ADAPTIVE LEARNING NETWORK APPROACH TO WEATHER FORECASTING: A SUMMARY

Andrew R. Barron [1]

Florence W. van Straten [2]

Roger L. Barron [1]

Introduction

Frequently, in solving performance prediction problems, rigorous mathematical or physical solutions are impractical or non-existent. Such is the case in the field of weather forecasting. Theoretical solutions to modeling meteorological activity can only be approximations. The atmosphere is subject to innumerable perturbations near the surface and is open-ended at the top. Empirical solutions to modeling weather are accurate only if they account for the remarkable diversity of weather and adapt to its ever-changing face.

The authors believe, and demonstrate preliminarily in this paper, that the Adaptive Learning Network (ALN) pattern recognition methodology is eminently suitable for improved solution of those classes of weather forecasting problems that involve density and pressure predictions. Forecasting of precipitation, temperature, visibility, humidity, winds, and severe storms should, in turn, be improved by better advance knowledge of density and pressure distributions. It is also possible that factors such as precipitation could be predicted directly and with greater accuracy via ALN techniques, but that has yet to be investigated.

Many examples of ALN use for empirical modeling and forecasting are contained in recent literature, including the preceding paper on steel shipment forecasting in this conference.

A specific atmospheric modeling problem had been previously subjected to ALN methodology: knowledge of the refractivity of the atmosphere between a radar and an observed aircraft had been provided by modeling a limited number of variables, including the radar metric parameters and the known height of the observed aircraft. [1, 2]

This earlier success prompted a desire to find if some aspects of weather forecasting can be solved by using ALN techniques. Furthermore, weather forecasting presents the challenge of being inherently a four-dimensional problem involving recognition of patterns in the three dimensions of space, a description of the evolution of these patterns with time, and prediction of future patterns. Weather forecasting is thus representative of a class of multivariate data processing problems in which patterns among the observed variables undergo continuous, dynamic variation. These problems must be approached from the standpoint of data-compressive transformations used to reduce the multiple measurements taken at each of a number of given locations into a few informative parameters, then map these parameters into dynamic three- (or at least two-) dimensional patterns. The latter may then be further parameterized by extraction of suitable features of the patterns. These features, along with their time derivatives (or, equivalently, the numerical values of the features at successive times) embody the requisite information for forecasting.

The vast number of atmospheric measurements routinely available to the forecaster for many altitudes and stations has required that this project begin modestly, and be followed, step-by-step, with more parameters and forecast elements used as success warrants.

Data Parameterization

In the present preliminary inquiry, only the first stage of parameterization has received close attention, viz., generation of data-compressive transformed variables indicative of conditions within essentially vertical filaments of the atmosphere probed by radiosondes (balloons) at the given sites. (These soundings are routinely made every 12 hours by the Government at a large number of locations throughout the United States and, to a lesser degree, at off-shore stations.) The authors show in this paper that transformed information communicated from as few as two known observation sites scattered within a radius of approximately 500 miles of Washington, DC, can be used to produce valid 12-hour forecasts of atmospheric density and pressure for that city.

In the first part of this study, computed values for the air density at 3 km. over each of twelve stations at midnight during the months of January 1971 and January 1972 were used to create ALN models for predicting the following midday density values over Washington, DC, during January of these or other years. Next, predictive ALN models were obtained for noon atmospheric surface pressure at Washington, DC, during January. [1]

Density was used as a modeling feature because it is a function of pressure, temperature, and relative humidity at a given height; not only does it transform several meteorological variables into one, but density has considerable physical significance. It is hypothesized by the authors that gradients between densities at neighboring stations determine the motion and activity of weather. On the other hand, predictive modeling of surface pressure was investigated because surface pressure is a readily monitored and universally recognized meteorological variable.

ALN Synthesis

The principal steps in creating preliminary ALN models for local forecasts of meteorological variables were:

(1) Data preprocessing and selection of variables for input to the ALN;

(2) Partitioning of the data base into three statistically-similar but independent subsets — fitting, selection, and evaluation;

[1] Adaptronics, Inc.

[2] Consultant in Atmospheric Physics.

[1] The month of January was chosen for this preliminary investigation partly because of the greater consistency of weather patterns during that time of the year and partly because of interest now focused on the winter months due to limited fuel supplies.

