Artificial Neural Networks and Long-Range Precipitation Prediction in California

DAVID SILVERMAN AND JOHN A. DRACUP

Civil and Environmental Engineering Department, University of California, Los Angeles, Los Angeles, California

(Manuscript received 18 August 1998, in final form 5 March 1999)

ABSTRACT

Artificial neural networks (ANNs), which are modeled on the operating behavior of the brain, are tolerant of some imprecision and are especially useful for classification and function approximation/mapping problems, to which hard and fast rules cannot be applied easily. Using ANNs, this study maps a 1-yr monthly (January–December) time series of the 700-hPa teleconnection indices and ENSO indicators onto the water year (October–September) total precipitation of California's seven climatic zones, with different lag times between the inputs and outputs. It was found that the pattern of rainfall predicted by the ANN model matched closely the observed rainfall with a 1-yr time lag for most California climate zones and for most years. This research shows the possibility of making long-range predictions using ANNs and large-scale climatological parameters. This research also extends the use of neural networks to determine important parameters in long-range precipitation prediction by comparing results gained using all the inputs with results from leaving an individual index out of the network training. This comparison gives insight into the physical meteorological factors that influence California's rainfall.

1. Introduction

Although California includes seven distinct climatic zones (Fig. 1) in which duration and intensity of precipitation vary markedly, all the zones have a well-defined wet season (November–March) separated by a long dry season. Activity in the El Niño–Southern Oscillation (ENSO) in the east Pacific and the 700-mb height anomaly over North America has been shown to be related to various phenomena in specific regions of California (Kahya and Dracup 1993; Redmond and Koch 1991; Ropelewski and Halpert 1989, 1987, 1986; Barnston and Livezey 1987; Wallace and Gutzler 1981). Because these activities appear to influence California precipitation, they continue to be studied, for by noting their regional and global patterns of climate variability [National Drought Mitigation Center (NDMC 1998)], scientists may be able to use their recurring patterns of changes in the height anomaly and ENSO as tools for long-range climate predictions.

The 700-mb circulation is a major force in the global climate and is a good measure of atmospheric circulation (speed and direction of the winds at about 3 km above sea level). Patterns in the 700-mb height anomaly may allow for regional precipitation predictions a year or more in advance based on the state of the system. The teleconnection indices for the 700-mb height are derived from a principal component analysis and as such represent the major 700-mb pressure patterns.

The relationship between future California precipitation and both ENSO and the 700-mb height anomaly is most likely nonlinear and not easily determined by ordinary statistical methods. Current methods have been unable to detect the patterns or governing modes necessary for long-range precipitation prediction (Navone and Ceccatto 1994; Venkatesan et al. 1997). Such failures may have their source in the noise or fuzziness inherent in the data, the overwhelming amount of data, in the nonlinear relationships involved, or in the limitations of current methods. Clearly, new methods that can analyze complex systems and their relationships need to be explored. Artificial neural networks (ANNs) may be powerful enough to identify such relationships.

This study is the beginning of a larger study with intertwined goals. The first goal is to verify the hypothesis that 700-mb circulation patterns are a driving force of regional precipitation in California. The second is that neural networks are able to determine patterns in the overwhelming amount of data. The third is to show that a trained neural network contains knowledge about the modeled process that can be extracted.

2. Artificial neural networks (ANNs)

a. Background

ANNs are based on biological models of the brain and the manner in which it recognizes patterns and

Corresponding author address: Dr. John A. Dracup, Civil and Environmental Engineering Department, University of California, Los Angeles, 5732 Boelter Hall, Los Angeles, CA 90095-1593.
E-mail: dracup@seas.ucla.edu

FIG. 1. California's seven climate zones.

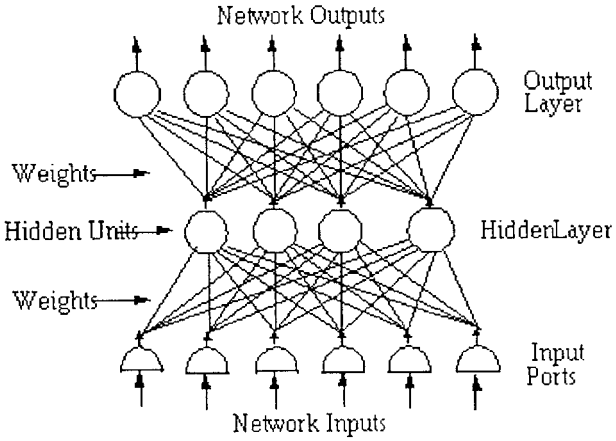

FIG. 2. The ANN feed forward–back propagation network is composed of an input layer of neurons, one or more hidden layers, and an output layer. Each layer comprises multiple units connected completely with the next layer, with an independent weight attached to each connection.

learns from example. The human brain contains more than a billion neurons with trillions of interconnections working simultaneously, enabling it immediately to pick an individual out of a crowd or to distinguish one voice out of a babble of voices and background noises—tasks that are difficult or even impossible for the most powerful supercomputers. Its construct also allows the brain to learn quickly from experience. An ANN comprises interconnected simple processing units that work in parallel, much as do the neuron networks of the brain, and which can discern patterns from input that is ill-defined, chaotic, and noisy. For an excellent layman's discussion on artificial neural networks, see Hinton (1992).

