

基于粒子群算法的神经网络短期降水预报建模研究

吴建生

(柳州师范高等专科学校 数学与计算机科学系,广西 柳州 545004)

摘要:用多样性粒子群算法优化神经网络的网络结构和连接权,获得神经网络集成个体;进一步用二次规划方法,计算各集成个体的最优非负权系数进行组合集成,生成神经网络集成的输出结论,进行短期降水预报建模研究.以广西全区的月降水量实例分析,结果表明该方法能有效提高系统的泛化能力.

关键词:神经网络集成;粒子群优化;二次规划

中图分类号: TP183 **文献标识码:** A **文章编号:** 1673-4785(2006)02-0067-07

Study on the short-time rainfall prediction model of neural ensemble based on PSO algorithms

WU Jian-sheng

(Department of Mathematics and Computer Science, Liuzhou Teachers College, Liuzhou 545004, China)

Abstract: This paper presents the evolving neural network architecture and connection weights based on Diversity-guided Particle Swarm Optimization algorithms. The ensemble strategy is carried out by using the quadratic programming to calculate the best non-negative weights. The weighted coefficient of each ensemble individual is obtained. This method can be used to establish the forecast model of the short-time rainfall. The applied example is built with the monthly mean rainfall in the whole area of Guangxi. The result shows that this method can effectively increase the generalization ability of neural network.

Key words: neural network ensemble; particle swarm optimization; quadratic program

旱涝灾害的气候预测问题是减灾防灾的重要研究课题,随着我国国民经济的高速发展,科技水平的日益提高,人们对灾害性气候的预测精度要求越来越高.在大气科学研究中,气候动力学方法虽然有了很大进展,但是天气系统的动力学模型难以客观描述和构造,目前天气预报业务中采用较多仍然是统计预报方法,从20世纪90年代以来,以神经网络方法为代表的非线性人工智能预报建模方法,已经应用在大气学科和气候分析等领域^[1-4],神经网络与传统的统计方法相比具有自适应和非线性映射等优良的性能得到广泛的认同,但是神经的初始连接权和网络结构选择缺乏定量的客观方法,这成为神经网络方法在旱涝灾害的天气业务应用中的重要技术障碍.

神经网络集成是用有限个神经网络对同一个问题进行学习,集成在某输入示例下的输出由构成集

成的各神经网络在该示例下的输出共同决定^[5-7].该方法可以显著地提高神经网络系统的泛化性能,是一种非常有效的工程化神经计算方法^[8-10].

粒子群优化(particle swarm optimization, PSO)是一种基于群体智能方法的进化计算技术,它是通过个体之间的互动协作来搜寻全局最优解,其概念简单、易于实现,既适合科学研究,又适合工程应用^[11].利用粒子群优化算法提高神经网络的泛化性能是一个十分活跃的研究领域^[12-13].文中利用粒子群算法优化神经网络的结构和连接权,生成集成个体,再用二次规划最优组合方法计算各集成个体的最优非负权系数进行组合集成,生成的输出结论,以此建立短期降水预测模型.

1 粒子群—神经网络集成的基本原理和方法

1.1 粒子群—神经网络的基本原理

BP算法是最普遍的神经网络训练算法,由于基于梯度下降的BP算法依赖于初始权值的选择^[14],

收稿日期:2006-04-28.

基金项目:广西省教育厅资助项目(200508234).

加之实际问题往往是极其复杂的多维曲面,所以 BP 算法收敛速度慢而且极易陷入局部最优;另外在神经网络应用中,网络结构的确定基本上依赖经验,主要是采用递增或递减的试探方法来确定的网络隐节点^[15],这些缺陷使得神经网络的训练样本和检测样本的输出具有不一致性和不可预测性,极大地限制了神经网络在实际气象预报中的应用。

粒子群算法的速度—位移搜索模型操作简单,计算复杂度低,并通过惯性权重协调全局搜索和局部搜索,能以较大概率保证最优解,克服 BP 算法局部最优的缺陷,又可以提高局部区域的收敛速度,避免局部搜索过程中的收敛停滞现象。为了保证种群的多样和过早收敛问题,文中采用 J. Riget 提出的多样性的粒子群算法^[16] (attractive and repulsive particle swarm optimizer, ARPSO) 来优化神经网络。