(3) Training the ALN on data from the fitting and selection subsets; and

(4) Evaluating the network by testing its ability to predict when interrogated on an independent evaluation data subset (January 1973 in this case).

Let us now consider each of these steps briefly.

Data preprocessing was critical in this project because the available input data were numerous. Candidate variables were chosen carefully to minimize the costs of future measurement, reporting, and computation tasks. However, once the dimensionality of the data was reduced, it was found that selection of variables for input to the ALN was not particularly critical, as the ALN synthesis algorithm automatically rejects nonessential information and poor performers among the candidate input parameters. For the preliminary models described in this summary, 12 observation stations, as shown in Figure 1, were retained as candidate observation sites in the data base. The densities at a geopotential altitude of 3 km. and surface pressures were chosen as input variables for the first and second preliminary models, respectively.

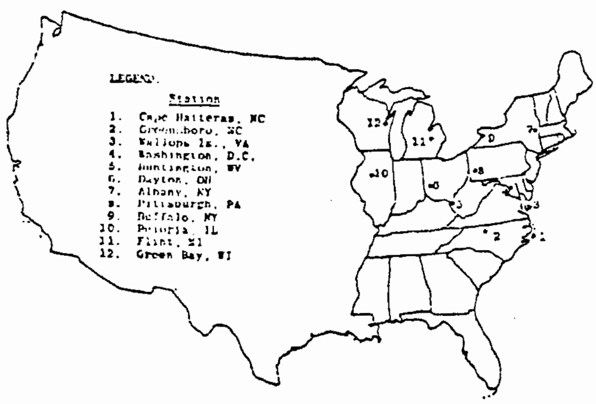

FIGURE 1: TWELVE CANDIDATE OBSERVATION STATIONS FOR FORECASTING WASHINGTON, DC, DENSITY AND PRESSURE

Partitioning of the data base was performed by incorporating observations from January 1971 in the fitting subset, January 1972 in the selection subset, and January 1973 in the evaluation subset. (No direct test, such as a clustering analysis, was performed to verify the statistical similarity of these subsets; however, the ALN's gave comparable performance on each, which is a good ad hoc indication.) Each observation consisted of midnight density or pressure values at each of the 12 candidate stations and the known value of the following noon density or pressure for Washington, DC. Observations were ignored when information for any day was incomplete; thus some gaps exist in the data base used. [1]

Network training was accomplished by creation of nonlinear algebraic networks of polynomial building-block elements. Each of the basic elements was a function of two input variables or of outputs from elements of previous layers of the network. Each element was of the basic form:

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_1x_2 + w_4x_1^2 + w_5x_2^2$$

where x_1 and x_2 denote inputs, the w's are constants determined during ALN synthesis, and y is the element output. These elements were connected in a series-parallel array via the ALN synthesis algorithm, as discussed in the references to the preceding paper.

The ALN January forecasting model for 3-km. density, which retained inputs from four of the candidate 12 stations, is shown in Figure 2. A six-station ALN January model for surface pressure forecasting is shown in Figure 3. Figure 4 shows that the accuracy of 3-km. density predictions increased as the synthesis algorithm was allowed to enlarge the number of observation stations used by the ALN. Network evaluation was accomplished using the January 1973 observations, as discussed below. (January 1973 data were not involved in training of the ALN's.)

