

GroupMeeting

2019.10.11

张慕琪

Papers

1. Downscaling temperature and precipitation: a comparison of regression-based method and artificial neural networks. J.T.SCHOOF and S.C.PRYOR, International Journal of Climatology, Int. J. Climatol. 21: 773-790(2001)
2. Spatial Interpolation of Surface Air Temperatures Using Artificial Neural Networks: Evaluating Their Use for Downscaling GCMs. SETH E. SNELL, SUCHARITA GOPAL and ROBERT K.KAUFMANN, 2 February, 1998
3. Artificial Neural Networks and Long-Range Precipitation Prediction in California. DAVID SILVERMAN and JOHN A.DRACUP, 18 August 1998.
4. How Neural Networks Learn from Experience. Geoffrey E. Hinton, Scientific American, September, 1992.

INTERNATIONAL JOURNAL OF CLIMATOLOGY

Int. J. Climatol. **21**: 773–790 (2001)

DOI: 10.1002/joc.655

DOWNSCALING TEMPERATURE AND PRECIPITATION: A COMPARISON OF REGRESSION-BASED METHODS AND ARTIFICIAL NEURAL NETWORKS

J.T. SCHOOF* and S.C. PRYOR

*Atmospheric Science Program, Department of Geography, Indiana University Student Building, 701 East Kirkwood Ave.,
Bloomington, IN 47405-7100, USA*

Received 9 May 2000

Revised 22 January 2001

Accepted 26 January 2001

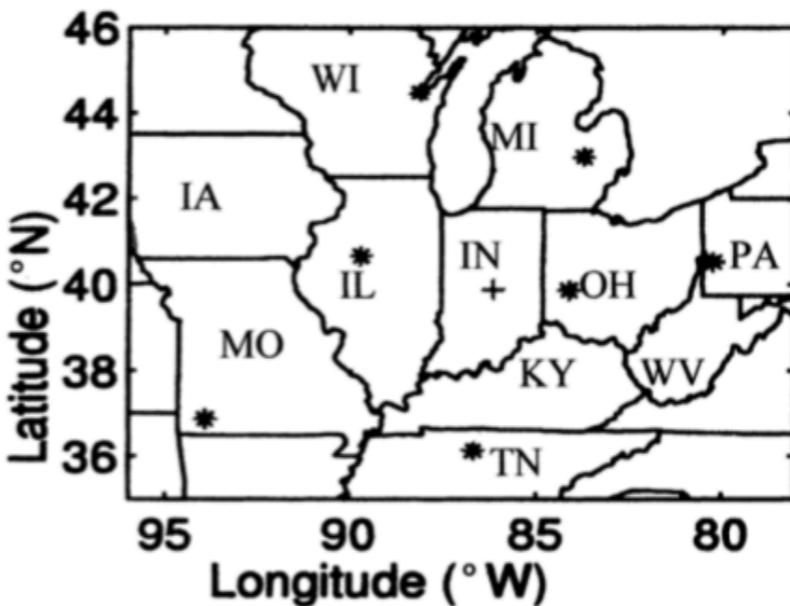
Introduction

研究农业生产所需要的天气尺度往往比现有模式的最高分辨率还要高。

高温持续时间的增加和低于正常水平的降水量很可能导致农作物的水汽环境改变（例如小麦和大豆），并因此对农业生产起到副作用。

Methods

1. 研究地区: from 35° to 45° N and from 80° to 95° W, includes Indiana (IN), Illinois (IL), Ohio (OH), Kentucky (KY), Tennessee (TN), Michigan (MI), Missouri (MO), western Pennsylvania (PA), and southeastern Wisconsin (WI)



该地区对于美国农业生产有很大作用

Figure 1. Map of study area showing rawinsonde stations (*) and Indianapolis, IN station (+)

Table I. Station information for rawinsonde stations

Station name	WMO station ID	Latitude (°N)	Longitude (°W)	Distance to Indianapolis (km)	Maximum length of data record (years)
Monett, MO	UMN 77349	36.88	93.90	720	26
Nashville, TN	BNA 72327	36.12	86.68	449	33
Flint, MI	ENT 72637	42.97	83.73	421	40
Peoria, IL	PIA 72532	40.67	89.68	325	37
Green Bay, WI	ORB 72645	44.48	88.13	529	51
Pittsburgh, PA	PIT 72520	40.53	80.23	520	32
Dayon, OH	DAY 72429	39.87	84.12	194	40

Methods

2. 研究时间：1973-1990（共18yr）

为确保研究中选取的这18年的时间段可以代表较长时期的特征（即18年是一个较长的时间范围），每个无线电探空台站地表温度和500hPa高度的数据都和更长时间序列的数据一一比较；

——使用Kolmogorov-Smirnov(K-S) test: 用以检验两个分布是否不同。

3. 研究变量：500hPa高度，700hPa高度，850hPa温度，海平面气压

(SLP)，柱平均相对湿度 (column average relative humidity, RH)

4. 数据来源：Solar and Meteorological Surface Observation

Network(SAMSON) dataset issued by the National Climatic Data

Center(NCDC/NREL, 1993)

Methods

1. Synoptic classification:

- 1) 去除年循环
- 2) 根据指标进行季节划分

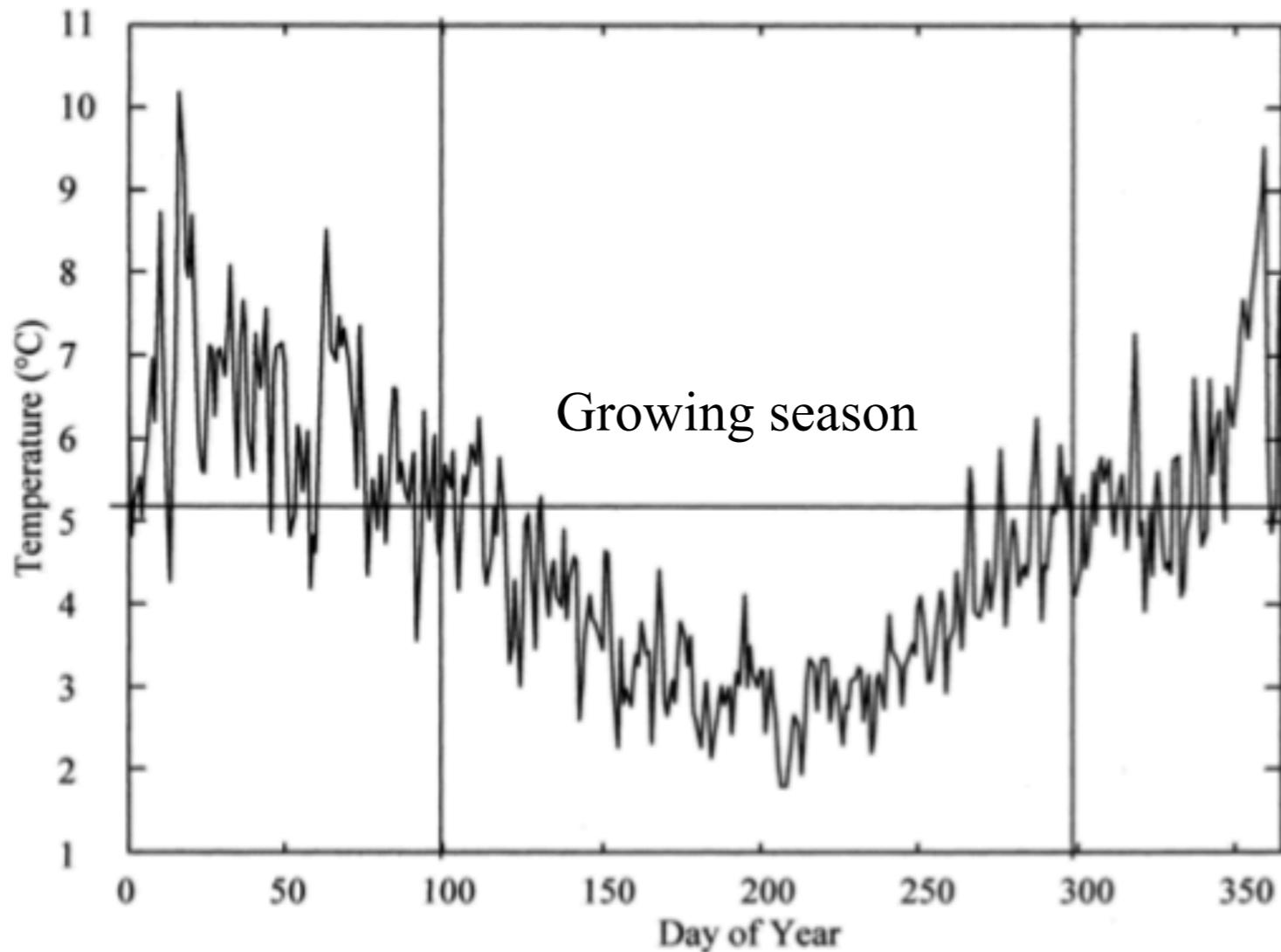


Figure 2. Standard deviation of daily surface air temperature for Indianapolis, IN. The horizontal line is the average standard deviation over the entire year. The seasons are determined by whether the mean daily standard deviation is above or below the annual mean

Methods

1. Synoptic classification:

3) 主成分分析 (PCA) : PCA的思想就是将n维特征映射到k维上 ($k < n$)，这k维是重新构造出来的全新维度特征，而不是简单的从n维特征去除 $n-k$ 维特征，这k维就是主成分。

2. Downscaling

1) Multiple regression

1. Multiple linear regression(MLR)用于日最高温和日最低温的降尺度，PCA得到的主成分为输入，温度变量为输出；

2. Poisson regression model用于降水的降尺度。

2) Artificial neural networks(ANNs)