粒子群—神经网络的优化问题数学描述如下:

$$\begin{cases} \min E(w, v, \dots) = \frac{1}{N_1} \sum_{k=1}^{N_1} \sum_{t=1}^n [y_k(t) - \hat{y}_k(t)]^2 < 1, \\ \hat{y}_k(t) = \sum_{j=1}^p v_{jk} \cdot f[\sum_{i=1}^m x_i \cdot w_{ij} + \theta_j] + \theta_k, \\ f(x) = \frac{1}{1 + e^{-x}}, \\ \text{s.t. } w \in R^{m \times p}, v \in R^{p \times m}, R^p, R^n. \end{cases} \quad (1)$$

式中: x 为训练样本, $\hat{y}_k(t)$ 为网络的实际输出, $y_k(t)$ 为网络的期望输出,利用粒子群优化算法求解的二次非线性规划问题。定义适度函数为

$$F(w, v, \dots) = \frac{1}{1 + \min E(w, v, \dots)}. \quad (2)$$

具体实现步骤如下:

1) 群体的位置和速度初始化,随机生成 L 个个体,每个个体由 2 部分组成,第 1 部分是群体的位置矩阵,第 2 部分对应粒子的速度矩阵;群体位置矩阵包括连接结构矩阵和权重系数矩阵,结构矩阵为二进制变量矩阵,对应的连接权存在则该变量为 1,否则为 0;权重系数矩阵为浮点数矩阵,取 $[-2, 2]$ 上的均匀分布随机数,它是控制网络的连接权值的大小。

2) 输入训练样本,依据式(1)计算每个粒子的适应度,并且初始化个体经历最好位置 $P_{best}(t)$,以及群体经历的最好位置 $P_{gbest}(t)$ 。

3) 对于每个个体,将其适应度与所经历的最好位置的适应度比较,若较好,则将其作为当前最好位置;并将其适应度与全局最好位置的适应度比较,若较好,则将其作为全局的最好位置。

4) 粒子的速度进化方程为

$$v_{ij}(t+1) = v_{ij}(t) + \text{dir} (c_1 r_1 (P_{best}(t) - x_{ij}(t)) + c_2 r_2 (P_{gbest}(t) - x_{ij}(t))). \quad (3)$$

$$\text{dir} = \begin{cases} -1, & \text{diversity} < 2, \\ 1, & \text{diversity} > 2. \end{cases} \quad (4)$$

$$\text{diversity}(L) = \frac{1}{|L| \cdot |R|} \cdot \prod_{i=1}^{|L|} \sqrt{\sum_{j=1}^{|M|} (p_{ij} - \bar{p}_j)^2}. \quad (5)$$

式中: $|L|$ 为种群所含微粒的个数, $|R|$ 为搜索半径, $|M|$ 为微粒的维数, p_{ij} 为第 i 个微粒的第 j 个分量, \bar{p}_j 为第 j 个分量的平均值。

5) 为保证连接结构矩阵进化后仍取 0 或 1,依据文献^[17],连接结构位置进化方程取为

$$x_{ij}(t+1) = \begin{cases} 0, & r > \frac{1}{1 + \exp(-v_{ij}(t+1))}, \\ 1, & r < \frac{1}{1 + \exp(-v_{ij}(t+1))}. \end{cases} \quad (6)$$

式中: r 为 $[0, 1]$ 均匀分布的随机数。连接权位置进化方程为

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1). \quad (7)$$

6) 反复进行 2) ~ 5),直到适应度满足要求或者达到总的进化代数(总的进化代数 K)。

7) 把进化后的最后一代 L 个体全部解码,得到 L 个神经网络的结构和网络连接权,以其作为神经网络的集成个体。

1.2 集成结论的生成

设 N 个训练样本的实际输出为 $\{y_t, t=1, 2, \dots, N\}$, L 个神经网络的输出: $\{\hat{y}_t(i), t=1, 2, \dots, N, i=1, 2, \dots, L\}$,每个神经网络输出被赋予权重 w_i ,则集成输出为 $\tilde{y}_t = \sum_{i=1}^L w_i \hat{y}_t(i)$, w_i 满足约束条件 $w_i \geq 0$,

