

# WeeklyNote

2019.10.06

張慕琪

## **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)

# Introduction

尽管加利福尼亚州有7个不同的气候区域（通过降水的周期和降水量划分），这7个区域有共同且好定义的雨季（11-3月）。

加利福尼亚地区的降水和ENSO以及700-mb高度异常很有可能是非线性的，很难用传统的统计方法解决。

因此，本文主要有三个目标。第一，确定700-mb环流是驱动加利福尼亚地区降水的主要因素；第二，神经网络可以判断所有数据中占主导地位的部分；第三，证明一个训练好的神经网络可以较好的进行预测。

# Methods

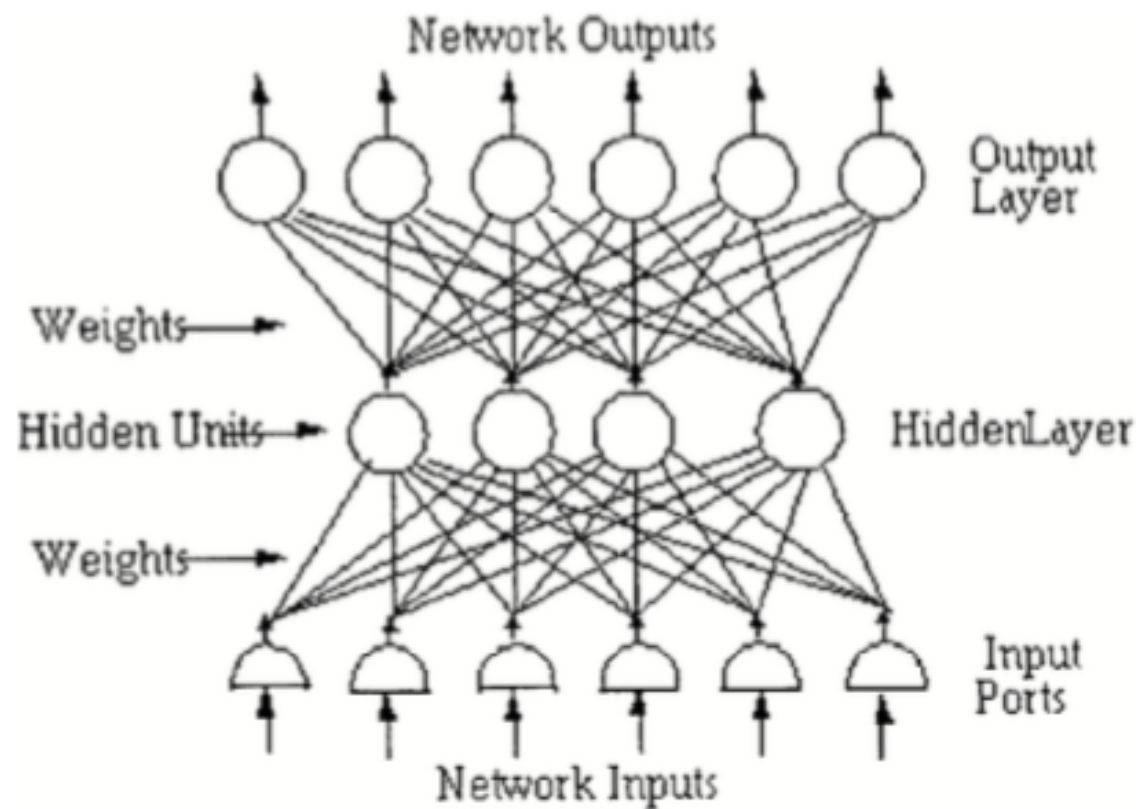


FIG. 2. The ANN feed forward-back propagation network is composed of an input layer of neurons, one or more hidden layers, and an output layer. Each layer comprises multiple units connected completely with the next layer, with an independent weight attached to each connection.

