

覃武, 林开平, 黄颖, 等. 基于遗传-神经网络的电网流域面雨量预报方法研究[J]. 灾害学, 2015, 30(3): 033-037.
[Qin Wu, Lin Kaiping, Huang Ying, et al. A prediction scheme with genetic neural network for area rainfall over power basin[J]. Journal of Catastrophology, 2015, 30(3): 033-037.]

基于遗传-神经网络的电网流域 面雨量预报方法研究*

覃武¹, 林开平², 黄颖³, 李勇⁴, 钟利华⁴, 罗小莉⁴

(1. 广西区气象局, 广西南宁 530022; 2. 广西区气象台, 广西南宁 530022;
3. 广西气象减灾研究所, 广西南宁 530022; 4. 广西气象服务中心, 广西南宁 530022)

摘要:以重点水力发电厂和大中型水库为主要考量, 并兼顾地形地貌和中小河流的分布特征, 将广西划分为23个电网流域, 研究了基于非线性的神经网络电网流域面雨量预报方法。以5-6月龙滩近库区、龙江流域等6个电网流域为例, 利用遗传算法优化BP神经网络的连接权和网络结构, 建立了各电网流域的遗传-神经网络电网流域面雨量预报模型。对独立样本的预报结果表明, 基于遗传-神经网络的电网流域面雨量预报模型的预报能力要优于传统的逐步回归预报模型, 也明显优于日本、德国数值模式预报产品所换算成的电网流域面雨量预报, 并与气象部门同期制作的综合面雨量预报产品能力相当, 因而, 遗传-神经网络面雨量集合预报模型有较好的业务应用前景。

关键词:遗传算法; 人工神经网络; 电网流域; 面雨量; 预报; 广西

中图分类号: P457; X4 **文献标志码:** A **文章编号:** 1000-811X(2015)03-0033-05

doi: 10.3969/j.issn.1000-811X.2015.03.007

广西地处低纬度东亚季风气候区, 雨量十分充沛。同时, 广西西邻云贵高原, 南濒北部湾, 有典型的喀斯特丘陵地貌, 区内中小河流域纵横交错, 水系尤为发达, 是我国水力电力资源最为丰富的省区之一, 水库遍布广西各地。水库的作用除了蓄水发电的功能外, 更重要的作用则是为了抗旱防洪。在全球气候变暖的背景下, 地球环境的许多要素也发生了变化。近几年来, 云南、贵州、广西天气气候极端异常事件出现频繁。干旱、洪涝等气象灾害频繁发生, 并呈加剧的趋势^[1-2]。暴雨既可能导致中小河流出现洪水, 有可能引发山洪、泥石流、滑坡等地质灾害^[3-5]。对各水库来说, 库区对应的流域降水少, 库区的蓄水不足, 使水电站的发电能力不足, 甚至无法发电。如果降水过多, 特别是持续性的强降水, 又会使水库水位猛增, 甚至超过警戒水位, 对库区的安全和库区下游群众的生命财产安全构成严重的威胁。因此, 对于库区调度来说, 气象部门如果能提供影响库区流域的精细化降水预报, 他们就能依据流域实时的水文资料和流域的面雨量预报, 科学地对库区的蓄水进行调度, 在保障防洪排涝的安全的前提下, 能充分地利用蓄水发电, 增加

水力发电的经济效益。

对于广西的面雨量预报, 已有一些研究^[6-7], 但主要是考虑广西的主要江河流域的面雨量预报问题, 很难满足水库和水力发电厂对中小流域面雨量精细化的要求。因此, 研究针对水库和水力发电厂中小河流域的面雨量预报方法很有必要。以往的面雨量预报多是采用相似预报法、MOS预报法、回归统计预报方法。但流域的面雨量与降水一样受众多的因素影响, 具有明显的非线性特征。人工神经网络是一种模拟人脑信息处理方法的非线性系统, 具有较强的处理非线性问题能力, 比较适合于一些信息复杂、知识背景不清楚和推理规则不明确问题的建模。遗传算法具有自适应性、全局优化性和隐含并行性等许多优点, 体现出很强的解决问题的能力^[8]。将遗传算法与BP神经网络相结合, 充分发挥它们各自的长处, 是目前一个十分活跃的研究课题。人工神经网络广泛应用于降水预报和台风强度、路径预报中, 取得了良好的效果^[9-11]。本文以经过细化的以主要水库和水电站为重要考量的中小河流域面雨量为研究对象, 对中小河流域面雨量非线性预报方法进行研究。