- 往往隐藏层的数量是预测变量数量的1.5倍
- 后向传播学习算法有两个参数——学习速率和动量因子；学习速率决定在一次学习中权重可以改变多少，动量因子决定当前时间片内有权重受到前一次权重的影响

$$w_{i,j}(\text{new}) = w_{i,j}(\text{old}) + \eta \delta_i o_j + \alpha [\Delta w_{i,j}(\text{old})] \quad \delta_i = (t_i - o_i)o_i(1 - o_i)$$

Results

1. PCA：根据主成分分析得出，在生长季节有7个主成分，而非生长季节有5个主成分，相应的，这些主成分分别可以解释探空台站数据集86%和83%的变化。

PC1和温度相关，其中在500hPa、700hPa高度和850hPa温度表现最佳；PC2和压力相关，在海平面气压变量上表现良好；PC3和湿度成分相关，因此在相对湿度变量上表现最佳；其余的四个成分同样和温度、压力和湿度相关，但是表现为不同方面。P.S. 非生长季节的数据并不重要，不做考虑。

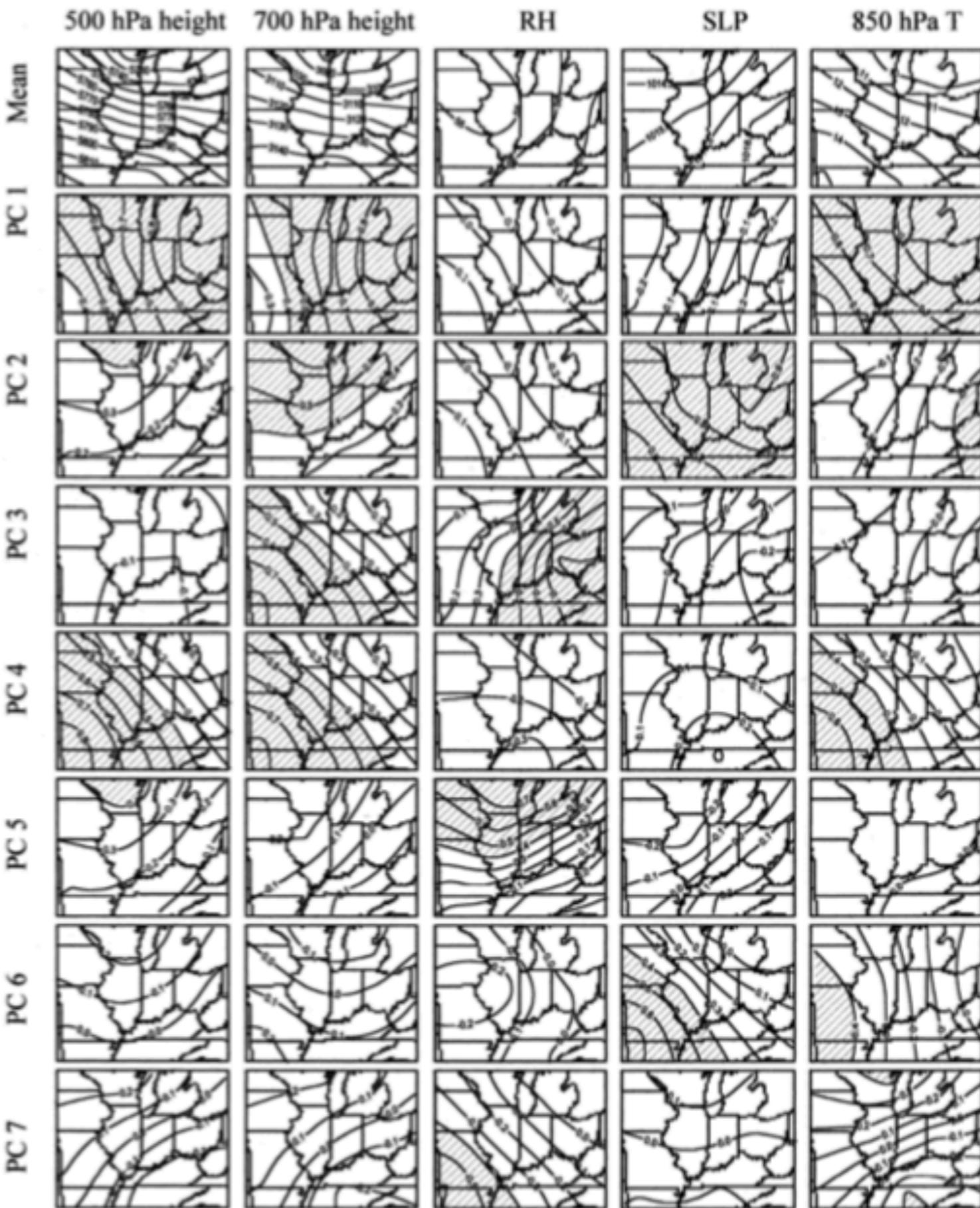


Figure 4. Growing season PC loadings. Values larger than ± 0.4 are shaded to aid in interpretation

Results

2. Downscaling: 75%的数据用于训练，25%用于测试。

1) Regression-based methods.

Table IV. Description of the regression models, input values and model forms^a

Predictand	Equation	Predictors
$T_{\max} \text{GS}$ (°C)	$T_{\max} \text{GS} = 1.86P_1 - 0.36P_2 + 0.63P_3 + 1.24P_4 - 0.25P_5 + 0.31P_6 - 0.83P_7 + 24.95$	00:00 GMT PC scores
$T_{\max} \text{GSAU}$ (°C)	$T_{\max} \text{GSAU} = 1.07P_1 - 0.41P_2 + 0.14P_3 + 0.52P_4 - 0.30P_5 + 0.07P_6 - 0.34P_7 + 0.76AU + 5.99$	00:00 GMT PC scores and an autoregressive term
$T_{\min} \text{GS}$ (°C)	$T_{\min} \text{GS} = 2.79P_1 - 0.54P_2 + 1.16P_3 + 1.18P_4 + 0.39P_5 - 0.66P_6 - 0.47P_7 + 13.70$	12:00 GMT PC scores
$T_{\min} \text{GSAU}$ (°C)	$T_{\min} \text{GSAU} = 1.17P_1 + 0.20P_2 - 0.15P_3 + 0.23P_4 + 0.54P_5 - 0.78P_6 + 0.21P_7 + 0.80AU + 2.74$	12:00 GMT PC scores and an autoregressive term
$T_{\max} \text{NS}$ (°C)	$T_{\max} \text{NS} = 1.93P_1 - 2.37P_2 - 0.16P_3 + 0.10P_4 - 0.53P_5 + 7.16$	00:00 GMT PC scores
$T_{\max} \text{NSAU}$	$T_{\max} \text{NSAU} = 1.22P_1 - 1.25P_2 - 0.22P_3 + 0.24P_4 - 0.10P_5 + 0.69AU + 2.30$	00:00 GMT PC scores and an autoregressive term
$T_{\min} \text{NS}$ (°C)	$T_{\min} \text{NS} = 3.80P_1 - 3.10P_2 - 0.15P_3 + 0.36P_4 - 0.32P_5 - 2.87$	12:00 GMT PC scores
$T_{\min} \text{NSAU}$	$T_{\min} \text{NSAU} = 2.15P_1 - 1.12P_2 - 0.21P_3 - 0.32P_4 + 0.11P_5 + 0.68AU - 0.94$	12:00 GMT PC scores and an autoregressive term
$\text{PREC}_d \text{GS}$ (mm)	$\text{PREC}_d \text{GS} = 0.04P_{12}1 - 0.30P_{12}2 + 0.71P_{12}3 - 0.09P_{12}4 + 0.20P_{12}5 - 0.10P_{12}6 - 0.45P_{12}7 - 0.10P_{00}1 - 0.07P_{00}2 + 0.19P_{00}3 + 0.14P_{00}4 - 0.09P_{00}5 + 0.02P_{00}6 + 0.04P_{00}7 + 0.84$	00:00 GMT PC scores ($P_{00} \#$) and 12:00 GMT PC scores ($P_{12} \#$)
$\text{PREC}_d \text{NS}$ (mm)	$\text{PREC}_d \text{NS} = 0.31P_{12}1 - 0.29P_{12}2 + 0.42P_{12}3 + 0.59P_{12}4 + 0.29P_{12}5 - 0.09P_{00}1 - 0.07P_{00}2 + 0.04P_{00}3 - 0.08P_{00}4 + 0.02P_{00}5 + 0.57$	00:00 GMT PC scores ($P_{00} \#$) and 12:00 GMT PC scores ($P_{12} \#$)

^a The predictors in the regression equations ($P \#$) indicate the PC scores. GS indicates the growing season and NS indicates the non-growing season. The term AU refers to the autoregressive term.

Results

2. Downscaling: 75%的数据用于训练，25%用于测试。
 1) Regression-based methods.

Table V. Model evaluation statistics for the regression models^a

Model	μ		σ		r	MAE	RMSE	RMSE _s	RMSE _u
	OBS	PRED	OBS	PRED					
$T_{\max}GS$	24.65	24.92	6.03	2.67	0.51	4.15	5.22	4.69	2.30
$T_{\max}GSAU$	24.65	24.88	6.03	5.25	0.85	2.31	3.18	1.58	2.76
$T_{\min}GS$	13.78	13.80	6.49	3.40	0.60	4.35	5.21	4.44	2.72
$T_{\min}GSAU$	13.78	13.85	6.49	5.52	0.88	2.39	3.08	1.63	2.61
$T_{\max}NS$	8.25	7.71	6.67	2.93	0.34	6.68	8.16	7.68	2.75
$T_{\max}NSAU$	8.25	7.46	8.67	6.64	0.81	4.04	5.17	3.39	3.90
$T_{\min}NS$	-3.09	-2.93	9.38	5.08	0.55	4.87	7.84	6.59	4.25
$T_{\min}NSAU$	-3.09	-3.66	9.38	9.59	0.80	3.23	6.09	1.83	5.81
PREC _d GS	2.69	0.66	7.19	1.12	0.36	2.97	7.16	7.09	1.04
PREC _d NS	2.76	0.49	6.06	0.99	0.50	2.72	6.07	6.01	0.86

^a The observed values are indicated by OBS, and the model predictions by PRE. The table includes the predicted and observed means (μ) and S.D.s (σ), Pearson product-moment correlation coefficient (r), MAE, RMSE, and the systematic and unsystematic components (RMSE_s and RMSE_u). The models are as described in Table IV.

