2.3

# RELATING WEST ANTARCTIC SURFACE METEOROLOGY TO THE LARGE-SCALE ATMOSPHERIC CIRCULATION

David B. Reusch* and Richard B. Alley
The Pennsylvania State University, University Park, PA

## 1. INTRODUCTION

A deeper understanding of Antarctic meteorology is one key to improved interpretation of the ever growing body of ice-core-based paleoclimate records from this region. Automatic weather stations (AWS) currently provide the only year-round, direct measurements of weather on the ice sheet. As the spatial coverage of the network has expanded year to year, so has our meteorological database. Unfortunately, many of the records are relatively short (less than 10 years) and/or incomplete (to varying degrees) due to the vagaries of the harsh environment. Recent developments in climate downscaling in temperate latitudes suggest it may be possible to use GCM-scale meteorological data sets (e.g., ECMWF reanalysis products) to both fill gaps in the AWS records and extend them back in time to create a uniform and complete database of West Antarctic surface meteorology (at AWS sites). Artificial neural network (ANN) techniques are used to predict known AWS data (e.g., temperature, pressure) using large-scale features of the atmosphere (e.g., 500 mb pressure height). Once trained, the ANN can predict from the GCM-scale data for periods where AWS data are not available.

## 2. METHODOLOGY

### 2.1 *Data*

**Automatic Weather Station Data:** The main source of direct meteorological data in West Antarctica is the network of AWS maintained by the University of Wisconsin-Madison since 1980 (Lazzara 2000). All stations provide air temperature and pressure, wind speed and direction; some stations also have relative humidity and the 3.0-0.5 m (above nominal snow surface) temperature difference. Data were selected to match ECMWF time-steps from the three hourly quality-controlled archive for Siple Dome (75.9° S, 84 °W) and Ferrell (77.91° S, 170.82° E) AWS.

**Gridded Meteorological Data:** The ECMWF 15-year reanalysis data product (ERA-15) is our source for GCM-scale meteorological data for the period 1979-1993 (ECMWF 2000). ECMWF operational analyses are used for 1994-present. The 2.5° horizontal resolution data sets are used in both cases. Potential problems have been noted with the ECMWF (re)analysis data over Antarctica, stemming in part from the flawed surface elevations used in these models (Genthon and Braun 1995). Elevation errors exceeding 1000 m exist in some areas of Queen Maud Land and the Antarctic Peninsula (e.g., Figure 3, Genthon and Braun 1995). Topography in West Antarctica is generally much better but errors from outside our study area will still have an influence on the reanalysis data (for example, an elevation error for Vostok station has broad effects on geopotential heights). Evaluations of several operational products (e.g., Bromwich et al. 1995; Bromwich et al. 2000; Cullather et al. 1998) and discussions with experienced polar meteorologists (D. Bromwich, J. Turner, pers. comm.) suggest that the ECMWF analyses are the best available data sets for Antarctica (see also Bromwich et al. 1998).

### 2.2 *Artificial Neural Networks*

At the simplest level, artificial neural networks (ANNs) are a computer-based problem solving tool inspired by the original, biological neural network – the brain. Because of their ability to generate non-linear mappings during training, ANNs are particularly well-suited to complex, real-world problems such as understanding climate (Elsner and Tsonis 1992; Tarassenko 1998). Meteorological examples include an improved understanding of controls on precipitation in southern Mexico (Hewitson and Crane 1994), prediction of summer rainfall over South Africa (Hastenrath et al. 1995) and northeast Brazil (Hastenrath and Greischar 1993), and extreme event analysis in the Texas/Mexico border region (Cavazos 1999). We have used the MATLAB® Neural Network Toolbox (Demuth and Beale 2000; Haykin 1999).