* 收稿日期: 2014-11-21 修回日期: 2015-01-06

基金项目: 广西科学研究与技术开发计划项目(桂科攻1355010-4); 广西自然科学基金北部湾重大专项项目(2011GXNSFE018006)

作者简介: 覃武(1963-), 男, 广西东兰人, 硕士, 高级工程师, 主要从事天气预报技术研究与管理。E-mail: gxqinwu@163.com.

通讯作者: 林开平(1960-), 男, 广西平南人, 博士, 教授级高级工程师, 硕士生导师, 研究方向为大气科学、人工智能应用。

E-mail: linkp0305@aliyun.com.

1 流域的划分和面雨量计算方法

1.1 流域的划分方案

广西的水力发电厂和各种类型的水库众多,遍布全区各地。为了更好地开展有针对性的气象服务,以广西的重点水力发电厂和大中型水库为主要考量,并兼顾地形地貌和中小河流的分布特征,划分了23个面雨量预报区间^[12],即:南盘江上游、南盘江下游、北盘江上游、北盘江下游、龙滩近库区、红水河中下游流域、右江流域上游、右江流域、左江流域、郁江流域、西津流域、沿海流域、桂东南流域、融江流域、龙江流域、柳江流域、洛清江流域、清水河流域、西江流域、桂东北流域、桂江流域、桂江中下游流域、贺江流域(图1)。



图1 广西流域面雨量预报区域分布图

从图1看到,南盘江上游、南盘江下游、北盘江上游、北盘江下游、龙滩近库区、融江流域等流域超出了广西地域,包含了部分云南、贵州区域。这样的划分虽然增加了气象部门制作面雨量预报的难度,但能满足水力发电厂和水库对蓄水调度的需求。

1.2 面雨量的计算

面雨量的计算方法有很多,如算术平均法、梯度法、泰森多边形法,各种算法都有优缺点,本文采用的是目前应用比较广泛的泰森多边形面雨量计算方法。对于流域的面雨量实况,我们采用了广西2000多个自动站和贵州、云南的800多个自动站点的降水资料,通过离散点的泰森多边形方法计算各流域的面雨量。而对于日本、德国的数值模式,则利用其输出的格点降水预报资料,用算术平均法计算各流域的面雨量,作为模式对各流域的面雨量预报结果。

2 遗传-神经网络的基本原理和方法

2.1 人工神经网络的基本原理

BP神经网络是目前应用最为有效神经网络模型之一,其广泛应用于医学诊断、图像处理和气象、水文、金融等领域的预报预测中。BP算法的学习过程是通过将输入因子沿网络正向传播,求得结果后计算误差值,然后根据误差来修改网络

权值,主要是利用均方差和梯度下降法来实现网络连接权值的修正,使所训练的神经网络的误差平方和达到最小或小于所设定的误差值。在神经网络中,每一层的神经元状态仅影响下一层神经元状态,从输入层输入样本的特征值,通过神经网络向下层传播,经过隐层逐层处理,最后经由输出层将结果输出。当神经网络的输出结果与期望输出结果的误差大于设定的误差标准时,神经网络就会将误差信号沿原来连接通路反向传播,同时对神经网络原来的连接权值进行修正,使得误差变小。如此反复训练,直到神经网络的输出结果与期望输出结果的误差小于设定的误差标准。