PC scores 和 $T_{\max}GS$ 和 $T_{\max}GSAU$ 变量有较高的相关性；其中， $T_{\max}GS$ 和 PC3 和 PC4 有更高的相关性， $T_{\max}GSAU$ 和 PC1 和 PC4 有更高的相关性。对于 $T_{\min}GS$ ，其和 PC1、PC3 有更大相关性。

Results

2. Downscaling: 75%的数据用于训练，25%用于测试。
2) ANNs.

Table VI. Learning rate, momentum factor, number of hidden nodes and sample sizes for the training and testing of the ANNs

Model	Learning	Momentum	# Hidden nodes	Train	Test
T_{\max} GS (°C)	0.009	0.9	10	1896	632
T_{\max} GSAU (°C)	0.01	0.5	12	1896	632
T_{\min} GS (°C)	0.012	0.9	10	2455	671
T_{\min} GSAU (°C)	0.012	0.8	10	2455	671
PREC _d GS (mm)	0.05	0.3	18	1813	604
PREC _m GS (mm)	0.3	0.05	12	68	22
T_{\max} NS (°C)	0.007	0.9	8	1316	435
T_{\max} NSAU (°C)	0.007	0.7	9	1316	435
T_{\min} NS(°C)	0.012	0.9	7	1995	611
T_{\min} NSAU (°C)	0.012	0.9	8	1995	611
PREC _d NS (mm)	0.05	0.3	15	1230	409
PREC _m NS (mm)	0.001	0.1	15	68	22

Results

2. Downscaling: 75%的数据用于训练，25%用于测试。

1) Regression-based methods.

Table VII. Model evaluation statistics for the ANNs. Entries are analogous to those in Table V

Model	μ		σ		r	MAE	RMSE	RMSE _s	RMSE _u
	OBS	PRED	OBS	PRED					
T_{\max} GS	24.77	24.84	6.07	3.12	0.58	3.86	4.98	4.27	2.55
T_{\max} GSAU	24.77	24.66	6.07	5.13	0.85	2.40	3.18	1.70	2.68
T_{\min} GS	13.69	13.73	6.51	3.59	0.63	4.24	5.08	4.25	2.79
T_{\min} GSAU	13.69	13.87	6.51	5.60	0.89	2.27	2.95	1.53	2.53
T_{\max} NS	7.63	7.60	8.16	3.10	0.36	6.12	7.61	7.03	2.89
T_{\max} NSAU	7.63	7.18	8.16	6.49	0.80	3.81	4.87	2.98	3.85
T_{\min} NS	-2.84	-2.73	7.59	5.00	0.64	4.52	5.81	4.37	3.83
T_{\min} NSAU	-2.84	-2.82	7.59	6.92	0.87	2.76	3.76	1.57	3.42
PREC _d GS	2.84	3.43	8.17	0.05	0.09	4.71	8.19	8.19	0.05
PREC _d NS	2.79	2.63	6.01	0.06	0.11	3.80	6.01	6.00	0.06
PREC _m GS	80.27	90.98	66.6	40.38	0.65	42.18	52.00	41.89	30.82
PREC _m NS	57.18	68.02	28.78	13.28	0.54	22.88	26.60	24.15	11.16

ANNs日降水无法和泊松回归的结果相比较
(无配图)。

神经网络的月平均降水
数据表现的比日降水模
型好很多!

Conclusions

1. 对于日最高温变量来说，ANNs的表现要比MLR模型表现的好（table V和table VII比较），ANNs中的预报值和用于预报的值之间有更高的相关系数；
2. 同样，ANNs的温度模型有更小的偏差，因此具有更大的预报性；并且，神经网络往往对于平均值和标准差的预估更加准确；
3. 在table V和table VII中，非生长季节的日最高温模型并没有生长季节的模型表现的好；然而，线性回归模型比神经网络能更好的预估日最高温的平均值；无论是线性回归模型还是神经网络模型都无法捕捉方差；
4. 关于PC scores的降尺度分析表明其同时含有两个季节的日最低温信息；
5. 神经网络的 $T_{\min GS}$ 变量要比线性回归模型更好；在测试集中，神经网络的相关系数要比线性回归模型略高一些；

Conclusions

6. 无论是线性回归还是神经网络都无法成功捕捉日降水信息，两个模型各由各自的缺点导致；
7. 2000年，Cannon and Lord提出了一种直方图均等过程 (histogram equalization process) 可能可以帮助神经网络中这类情况的提高；
8. ANNs的月平均降水要远比日平均降水好很多；这说明ANNs可以学习到数据中普遍存在的规律，但无法捕捉到与数据相关的规律（这些规律偏离了月降水量的平均值）。

Spatial Interpolation of Surface Air Temperatures Using Artificial Neural Networks: Evaluating Their Use for Downscaling GCMs

SETH E. SNELL

Department of Geography, University of New Mexico, Albuquerque, New Mexico

SUCHARITA GOPAL AND ROBERT K. KAUFMANN

Department of Geography, Boston University, Boston, Massachusetts

(Manuscript received 2 February 1998, in final form 28 September 1998)

Introduction

如今GCMs的精度范围为 2.5° 2.5° ~ 8° (lat and long)。

仍有几个因素主导了GCMs进行降尺度的难度：气象要素例如温度和降水会随着地形而改变；大多数的变量在空间中并非是一阶或者二阶稳定的；大多数气象变量会受到其他气象条件的影响，例如湿度、压力场等。

因此，对GCMs进行传统的空间插值方法效果较差。

Methods

1. 研究地区：美国中部地区的11个台站
2. 研究时间：1931-1993 (63-yr)
3. 数据来源：National Oceanic and Atmospheric Association(NOAA) ground weather stations(NCDC 1994)
4. 研究变量：Tmax
5. 使用模型：多层前反馈后向传播ANN(multilayer feedforward backpropagation ANN)

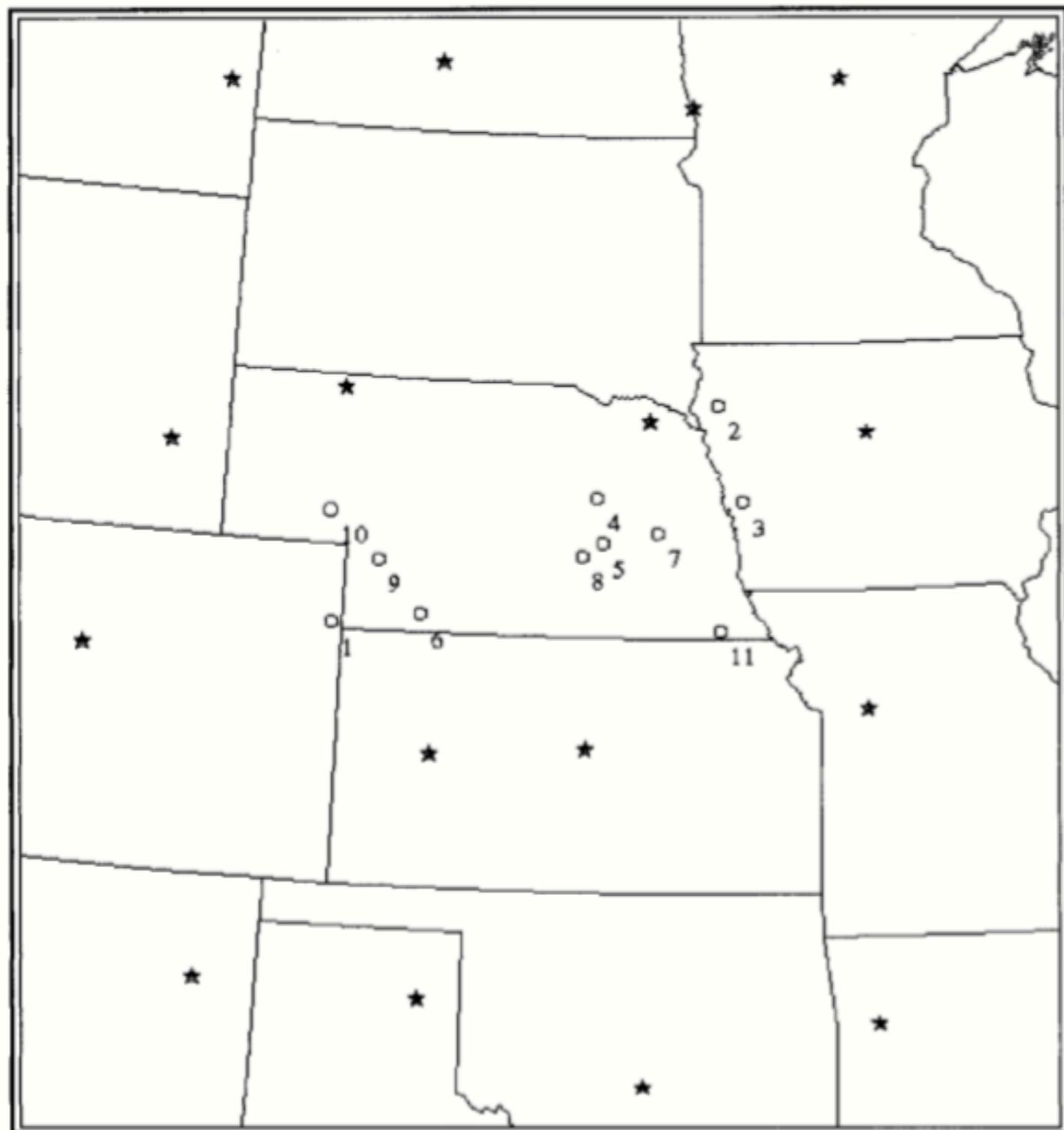


FIG. 1. Map of the study area. Stars depict grid of input NOAA weather stations. Open circles are 11 NOAA weather stations within the grid at which estimates are made.