Multilayer feed-forward ANNs were chosen to follow previous work with climate downscaling in the literature (e.g., Cavazos 1999). These ANNs consist of a large number of highly interconnected, simple processing nodes (a.k.a. neurons) organized into three layers (Figure 1). The input layer serves to receive input data, with one node for each input variable. The hidden layer consists of nodes with inputs from each node in the input layer. The number of hidden nodes is problem dependent and is a significant factor in how well the ANN works. The output layer receives intermediate results from the hidden layer and translates them to the desired output format.

ANNs need to be taught to produce the desired outputs (AWS observations) from the inputs (ECMWF data) before they can be used for predictions, a task done iteratively in three main phases: training, testing and validation. The training phase adjusts the connection weights using an optimization function that reduces the

* *Corresponding author address:* David B. Reusch, Department of Geosciences and EMS Environment Institute, The Pennsylvania State University, University Park, PA 16802; e-mail: dbr@geosc.psu.edu

error in the network's results. The training error is calculated by comparing the network's output prediction to the AWS observations. Weights are adjusted based on the cumulative error from one pass through the complete training set. Testing uses a second subset of the data to evaluate training performance. Validation is used to avoid overfitting the training data and tests the network with data distinct from the training and testing samples. The cycle then repeats until the desired output is achieved (or the error cannot be further reduced). In our work so far, training, testing and validation have used 70%, 20% and 10% of one year of data, respectively.

### 2.3 *Training and Testing*

Separate ANNs are currently used for each AWS variable at each site due to the different physical controls and geography involved. Selection of a "best" ANN involves numerous dimensions of possible parameters. In the physical domain, a variety of predictor variables have been tested, including geopotential height, thickness, wind speed and direction, and temperature advection, from up to three different pressure levels (e.g., 850, 700 and 500 mb). Selection of appropriate ECMWF grid points adds a second physical dimension, though results have not been particularly sensitive to our choices. One highly influential parameter is the number of hidden nodes in the ANN, as is expected from ANN theory. Reasonable values can be found through iterative testing over a range with validation checking to avoid overfitting. Lastly, because of the complexity of the error surfaces and the non-zero probability of training into inappropriate local minima, each ANN configuration is trained repeatedly (50 to 100 times) using different, randomly chosen training sets. The best performer is then saved for production use.

## 3. RESULTS

While significant progress has been made in understanding how best to apply ANNs to this problem, our research is still emphasizing exploration of the many different parameters available. The basic framework has been covered but many variations remain. As such, our results are still quite preliminary.

### 3.1 *Ferrell AWS (77.91° S, 170.82° E)*

Ferrell AWS was installed on the Ross Ice Shelf in December 1980, making it among the longest-running records in this region (Byrd Station, 80.01° S, 119.40° W, is the oldest AWS still running in West Antarctica). Training to date has focused on 1987, a year with a low percentage of missing data (only about 3%, where the average for 1981-1998 is 12.4%). The best surface pressure-predicting ANN for this year used 850 and 700 mb geopotential, 700-850 thickness and 850 mb winds to yield an $r^2$ of 0.96 and an RMS error of 2.1 mbar. The latter amounts to 3% of the range of observations during this year (70 mb). This ANN was used to produce a (sample) complete record for 1979-1993 (Figure 1). In the observational data, gaps range from 2-33% of each year for 1981-1993 and no data are available for 1979 and 1980. The most significant portions of the ANN-predicted record are the two new years (1979 and 1980) and the long (multi-week) fills during 1983, 1984, 1985, 1991 and 1992.

The best results for temperature prediction have used 850 mb temperature and temperature advection plus 500-850 mb thickness for an $r^2$ of 0.84 and RMS error of 4.7 °C (~8% of the 58 °C range). The lower accuracy for temperature likely reflects the greater complexity of controls on this variable. The ANN's predictive skill should improve as further variables are explored. The greater general interest in temperature (as compared to pressure) warrants continued research.