The advantages of using ANNs include the following (French et al. 1992; Raman and Sunilkumar 1995).

- A priori knowledge of the underlying process is not required.
- Existing complex relationships among the various aspects of the process under investigation need not be recognized.
- Solution conditions, such as those required by standard optimization or statistical models, are not preset.
- Constraints and a priori solution structures are neither assumed nor enforced.

Disadvantages (Masters 1993) include exponential increase of training time with increased dataset size, uncertainty about network and training set size requirements, and uncertainty about the relationships used by the network to produce the network's output.

b. Structure

An ANN is composed of an input layer of neurons, one or more hidden layers, and an output layer. Each layer comprises multiple units connected completely with the next layer (Fig. 2), with an independent weight attached to each connection. The number of nodes in the output layer is equal to the number of values to be predicted. In our network, we have one output node for the yearly total precipitation. The number of nodes in the hidden layer(s) is dependent on the process being modeled and must be determined by trial and error. The number of nodes in the input layer is equal to the number of input parameters. For our network, it is 12 months of the 17 teleconnection indices, or 204 input nodes.

Known values (e.g., monthly teleconnection index values) are presented to the input nodes. As the input nodes pass these values to the next level, the values are multiplied by the weight of the connection. At the next level, each processing unit (neuron) sums all of its weighted inputs and then applies an activation function to determine whether the neuron fires or remains dormant. The connection weights/neuron group, by firing or remaining dormant, determines the importance of that particular input to the overall prediction.

The activation function is normally the *s*-shaped sigmoid function of the form such as

$$\frac{1}{1 + e^{-x}},$$

where x = the sum of all weighted inputs coming to that node (Fig. 3). The function is scaled such that the asymptotes are either 0 and 1 or -1 and 1. This process is repeated for each successive layer, including the output layer.

During the training process, the error between the desired output and the calculated output is propagated back through the network (hence ''back propagation networks''). At the beginning of the training, all weights are set to independent random values. The back propagation attempts to minimize iteratively the average squared error

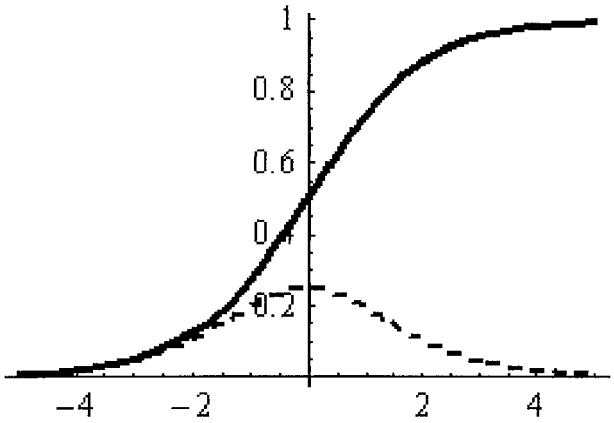

FIG. 3. The sigmoid neuron activation function "$(1 + e^{-x})^{-1}$" (solid) and its derivative (dashed). The sigmoid function determines the activation level of a neuron based on the weighted inputs. The derivative is used during network development to determine the level of training a weight connection receives.

$$E = \frac{1}{2}\sum_{k=1}^{K}(y_k - d_k)^2,$$

where E = error for a given training case, K = number of output nodes, y_k = calculated output, and d_k = desired output. This is done by computing the gradient for each element on the output layer

$$\delta_k = d\sigma_k \cdot (y_k - d_k),$$

where $d\sigma_k$ = the derivative of the sigmoid function applied at y_k determined for each output node k. For the next layer back, the gradient for the hidden layer is

$$\delta_j = d\sigma_j \sum_{k=1}^{K} \delta_k w_{jk},$$

where $\delta\sigma_j$ = the derivative of the sigmoid function, and w_{jk} = weight from hidden node j to output node k. The weights are updated by

$$\Delta w_{jk} = w_{jk} - \eta\delta_k y_j$$

for the output layer and

$$\Delta w_{ij} = w_{ij} - \eta\delta_j y_i,$$

where η = the learning rate, for each layer back.

Input sets are chosen randomly, and the weights are adjusted by the above procedure until the error is minimized or reduced to a predefined limit or the desired number of training periods, or epochs, is achieved.

3. Teleconnections

A teleconnection is a recurring and persistent relationship between anomalous climatic phenomena in two discrete areas of the earth. Activity in one part of the earth can lead to climatic phenomena occurring in another part of the globe. ENSO refers to a large-scale circulation pattern that is influenced by ocean–atmospheric phenomena in the eastern tropical Pacific. The Southern Oscillation, the atmospheric component, is gauged by the Southern Oscillation index (SOI), which is the measure of the pressure difference between Tahiti and Darwin, Australia, and which is known to affect the strength of the tropical Pacific trade winds. El Niño represents the oceanic component and is measured by the sea surface temperature (SST) anomaly in the tropical Pacific Ocean.