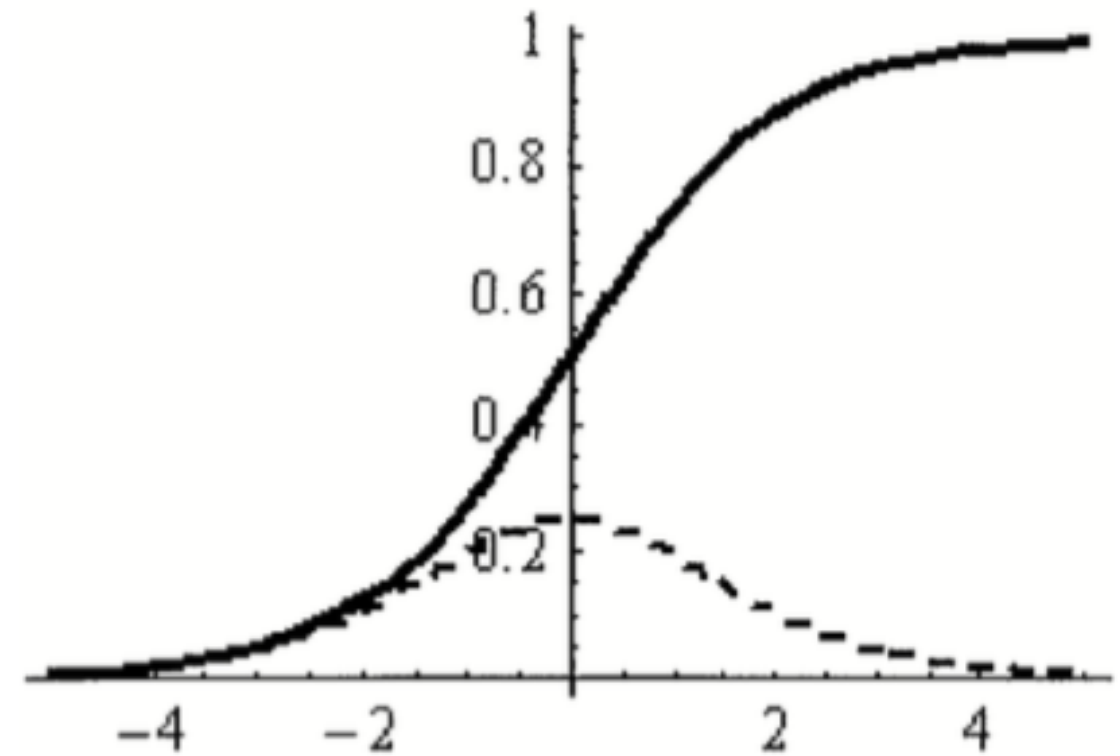


FIG. 3. The sigmoid neuron activation function " $(1 + e^{-x})^{-1}$ " (solid) and its derivative (dashed). The sigmoid function determines the activation level of a neuron based on the weighted inputs. The derivative is used during network development to determine the level of training a weight connection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

# Methods

1. 研究地区：美国加利福尼亚地区的7个区域。
2. 研究时间：1951-1987(train) & 1988-1996(test)
3. 数据来源：NOAA Climate Analysis Center.
4. 研究变量：Precipitation
5. 使用模型：输出为全年降水总量，输入12（月）\*17（遥相关变量）个节点，隐藏层节点的数量由RMSE决定。



FIG. 1. California's seven climate zones.

# Methods

关于700-mb高度数据：

NOAA使用旋转主成分分析方法(rotated principal component analysis)对于北半球的700-mb高度数据最终得到了13个遥相关变量：

