1 UNTIL K DO
           A[I,J]←A[I,J]+X[K+2,L]×X[I,L]×X[J,L];
     END;
           GO TO P3;
P4:        FOR I←1 STEP 1 UNTIL K DO
           FOR J←I STEP 1 UNTIL K DO
```

```
            A[I,J]←A[I,J]−EXS[I]×EXS[J]/SES;
            FOR I←1 STEP 1 UNTIL K DO
            FOR J←1 STEP 1 UNTIL K DO
            A[J,I]←A[I,J];
            MATPROD(K,K,K,S,A,C);
            FOR I←1 STEP 1 UNTIL K DO
            E[I,1]←1.0;
L12:        MATPROD(K,K,1,C,E,F);
            EQ←0.0;
            FOR I←1 STEP 1 UNTIL K DO
        BEGIN
            G[I]←F[I,1]/F[1,1];
            EQ←EQ+ABS(E[I,1]−G[I]);
            E[I,1]←G[I];
         END;
            IF EQ≥0.01 THEN GO TO L12;
            CP←0.0;
            FOR I←1 STEP 1 UNTIL K DO
            CP←CP+EXS[I]×E[I,1];
            CP←CP/SES;
            WRITE (PKGOUT,F4,LST2);
            IF N[M]<NL THEN GO TO P6;
            N[2×M]←0;
            N[2×M+1]←0;
            L←IM;
P5:         IF L=NM−IL THEN GO TO P6;
            TL←0.0;
            L←L+1;
            IF IX[L]≠M THEN GO TO P5;
            FOR I←1 STEP 1 UNTIL K DO
            TL←TL+E[I,1]×X[I,L];
            IF TL<CP THEN GO TO L16;
            IX[L]←2×M+1;
            N[2×M+1]←N[2×M+1]+1;
            GO TO P5;
L16:        IX[L]←2×M;
            N[2×M]←N[2×M]+1;
            GO TO P5;
L5:         SES←0.0;
            L←IM;
```

```
P7:          IF L=NM-IL THEN GO TO L19;
             L←L+1;
             IF IX[L]≠M THEN GO TO P7;
             P←0.0;
             FOR I←1 STEP 1 UNTIL K DO
             P←P+B[I,1]×(X[I,L]-XM[I]);
             X[K+2,L]←(X[K+1,L]-YM-P)*2;
             SES←SES+X[K+2,L];
             GO TO P7;
L19:         WRITE (PKGOUT,F5,LST3);
P6:          M←M+1;
             MP←M/2;
             MP←ENTIER(MP);
             IF N[MP]≥NL THEN GO TO L20;
             N[M]←0;
             IF M<ML THEN GO TO P6;
        END;
        END.
```

APPENDIX B

ALGOL PROGRAM FOR GENERATING ARBITRARY DATA

```
          BEGIN
COMMENT        ARBITRARY DATA FOR NON LINEAR ANALYSIS
               P K GEORGE;
REAL           D,TH,TL;
INTEGER        C,T,I;
REAL ARRAY     X[0:4,0:1000];
LABEL          START;
FILE IN        PKGIN(2,10);
FILE OUT       PKGOUT 0(2,15);
FORMAT         F1(4(X4,F8.3));
LIST           LST1(FOR T←1 STEP 1 UNTIL 1000 DO FOR I←1 STEP 1
               UNTIL 4 DO X[I,T]);
               REAL PROCEDURE R;
          BEGIN
               DOUBLE(C,0,D,0,×,←,TH,TL);
               DOUBLE(TH,TL,ENTIER(TH),0,-,←,D,TL);
               R←D;
           END;
START:         C←5497558138885;
               D←C/8*13;
               FOR T←1 STEP 1 UNTIL 1000 DO
          BEGIN
               X[1,T]←R×10;
               X[2,T]←R×10;
               X[3,T]←R×5;
               X[4,T]←X[1,T]*2+X[2,T]*2+X[3,T]*2;
           END;
               WRITE (PKGOUT,F1,LST1);
           END.