The term ENSO embraces the interaction of atmosphere and ocean. ENSO activity, as well as that of other teleconnections such as the Madden–Julian oscillation (Lau et al. 1988), is forced by changes in the tropical SSTs and tropical convection (Mo and Livezey 1986; Barnston and Livezey 1991), which also affect jet stream locations and intensity, storm tracks, temperature, and rainfall. For example, as the water temperature increases off the South American coastline, more water is released into the atmosphere in the western hemisphere, providing a huge increase in the source of water available in the West for precipitation. By extension, tracking the SSTs in the tropical Pacific can foreshadow climatic conditions in parts of California and the northwestern United States.

The 700-mb height anomaly embraces a number of persistent large-scale pressure patterns that cover the globe and reflect the dynamics of the chaotic atmospheric system, including large-scale changes in the atmospheric wave and jet stream patterns. These indices represent a relationship in atmospheric pressure similar to that found for the SOI. Each index is a pressure oscillation between two distant points in the earth's atmosphere. For purposes of this study, we are concerned with all 13 700-mb level patterns that lie over the Northern Hemisphere. Full descriptions and example pressure pattern pictures can be found through the U.S. National Oceanic and Atmospheric Administration (NOAA) Climate Analysis Center Web site http://www.cpc.ncep.noaa.gov/data/teledoc.

4. Background

Although few studies applying ANNs specifically to rainfall prediction have been reported (Thirumalaiah and Deo 1998), the use of ANNs in the field of water resources overall in the past decade has grown extensively. ANN results are found to be superior when compared with those of other models using such methods as linear regression, persistence, or nowcasting. The mathematical process used by ANNs is well documented, but the physical meaning of each step used by ANNs to manipulate data to arrive at their results is not known and is constantly being researched, nor does the continued development of the ANN model give us any additional insight into the real process being studied. Yet the research referred to below indicates that ANNs have indeed great potential as an effective modeling tool.

Venkatesan et al. (1997) compared three neural network models to predict river elevations for total summer and two days in advance using upstream river elevation data and the river elevation data at the gauge in question for the training. The network was able to accurately (r = 0.88–0.98) predict river elevations at the gauge in question. It was found that omitting the upstream data did not significantly decrease prediction results.

Muttiah et al. (1997) compared ANN predictions of 2-yr peak stream discharges with the output from multivariate linear regression models for 17 river basins in the continental United States. Results for the ANNs were comparable with those from the regression model.

Probably the most studied application of ANNs to the field of hydrology is that for modeling the rainfall–runoff process. All studies are similar in methodology and results. For example, Hsu et al. (1995) developed an ANN model to study the rainfall–runoff process in the Leaf River basin, Mississippi. The network was compared with conceptual rainfall–runoff models, such as Hydrologic Engineering Center (HEC)-I and the Stanford Watershed Model, and to linear time series models. The ANN was found to give the best one-step ahead predictions. The authors emphasize that the ANN has an unknown basis in the physical properties of a watershed and that, from a conceptual standpoint, the modeling technique may not be desirable.

Ozelkan et al. (1996) used 500-mb atmospheric circulation patterns to predict precipitation one month in advance. Their study used a fuzzy rule-based model and circulation pattern indices developed by the researchers for the prior 12 months to predict a month's precipitation in various areas of Arizona. When results were compared with those from a regression model, the fuzzy rule-based model gave better results than did the regression model. January prediction was better than July, probably because July is the monsoon season in the study area and local factors may control.

These studies show that artificial neural networks have a wide range of applicability and ability in the area of water resources and that improvement over many accepted models and techniques is possible.

5. Data

Data in this study are based on two types of teleconnections: ENSO and the 700-mb height anomaly. Complete datasets for the indices are available from 1951 to the present.

a. ENSO

ENSO data in this study were obtained from the NOAA Climate Analysis Center. The data on SOI are measured following the method developed by Ropelewski and Jones (1987). The El Niño data included in this study are from the Niño-1+2 lying in 0°–10°S, 90°–80°W, and the Niño-3.4 lying in 5°N–5°S, 170°–120°W. All data are available on the Internet (NOAA 1998a).

b. 700-mb height data

NOAA has identified 13 teleconnection patterns in the Northern Hemisphere using a rotated principal component analysis of the 700-mb height data (NOAA 1998b). Following is a list of the prominent teleconnection patterns and the regions and months of effect.

Prominent patterns over the North Atlantic:

- North Atlantic Oscillation pattern (NAO) exists in all months.
- East Atlantic pattern (EA) exists from September to April.
- East Atlantic jet pattern (EA Jet) exists from April to August.

Prominent patterns over Eurasia:

- East Atlantic–Western Russia pattern (EATL–WRUS) exists from September to May.
- Scandinavia pattern (SCAND) exists in all months except June and July.
- Polar–Eurasia pattern (POL) exists from December to February.
- Asian summer pattern (ASU) exists from June to August.