Methods

1. 独立的两个网络 ($4 \rightarrow 1$, $16 \rightarrow 1$)
2. 激活函数使用sigmoid;
3. 隐藏层设置: 4-point ANN 有30个节点, 16-point ANN有54个隐藏节点;
4. 数据划分: 80% (14624 days)训练集+20% (2924 days)验证集, 随即产生, 每个模型分别运行5次。

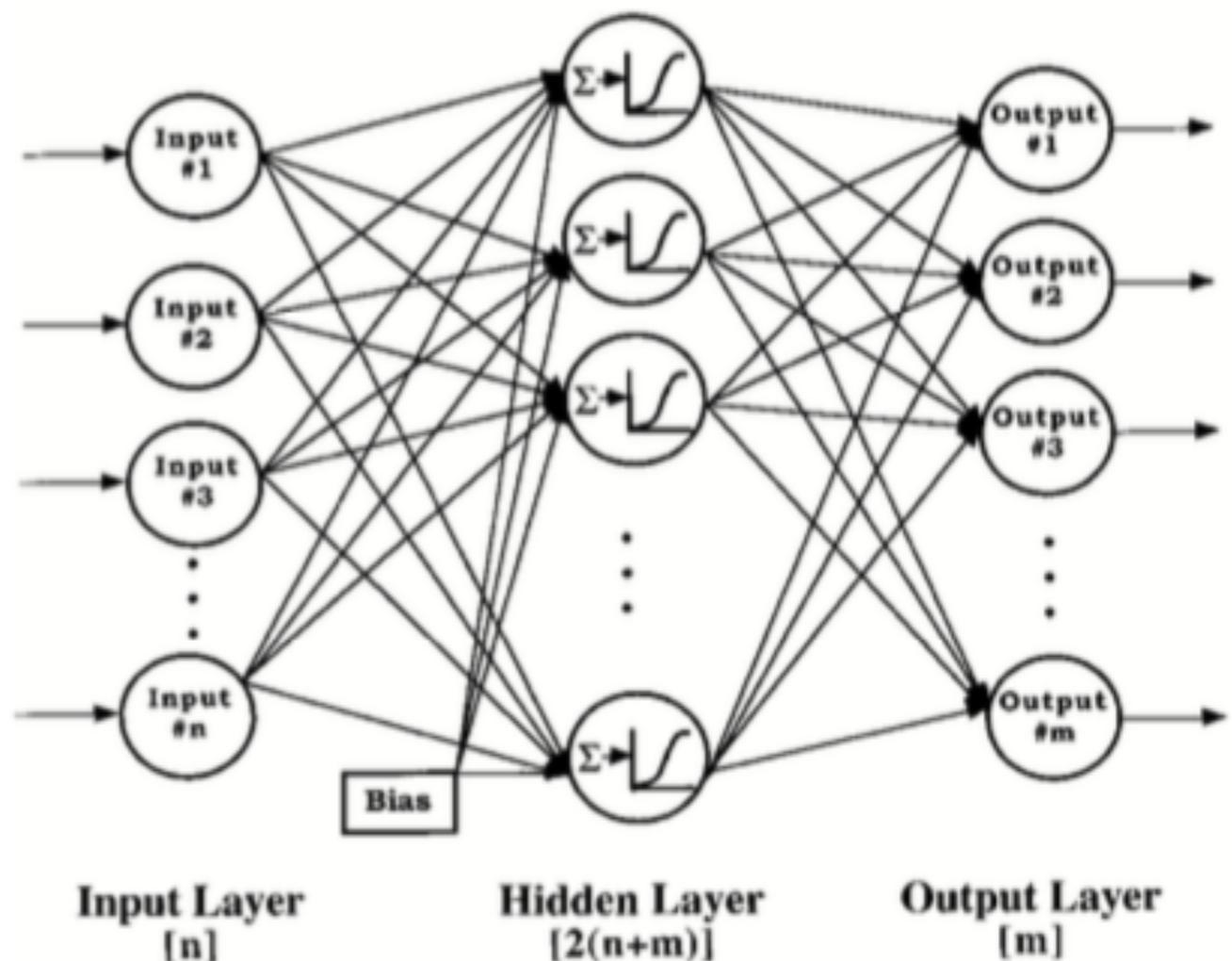


FIG. 2. ANN architecture used to interpolate maximum daily surface air temperature.

Results

16-point在station 6表现最差，而在station 2的表现最佳；
4-point对于station 6的表现也最差；
两个模型RMSE的表现和R²大体一致。

TABLE 1. Overall performance measures. RMSE and R² for test dataset using the 4-point and 16-point ANNs by station and on average across stations.

	4-point ANN		16-point ANN	
	rmse	R ²	rmse	R ²
Station 1	6.123	0.916	5.889	0.922
Station 2	5.230	0.954	4.131	0.971
Station 3	5.409	0.947	4.125	0.969
Station 4	6.741	0.917	6.036	0.934
Station 5	4.905	0.956	4.664	0.960
Station 6	6.929	0.905	6.697	0.911
Station 7	6.158	0.932	5.340	0.949
Station 8	5.313	0.949	5.110	0.953
Station 9	4.994	0.951	4.761	0.955
Station 10	5.566	0.932	5.064	0.944
Station 11	5.268	0.946	4.488	0.961
Average	5.694	0.937	5.119	0.948
Minimum	4.905	0.905	4.125	0.911
Maximum	6.929	0.956	6.697	0.971

Results

TABLE 2. Test results for systematic bias. Regression coefficients from Eq. (1) for each station for the predicted output from the 4-point and 16-point ANNs. Bracketed values are p values for t statistics testing significant differences for α and β from 0 and 1, respectively. Values significant at $p < 0.05$ are in boldface.

	4-point ANN		16-point ANN	
	α	β	α	β
Station 1	0.464 [0.243]	0.992 [0.156]	0.693 [0.068]	0.989 [0.043]
Station 2	-1.380 [0.000]	1.007 [0.100]	-0.163 [0.434]	1.004 [0.261]
Station 3	-0.514 [0.076]	0.993 [0.123]	0.273 [0.213]	0.999 [0.819]
Station 4	-0.504 [0.166]	0.988 [0.029]	0.211 [0.518]	0.997 [0.564]
Station 5	-1.735 [0.000]	1.003 [0.459]	-0.148 [0.566]	1.001 [0.811]
Station 6	0.091 [0.828]	0.991 [0.151]	-0.063 [0.876]	0.998 [0.750]
Station 7	-0.343 [0.296]	0.986 [0.004]	0.554 [0.052]	0.993 [0.110]
Station 8	-0.805 [0.005]	0.999 [0.746]	0.351 [0.205]	0.996 [0.307]
Station 9	-0.299 [0.313]	1.001 [0.726]	0.242 [0.388]	0.996 [0.329]
Station 10	-0.080 [0.814]	0.994 [0.193]	0.477 [0.119]	0.992 [0.087]
Station 11	-1.069 [0.001]	1.002 [0.694]	-0.137 [0.602]	1.005 [0.144]

$$T_{\max_{it}} = \alpha + \beta \hat{T}_{\max_{it}} + \mu_{ij},$$

对于16-point，仅有一个未通过t-test；而4-point ANN有很明显的偏差，其截距在station2, 5, 8和11远小于零，斜率在station4和7明显小于1；因此16-point要明显优于4-point。

Conclusions

1. 采用16-point ANN的网络要比4-point的网络表现更好；
2. 这也可以表明使用ANNs对于GCMs的表面温度进行降尺度是一个有效的方法，ANNs可以抓住非线性系统中的复杂联系并学习空间中的相关之处。

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且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，神经网络可以判断输入的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

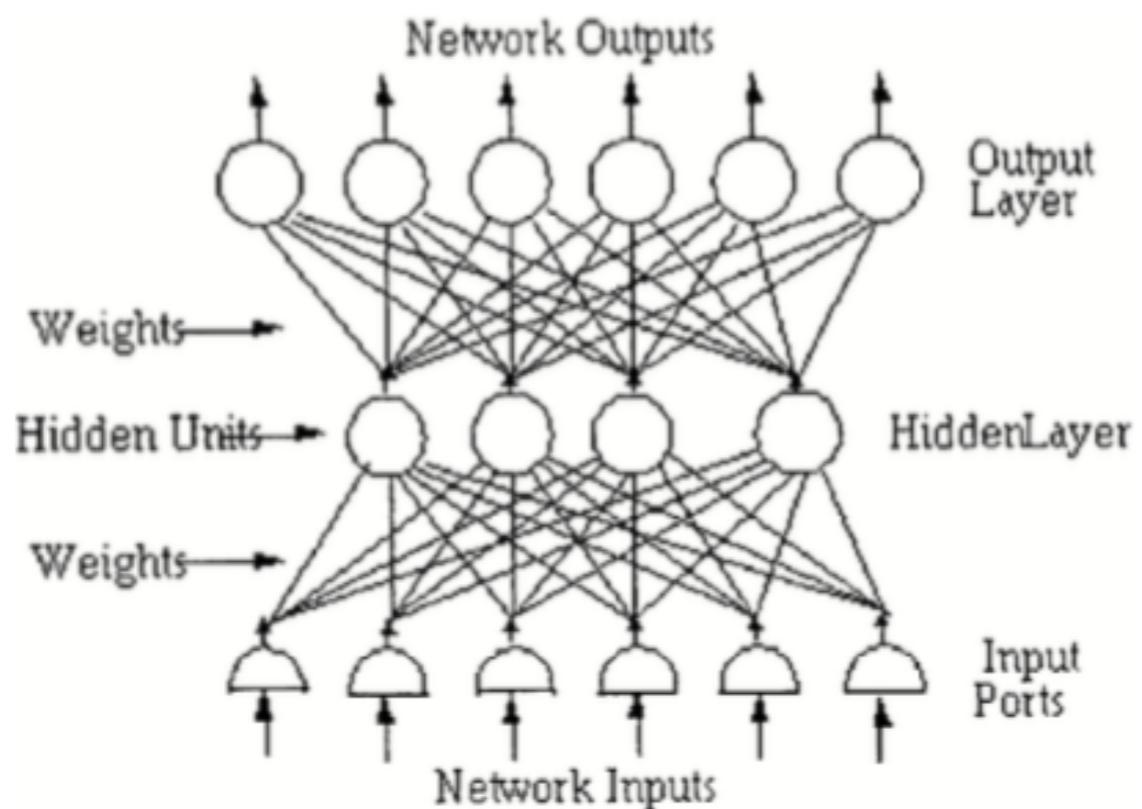


FIG. 2. The ANN feed forward-back propagation network is composed of an input layer of neurons, one or more hidden layers, and an output layer. Each layer comprises multiple units connected completely with the next layer, with an independent weight attached to each connection.