全月均存在的北大西洋涛动（NAO）；

9-4月的东大西洋模式（EA）；

4-8月东大西洋jet pattern；

9-5月的东大西洋-俄罗斯西部模式（EATL-WRUS）；

除去6月和7月的斯坎蒂维娅模式（SCAND）；

12-2月的极地-欧亚大陆模式（POL）；

6-8月的亚洲夏季模式（ASU）；

全月均存在的西太平洋模式（WP）；

除8月和9月均存在的东太平洋模式（EP）；

3-7月的北太平洋模式（NP）；

除6月和7月存在的太平洋-北美模式（PNA）；

11-1月的热带-北半球模式（TNH）；

5-8月的太平洋过渡模式（PT）。

# Methods

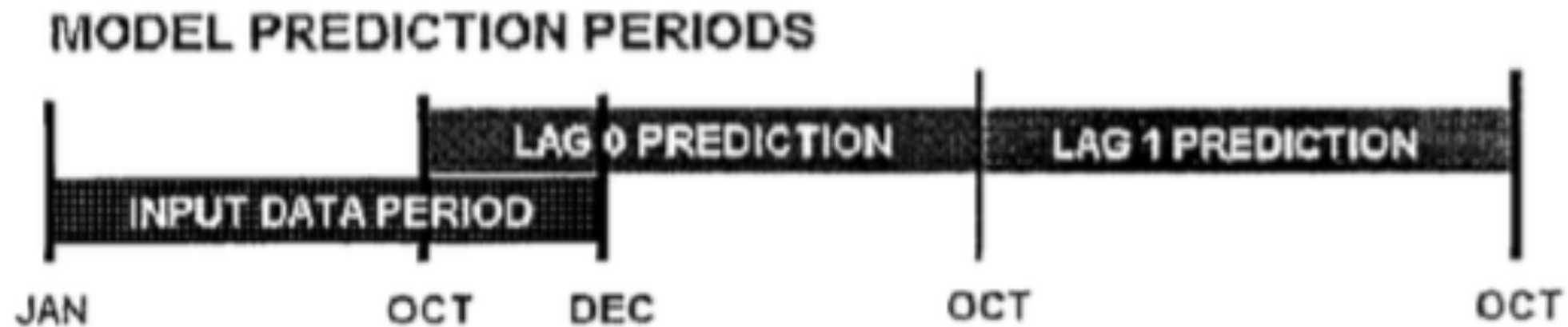


FIG. 4. Prediction periods used in the development of the neural networks. The input data are monthly index values for the calendar year. The predicted output is the total precipitation for the water year, Oct–Sep.

两种网络：lag0，lag1

加利福尼亚地区具有重要的水源供应作用，针对Zone3的lag1训练网络，最终在应用到其他6个区域中。

隐藏节点的数量由2，4，8，12，16，20，50逐渐增加。节点的增加意味着预报能力的微小提升。在增加的过程中发现，在12个节点后，网络能力的提升几乎没有太大改变，因此之后的隐藏层节点均设置为12。

# Meth

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

| Year | Lag 0 |   |   |   |   |   |   | Lag 1 |   |   |   |   |   |   |
|------|-------|---|---|---|---|---|---|-------|---|---|---|---|---|---|
|      | 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 |   |   |   |   |

数据集的  
随机选择



# Results

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

| Lag | Zone |      |      |      |      |      |      |
|-----|------|------|------|------|------|------|------|
|     | 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 |

lag1模型中的高分表明ANNs有能力进行长期的降水预测；尽管有些地区lag1和lag0表现相同，但除了第7区的其他所有的地区在lag1的预测比lag0好，这也表明ANNs具有较好的长期预测降水能力；第7区市加利福尼亚的沙漠地区，在短期预测上表现更佳，表明沙漠气候系统有更多的变化，而其他地区的局地改变较少。

# Results

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    |
|-----------|-----|------|------|------|------|------|------|------|
| No ENSO   | 0   | 0.62 | 0.62 | 0.67 | 0.62 | 0.57 | 0.56 | 0.58 |
|           | 1   | 0.80 | 0.74 | 0.73 | 0.70 | 0.20 | 0.56 | 0.47 |
| Only ENSO | 0   | 0.12 | 0.15 | 0.32 | 0.16 | 0.13 | 0.10 | 0.14 |
|           | 1   | 0.55 | 0.59 | 0.63 | 0.49 | 0.48 | 0.24 | 0.32 |

相关性表格同样表示lag1受ENSO的影响比lag0更大，暗示了厄尔尼诺对加利福尼亚时间之后的影响。

# Results

可以看出，大部分地区预测相位和实际观测相位相同，但是ANNs对降水幅度的预测较差。ANNs总是可以预测什么时候降水总量增加或减少，但并不能总是较好预测具体增加或减少多少。（Zone6, lag1非常清楚地反映了这点）