Prominent patterns over North Pacific–North America:

- West Pacific pattern (WP) exists in all months.
- East Pacific pattern (EP) exists in all months except August and September.
- North Pacific pattern (NP) exists from March to July.
- Pacific–North American pattern (PNA) exists in all months except June and July.
- Tropical–Northern Hemisphere pattern (TNH) exists from November to January.
- Pacific transition pattern (PT) exists from May to August.

The East Atlantic–Western Russia pattern and the Scandinavia pattern are referred to by Barnston and Livezey (1987) as the Eurasia-2 and Eurasia-1 patterns, respectively. The data run from 1950 to the present. The index values are available at http://www.cpc.ncep.noaa.gov/data/teledoc.

6. Methodology

This study is unique with regard to the input data used, so no other model results are available with which to compare accuracy or improvement. Model results can only be compared with the observed values. There are too many inputs and not a sufficiently long dataset for a linear regression model.

Even though the network produces predictions, another goal is to determine which of the teleconnections

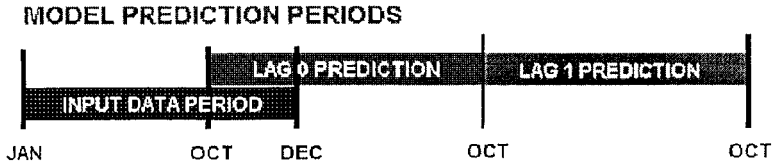

FIG. 4. Prediction periods used in the development of the neural networks. The input data are monthly index values for the calendar year. The predicted output is the total precipitation for the water year, Oct–Sep.

are the most important in each zone for precipitation prediction and which interactions, if any, among the teleconnections are dominant. Because of this second goal, some aspects of the network training differ from traditional network training to enhance the ability to extract useful information from the trained network. All the recognized Northern Hemisphere teleconnection indices are included in the study to see if the network is able to discern useless information from important parameters.

a. Development of a predictive neural network

The first step is to justify the use of neural networks and the teleconnection indices for precipitation prediction. The networks consist of an input layer with nodes for each monthly value of the teleconnection indices, a hidden layer, and an output layer of a single node for the predicted (October–September) precipitation. The ''lag-0'' network is trained using the monthly data for a given year (January–December) to predict the following water year (October–September) precipitation; for example, the input of January–December 1990 predicts the precipitation for October 1990–September 1991. The ''lag-1'' network is trained to predict the next water year precipitation; that is, input of January–December 1990 predicts the precipitation for October 1991–September 1992 (Fig. 4).

The general procedure for network development is to choose a subset of the data containing the majority of the cases, train a network, and test against the remaining cases (Masters 1993). The first networks were trained using data for 1951–1987 and tested against 1988–1996. The number of hidden nodes were varied to determine the best network structure. An ANN is trained until the root-mean-square error is below a small error tolerance, in this case 0.0001 in. of rainfall. Because of its importance to California's water supply, we chose Zone 3 (Fig. 1) as a test case. The number of hidden nodes was varied from 2, 4, 8, 12, 16, and then 20. Each increase in the number of hidden nodes showed a slight increase in predictive ability. After 12 nodes, the rate of increase decreased. Increasing from 20 to 50 nodes showed only slight improvement, so 12 nodes were chosen as the network size for the rest of the study. After training, the network for each zone–lag combination was run with the complete dataset and the predictions for each year were recorded and compared with observed values.

b. Development of the networks for data extraction

In this study, again two separate neural networks for each zone are trained using different lag times to predict the total water year precipitation. The network architecture and prediction periods are the same as before. The purpose of this network's development is to contain as much information as possible determined from as few examples as possible.

The first step before training is to determine which dataset will be used for developing the ANN. Normally, many examples of the process being modeled are available, so a subset of the available examples is taken at random for training, with the hope that examples of all important patterns are included. With only 45 yr of complete data for the inputs, however, using a randomly selected training set does not guarantee that all or even most patterns will be presented during the training period. To determine that the training set is fully comprehensive without adding too much personal bias to the training (Masters 1993), small subsets of 5 yr of consecutive data (1951–55, 1956–60, 1961–65, etc.) are used to train a series of networks. We again chose climate Zone 3 (Fig. 1) for developing the 5-yr subset models. The years successfully predicted by each mini-network are recorded. The process is repeated to ensure that the ANN is finding patterns in the data and not giving random results. The results show that the first 15 yr (1951–55, 1956–60, and 1961–65) of data predicted all of the remaining years except for some years that were never predicted correctly by any subset. The training set was set to these 15 yr (1951–65) plus the anomalous yr (1974, 1975, 1980, 1981, 1984, 1985, 1992, 1993) that were never predicted correctly. Twenty-three years of data were used to train the network to predict 45 yr of precipitation. Other combinations of the 5-yr training sets could also predict all years correctly, but the set chosen used the smallest number of years.

Again, we varied the number of hidden nodes to determine the best network structure with the new training set. Again, using Zone 3 as a test case, the number of hidden nodes was varied from 2, 4, 8, 12, 16, and then 20. Each increase in the number of hidden nodes showed a slight increase in predictive ability. Twelve nodes were chosen as the network size for the study. After training, the network for each zone–lag combination was run with the complete dataset and the predictions for each year were recorded and compared with observed values. In

TABLE 1. Individual training years for each lag and zone. The Xs represent years to be included in the training set.