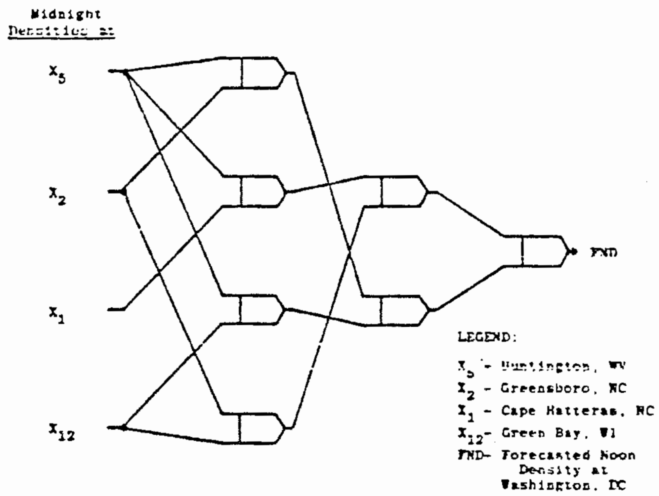

FIGURE 2: ALN USING FOUR OBSERVATION STATIONS FOR FORECASTING OF 3-KM. DENSITY AT WASHINGTON, DC

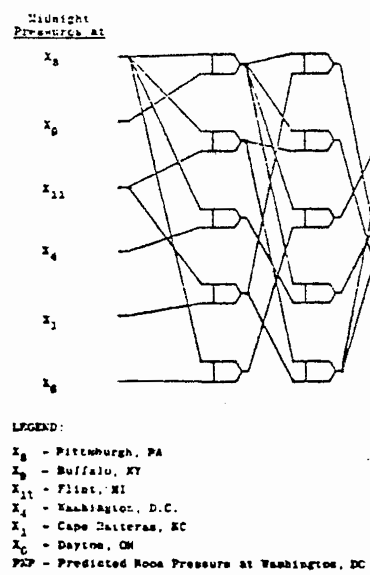

FIGURE 3: ALN USING SIX OBSERVATION STATIONS FOR FORECASTING SURFACE PRESSURE AT WASHINGTON, DC

[1] This points up the need for having a family of alternative ALN's that can function without information from "missing" sites.

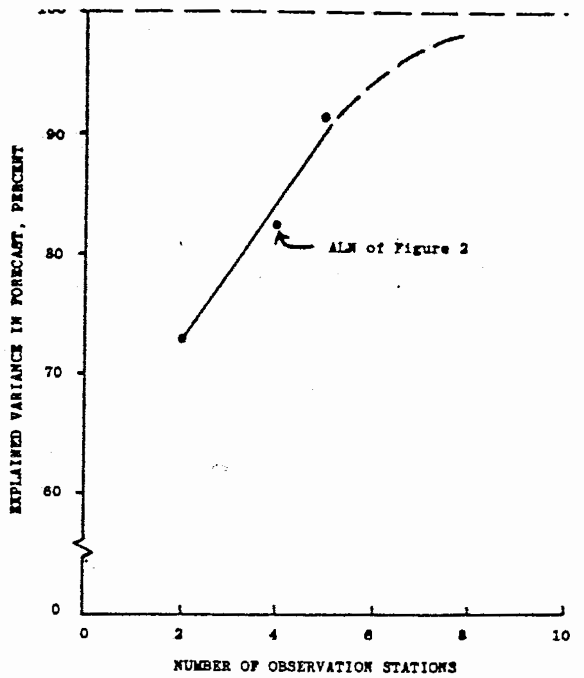

FIGURE 4: EXPLAINED VARIANCE OF 3-KM. DENSITY FORECASTS VS. NUMBER OF OBSERVATION STATIONS USED BY ALN's.

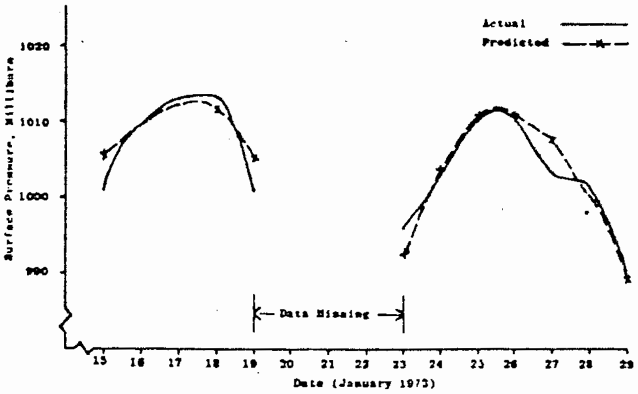

FIGURE 6: ACTUAL AND PREDICTED SURFACE PRESSURES AT WASHINGTON, DC, IN JANUARY 1973 (EVALUATION DATA SUBSET)