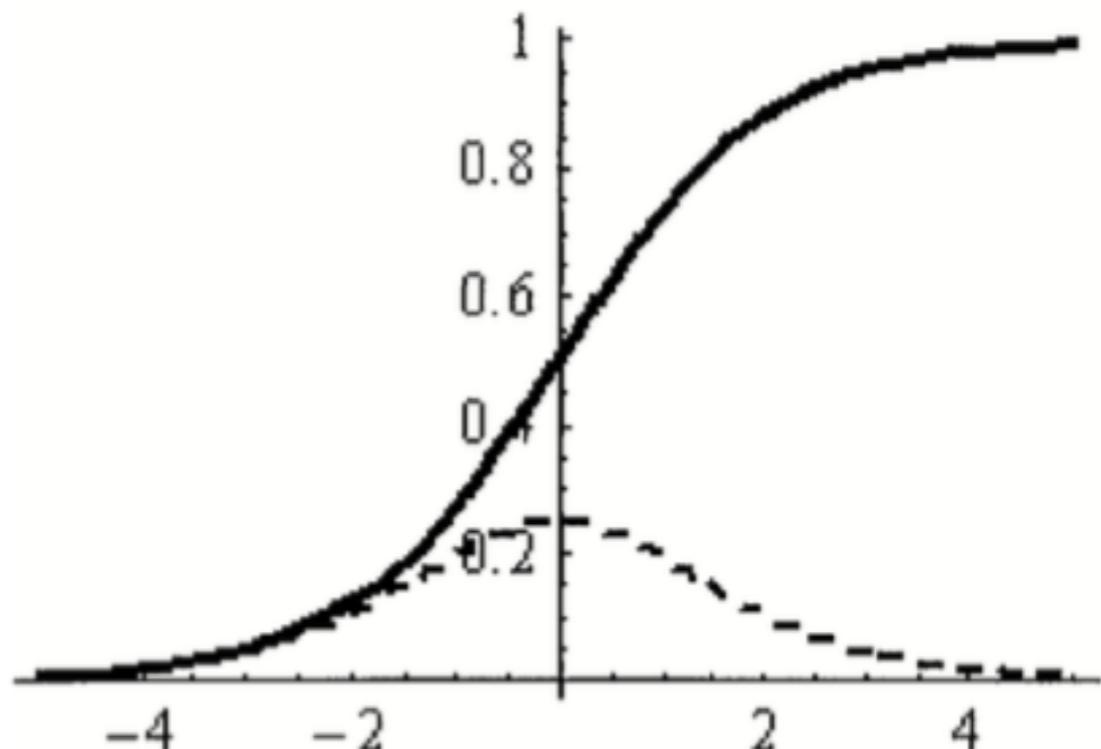


FIG. 3. The sigmoid neuron activation function " $(1 + e^{-x})^{-1}$ " (solid) and its derivative (dashed). The sigmoid function determines the activation level of a neuron based on the weighted inputs. The derivative is used during network development to determine the level of training a weight connection 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，隐藏层节点的数量具体由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：

1. 全月均存在的北大西洋涛动 (NAO) ；
2. 9-4月的东大西洋模式 (EA) ；
3. 4-8月东大西洋jet pattern；
4. 9-5月的东大西洋-俄罗斯西部模式 (EATL-WRUS) ；
5. 除去6月和7月的斯坎蒂维娅模式 (SCAND) ；
6. 12-2月的极地-欧亚大陆模式 (POL) ；
7. 6-8月的亚洲夏季模式 (ASU) ；
8. 全月均存在的西太平洋模式 (WP) ；
9. 除8月和9月均存在的东太平洋模式 (EP) ；
10. 3-7月的北太平洋模式 (NP) ；
11. 除6月和7月存在的太平洋-北美模式 (PNA) ；
12. 11-1月的热带-北半球模式 (TNH) ；
13. 5-8月的太平洋过渡模式 (PT) 。

Methods

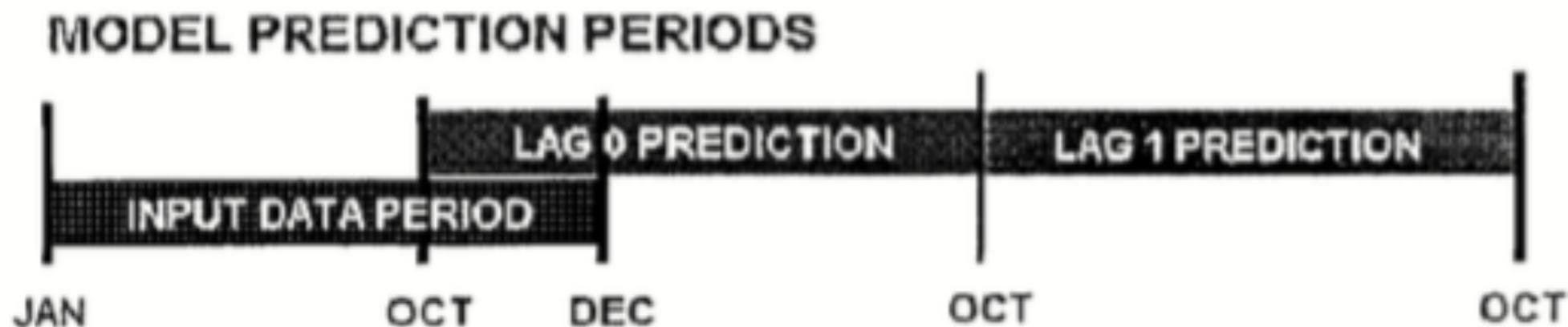


FIG. 4. Prediction periods used in the development of the neural networks. The input data are monthly index values for the calendar year. The predicted output is the total precipitation for the water year, Oct–Sep.

两种网络：lag0, lag1（如上图）

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

隐藏节点的数量设置为由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，在到达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	X
1955	X			X	X					X	X	X		X
1956	X		X	X	X	X	X			X	X	X	X	X
1957	X	X	X	X	X	X	X			X	X	X	X	X
1958	X		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		X
1961					X				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	X
1965					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
1972	X			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	X
1975		X	X	X	X	X	X		X	X	X	X	X	X
1976	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	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	X
1983	X	X	X	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	X	X
1987	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1988		X	X	X	X	X	X	X	X	X	X	X		X
1989		X	X	X	X	X	X	X	X	X	X	X		X
1990	X	X	X	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表现相近，但除了第7区的其他所有的地区在lag1的预测比lag0好，这也同样表明了ANNs具有较好的长期预测降水能力；第7区是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，更难预报，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

针对ENSO的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滞后的影响。

Results

可以看出，大部分地区的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）

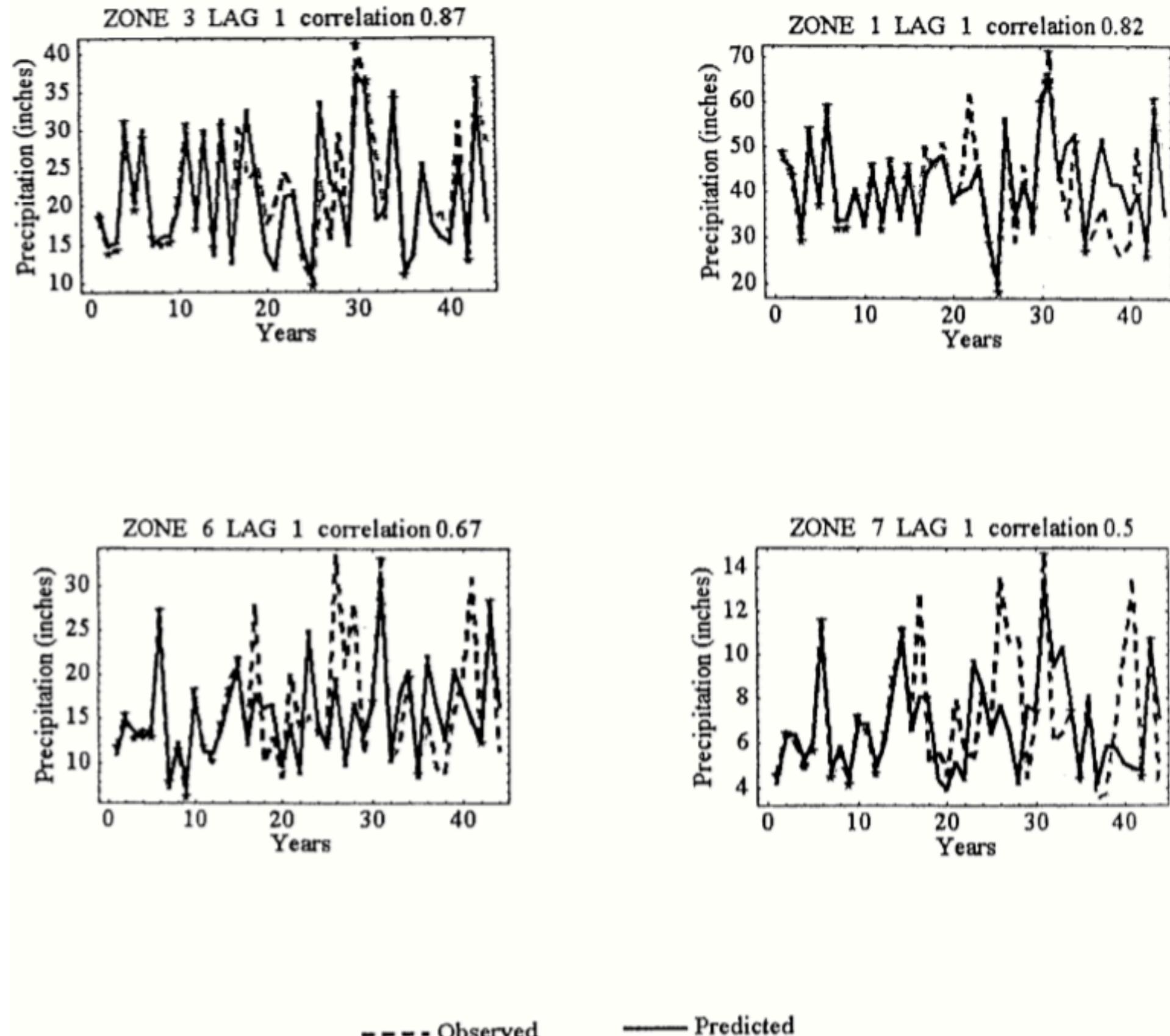


FIG. 5. Example prediction results.