BP神经网络的数学模型是求解如下函数最优解:

$$\min_{a \leq \omega \leq b} E(\tilde{\omega}) = \frac{1}{2} \sum_{k=1}^n (y_k - \hat{y}_k)^2; \quad (1)$$

$$\hat{y}_k = f\left(\sum_{i=1}^n (x_i \tilde{\omega}_i + \theta_i)\right). \quad (2)$$

f 取sigmoid函数 $f(x) = \frac{1}{1 + e^{-x}}$,式中: N 为样本

数, $\tilde{\omega}$ 为BP网络权值, θ 为网络阈值, x_i 为训练样本的输入, \hat{y}_k 实际输出, y_i 期望输出。由式(2)可以求出神经网络的最终输出值。

2.2 利用遗传算法优化神经网络的网络结构和连接权

遗传算法具有全局性搜索的特点,通过遗传算法寻找最为合适的神经网络的网络结构和连接权,从而优化神经网络是一种快捷而可行的方法。为了简单起见,这里以一个包含输入、输出和隐层三层的BP网络为例,来分析如何利用遗传算法对神经网络预报模型的连接权和网络结构进行优化。

设神经网络预报模型的全体样本集为:

$$\phi = \{(x_i, y_i), i = 1, 2, \dots, n\}. \quad (3)$$

式中: x_i 是输入变量, y_i 是输出结果, n 为样本总数。我们可以将样本集分为训练样本 ϕ_1 和检测样本 ϕ_2 两个部分:

$$\phi_1 = \{(x_i, y_i), i = 1, 2, \dots, m, m < n\}; \quad (4)$$

$$\phi_2 = \{(x_i, y_i), i = m + 1, m + 2, \dots, n\}. \quad (5)$$

先设定网络的训练次数及网络训练误差 ε_a ,利用训练样本 ϕ_1 对网络进行训练。一般来说,只要训练次数足够多,BP网络往往都可能取得有很好的拟合效果,但对于试验样本或实际情况的预报效果却差强人意,这就是所谓的“过拟合”现象。

由于神经网络的预报能力受众多参数及网络结构等多方面的影响,为此我们还需要进一步考虑BP网络对检测样本的预报能力。通过计算检测样本

ϕ_2 的平均相对误差 $\varepsilon_b = \frac{1}{n-m} \sum_{i=1}^{n-m} \frac{|y_i - \hat{y}_i|}{|y_i|}$,将训练误差 ε_a 和检测集平均相对误差 ε_b 同时作为网络的目标函数进行分析,并以 $[0, 1]$ 作为连接权的初步解空间,即在这个区间搜索合适的权值、阈值。对于隐节点的个数也先给出初步的搜索区间。依据经

验,隐节点的最佳个数大多在输入节点个数附近,因而以输入节点 0.5 ~ 1.5 倍作为隐节点个数的搜索空间。

遗传-神经网络算法的具体实现步骤可归结为:

(1) 采用二进制对神经网络结构和连接权进行编码。每个遗传个体由一个二进制码串组成,它与某组神经网络权值、阈值、隐节点一一对应,是一种可能的优化个体。二进制码串由两部分组成,前面是神经网络结构的编码,控制隐节点的个数;后面部分是神经网络权重系数码。

(2) 在编码空间中,随机生成一个初始群体。

(3) 计算当前群体中所有遗传个体的适应度时,首先将个体二进制码串解码为神经网络的连接权、隐节点,输入训练样本集,按照适应度函数计算每个遗传个体的适应度。对训练样本集合 $\phi_1 = \{(x_i, y_i), i = 1, 2, \dots, m, m < n\}$, 设被训练的网络为 $net, \hat{y}_i = net(x_i)$ 为实际输出,定义适应度函数为:

$$F(x) = \frac{1}{\sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}} \quad (6)$$

(4) 根据个体的适应度,对群体进行遗传操作。其中选择采用赌轮法,交叉则采用多点交叉,在变异时,当某个神经元被变异运算删除时,相应的有关权重系数编码被置为 0,而当变异运算增加某个神经元时,则随机产生有关的初始化权重系数编码。