### 3.2 *Siple Dome AWS (75.9° S, 84 °W)*

Because the station at Siple Dome was only installed in 1997, training data has been more limited at this site. We are also forced to use ECMWF operational analyses, rather than the ERA-15 data, which exposes us to changes in the data due strictly to model changes. Completion of the ECMWF 40-year reanalysis (Gibson et al. 1999) would therefore benefit our work. As with Ferrell, geopotential heights and thickness are used for pressure prediction but at the 700 and 500 mb levels due to the higher elevation of Siple Dome. Training and testing at this site are still in progress. Production ANNs will be used to generate entirely new records for the period 1979-1996 and to fill gaps thereafter.

## 4. SUMMARY

ANNs provide a potentially powerful tool for repairing and extending the AWS surface meterological record in West Antarctica. Early results have been encouraging in many respects but much research remains. The potential benefits to the ice coring community towards improved interpretations of paleoclimate proxies are but one reason of many to continue.

## 5. REFERENCES


Bromwich, D. H., R. I. Cullather, et al., 1998: Antarctic precipitation and its contribution to the global sea-level budget. *Annals of Glaciology*, **27,** 220-226.

———, F. M. Robasky, et al., 1995: The atmospheric hydrologic cycle over the Southern Ocean and Antarctica from operational numerical analyses. *Monthly Weather Review*, **123,** 3518-3539.

———, D. H., A. N. Rogers, et al., 2000: ECMWF Analyses and Reanalyses Depiction of ENSO Signal in Antarctic Precipitation. *Journal of Climate*, **13,** 1406-1420.

Cavazos, T., 1999: Large-scale circulation anomalies conducive to extreme events and simulation of daily rainfall in northeastern Mexico and southeastern Texas. *Journal of Climate*, **12,** 1506-1523.

Cullather, R. I., D. H. Bromwich, et al., 1998: Spatial and temporal variability of Antarctic precipitation from atmospheric methods. *Journal of Climate*, **11,** 334-367.

Demuth, H. and M. Beale, 2000: *Neural Network Toolbox*. Mathworks, Inc.

ECMWF, cited 2001: ERA-15. [Available online from http://wms.ecmwf.int/research/era/Era-15.html.]

Elsner, J. B. and A. A. Tsonis, 1992: Nonlinear Prediction, Chaos, and Noise. *Bulletin of the American Meteorological Society*, **73,** 49-60.

Genthon, C. and A. Braun, 1995: ECMWF Analyses and Predictions of the Surface Climate of Greenland and Antarctica. *Journal of Climate*, **8,** 2324-2332.

Gibson, J. K., M. Fiorino, et al., 1999: The ECMWF 40 Year Re-analysis (ERA-40) Project - Plans and Current Status. *10th Global Change Studies*, American Meteorological Society, 369-372.

Hastenrath, S. and L. Greischar, 1993: Further Work on the Prediction of Northeast Brazil Rainfall Anomalies. *Journal of Climate*, **6,** 743-758.

———, ———, et al., 1995: Prediction of the Summer Rainfall over South Africa. *Journal of Climate*, **8,** 1511-1518.

Haykin, S. S., 1999: *Neural networks : a comprehensive foundation*. 2nd ed. Prentice Hall, 842 pp.

Hewitson, B. C. and R. G. Crane, 1994: Precipitation controls in southern Mexico. *Neural Nets: Applications in Geography*, B. C. Hewitson and R. G. Crane, Eds., Kluwer Academic, 121-143.

Lazzara, M. A., cited 2000: Antarctic Automatic Weather Stations Web Site Home Page. [Available online from http://uwamrc.ssec.wisc.edu/aws/.]

Tarassenko, L., 1998: *A Guide to Neural Computing Applications*. John Wiley & Sons, Inc., 139 pp.