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模型结果表明ENSO和其影响具有延迟关系；
3. ANN模型结果也暗示不同变量对不同区域有不同的影响；
4. ANN模型结果表明其可以对相位有较好的预测。

How Neural Networks Learn from Experience

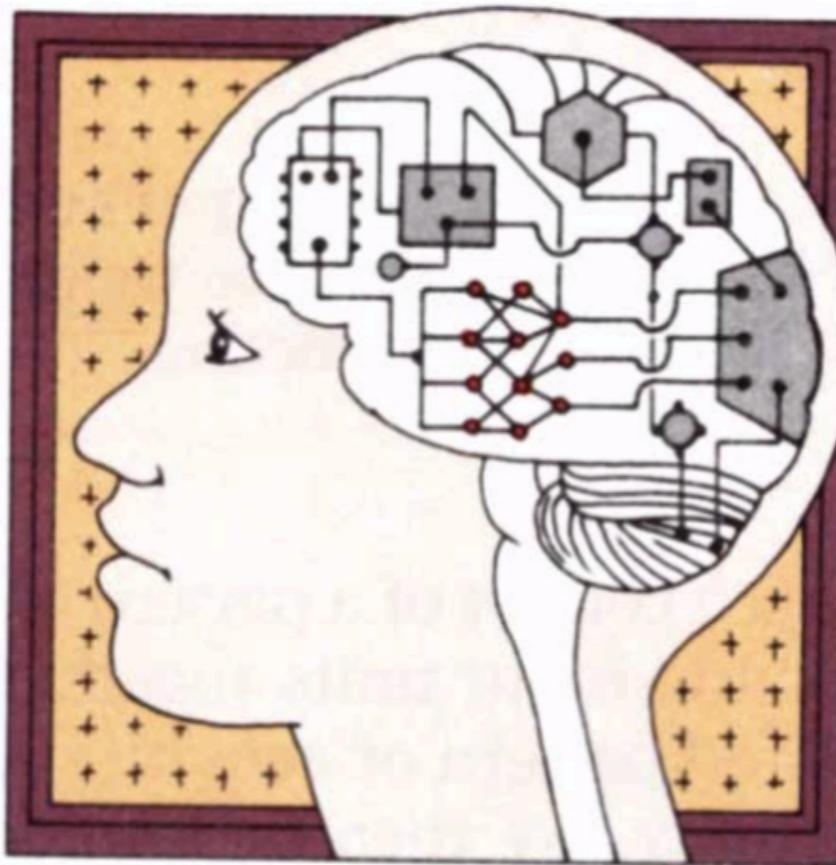
Networks of artificial neurons can learn to represent complicated information. Such neural networks may provide insights into the learning abilities of the human brain

by Geoffrey E. Hinton

Introduction

The brain is a remarkable computer。

- 快速解析场景中的模糊信息；
- 在嘈杂的环境中分辨细微的声音；
- 最重要的是，大脑可以在没有明确指示的情况下进行学习！



在现实世界中是否可以模拟脑内的学习过程？

Introduction

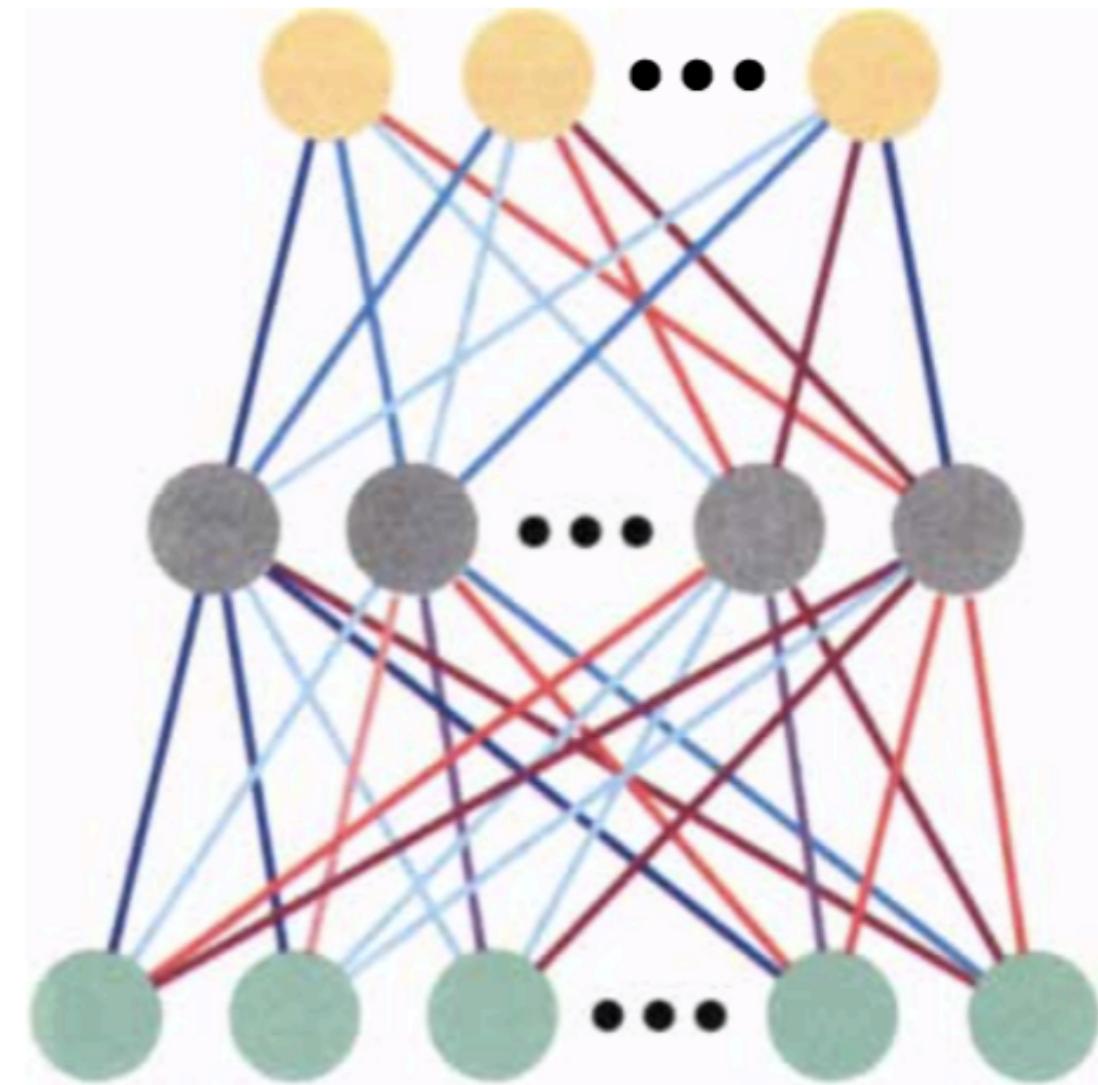
- 在人类大脑中，神经细胞专门负责收集信号；
- 神经元之间通过无数个连接的轴突传递电信号；
- 当轴突收到的信号高于原本的抑制信号时，它会沿着轴突继续向下传递电信号。



人工神经网络相应的通过“节点”、“权重”对人类大脑的神经细胞进行模拟。

Introduction

- 人工神经网络的能力取决于输入和输出间的传递函数和权重两方面；
- 传递函数一般分为三类：线性型、阈值型、sigmoid；
- 人工神经网络包含三层：和隐藏层相连的输入层，和输出层相连的隐藏层；
- 输出层的能力取决于隐藏层和输出层间的权重设计。



COMMON NEURAL NETWORK consists of three layers of units that are fully connected. Activity passes from the input units (*green*) to the hidden units (*gray*) and finally to the output units (*yellow*). The reds and blues of the connections represent different weights.