	Lag 0							Lag 1						
Year	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1951	X		X	X						X		X		X
1952	X			X			X	X		X		X		X
1953	X			X						X	X	X		X
1954	X	X		X						X		X	X	X
1955	X			X	X					X		X		X
1956	X		X		X	X	X			X	X	X	X	X
1957	X	X	X	X	X	X	X			X		X	X	
1958	X		X	X		X	X			X		X		
1959	X		X			X	X			X		X	X	X
1960	X		X	X	X	X	X		X	X		X		
1961					X				X	X	X	X	X	
1962						X	X	X	X	X	X	X		
1963				X						X	X	X	X	
1964										X	X	X		X
1965				X	X	X	X		X	X	X	X	X	X
1966	X												X	
1967	X		X			X	X						X	X
1968			X	X			X							
1969		X		X										
1970					X								X	
1971			X	X	X	X		X	X		X		X	X
1972	X			X	X	X		X	X		X		X	
1973	X			X	X	X	X	X	X		X		X	
1974				X	X	X		X	X	X	X	X	X	
1975		X		X	X	X		X	X	X	X	X	X	
1976	X	X	X	X		X	X	X	X				X	X
1977	X	X		X			X	X			X		X	X
1978	X	X		X			X	X					X	X
1979	X	X	X	X				X						X
1980	X	X		X	X	X		X		X	X	X	X	X
1981	X	X	X		X	X	X	X	X	X	X	X	X	X
1982	X	X	X	X	X	X	X	X	X		X	X	X	
1983	X	X	X		X	X	X	X	X		X	X		
1984	X	X	X		X	X	X	X	X	X	X	X	X	X
1985	X	X	X		X	X	X	X	X	X	X	X	X	X
1986	X	X	X	X	X	X	X	X	X		X		X	X
1987	X	X	X		X	X	X	X	X		X			X
1988		X	X	X	X	X	X	X	X		X		X	X
1989		X	X	X	X	X	X	X	X		X			X
1990	X	X	X		X	X	X	X	X		X			X
1991														
1992										X				
1993										X				

this case, the correlation value is a proxy for the amount of information contained in the network, so the value is the correlation between the observed and predicted values for all years.

ANNs were also trained for each climate zone, omitting one variable at a time (i.e., using each combination of 16 of the 17 inputs). In addition to leaving out one variable, networks were also trained by, first, leaving out all ENSO indicators (SOI, Niño-1+2, and Niño-3.4) and, second, by using only ENSO indicators.

Training by the standard back propagation method is computationally slow, and the lack of time was preventing an exploration of different combinations of inputs. By modifying the network training method to use the conjugate gradient method (Masters 1993), training time for the 14 cases decreased by 95%. With this increased performance, it was possible to determine a different training set using the 5-yr subset method described earlier for each zone and lag. Taking these individualized training sets, networks were again trained for each zone and lag. The individualized training sets are shown in Table 1.

7. Results and discussion

a. Development of a predictive neural network

The correlations between the observed and predicted precipitation for the verification cases for the 12-node network were 0.66, 0.73, 0.55, 0.39, 0.68, 0.61, and 0.42 for the seven zones, respectively, for lag 1. The results for lag 0 were −0.07, 0.16, 0.17, 0.11, 0.6, 0.73, and 0.84. The lag-1 results show that the network can extract information necessary for prediction from the

TABLE 2. Correlation between observed and predicted precipitation with the training set determined using Zone 3, lag 1 as the basis.

	Zone						
Lag	1	2	3	4	5	6	7
0	0.59	0.63	0.65	0.64	0.55	0.64	0.63
1	0.85	0.83	0.85	0.78	0.80	0.64	0.60

indices and that a long-range predictive model is possible. The lag-0 results show that the 700-mb atmospheric circulation is not, in general, a dominant factor in shorter-term prediction. This difference between the results for the different lag periods was expected. The 700-mb circulation on a global scale brings weather to a region. As the prediction period shortens, local conditions begin to dominate and, if strong enough, override the weather patterns brought in by the 700-mb circulation.

b. Development of the networks for data extraction

The ANNs were easily able to learn the new training set and gave consistent results over the three training attempts. Table 2 shows the correlation between the best network predictions of the three developed for each zone and the observed precipitation for each zone for the different lag times. Example graphs of the results are in Fig. 4. The high lag-1 prediction correlations again show the ability of ANNs to produce good, long-range predictions. All zones except Zone 7 predict as well or better with lag 1 versus lag 0, although some zones were about equally good at short- and long-term prediction. Zone 7 is the desert area of California and gives a better short-term prediction, showing that the desert climate system has a more variable system that is easily changed by local changes in the atmosphere, while the other zones are more persistent.

The improved lag-0 predictions using the new training set show that for some years the 700-mb circulation is still a dominant force in the short term, but there are other important parameters that have not been identified. The change in the results show that in the short term it is common for local conditions to overcome the short-term 700-mb conditions. The years used in producing the refined training set must therefore contain some important information about California's precipitation that becomes muddied when most years are included in training because the unidentified parameters that affect the short-term prediction are not included in the training.