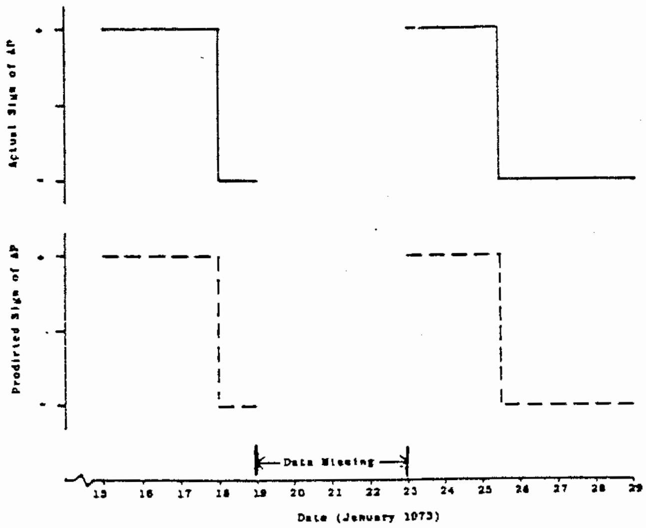

FIGURE 7: SIGNS OF ACTUAL AND PREDICTED WASHINGTON, DC, SURFACE PRESSURE CHANGES (EVALUATION DATA SUBSET)

Preliminary Results

Figure 5 shows a comparison between the fitting-data subset, six-station ALN, Washington, DC, noon, surface-pressure forecasts for January 1971 and the actual pressure values during that month. Figure 6 presents a similar comparison for data in the evaluation subset, which reveals how the ALN would have predicted fluctuations in surface pressure on an independent test during January 1973. Note that in both figures the predicted values were close to the actual, and that, furthermore, the predicted values correctly anticipated whether the actual pressure would rise or fall. The ability to forecast pressure changes correctly is seen most graphically in a plot of the sign of the pressure changes, sgn(Δp), that occurred between successive observations of the actual and predicted pressures. Such a plot for the January 1973 data is shown in Figure 7.

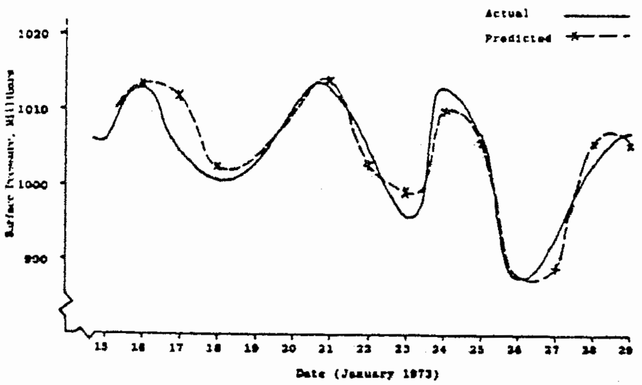

FIGURE 5: ACTUAL AND PREDICTED SURFACE PRESSURES AT WASHINGTON, DC, IN JANUARY 1971 (FITTING DATA SUBSET)

Results obtained with ALN forecast models for pressure and density are also compared to three baseline forecasting techniques in Tables 1 and 2. The "constant" forecast assumed that the Washington, DC, noon pressure (or density) for a given observation was equal to the average noon pressure (or density) at Washington, DC, in January. [1] The "noon persistence" forecasts shown in the tables assumed that the noon Washington, DC, pressure (or density) for a given day was equal to its own value measured at the previous noon. The "midnight persistence" forecasts assumed persistence of conditions observed at midnight until the following noon. These three techniques, none of which could predict changes in pressure or density, provide means for judging accuracy of the ALN forecasts (which are superior for both pressure and density and for both the training and evaluation data subsets). It is also significant that ALN accuracies, as measured on the evaluation data, were substantially the same as on the training data. This indicates that the ALN's generalized correctly on their limited training experience and therefore had not been overfitted during training.

[1] To compute this average, the authors assumed that the mean January pressure (or density) for the years '71–'73 equalled the historical (pre '71) January mean.