```

APPENDIX C

ALGOL PROGRAM FOR COMPARING LINEAR AND NON-LINEAR PREDICTIONS

```
         BEGIN
COMMENT           PROGRAM FOR COMPARING THE RESULTS OF MULTIPLE LINEAR
                  REGRESSION AND PIECEWISE LINEAR REGRESSION
                  P K GEORGE;
INTEGER           K,ILM,NG,NH;
                  K←6;
                  ILM←79;
                  NG←83;
                  NH←ENTIER(NG/2);
         BEGIN
REAL              YCP,MS;
INTEGER           I,J,M;
REAL ARRAY        YHAT[0:ILM],X[0:ILM,0:K+1],XM[0:NG,0:K+1],
                  B[0:NG,0:K],E[0:NH,0:K+1],RS[0:ILM];
INTEGER ARRAY     N[0:NG];
LABEL             START,L1,L2,L3,L4;
FILE IN           PKGIN (2,10);
FILE OUT          PKGOUT 6(2,15);
FORMAT            F1(7(X4,F6.2)),
                  F2(I1),
                  F3(7(X2,F9.2)),
                  F4(6(X2,F9.2)),
                  FR1(9(X2,F7.2)),
                  FR2(//,X10,"ERROR MEAN SQUARE  =",F10.3);
LIST              LST1(FOR I←1 STEP 1 UNTIL ILM DO
                  FOR J←1 STEP 1 UNTIL K+1 DO X[I,J]),
                  LST2(FOR I←1 STEP 1 UNTIL NG DO N[I]),
                  LST3(FOR I←1 STEP 1 UNTIL NG DO
                  FOR J←1 STEP 1 UNTIL K+1 DO XM[I,J]),
                  LST4(FOR I←1 STEP 1 UNTIL NG DO
                  FOR J←1 STEP 1 UNTIL K DO B[I,J]),
                  LST5(FOR I←1 STEP 1 UNTIL NH DO
```

```
               FOR J←1 STEP 1 UNTIL K+1 DO E[I,J]),
               LSTR1(FOR I←1 STEP 1 UNTIL ILM DO
               [FOR J←1 STEP 1 UNTIL K+1 DO X[I,J],YHAT[I],RS[I]]),
               LSTR2(FOR J←1 STEP 1 UNTIL K+1 DO X[I,J],YHAT[I],RS[I]);
START:         WRITE (PKGOUT [NO]);
               READ (PKGIN,F1,LST1);
               READ (PKGIN,F2,LST2);
               READ (PKGIN,F3,LST3);
               READ (PKGIN,F4,LST4);
               READ (PKGIN,F3,LST5);
               CLOSE (PKGIN,RELEASE);
               FOR I←1 STEP 1 UNTIL ILM DO
          BEGIN
               YHAT[I]←0.0;
               FOR J←1 STEP 1 UNTIL K DO
               YHAT[I]←YHAT[I]+B[1,J]×(X[I,J]-XM[1,J]);
               YHAT[I]←YHAT[I]+XM[1,K+1];
           END;
               FOR I←1 STEP 1 UNTIL ILM DO
               RS[I]←(X[I,K+1]-YHAT[I])*2;
               WRITE (PKGOUT,FR1,LSTR1);
               MS←0.0;
               FOR I←1 STEP 1 UNTIL ILM DO
               MS←MS+RS[I];
               MS←MS/ILM;
               WRITE (PKGOUT,FR2,MS);
               WRITE (PKGOUT [PAGE]);
               MS←0.0;
               I←0;
L1:            I←I+1;
               IF I=ILM+1 THEN GO TO L4;
               M←1;
L2:            IF N[M]≠1 THEN GO TO L3;
               YCP←0.0;
               FOR J←1 STEP 1 UNTIL K DO
               YCP←YCP+E[M,J]×X[I,J];
               YCP←YCP-E[M,K+1];
               IF YCP<0 THEN
               M←2×M ELSE
               M←2×M+1;
```

```
                GO TO L2;
L3:             YHAT[I]←0.0;
                FOR J←1 STEP 1 UNTIL K DO
                YHAT[I]←YHAT[I]+B[M,J]×(X[I,J]-XM[M,J]);
                YHAT[I]←YHAT[I]+XM[M,K+1];
                RS[I]←(X[I,K+1]-YHAT[I])*2;
                WRITE (PKGOUT,FR1,LSTR2));
                MS←MS+RS[I];
                GO TO L1;
L4:             MS←MS/ILM;
                WRITE (PKGOUT,FR2,MS));
            END;
                END.