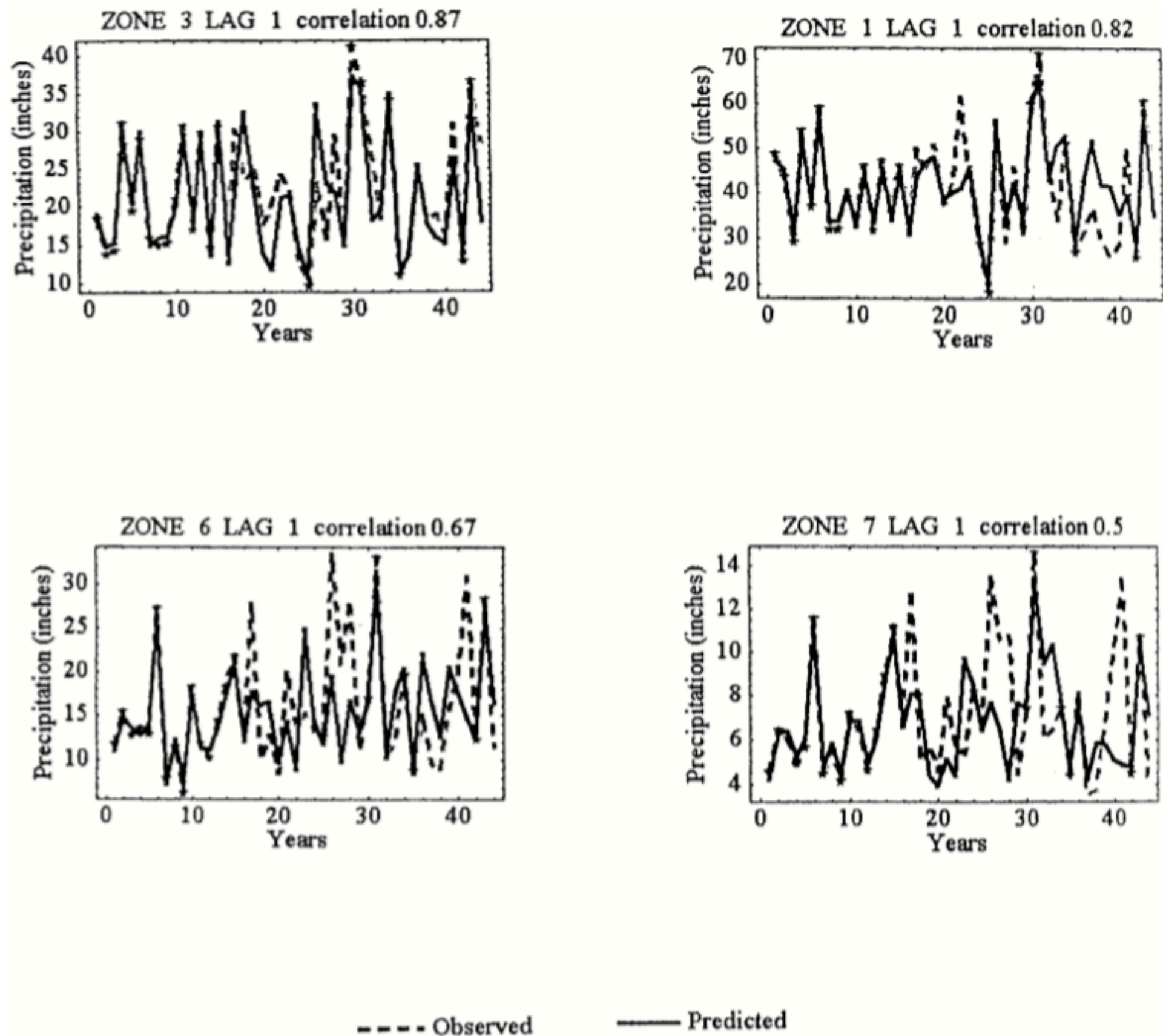


FIG. 5. Example prediction results.

# Results

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

| Lag | Zone: |      |      |      |      |      |      |
|-----|-------|------|------|------|------|------|------|
|     | 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 |

对每一个区域使用不同的优化数据集进行训练会有更好的表现。第7区在lag1上表现明显提升表明相较于其他地区，沙漠地区的气候在处理上更独立和特殊。

# Conclusions

1. ANN模型结果表明了ENSO确实存在重大影响；
2. ANN模型表明 $Ni\tilde{n}o-1+2$ 对加利福尼亚降水没有影响；
3. ANN模型结果表明ENSO和其影响具有延迟关系；
4. ANN模型结果也暗示不同变量对不同区域有不同的影响；
5. ANN模型结果表明其可以对相位有较好的预测。

谢谢