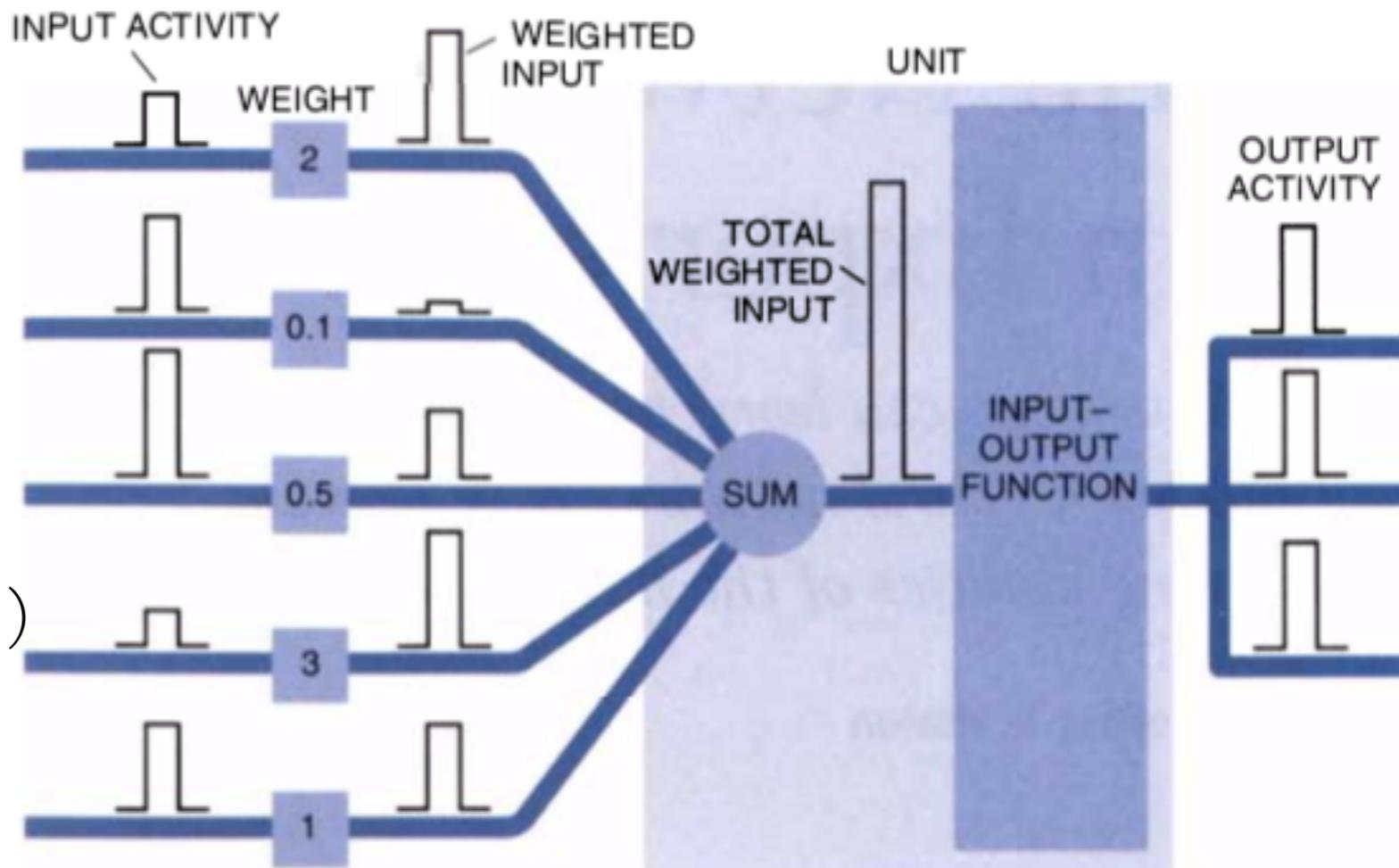
Introduction

如何判断网络已经被调节至最优?

——均方根误差 (RMSE)

但如何调整网络已达到最佳?

——Back Propagation (BP算法)



Back propagation

- Get the gradient: back propagation
- We can represent the computation with the computation graphs

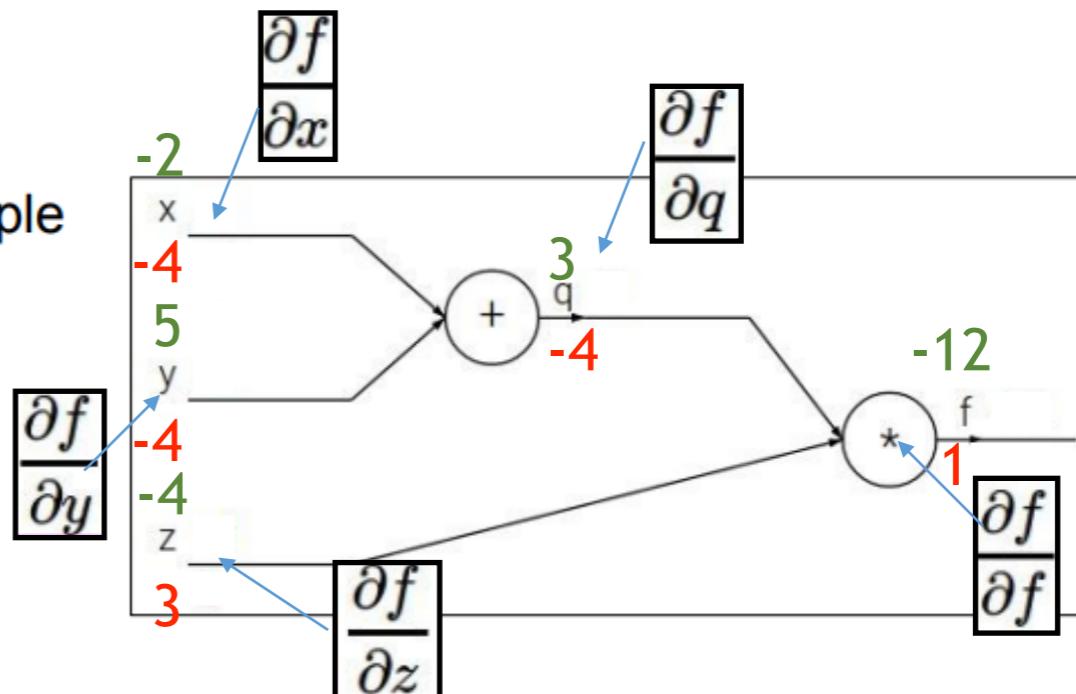
Backpropagation: a simple example

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

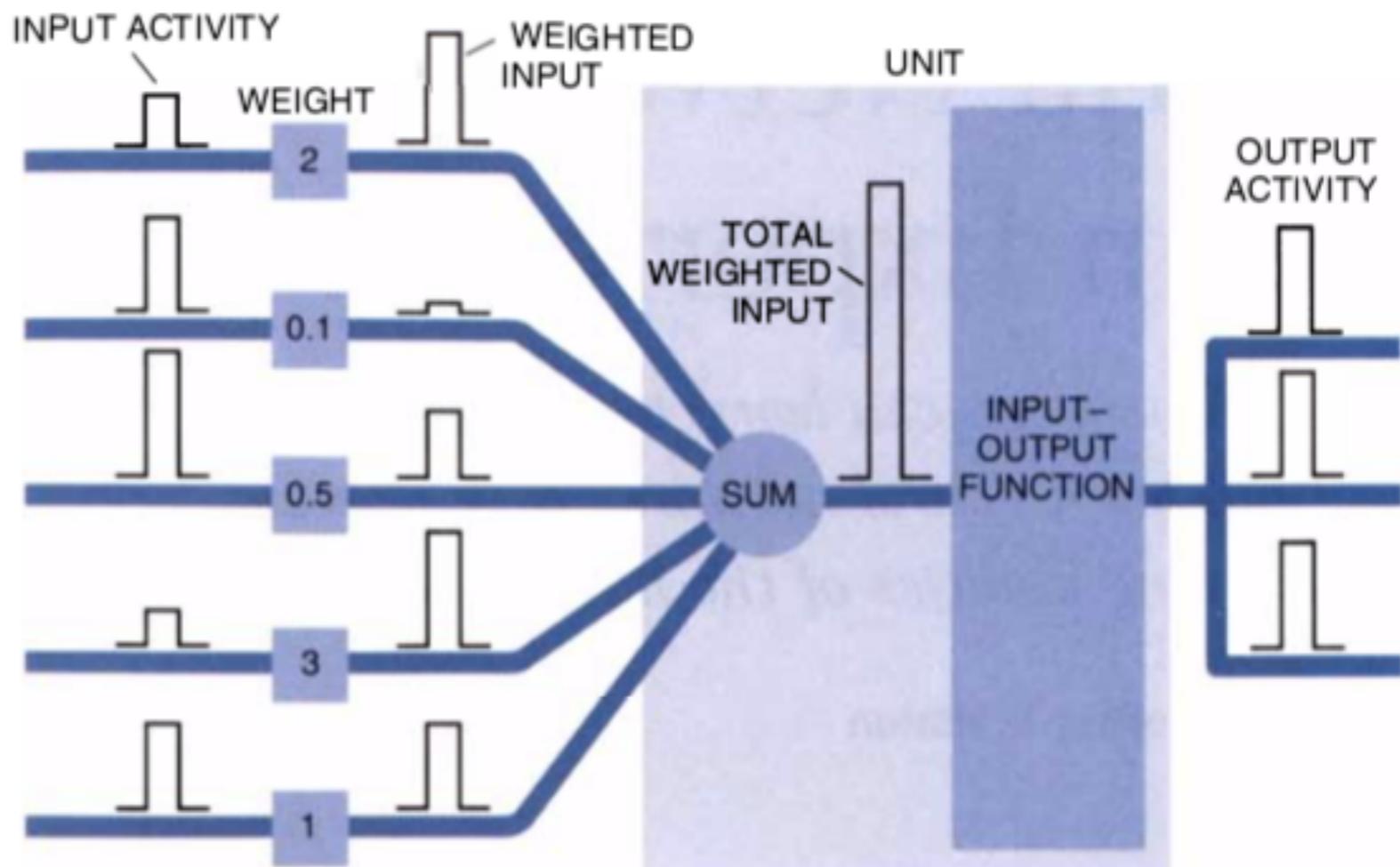
$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$