(5) 生成新一代群体。

(6) 反复进行(3) ~ (5),每进行一次,群体就进化一代,一直进化到第 N 代,并在进化过程中,每代保留适应度最大的个体。

最终进化到 N 代时(N 为总的进化代数),全部进化计算结束。这样共挑选出 N 个个体,比较这些个体的适应度,把其中适应度最佳个体予以保留。把最佳个体解码,得到神经网络的连接权和隐节点,输入检测样本进行预报。其中,步骤(6)中全局性是 GA 搜索机理提供,GA 能以较大概率进化至全局解的区域,但不能明确是哪个解,因此不能简单地认为最后一代适应度最高的个体就是全局最优解,每代保留适应度最高的个体,直到进化结束,这样就大大增加了包含最优解的概率。

3 流域面雨量预报模型及预报结果分析

本文以广西区域前汛期 5-6 月份的各主要水力发电厂和大中型水库所对应的中小河流域的面雨量为预报对象,对基于遗传-神经网络的中小河流域面雨量预报方法进行研究。

3.1 预报对象处理和预报因子的选取

考虑到既要研究重点的库区流域,又要兼顾到有代表性的库区,在这里我们选取了龙滩近库区、龙江流域、桂江中下游流域、左江流域、右

江流域和清水河流域等 6 个流域的作为研究对象。

从多年的降水研究和实际天气预报经验来看,华南西部的降水成因很复杂,降水既受大气环流的影响,也受到各种天气系统的制约和它们之间的配置和相互作用影响,此外,水汽供应条件、辐合辐散气流、上升气流和地形也是影响降水的重要因素。因此,在预报因子初选时,重点是考虑地面、850 hPa、700 hPa、500 hPa 各层的温度、湿度、气压(高度)、风向风速等各种气象要素和涡度、散度、水汽通量散度、垂直速度等物理量。从预报实践中发现,针对广西的降水,欧洲中心的数值预报模式的预报性能更好。因此,选取欧洲中心的数值预报模式的预报产品作为备选场。

考虑到一般台站获取数值预报产品的滞后性,本文所选用的欧洲中心数值预报产品均为 48 h 预报场,范围为: $100^\circ \sim 120^\circ \text{E}$, $15^\circ \sim 30^\circ \text{N}$, 格距为 $1^\circ \times 1^\circ$, 这样,所选区域的格点数为 336。选取的产品包括各标准层的各个常规气象要素预报场及物理量预报场。

对 2009-2012 年 5 月、6 月欧洲中心数值预报产品场与预报对象进行场相关普查,通过普查,发现不同的流域面雨量对同一层次、同一要素场的相关程度差别很大。因此,我们将龙江流域等 6 个流域分别对上述所选区域各标准层各个常规气象要素预报场及物理量预报场进行相关普查,将成片稳定(置信水平高于 0.05)的高相关格点作为预报因子的选择区,在选择区内选择相关系数最大的 2 个相邻的格点的平均值作为该相关区的代表值,作为待选因子。另外对与预报对象相关好但符号相反的两个相邻或相近区域,将这两个区域的代表值相减,获得组合预报因子。预选因子时,以达到或超过 0.01 置信度水平为入选标准,最终得到的各流域各月的预报因子见表 1。这些待选因子中包含了数值预报产品的要素预报场和各种物理量。各预报区最终入选的因子,既含有与降水有关的大尺度形势场,又有与降水密切相关的物理量场。

表 1 各流域面雨量预报因子

	龙滩近库区	龙江流域	桂江中下游	左江流域	右江流域	清水河流域
5 月	11	12	7	10	9	11
6 月	9	16	9	12	10	12

3.2 模型试验及结果分析

3.2.1 逐步回归预报方法建模试验

根据上述所选定的 6 个流域的预报因子,采用逐步回归建模方法,分别建立各流域的 5 月、6 月流域面雨量逐步回归预报模型。利用表 1 给出的各流域 5 月、6 月份预报因子作为备选因子,将 5 月和 6 月的样本分为建模样本和独立样本两部分。本文规定,对于各流域 5 月、6 月的各个样本序列,均把最后 30 个样本作为独立样本,其余的样本作为建模样本。这样,各流域 5 月份的建模样本 124 个,独立样本 30 个,6 月份的建模样本为 120 个,