且 $\sum_{i=1}^L w_i = 1$ 。记 e_t 为第 t 个样本的误差,即

$$e_t = y_t - \hat{y}_t = y_t - \sum_{i=1}^L w_i \hat{y}_t(i) = \sum_{i=1}^L w_i (y_t - \hat{y}_t(i)). \quad (8)$$

所有训练样本的误差记为 $Q = \sum_{t=1}^N (e_t)^2$,以其最小的非负权重系数 w_i ,即求下式的解:

$$\begin{cases} \min Q = W^T E W, \\ \text{s.t. } R^T W = 1, \\ W \geq 0. \end{cases} \quad (9)$$

式中: $W = (w_1, w_2, \dots, w_L)^T$, $R = (1, 1, \dots, 1)^T$, $e_{ij} = \sum_{t=1}^N (y_t - y_t(i))(y_t - y_t(j))$, $i, j = 1, 2, \dots, L$,

$$E = \begin{pmatrix} e_{11} & \dots & e_{1L} \\ \dots & \ddots & \dots \\ e_{L1} & \dots & e_{LL} \end{pmatrix}, \text{ 矩阵 } E \text{ 称为信息阵. 由文献}$$

[18]可知式(9)存在唯一非负最优组合权重向量,而且非负权重最优组合方法的误差平方和不大于参加组合的各个单项的误差平方和中最小者. 根据非线性规划理论^[19],求得式(9)的最优解,以其作为各集成神经网络个体的权重.

以上的方法可以归纳为:先通过粒子群算法优化得到 L 个神经网络,再利用二次规划最优组方法计算各集成个体的最优非负权系数进行组合集成,生成神经网络的输出结论.

2 建模前的数据预处理

神经网络的输入矩阵质量一定程度上影响神经网络泛化能力,但是神经网络本身并不提供如何构造神经网络学习矩阵的方法. 而如何建立合理的神经网络输入矩阵是神经网络建模的另一项关键技术问题,是保障预测模型具有良好泛化能力和进行实际应用的重要条件. 而且气象资料在收集过程中受许多人为因素影响,数据本身不可避免包含噪声,由此所建立的预测模型会失真,预测结果会出现偏差. 为了提高预测的准确率,需要尽量有效地减少样本序列中噪声的影响.

文中把降水量看作一个时间序列,建立时间序列的外推预测模型,首先采用奇异谱分析(singular spectrum analysis, SSA)方法^[20]对原始降水时间序列重构,并用均生函数(mean generating function, MGF)方法^[21]对重构序列构造均生函数延拓矩阵,以其作为自变量,原始降水序列作为因变量,再利用偏最小二乘(partial least-squares regression, PLS)方法^[22]进行处理,提取对因变量解释最强的综合变量作为神经网络的输入因子,原始时间序列作为输出因子.

2.1 奇异谱分析

奇异谱分析是 Karhunen-Loeve 分解理论的发展与应用,最早应用于数字信号处理,其后被推广到海洋学、非线性动力学领域,近年来又开始应用于气候诊断和预测中. 它可以从包含噪声数据序列中提取尽可能多的可靠信息,并且有效利用周期分量重建序列预测模型. 它的好处是能够提炼出主要成分,滤去非周期性的异常现象.

SSA 分析的对象是中心化的一维时间序列,记为 $\{x(t), t=1, 2, \dots, N\}$, 其 M 阶延迟得到矩阵 X

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_{N-M+1} \\ x_2 & x_3 & \dots & x_{N-M+2} \\ \dots & \dots & \dots & \dots \\ x_M & x_{M+1} & \dots & x_N \end{pmatrix} \quad (10)$$

$$\begin{pmatrix} X_{10} & X_{11} & \dots & X_{1, N-M} \\ X_{20} & X_{21} & \dots & X_{2, N-M} \\ \dots & \dots & \dots & \dots \\ X_{M0} & X_{M1} & \dots & X_{M, N-M} \end{pmatrix}$$

X 的第 i 个状态向量为

$$X_i = \begin{pmatrix} x_{i+1} \\ x_{i+2} \\ \dots \\ x_{i+M} \end{pmatrix} = \begin{pmatrix} X_{1i} \\ X_{2i} \\ \dots \\ X_{Mi} \end{pmatrix}, i = 0, 1, \dots, N - M. \quad (11)$$

