ADAPTIVE LEARNING NETWORK APPROACH TO WEATHER FORECASTING: A SUMMARY

Andrew H. Barron 1/
Florence N. van Straten 2/
Roger L. Barro 3/

Introduction

Frequently, in solving performance prediction problems, rigorous mathematical or physical solutions are impractical or nonexistent. 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 immeasurable perturbations near the surface and is capped-at the top. Empirical solutions to model-ral 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 currently suitable for improved solution of these 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 work yet to be investigated.

The example 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 reactivity of the atmosphere between a radar and an observed aircraft had been provided by modeling a limited number of variables, including the radar returns, parameters and the known height of the observed aircraft.

This earlier success prompted a desire to find if some aspects of weather forecasting could be solved by using ALN techniques. Furthermore, weather forecasting presents characteristics that are 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 approaches 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 significant features of the patterns. These features, along with their time derivatives (or, equivalently, the material values

1/ Adaptronics, Inc.
2/ Consultant in Atmospheric Physics.

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 measurements and forecast elements used as success warrants.

Data Parameters

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 radar echoes (echoons) 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, as off-site stations.)

The authors note in this paper that transformed information resulting 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 48-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 midsagittal density values over Washington, DC, during January of those or other years. Next, predictive ALN models were obtained for sea asperometry surfacic pressure at Washington, DC, during January. These density was used as a modeling feature because it is a factor of high energetic significance. The air asperometry density 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 readily monitored and universally recognized meteorological variable.

ALN Network

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

1. Data preprocessing and selection of variables for input to the ALN.

2. Partitioning of the data into three statistically-similar but independent subsets — fitting, selection, and evaluation.

The study 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 ANN on data from the fitting and selection subsets; and
(4) Evaluating the network by testing its ability to predict wind speed 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 inexact. 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 ANN was not particularly critical, as the ANN synthesis algorithm automatically recombines monometric information and poor performance among the candidate input parameters. For the preliminary models described in this summary, 32 observation stations, as shown in Figure 1, were retained as candidate observation sites in the data base. The densities at a geodetic altitude of 3 ft. and surface pressures were chosen as input variables for the first and second preliminary models, respectively.

![Figure 1: Twelve Candidate Observation Stations for Forecasting of 3-ft. Density at Washington, DC](image)

Partitioing of the data base was performed by incorporating observations from January 1973 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 ANN's are comparable performance on each, which is good of the indication.) Each observation obtained from the density or pressure values at each of the 32 candidate stations and the known values of the following noon density or pressure for Washington, DC. Observations were ignored when information for any day was incomplete; this same gap did not exist in the data base used.

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

\[ y = \sum w_i x_i - \sum w_i x_i^2 - \sum w_i x_i^3 + \sum w_i x_i^4 + \sum w_i x_i^5 + \sum w_i x_i^6 \]

where \( x_i \) and \( w_i \) denote inputs, the \( w_i \) are constants determined during ANN synthesis, and \( y \) is the element output. These elements were connected in a serial-parallel array via the ANN synthesis algorithm, as discussed in the references to the preceding paper.

The ANN January forecasting model for 3-ft. density, which retained inputs from four of the candidate 32 stations, is shown in Figure 2. A six-station ANN January model for surface pressure forecasting is shown in Figure 3. Figure 4 shows that the accuracy of 3-ft. density predictions increased as the synthesis algorithm was allowed to enlarge the number of observation stations used by the ANN. Network evaluation was accomplished using the January 1973 observations, as discussed below. January 1973 data were not involved in training of the ANN's.

![Figure 2: ANN Using Four Observation Stations for Forecasting of 3-ft. Density at Washington, DC](image)

![Figure 3: ANN Using Six Observation Stations for Forecasting Surface Pressure at Washington, DC](image)
Figure 4: Explained variance of 30-day density forecasts vs. number of observation stations used by ALN's.

Preliminary Results

Figure 5 shows a comparison between the fitting-data subset, six-station ALN, Washington, DC, noon, surface-pressure forecasts for January 1973 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 in both most graphically in a plot of the size of the pressure changes, on line, 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 1973 (Fitting Data Subset)

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)

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. 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 accuracy, as evidenced on the evaluation data, was 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.

To compute this average, the authors assumed that the mean January pressure (or density) for the years '72-'73 equaled the historical (pre '71) January mean.
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 the data from a few short-duration time series, and (potentially) forecasting based on these data. As to this work to date, the attendant advantages of AIR 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 AIR forecasting models using these data. These models require very few inputs to provide useful 12-hour predictions of surface pressure and 3-4-m density and to anticipate accurately whether the surface pressure will rise or fall.

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

Acknowledgments

The authors are grateful to the U.S. Naval Weather Service Center (Naval Oceanic, with the cooperation of the National Oceanic and Atmospheric Administration, applied the data analysis) and to the Applied Resonance, Inc. for financial support of this investigation and use of proprietary computer programs.

References


Bibliographical Note