When the networks were trained for each climate zone by omitting one variable, the correlations for these networks were usually close to the full input networks (Table 3). The average absolute change in correlations by leaving out a variable during training was 0.023; the average change was 0.009. Table 4 shows those omitted variables that caused a change in correlations of one (***) or two (XXX) standard deviations from the total correlation change distribution. The changes in correlation do not appear to be random. Looking at the EP pattern, every zone except Zone 1 had more than one standard deviation change in error for lag 0, but no effect on the lag 1. Leaving out most variables affected only one of the lag periods, showing that the ANN was finding distinct influences of the indices for different zones and lags.

It also appeared that Niño-1+2 has no relationship to the California climate, which was expected because of Niño-1+2's very small area.

The correlations for the ANNs trained by leaving out all ENSO indicators or by including only ENSO indicators (Table 5) show that these variables alone are not enough for a complete prediction. The correlations also show that lag 1 is influenced more than lag 0 by ENSO, indicating the acknowledged time lag between the onset of El Niño and its influence in California.

Looking at Fig. 5, we see that prediction phase (the up–down motion) follows the observed pattern in most cases, but the magnitude is not predicted as well. The network usually predicts when there will be an increase or decrease in rainfall totals, but does not always predict

TABLE 3. Correlation between observed and predicted precipitation with the listed variable removed from the training set.

		Performance correlation with variable left out of training																	
Zone	Lag	NONE	SOI	N12	N34	NAO	EA	EAJ	WP	EP	NP	PNA	EAW	SCA	TNH	POL	PT	SZ	ASU
1	0	0.59	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
1	1	0.85	0.8	0.9	0.9	0.9	0.8	0.8	0.8	0.9	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.9	0.9
2	0	0.63	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.7	0.6	0.7
2	1	0.83	0.8	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
3	0	0.65	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.6	0.7	0.7	0.7	0.7	0.6	0.7	0.7	0.6	0.7
3	1	0.85	0.8	0.8	0.8	0.9	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.8	0.9	0.8	0.8
4	0	0.64	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.6	0.7	0.7	0.7
4	1	0.78	0.8	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.7	0.8
5	0	0.55	0.6	0.6	0.6	0.6	0.5	0.6	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
5	1	0.80	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
6	0	0.64	0.6	0.7	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6
6	1	0.64	0.6	0.6	0.6	0.6	0.5	0.6	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.7
7	0	0.63	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
7	1	0.60	0.6	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

TABLE 4. Correlation between observed and predicted precipitation with the listed variable removed from the training set; *** represents changes in correlation between one and two standard deviations, and XXX represents changes in correlation of more than two standard deviations.

		Performance correlation error with variable left out of training																	
Zone	Lag	NONE	SOI	N12	N34	NAO	EA	EAJ	WP	EP	NP	PNA	EAW	SCA	TNH	POL	PT	SZ	ASU
1	0	0.59				***				***	***			***		***			***
1	1	0.85													***				
2	0	0.63								***	***			***					***
2	1	0.83	***				***						***						
3	0	0.65				***	***			XXX				***					
3	1	0.85	***						***				XXX			***			
4	0	0.64				***	***			XXX									
4	1	0.78	***		XXX		XXX	***	XXX				***			***		***	
5	0	0.55	***		***	***		***	***	***			XXX	XXX	***	***	***	***	
5	1	0.80	***				***		***	***		***			***				
6	0	0.64	***		***		***			XXX		XXX			***				
6	1	0.64			XXX	***	XXX	***	XXX	***			***	***				***	
7	0	0.63			XXX		XXX			XXX		***	***					***	***
7	1	0.60	***		XXX	***	XXX	***	XXX	***		***	***	***		***	***	***	

the magnitude of increase or decrease. Zone 6, lag 1 shows this clearly. This indicates that the amount of moisture available for precipitation is not being predicted correctly. In this study, the only indicator of moisture available for precipitation is the tropical Pacific SST (El Niño). Also, since the lag-1 predictions are made before an El Niño is fully developed, the true magnitude and, therefore, the effect, of the El Niño event cannot be accurately predicted. Including other sources, such as North Pacific SST, and measures of moisture, such as the Palmer Drought Severity Index, might increase the accuracy. The lag-1 prediction for Zone 6 for the period October 1997–September 1998 is 24.9 in. or about 10 in. above the average. The prediction that would have been available January 1997 and made prior to the development of the 1997–98 El Niño, one of the largest on record, would have predicted much greater than average rainfall. The actual amount received was 34.4 in., an amount even greater than the prediction. The 1997–98 El Niño was a great source of atmospheric water, and although the ANN indirectly predicted the El Niño, it underpredicted its effects.

Another point that must be considered is the effect of the choice of the training set. The analysis thus far uses the networks trained on the training set developed for Zone 3, lag 1. If the important patterns for prediction for the other zones are not represented in that training set, then results may be poor. Using an optimized training set for each zone and lag improves the overall performance (Table 6). Except for Zone 7, no significant improvements in lag-1 predictions were observed, showing that long-range prediction patterns are fairly homogeneous and leading us to believe that a simple relationship between the indices and precipitation exists. The significant lag-1 increase in Zone 7 indicates that it has a climate quite different from those of the other six zones and therefore may need to be handled individually.