共 $N - M + 1$ 个状态, X 称相空间中的轨迹矩阵,矩阵 X 中的元素与原序列对应关系为

$$X_{ji} = x_{j+1}. \quad (12)$$

X 的协方差矩阵记为 T_x , 它是一个非负的对称矩阵,其特征根也是非负的. 将这些特征根按降序排列 $e_1, e_2, \dots, e_M, 0$. 矩阵 T_x 的特征根 e_k 对应的特征向量 E^k 称为时间经验正交函数(time empirical orthogonal function, T-EOF),第 k 个时间主成分(time principal component, T-PC)定义为原始序列 $\{x_i\}$ 在第 k 个时间经验正交函数上正交投影系数:

$$a_i^k = \sum_{j=1}^M x_{i+j} E_j^k, 0 \leq i \leq N - M; 1 \leq k \leq M. \quad (13)$$

任意 T-EOF 的 M 个分量构成一个时间序列,反映原始序列中的时间演变型,时间主成分 a_i^k 是 E^k 表示的时间型在原始序列的 $[x_{i+1}, x_{i+2}, \dots, x_{i+M}]^T$ 时段的权重.

SSA 的重要功能由重建成分(reconstruction components, RC)实现,用于在分析和预报中提取感兴趣的信息,过滤噪声,它是利用 T-EOF 和 T-PC 重建一个长度为 N 的序列. 由第 k 个 T-EOF 和 T-PC 重建 x_i 的成分记为 x_i^k ,即

$$x_i^k = \begin{cases} \frac{1}{M} \sum_{j=1}^M a_{i-j}^k E_j^k, & 0 \leq i \leq N - M + 1, \\ \frac{1}{i} \sum_{j=1}^i a_{i-j}^k E_j^k, & 1 \leq i \leq M - 1, \\ \frac{1}{N - i + 1} \sum_{j=i-N+M}^M a_{i-j}^k E_j^k, & N - M + 2 \leq i \leq N. \end{cases} \quad (14)$$

RC 具有叠加性,所有 RC 之和等于原始序列:

$$x_t = \sum_{k=1}^M x_t^k \quad t = 1, 2, \dots, N. \quad (15)$$

在实际应用中,只需要用前 m 个主成分即可重建原序列,降低噪声的干扰,达到提炼主要趋势的目的,即有:

$$x_t = \sum_{k=1}^m x_t^k, m < M, t = 1, 2, \dots, M. \quad (16)$$

2.2 均生函数方法

均生函数方法是將一维时间序列观测值按一定的时间间隔计算均值而得到的均值生成函数,是曹鸿兴等在 20 世纪 90 年代初提出的一种预测方法,它拓广了数理统计中的算术平均值的概念,建立具有多步预测能力的数学模型,能从数据中提取“自然”周期,因此能很好地与原序列拟合。

设经过标准化的时间序列 $\{x_t, t = 1, 2, \dots, N\}$, MGF 计算方法如下:

$$x_l(i) = \frac{1}{N_l} \sum_{j=0}^{N_l-1} x(i+jl), i = 1, 2, \dots, l, 1 \leq l \leq Q. \quad (17)$$

式中: $N_l = \text{INT}(\frac{N}{l})$, $Q = \text{INT}(\frac{N}{2})$, l 为均生函数的周期, Q 为最大周期的长度, INT 表示取整,生成如下上三角矩阵:

$$X^* = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_Q(1) \\ & x_2(2) & \dots & x_Q(2) \\ & & \ddots & \dots \\ & & & x_Q(Q) \end{bmatrix}. \quad (18)$$

再对其作周期性延拓,得到外延序列:

$$f_l(t) = x_l[t - l \cdot \text{INT}(\frac{t-1}{l})], t = 1, 2, \dots, N + P. \quad (19)$$

式中: P 为预报步数,从而得到外延均生函数序列矩阵。

2.3 偏最小二乘回归方法

1983 年 S. wold 及 C. Albano 等人首次提出了偏最小二乘回归,近几十年来,它在理论方法和应用方面都得到迅速发展,许多统计学家开始致力于其理论研究,它主要是针对多因变量对多自变量的回归建模方法,在当因变量只有一个时,称其为 PLS1 回归。在自变量之间存在较高相关性时,特别是样本个数较少,甚至样本个数小于自变量个数时,该方法已经被证明是一种非常有效的方法。