独立样本也为 30 个。在建立的各个逐步回归方程中,都要求能通过 $F=2.0$ 的显著性检验。以桂江中下游流域为例,采用逐步回归方法建立了 5 月、6 月逐步回归预报方程进行独立样本预报试验,对每个独立样本,建立一个回归预报方程,之后把用过的独立样本追加到建模样本,使下一个独立样本的回归预报方程的建模样本数增加,因此,某一组的 30 个独立样本,就有 30 个不同的回归方程。对 30 个独立样本的预报绝对误差求平均,就得到了 5 月份的面雨量预报平均绝对误差为 5.83 mm,而 6 月份的平均绝对误差为 7.84 mm。表 2 给出了各流域 5 月和 6 月份的逐步回归模型对独立样本的预报情况。

表 2 各流域 5 月、6 月的面雨量逐步回归预报模型对独立样本预报平均绝对误差 mm

	龙滩近 库区	龙江 流域	桂江中 下游	左江 流域	右江 流域	清水河 流域
5 月	8.34	10.89	5.83	7.12	7.10	9.90
6 月	8.16	11.04	7.84	6.40	8.58	8.4

3.2.2 遗传-神经网络预报模型建模试验

以逐步回归模型相同的样本,采用遗传算法与神经网络相结合的方法,分别建立 6 个流域 5 月、6 月的流域面雨量遗传-神经网络预报模型。为了便于与回归模型进行比较,在建立各流域面雨量的遗传-神经网络预报模型时,所选的预报因子与逐步回归模型所选的因子完全相同。建模试验中,神经网络训练次数为 200 次,学习因子和动量因子分别取 0.5;对进化计算的遗传算法,设置遗传算法的隐节点搜索空间范围设定为输入节点的 0.5 倍到 1.5 倍之间,交叉概率为 0.6,变异概率为 0.05,网络连接权的解空间设定为 $[-2, 2]$,初始遗传种群数为 50,进化代数为 50 代。进化计算结束后,对遗传种群的 50 个遗传个体解码,得到 50 个神经网络预报个体,并对每个预报个体赋以的相同的权重,分别建立了 6 个流域 5 月、6 月共 12 个遗传-神经网络集合预报模型。

利用所建立的 12 个流域面雨量遗传-神经网络集合预报模型,对独立样本的预报时采用与 3.2.1 节中相同的方法,即对每个独立样本,建立一个遗传-神经网络集合预报模型,之后把用过的独立样本追加到建模样本,使下一个独立样本的遗传-神经网络集合预报模型的建模样本数增加,因此,某一组的 30 个独立样本,就有 30 个不同的遗传-神经网络集合预报模型。表 3 给出了各流域 5 月和 6 月份的遗传-神经网络集合预报模型对独立样本的预报情况。

表 3 各流域 5、6 月面雨量遗传-神经网络集合预报模型对独立样本预报平均绝对误差 mm

	龙滩近 库区	龙江 流域	桂江中 下游	左江 流域	右江 流域	清水河 流域
5 月	6.32	8.41	5.90	6.95	7.13	9.08
6 月	6.67	8.52	7.69	5.80	8.08	8.6

3.2.3 遗传-神经网络预报模型与逐步回归预报模型的比较

表 2、表 3 中分别给出了面雨量逐步回归预报模型和遗传-神经网络集合预报模型对相同的独立样本的预报平均绝对误差,从预报结果来看,两种预报模型对 6 个流域的面雨量都有较好的预报能力。图 2、图 3 中分别给出了回归预报模型和遗传-神经网络集合预报模型对各流域 5 月、6 月独立样本的预报结果对比。从图中可以看到,不管是 5 月份还是 6 月份,对于绝大多数流域来说,遗传-神经网络集合预报模型对独立样本预报的平均绝对误差要明显小于逐步回归预报模型,尤其是对处于广西西北部的龙滩近库区流域、龙江流域,遗传-神经网络集合预报模型的预报能力明显优于逐步回归预报模型。但对于地处桂东的桂江流域和清水河流域,遗传-神经网络集合预报模型的预报能力与逐步回归预报模型相当,甚至略逊于逐步回归预报模型。