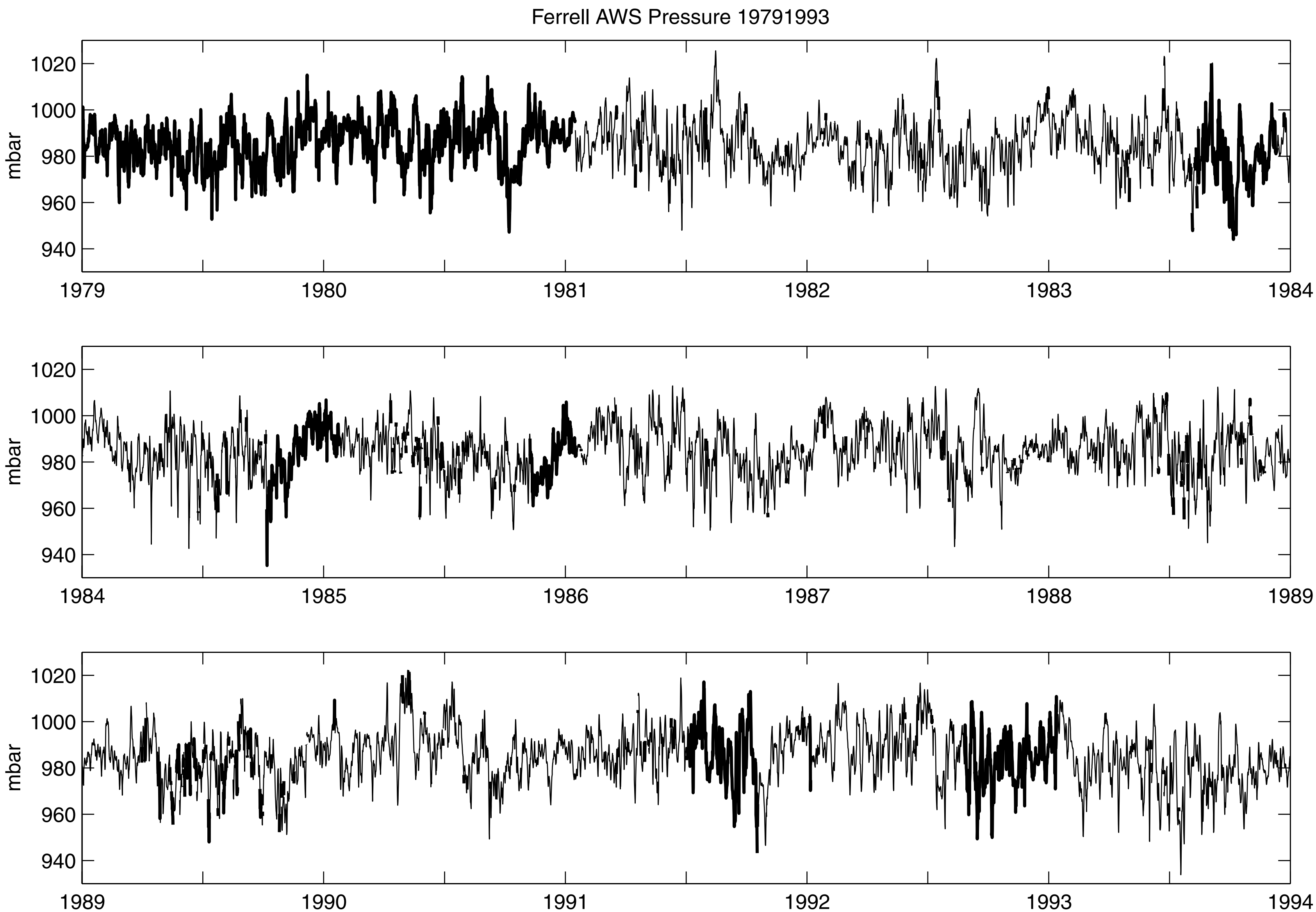


Figure 1. Preliminary reconstructed surface pressure at Ferrell AWS for 1979-199 (original observations as thin line, ANN-modeled values as thick line). The ANN was trained with 1987 data. ECMWF data were used to fill gaps and extend record back to 1979. Ferrell AWS was installed on the Ross Ice Shelf (77.91° S, 170.82° E) in December 1980.