偏最小二乘回归方法与主成分分析回归建模方法的基本思路相同,主要区别在信息综合与筛选过程中,它不但考虑自变量的降维与信息综合,而且要考虑新的信息对因变量具有最佳的解释能力。可以

说偏最小二乘回归方法集多元线性回归分析、典型相关分析和主成分分析的基本功能为一体,将数据分析方法和对模型认识分析方法有机结合。以 PLS1 过程说明偏最小二乘的具体计算步骤^[30]:

设自变量矩阵为

$$X_0 = (X_{ij})_{n \times m} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}. \quad (20)$$

因变量矩阵为

$$Y_0 = (y_{il})_{n \times 1} = \begin{bmatrix} y_{11} \\ y_{21} \\ \dots \\ y_{n1} \end{bmatrix}. \quad (21)$$

1) 标准化自变量矩阵和因变量矩阵,得到标准化数据 X_0^*, Y_0^* .

2) 计算主轴

$$w_i = \frac{X_{i-1}^* Y_{i-1}^*}{X_{i-1}^* Y_{i-1}^*}, i = 1, 2, \dots, T. \quad (22)$$

相应地得到第 k 个综合变量 $F_i = X_{i-1}^* w_i$, F_i 和 X_{i-1}^* 进行普通最小二乘回归估计,回归系数 $p_i = \frac{X_{i-1}^* F_i}{t_i}$,并计算残差矩阵 $X_i^* = X_{i-1}^* - F_i p_i$.

3) 检验交叉有效性,若 $Q^2 > 0.0975$,继续计算,否则停止。

4) 提取 T 个成分 F_1, F_2, \dots, F_T ,建立 X_0^* 和 Y_0^* 在其上的回归方程。由于 F_1, F_2, \dots, F_T 均是 X_0^* 的线性组合,故最终可以改写成 Y_0 关于 X_0 的回归方程。

3 应用实例及结果分析

以广西全区 6~8 月(主汛期)降水量作为预测对象进行短期降水预报建模研究,样本长度为 1957~2005 年共 49 个,其中用 1957~1995 年共 39 个作为训练样本,留取 1996~2005 年共 10 个作为检测样本,检验预测模型实用效果。采用逐年预报形式预测 10 个检测样本的结果,即用 1957~1995 年 6 月份降雨量建立的预报模型,预测 1996 年 6 月份雨量;再用 1957~1996 年 6 月份降雨量建立的预报模型,预测 1997 年 6 月份雨量;依次类推直到 1957~2004 年 6 月份降雨量建立的预报模型,预测 2005 年 6 月份雨量。

3.1 基于 SSA-MGF 的数据预处理

以 1957~1995 年 6 月份降雨量建立的预报模型,预测 1996 年 6 月份雨量为例,说明建立模型的

过程,首先利用 SSA-MGF 方法对 39 个原始降水序列重构,选取延迟阶数 $M = 28$,得到延迟矩阵 $X_{28 \times 12}$,利用式 (10) ~ (16) 选取方差积累贡献率 80% 的值,得到重构序列. 重构结果如图 1 所示.

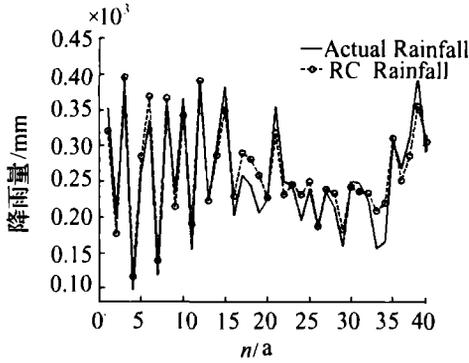


图 1 原始数据和重构数据图

Fig. 1 Actual data and reconstruction data

利用式 (17)、(18) 生成均生函数矩阵 $X_{39 \times 19}$,再依照式 (19) 对其外延 1 步,生成自变量矩阵 $X_{40 \times 19}$,原始降水序列看作因变量 $Y_{40 \times 1}$,将自变量利用偏最小二乘回归处理,提取对因变量影响强的成分,在交叉检验有效时,共提取到 6 个综合变量 F_1, F_2, \dots, F_6 ,以其建立偏最小二乘回归模型(记为 PLS)为

$$Y = 0.366 F_1 + 0.105 F_2 + 0.097 F_3 + 0.165 F_4 + 0.084 F_5 + 0.036 F_6 - 0.0018. \quad (23)$$

并以这 6 个变量建立传统的 BP 模型(记为 BP)和基于粒子群算法进化神经网络集成预报模型(记为 PSO-ANN),分别对 39 个样本拟合和对 1 个样本预报,比较结果来考察模型的效果.