```

APPENDIX D

PIECEWISE LINEAR REGRESSION ON ARBITRARY DATA

ACTUAL COMPUTER OUTPUT

SSY=6.20364@+06

GROUP	NU	SIGMA RES	MEANS Y	X1	X2	X3
				B1	B2	B3
			CP	E1	E2	E3
1	900	1.07233@+05	75.42	5.08	4.99	2.39
				10.09	9.94	4.71
			4.62	1.00	-0.67	1.42
2	428	1.73534@+04	73.39	3.31	6.92	1.63
				7.32	13.40	1.81
			-3.22	1.00	-1.59	2.64
3	472	1.80065@+04	77.26	6.68	3.23	3.07
				12.64	7.08	7.37
			9.58	1.00	-1.02	1.95
4	199	3.80738@+03	87.70	2.80	8.37	1.14
				6.74	16.86	1.92
			-3.19	1.00	-1.27	4.03

5	229	7.62938@+03	60.96	3.76	5.67	2.06
				8.62	11.29	4.22
			0.80	1.00	0.16	-1.96
6	268	8.87591@+03	75.55	6.15	4.34	2.62
				11.66	8.29	5.89
			6.54	1.00	-1.20	2.09
7	204	5.80841@+03	79.51	7.37	1.77	3.67
				13.40	4.92	8.27
			32.18	1.00	6.17	3.11
8	100	1.84288@+03	90.34	2.30	8.79	0.67
				6.44	17.03	0.99
			-0.27	1.00	-0.59	2.80
9	99	1.88183@+03	85.03	3.31	7.95	1.61
				7.29	16.17	3.13
			-4.82	1.00	-1.69	3.30
10	127	1.46713@+03	54.69	2.38	5.84	2.89
				6.84	10.39	6.46
			-0.50	1.00	-0.86	0.68
11	102	2.67697@+03	68.76	5.47	5.46	1.03
				8.07	13.85	-3.20
			15.61	1.00	1.53	1.45
12	126	3.89071@+03	83.43	6.44	5.19	2.21
				12.21	8.78	5.81
			6.00	1.00	-0.22	0.57

13	142	4.48183@+03	68.56	5.90	3.58	2.98
				11.45	7.92	6.09
			8.87	1.00	-1.14	2.35
14	158	3.92757@+03	74.25	7.34	1.03	3.46
				13.33	-380.61	-248.43
			-26295.21	1.00	-7892.68	-5259.64
15	46	5.03605@+00	97.54	7.47	4.32	4.38
				14.27	9.60	9.74
16	63	1.26822@+01	78.65	1.17	8.54	0.60
				3.58	16.78	0.04
			-6.81	1.00	-1.15	2.23
17	37	6.62932@+02	110.24	4.23	9.21	0.78
				10.65	15.28	14.02
18	50	9.46308@+02	86.98	3.44	8.08	1.28
				7.30	16.40	2.74
			-4.56	1.00	-1.60	3.82
19	49	9.15670@+02	83.03	3.18	7.81	1.96
				6.45	17.29	1.02
20	77	9.72815@+00	64.24	2.15	6.81	3.04
				4.96	13.48	6.22
			-3.96	1.00	-3.42	5.36
21	50	3.01375@+00	39.97	2.75	4.35	2.67
				10.59	8.89	4.75
			0.61	1.00	-2.24	3.47

22	46	1.20952@+00	35.10	4.01	3.83	0.67
				7.88	8.11	1.75
23	56	6.42150@+00	96.41	6.66	6.79	1.33
				12.40	13.83	3.01
			3.05	1.00	-0.86	1.58
24	52	6.59191@+02	54.16	4.26	4.53	3.01
				7.87	8.75	5.68
			5.27	1.00	-0.87	1.64
25	74	1.07868@+03	104.00	7.98	5.66	1.65
				14.60	9.96	4.54
			6.33	1.00	-0.70	1.34
26	72	2.06043@+03	69.93	6.26	3.57	2.49
				12.70	6.99	8.11
			7.67	1.00	-1.19	2.29
27	70	2.25914@+03	67.15	5.52	3.58	3.49
				11.68	7.30	7.65
			10.10	1.00	-1.12	2.47
28	68	1.98842@+03	62.37	6.53	1.14	3.29
				12.14	3513.59	2345.75
			4077.50	1.00	1218.53	815.02
29	90	1.71795@+03	83.23	7.95	0.94	3.59
				13.67	-452.00	-296.09
			-5516.68	1.00	-1658.74	-1104.32

32	34	1.21407@+00	97.08	1.10	9.66	0.52
				3.47	-43.12	-40.72
33	29	2.63930@-01	57.05	1.27	7.23	0.71
				4.44	16.08	2.42
36	21	4.16541@+02	84.58	3.17	8.10	1.18
				5.89	18.04	-1.73
37	29	4.67709@+02	88.72	3.63	8.06	1.35
				3.83	22.33	-10.42
40	49	5.93628@-01	68.17	2.66	7.11	2.60
				4.75	14.03	5.19
41	28	3.08570@-01	57.37	1.24	6.27	3.80
				-184.95	-50.72	7.24
42	14	1.33869@-02	82.60	4.01	7.26	3.40
				62.38	1.16	-14.40
43	36	6.08628@-01	23.40	2.26	3.22	2.39
				132.83	49.51	4.92
46	33	2.59394@-01	104.37	6.89	7.20	1.17