Introduction

Back Propagation (BP算法)
还可以通过训练网络判断癌变细胞的存在。

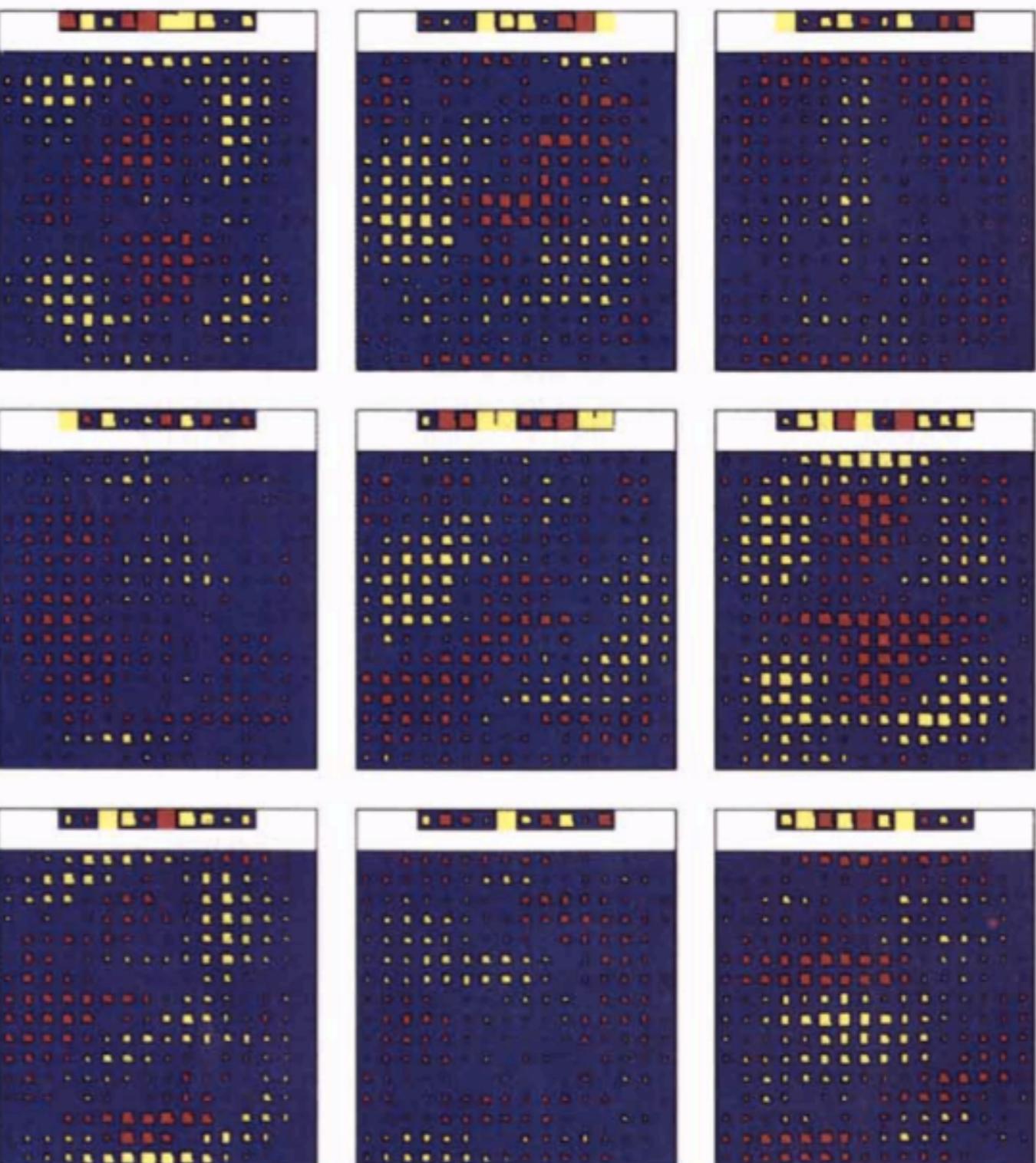
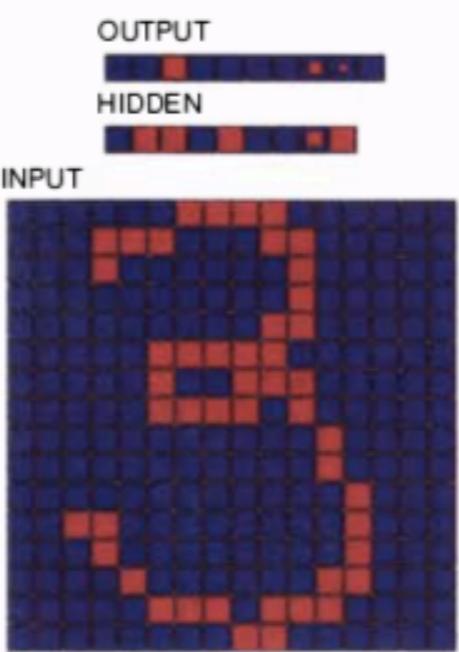


Introduction

For example:

How a Neural Network Represents Handwritten Digits

A neural network—consisting of 256 input units, nine hidden units and 10 output units—has been trained to recognize handwritten digits. The illustration below shows the activities of the units when the network is presented with a handwritten 3. The third output unit is most active. The nine panels at the right represent the 256 incoming weights and the 10 outgoing weights for each of the nine hidden units. The red regions indicate weights that are excitatory, whereas yellow regions represent weights that are inhibitory.



Introduction

几乎所有的非监督学习都可以被视为实现最小的代码成本和重建（reconstruction cost）成本的方法。

—代码成本即描述隐藏层活动所需的代码；

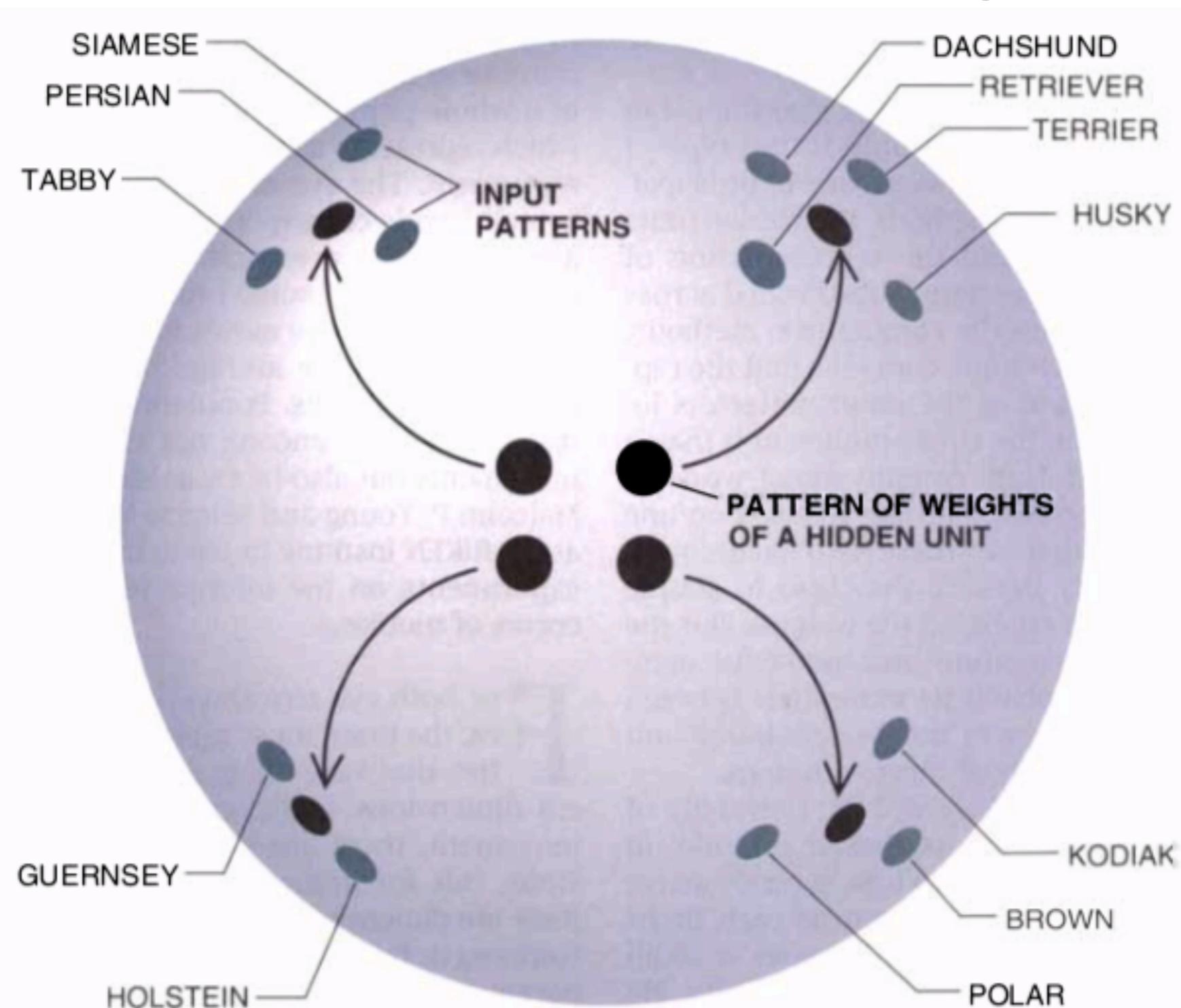
—重建成本即描述原始输入和从隐藏层中重构的对原始输入最佳近似的复制品之间的误差（The reconstruction cost is the number of bits required to describe the misfit between the raw input and the best approximation to it that could be reconstructed from the activities of the hidden units.）；

—重建成本和原始输入和其重建的复制品间的方差成正比。

Introduction

主成分分析方法的提出：如果成对的输入变量相关联，那么对其分别描述是一种资源浪费。

Competitive learning



COMPETITIVE LEARNING can be envisioned as a process in which each input pattern attracts the weight pattern of the closest hidden unit. Each input pattern represents a set of distinguishing features. The weight patterns of hidden units are adjusted so that they migrate slowly toward the closest set of input patterns. In this way, each hidden unit learns to represent a cluster of similar input patterns.

More...

- Population coding;
- Bumps of activity.

计算机的神经网络需要提供正确的信息来完成训练，相反的，人们大多数时候的学习并不会得到老师的指导，我们在没有任何直接帮助的情况下学会了解语句和视觉场景。人工神经网络如何模拟这一步？

谢谢