为了定量比较 3 种模型的效果,依据文献[23]引入以下 4 种误差:平均相对误差(the mean absolute percentage error, MAPE)、均方根误差(the mean squares error, MSE)、平均绝对误差(the mean absolute error, MAE)、Pearson 相关系数(pearson relative coefficient, PR).

独立训练 10 个 BP 网络,每个网络都有 6 个输入神经元,1 个输出神经元,隐层神经元个数 6 个,训练参数设置:训练次数 1 000,学习因子为 0.9,动量因子为 0.7,总体误差为 0.001,训练完成后以预测结果最好者作为和 PLS、PSO-ANN 集成模型的对比实验. PSO-ANN 集成模型中的参数设置为:进化代数 100,群体个数 40.

在后续的逐年预报中,建模样本的重构是选取方差积累贡献率 80% 的值,得到重构序列,均生函数外延每次外延 1 步,利用偏最小二乘回归提取综合变量建模因子时,以交叉有效为依据确定建模因

子个数.分别建立偏最小二乘回归模型. BP 模型和 PSO-ANN 集成模型, BP 模型的输入节点依据因子个数确定,输出节点 1 个,隐层神经元个数和输入节点个数一样多,训练参数设置和以上相同.

3.2 结果对比分析

图 2 为 PSO 训练阶段适应度随进化次数的变化曲线,从图可以看出适应度的最大值、平均值、最差值随进化次数增加将趋于稳定并迅速收敛.

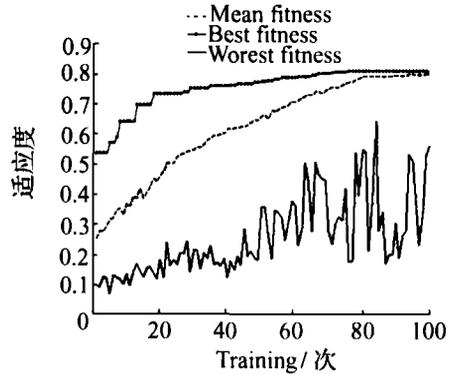


图 2 适应度变化曲线

Fig. 2 Curves of fitness in the training stage.

PLS 模型、BP 模型和 PSO-ANN 模型对 39 个训练样本拟合和 10 个预测样本的各种统计指标结果见表 1,拟合和预测效果分别见图 3 和图 4. 从对比统计指标和图 3 可以看出 BP 模型拟合效果最好,它的相对误差 1.48,相关系数 0.998 3,可以说完全反映了训练样本的情况,其次是 PSO-ANN 模型和 PLS 模型. 评价一个模型的优劣看其拟合效果是一个方面,但更重要的是看其预测效果的优劣,即神经网络的泛化能力.

表 1 3 种模型拟合和预测结果的统计评价
Table 1 The fitting and forecasting evaluate index about 39 samples of three prediction models

模型	MAPE	MSE	MAE	PR
PLS 拟合	21.22	58.20	50.00	0.896 1
PLS 预测	27.34	86.19	77.23	0.910 2
BP 拟合	1.48	4.41	3.26	0.998 3
BP 预测	31.03	121.20	90.97	0.850 3
PSO-ANN 拟合	7.23	22.16	17.67	0.964 4
PSO-ANN 预测	10.92	45.17	34.30	0.927 1

这 3 种模型的都是用 SSA-MGF 对原始降水量序列预处理得到的建模因子,再经过 PLS 方法对建模因子的数据进行分解和筛选,提取对原始序列解释性最强的综合变量建立的非线性模型;对比表 1 的拟合和预测结果,可以看出 PLS 模型对训练样本