				12.00	14.32	2.11
47	23	3.33357@-01	84.99	6.34	6.21	1.56
				13.16	13.08	4.32

48	26	3.31367@+02	56.38	4.28	4.86	2.90
				9.33	7.93	7.56
49	26	2.99813@+02	51.95	4.24	4.20	3.13
				4.94	11.32	0.90
50	34	5.08948@+01	110.31	7.53	6.83	1.95
				10.15	12.32	3.56
51	40	5.10607@+02	98.64	8.36	4.66	1.39
				13.28	10.78	1.57
52	33	9.74156@+02	72.23	6.81	3.50	2.02
				10.82	9.41	2.88
53	39	9.85283@+02	67.99	5.80	3.64	2.88
				13.68	5.79	11.02
54	37	1.11965@+03	69.82	6.03	3.54	3.07
				18.19	0.09	22.50
55	33	8.75473@+02	64.16	4.95	3.62	3.96
				10.77	7.83	8.87
56	38	1.05799@+03	59.59	6.46	1.42	2.88
				10.27	-1684.27	-1124.60
57	30	7.48215@+02	65.90	6.63	0.79	3.82
				11.98	13907.01	9275.47

58	20	2.50858@+02	40.70	4.71	0.87	3.70
				9.96	-8461.00	-5638.90
59	70	7.98056@+01	95.38	8.87	0.96	3.55
				16.24	-762.40	-502.32
			13293.20	1.00	3982.51	2658.02
118	28	2.73349@+00	93.76	9.11	1.60	2.60
				15.23	-727.66	-481.64
119	42	5.67790@+00	96.46	8.71	0.54	4.19
				16.42	-53.25	-27.82

BIBLIOGRAPHY

1. Allen, R. A. and Vernon, E. M., "Objective Weather Forecasting," *Compendium of Meteorology* (Thomas F. Malone, Editor) American Meteorological Society, Boston, Massachusetts, 1951, pp. 797-801.

2. Aubert, E. J., Lund, I. A. and Thomasell, A., "Some Objective Six-hour Predictions Prepared by Statistical Methods," *Journal of Meteorology*, Vol. 16, No. 4, August 1959, pp. 436-446.

3. Wadsworth, George P., "Application of Statistical Methods to Weather Forecasting," *Compendium of Meteorology* (Thomas F. Malone, Editor), American Meteorological Society, Boston, Massachusetts, 1951, pp. 849-855.

4. Gringorten, Irwing I., "Methods of Objective Weather Forecasting," *Advances in Geophysics, Volume II* (H. E. Landsburg, Editor), Academic Press, Inc., Publishers, New York, 1955, pp. 57-92.

5. Hanken, Albert F., "A Method and Model for the Analysis and Description of Car-following Performance," *Report No. 202-B-2*, Engineering Experiment Station, The Ohio State University, Columbus, Ohio, June 1965.

6. Lorenz, Edward N., "Prospects for Statistical Weather Forecasting," *Final Report—Statistical Weather Forecasting Project*, Massachusetts Institute of Technology, January 1959.

7. Miller, Robert G., "Statistical Prediction by Discriminant Analysis," *Meteorological Monographs*, American Meteorological Society, Vol. 4, No. 25, October 1962.

8. O'Neill, Thomas H. R., "A Historical Survey of Statistical Weather Prediction," *Research Report*, U. S. Navy Weather Research Facility, Norfolk, Virginia, 1963.

9. Panofsky, H. A. and Brier, G. W., *Some Applications of Statistics to Meteorology*, The Pennsylvania State University, Pennsylvania, 1958.

10. Schumann, T. E. W., "An Inquiry into the Possibilities and Limits of Statistical Weather Forecasting," *Quarterly Journal of the Royal Meteorological Society*, Vol. 70, No. 305, July 1944, pp. 181-195.

11. U. S. Department of Commerce, Weather Bureau, Asheville, North Carolina, "Local Climatological Data for Atlanta, Raleigh, Mobile and Nashville, January 1965 to June 1967."