TABLE 1

SUMMARY OF PRELIMINARY RESULTS FOR WASHINGTON, DC, FORECASTS OF SURFACE PRESSURE AT NOON

TYPE OF FORECAST	TRAINING DATA		EVALUATION DATA	
	Average Absolute Difference	% Average Absolute Difference	Average Absolute Difference	% Average Absolute Difference
Constant (1107.0 mBars)	6.55	.650	7.55	.750
Noon Persistence	8.74	.868	6.56	.651
Midnight Persistence	4.43	.440	3.30	.328
6-Station ALN	2.19	.217	2.44	.242

TABLE 2

SUMMARY OF PRELIMINARY RESULTS FOR WASHINGTON, DC, FORECASTS OF 3-KM. DENSITY AT NOON

TYPE OF FORECAST	TRAINING DATA		EVALUATION DATA	
	Average Absolute Difference	% Average Absolute Difference	Average Absolute Difference	% Average Absolute Difference
Constant (.92197 kg/m^3)	.0153	1.66	.0126	1.37
Noon Persistence	.0143	1.55	.0079	.86
Midnight Persistence	.0087	.94	.0064	.69
5-Station ALN	.0055	.59	.0059	.64

Plans For Future Investigation

The next step in this study will investigate the feasibility of including wind direction and speed information in the density and pressure models. A weighted-mean wind vector for each radiosonde observation will be computed by finding the vector sum of wind values at the various heights. The vector at each height in this summation will be weighted according to density at that height and the difference in heights between adjacent wind measurements. Additionally, the vertical density and pressure profiles will each be expressed with only two parameters, e.g., their values measured at ground level and the exponential rates of decay of these quantities with increasing altitude.

Each station, with its transformed wind, density, and pressure information (and possibly temperature and humidity information 1/) will thus be represented by a point in a two-dimensional data field (perhaps seven or eight data fields in all will be used). Parameters (features) that characterize the entire fields at given times will then be extracted via pattern recognition techniques. Each of these parameters will form a sequence over successive times of observation. Characteristics of these time series will become candidate inputs to ALN's.

1/ Note that density, pressure, temperature, and humidity are related via the Gas Law; thus one of these four variables may be omitted, in general, with no loss of completeness in characterizing the state of the atmosphere at that point.

It is expected that the advantages of this approach will include further reduction of the number of observation stations needed in the data field, further compression of large quantities of meteorological data into a few short-duration time series, and (potentially) forecasting of factors such as rainfall.

As in this work to date, the attendant advantages of ALN modeling will be:

- Identification of the most relevant parameters,
- Adaptive evolution of both the structure and the coefficients of the models,
- Modeling of nonlinear as well as linear interactions between parameters,
- Avoidance of data overfitting,
- Rapid computations with the trained models.

Conclusions

Techniques have been investigated for parameterizing multi-station, multi-level meteorological data and for synthesizing ALN forecasting models using these data. These models require very few inputs to provide useful 12-hour predictions of surface pressure and 3-km. density and to anticipate accurately whether surface pressure will rise or fall.

Further investigations should be conducted to enlarge upon the findings of this preliminary study.

Acknowledgements

The authors are grateful to the U. S. Naval Weather Service Command, who, with the cooperation of the National Oceanic and Atmospheric Administration, supplied the data. The authors also express appreciation to Adaptronics, Inc. for financial support of this investigation and use of proprietary computer programs.

References

1. van Straten, R. W., R. L. Barron, and R. F. Snyder, Refractive Index Measurements by Adaptive Techniques, Final Technical Report prepared for Office of Naval Research by Adaptronics, Inc. under Contract Number N00014-73-C-0103, June 30, 1974.

2. Barron, R. L., F. W. van Straten, and R. F. Snyder, "Inference of Refractivity Structure and Generation of Ray Traces by Analytical and Adaptive Methods," Proc. of IEEE 1975 National Aerospace and Electronics Conference, pp. 479-486.

Bibliographical Item

Mucciardi, A. N., "New Developments in Water Shed Pollution Forecasting Using Adaptive Trainable Networks," Proc. of IEEE 1974 International Conference on Systems, Man and Cybernetics, pp. 373-378.

PROCEEDINGS

INTERNATIONAL CONFERENCE ON CYBERNETICS AND SOCIETY

September 19-21, 1977

Sponsored by:

IEEE Systems, Man and Cybernetics Society

With the cooperation of:

College of American Pathologists
Human Factors Society
IEEE Computer Society
IEEE Engineering in Medicine and Biology Society

Mayflower Hotel, Washington, D.C.