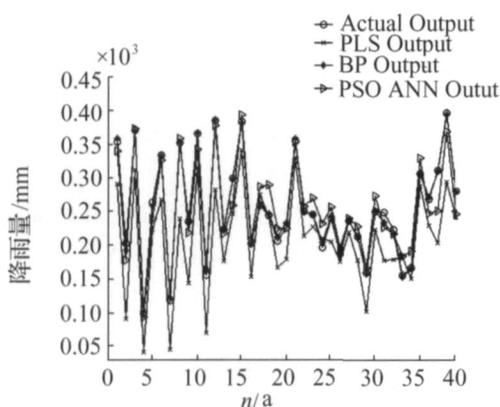


图 3 3 种预报模型的拟合效果

Fig. 3 Fitting results of the three models

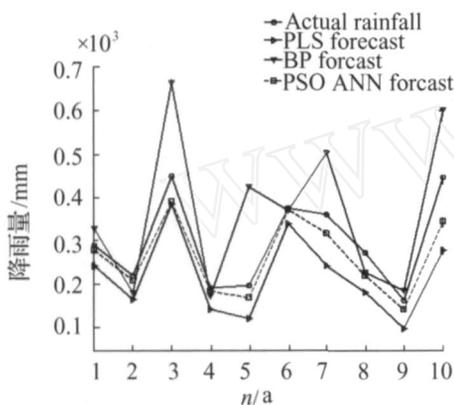


图 4 3 种预报模型的预测效果

Fig. 4 Forecasting results of the three models

和检测样本的平均相对误差分别为: 21.22 和 27.34;BP 模型对训练样本和检测样本的平均相对误差分别为: 1.48 和 31.03;PLS-ANN 模型对训练样本和检测样本的平均相对误差分别为: 7.23 和 10.92,从表 1 和图 4 可以看出,在建模样本相同,预报因子相同的条件下,PSO-ANN 模型对 10 个样本的预报精度明显优于 PLS 模型、BP 模型。

BP 模型拟合最好而预测最差,说明 BP 模型出现“过拟合”现象;PLS 模型拟合与预测比较接近,说明统计模型比较稳定,PSO-ANN 模型的拟合和预测相差 3.69,相对来说该模型也是比较稳定,但是其预测能力远远好于 BP 模型和 PLS 模型。

同时也对广西全区的 7~8 月的降雨量分别利用上述方法建模分析,结果同样表明,PSO-ANN 模型的预报效果优于传统的 PLS 模型和 BP 模型。

4 结 论

降水系统是气候系统中最为活跃、相互作用最

为复杂的子系统之一,并且由于降水变化受众多因素的影响,预报十分困难.本文利用 SSA-MGF 方法对原降水序列重构并延拓,以延拓矩阵作为自变量,原序列作为因变量,再利用 PLS 方法提取对系统解释最强的综合变量作为神经网络的输入因子,原始降水序列作为输出因子,建立基于粒子群算法的神经网络集成预报模型,通过对广西全区主汛期(6~8 月)降水量的实例计算对比表明,该方法具有:

1) 利用 SSA-MGF 方法对原始数据降噪和重构,并利用 PLS 处理,提取对系统解释性最强的综合变量,克服了变量之间的多重相关性,提高模型精度和可靠性;又对神经网络的输入矩阵降维,使得网络结构规模变小,增强网络的稳定性。

2) 利用多样性粒子群优化算法的优化神经网络的结构和连接权并集成,它有效结合神经网络的非线性映射能力和粒子群算法的全局搜索能力,预报效果明显优于 PLS 模型和 BP 模型,且预报结果稳定,具有一定的普遍适用性。

3) 粒子群算法优化神经网络结构和连接权的集成过程中,如何更加有效选择集成个体,提高系统的泛化能力,将是进一步工作。

参考文献:

[1] DEAN A, ANDREW R, BRIAN H. Forecasting warm season burn-off low clouds at the San Francisco international airport using linear regression and a neural network [J]. *Apply Meteor*, 2002, 41(6): 629 - 639.

[2] HSIEH W W. Nonlinear canonical correlation analysis of the tropical pacific climate variability using neural network approach[J]. *Apply Meteor*, 2001, 14(12): 2528 - 2539.

[3] GRIORGIO CORANI, GIOGIO GUARIO. Coupling fuzzy modeling and neural networks for river flood prediction [J]. *IEEE Transactions on Systems, Man, And Cybernetic-Part C: Applications and Reviews*, 2005, 25(3): 382 - 388.