Lag-0 prediction improvements show that very different patterns are necessary for long- and short-range predictions. The training set developed for Zone 3, lag 1 is sufficient for Zones 1–6, lag 1. Further research will be needed to ascertain if a single training set can be used for all the zones with lag 0 except, perhaps, Zone 7.

When considering the overall performance of the ANN models, we find that those with the larger number of hidden nodes predict better. A larger number of hidden nodes enable the ANN to separate more patterns, although the difference between patterns becomes less and less distinct as the number of nodes increases. At the limit, the network could specialize and match only the exact patterns presented. The difficulty lies in balancing over- and undergeneralization to give an accurate prediction.

TABLE 5. Correlation between observed and predicted precipitation with ENSO variables removed from the training set and with only ENSO variables used for training.

	Lag	1	2	3	4	5	6	7
	0	0.62	0.62	0.67	0.62	0.57	0.56	0.58
No ENSO	1	0.80	0.74	0.73	0.70	0.20	0.56	0.47
	0	0.12	0.15	0.32	0.16	0.13	0.10	0.14
Only ENSO	1	0.55	0.59	0.63	0.49	0.48	0.24	0.32

8. Conclusions

The results show the successful application of ANNs to long-range prediction of precipitation. The results so far indicate the following.

1) The ANN model results show agreement with the accepted views on the significant influence of ENSO.
2) The ANN model results agree that Niño-1+2 has no effect on California precipitation.
3) The ANN model results agree that a lag relationship exists between ENSO and its effects.

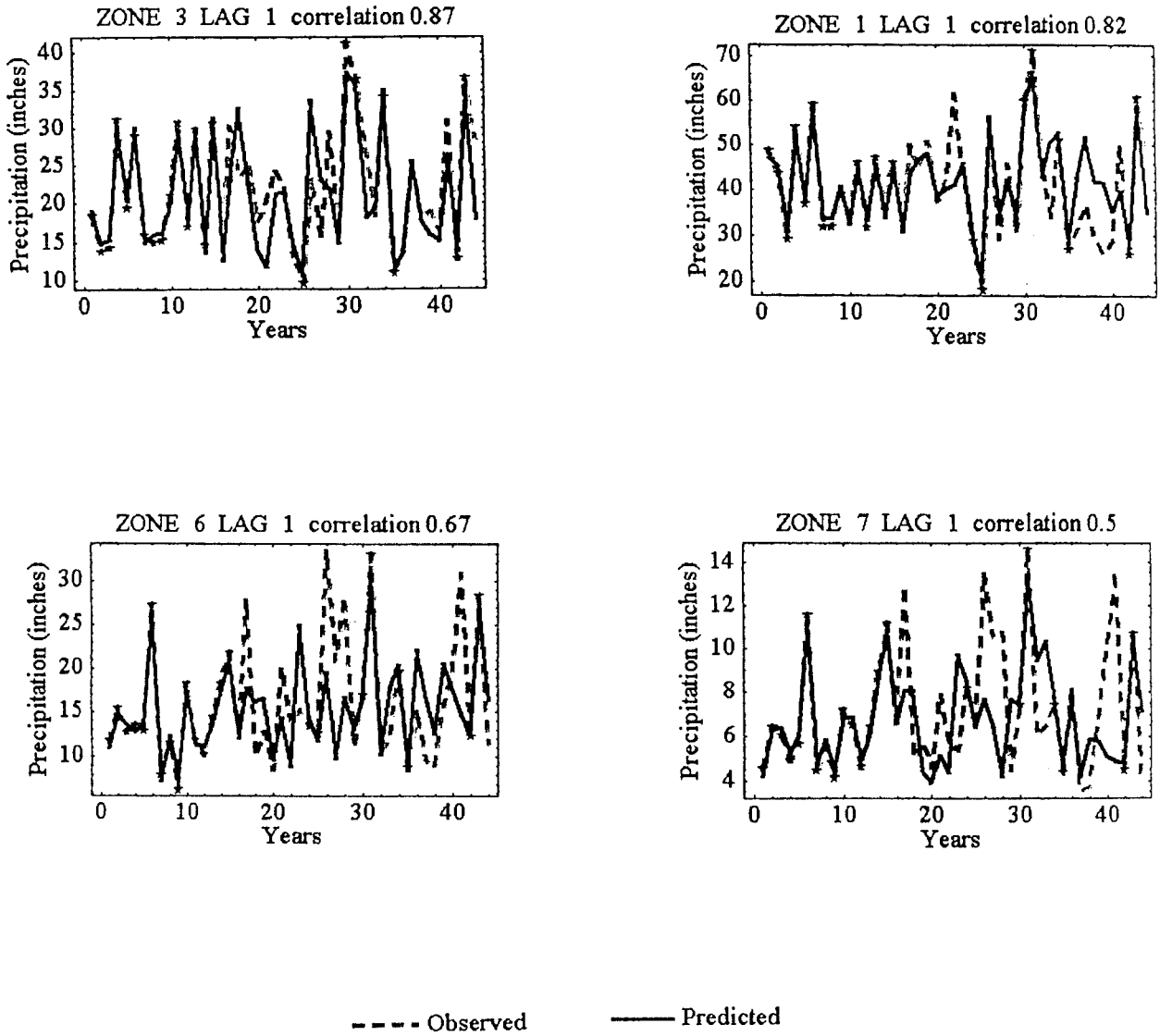

FIG. 5. Example prediction results.