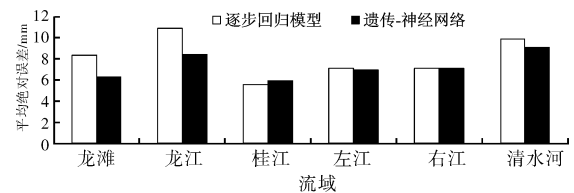


图2 5月份各流域回归模型和遗传-神经网络模型预报平均绝对误差(mm)比较

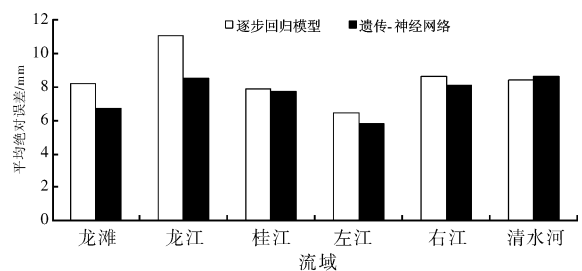


图3 6月份各流域回归模型和遗传-神经网络模型预报平均绝对误差(mm)比较

3.2.4 遗传-神经网络预报模型与其它数值预报模

为了研究各流域面雨量的遗传-神经网络预报模型的预报性能,将遗传-神经网络预报模型对独立样本的预报结果与同期的日本、德国数值预报模式对同样独立样本的预报进行比较。具体方法是将日本、德国的数值预报模式 24 h 降水预报格点值,通过泰森多边形面雨量计算方法,分别求得日本数值模式和德国数值模式对各流域的 24 h 面雨量预报。这里,仅以 6 月各流域的面雨量预报为例进行比较。图 4 中给出了 6 月份遗传-神经网络集合预报模型与日本数值预报模式、德国数值预报模式对各流域独立样本的预报结果对比。从图中可以看到,日本数值预报模式对各流域的独立样本的面雨量预报能力要优于德国的数值模式,而遗传-神经网络集合预报模型的预报能力

则远远优于日本数值预报模式和德国的数值预报模式。为了便于在实际的预报业务应用和为气象服务提供参考, 这里还将遗传-神经网络集合预报模型对独立样本的预报能力与同期广西气象部门的气象服务中心提供给各流域水电站的综合面雨量预报结果进行比较(图4), 从图4中可以看到, 遗传-神经网络集合预报模型的预报能力与气象业务部门提供给各流域水电站的综合面雨量预报能力相当, 遗传-神经网络集合预报模型在龙江流域和龙滩近库区流域的面雨量预报平均绝对误差还略小于气象业务部门的综合预报。

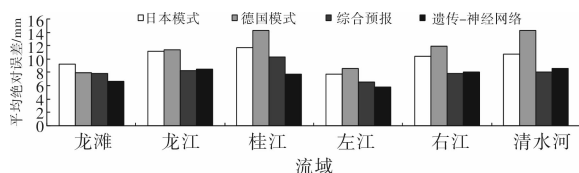


图4 6月份各流域遗传-神经网络模型与其它数值模型预报平均绝对误差(mm)比较

4 结论与讨论

(1)居于非线性方法的遗传-神经网络预报模型与线性的逐步回归预报模型不同, 前者能显示出“学习”能力。遗传-神经网络预报模型对广西各流域5月、6月的面雨量预报能力明显优于逐步回归预报模型, 如龙滩近库区流域和龙江流域遗传-神经网络预报模型对5月和6月份面雨量的预报精度比逐步回归预报模型的精度提高了18.3%~24.2%。