[4] 吴建生,金龙,农吉夫. 基于遗传算法的神经网络建模研究[J]. *数学的实践与认识*, 2005, 35(1): 83 - 88. WU Jiansheng, JIN long, LONG Jifu. Study on model of neural network based on genetic algorithm [J]. *Mathematics in Practic and theory*, 2005, 35(1): 83 - 88.

[5] SOLLICH P, KROGH A. Learning with ensembles: how over-fitting can be useful[A]. In *Proceeding of Advances in Neural Information Processing Systems 8*[C]. Cambridge, MA: MIT Press, 1996.



- [6] HANSEN L K, SALAMON P. Neural network ensembles [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(10): 993 - 1001.
- [7] 周志华, 陈世福. 神经网络集成[J]. 计算机学报, 2002, 25(1): 1 - 8.
ZHOU Zhihua, CHEN Shifu. Neural network ensemble [J]. Chinese Journal of Computer, 2002, 25(1): 1 - 8.
- [8] MAO J. A case study on bagging boosting and basic ensembles of neural networks for OCR[A]. In Proceedings of International Joint conference on Neural Networks 1998[C]. Anchorage, AK, 1998:1828 - 1833.
- [9] GUTTA S, WECHSLER H. Face recognition using hybrid classifier systems[A]. In Proceedings of International Conference Neural Network 1996[C]. Washington, DC, 1996:1017 - 1022.
- [10] SOLLICH P, INTRATOR N. Classification of seismic signals by integrating ensembles of neural networks [J]. IEEE Transactions Signal Processing, 1998, 46(5): 1194 - 1021.
- [11] BONABEAU E, DORIGO M, G THERAULAZ. Inspiration for optimization from social insect behavior [J]. Nature, 2000, 406(6): 39 - 42.
- [12] KENNEDY J, EBERHART R C. Swarm intelligence [M]. San Francisco: Morgan Kaufmann Publishers, 2001.
- [13] 高海兵, 高亮, 周驰, 等. 基于粒子群优化的神经网络训练算法研究[J]. 电子学报, 2004, 32(9): 1572 - 1574.
GAO Haibing, GAO Liang, ZHOU Chi, et al. Particle swarm optimization based algorithm for neural network learning[J]. Acta Electronic Sinica, 2004, 32(9): 1572 - 1574.
- [14] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back propagating errors. Nature, 1986, 323(11): 456 - 466.
- [15] REED R. Pruning algorithms—a survey [J]. IEEE Transactions on Neural Networks, 1993(4): 740 - 747.
- [16] JACQUES RIGET, JAKOB VESTERSTR. A Diversity-guided Particle Swarm Optimizer—the ARPSO [A]. Development and Practice of Artificial Intelligence Techniques[C]. Durban, South Africa, 1999, 41 - 45.
- [17] KENNEDY J, SPEARS W. Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimode problem generator[A]. In Proceedings of IEEE International Conference on Evolutionary Computation[C]. Anchorage, Alaska, USA, 1998.
- [18] 马永开, 唐小我, 杨桂元. 非负权重最优组合预测方法的基本理论研究[J]. 运筹与管理, 1997, 6(2): 1 - 8.
MA Yongkai, TANG Xiaowo. YANG Guiyuan. A study on basic theory of the optimal combined prediction method of non negative weights[J]. Operations Research and Management Science, 1997, 6(2): 1 - 8.
- [19] 马振华. 运筹学与最优化理论[M]. 北京:清华大学出版社, 1998.
- [20] VAUTARD. SSA: a toolkit for noisy chaotic signals [J]. Physica D, 1992, 58: 95 - 126.
- [21] 魏凤英, 曹鸿兴. 长期预测的数学模型及应用[M]. 北京:气象出版社, 1990.
- [22] 王惠文. 偏最小二乘回归方法及其应用[M]. 北京:国防工业出版社, 1999.
- [23] 金龙. 神经网络气象预报建模理论与应用[M]. 北京:气象出版社, 2004.

作者简介:



吴建生,男,1974年生,硕士,讲师,主要研究方向为神经网络应用及智能优化算法研究. E-mail: wjsh2002168@163.com.