4) The ANN model results indicate which zones are affected by which indices.
5) The ANN model results indicate the different effects of different lags.
6) The ANN model results predict phases well.

Although the phase of the precipitation was well predicted, the magnitude was only sometimes predicted. This may be because the magnitude of water sources such as El Niño had not been fully developed. Data on other water sinks and sources beside those from the tropical Pacific may need to be incorporated. In general, teleconnections are good indicators of precipitation. The larger number of hidden nodes in the ANN does give better results, but training time must then be increased and accuracy is decreased.

TABLE 6. Correlation between observed and predicted precipitation with the training set determined for each zone and lag independently.

	Zone:						
Lag	1	2	3	4	5	6	7
0	0.82	0.74	0.75	0.74	0.76	0.87	0.89
1	0.86	0.74	0.87	0.76	0.77	0.71	0.81

REFERENCES

Barnston, A., and E. J. Livezey, 1987: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115,** 1082–1126.

——, and ——, 1991: Statistical prediction of January–February mean Northern Hemisphere lower-tropospheric climate from the 11-year solar cycle and the Southern Oscillation for west and east QBO phases. *J. Climate*, **4,** 249–262.

French, M. N., W. F. Krajewski, and R. R. Cuykendall, 1992: Rainfall

forecasting in space and time using a neural network. *J. Hydrol.*, **137,** 1–31.

Hinton, G. E., 1992: How neural networks learn from experience. *Sci. Amer.*, (9), 144–151.

Hsu, K., H. V. Gupta, and S. Sorooshian, 1995: Artificial neural network modeling of the rainfall–runoff process. *Water Resour. Res.*, **31,** 2517–2530.

Kahya, E., and J. A. Dracup, 1993: U.S. streamflow patterns in relation to the El Niño–Southern Oscillation. *Water Resour. Res.*, **29,** 2491–2503.

Lau, N.-C., I. M. Held, and J. D. Neelin, 1988: The Madden–Julian oscillation in an idealized general circulation model. *J. Atmos. Sci.*, **45,** 3810–3832.

Masters, T., 1993: *Practical Neural Network Recipes in C++*. Academic Press, 493 pp.

Mo, K. C., and R. E. Livezey, 1986: Tropical–extratropical geopotential height teleconnections during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **114,** 2488–2515.

Muttiah, R. S., R. Srinivasan, and P. M. Allen, 1997: Prediction of two-year peak stream discharges using neural networks. *J. Amer. Water Resour. Assoc.*, **33,** 625–650.

Navone, H. D., and H. A. Ceccatto, 1994: Predicting Indian monsoon rainfall: A neural network approach. *Climate Dyn.*, **10,** 305–312.

NDMC, cited 1998: Predicting drought. [Available online at http://enso.unl.edu/ndmc/enigma/predict.htm.]

NOAA, cited 1998a: CPC-data: Current monthly atmospheric and SST index values. [Available online at http://www.cpc.ncep.noaa.gov/data/indices/.]

——, cited 1998b: Teleconnection pattern calculations. [Available online at http://www.cpc.ncep.noaa.gov/data/teledoc/telepatcalc.html.]

Ozelkan, E. C., F. Ni, and L. Duckstein, 1996: Relationship between monthly atmospheric circulation patterns and precipitation: Fuzzy logic and regression approaches. *Water Resour. Res.*, **32,** 2097–2103.

Raman, H., and N. Sunilkumar, 1995: Multivariate modeling of water resources time series using artificial neural networks. *Hydrol. Sci. J.*, **40,** 145–163.

Redmond, K. T., and R. W. Koch, 1991: Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices. *Water Resour. Res.*, **27,** 2381–2399.

Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño–Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114,** 2352–2362.

——, and ——, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115,** 1606–1626.

——, and P. D. Jones, 1987: An extension of the Tahiti–Darwin Southern Oscillation index. *Mon. Wea. Rev.*, **115,** 2161–2165.

——, and M. S. Halpert, 1989: Precipitation patterns associated with the high index phase of the Southern Oscillation. *J. Climate*, **2,** 268–284.

Thirumalaiah, K., and M. C. Deo, 1998: River stage forecasting using artificial neural networks. *J. Hydrol.*, **3,** 26–31.

Venkatesan, C., S. D. Raskar, S. S. Tambe, B. D. Kulkarni, and R. N. Keshavamurty, 1997: Prediction of all India summer monsoon rainfall using error-back-propagation neural networks. *Meteor. Atmos. Phys.*, **62,** 225–240.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the potential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109,** 784–812.