(2)遗传-神经网络集合预报模型对广西各流域5月、6月的面雨量的预报能力远远优于日本数值预报模式和德国的数值预报模式, 并与气象部

门同期制作的综合面雨量预报产品能力相当, 因而, 遗传-神经网络面雨量集合预报模型有较好的业务应用和推广前景。

(3)由于所用的建模样本和检验样本年限只有5年, 样本长度相对较短, 因此, 所建的各流域5月、6月面雨量遗传-神经网络预报模型的稳定性还有待进一步的验证。

参考文献:

- [1] 秦大河, 陈振林, 罗勇, 等. 气候变化科学的最新认知[J]. 气候变化研究进展, 2007, 3(2): 63-73.
- [2] 郑建萌, 任菊章, 张万诚. 云南近百年来温度雨量的变化特征分析[J]. 灾害学, 2010, 25(3): 24-31.
- [3] 胡娟, 闵颖, 李华宏, 等. 云南省山洪地质灾害气象预报预警方法研究[J]. 灾害学, 2014, 29(1): 62-66.
- [4] 陈静静, 姚蓉, 文强, 等. 湖南省降雨型地质灾害致灾雨量阈值分析[J]. 灾害学, 2014, 29(2): 42-47.
- [5] 倪化勇, 王德伟. 基于雨量(强)条件的泥石流预测预报研究现状、问题与建议[J]. 灾害学, 2010, 25(1): 124-128.
- [6] 林开平, 孙崇智, 郑凤琴, 等. 丘陵区地面雨量计算方法及应用[J]. 气象, 2003, 29(10): 8-12.
- [7] 林开平, 孙崇智, 陈冰廉, 等. 广西主要江河流域的面雨量合成分析与洪涝的关系[J]. 热带地理, 2003, 23(3): 222-225.
- [8] Lakshmanan V. Using a genetic algorithm to turn a bounded weak echo region detection algorithm[J]. Journal of Applied Meteorology, 2000, 39(2): 222-230.
- [9] 姚才, 金龙, 黄明策, 等. 遗传算法与神经网络相结合的热带气旋强度预报方法试验[J]. 海洋学报, 2007, 29(4): 11-19.
- [10] 林开平. 人工神经网络的泛化性能与降水预报的应用研究[D]. 南京: 南京信息工程大学, 2007.
- [11] 林健玲, 金龙, 彭海燕. 区域降水数值预报产品人工神经网络释用预报研究[J]. 气象科技, 2006, 34(1): 12-17.
- [12] 钟利华, 钟仕全, 李勇, 等. 广西电网流域面雨量监测、预报、报警系统[J]. 气象研究与应用, 2013, 34(3): 111-112.

A Prediction Scheme with Genetic-Neural Network for Area Rainfall over Power Basin

Qin Wu¹, Lin Kaiping², Huang Ying³, Li Yong⁴, Zhong Lihua⁴ and Luo Xiaoli⁴

(1. Guangxi Meteorological Service, Nanning 530022, China; 2. Guangxi Meteorological Observatory, Nanning 530022, China; 3. Guangxi Research Institute of Meteorological Disasters Mitigation, Nanning 530022, China; 4. Guangxi Meteorological Science and Technology Service Center, Nanning 530022, China)

Abstract: Taking the key hydropower plant and large and middle reservoirs as a major consideration, 23 power basin in Guangxi are divided based on the landform and physiognomy and the distribution of medium and small rivers, and a nonlinear neural network prediction method for area rainfall over power basin has been developed. For the six power basin in May and June, such as Long Tan and Long Jiang power basin, the genetic neural network prediction models for area rainfall over power basin are established, and the connection weight and structure of the BP neural network is optimized using the genetic algorithm. Results of independent samples show that the genetic neural network prediction model is better than the traditional stepwise regression method for area rainfall over power basin, and is superior to the predictions converted from Japan and Germany numerical prediction products, moreover, the prediction capacity of the genetic neural network model is the same as that of the corresponding integrated area rainfall prediction products from meteorological department. Therefore, the genetic neural network model for area rainfall over power basin opens up a vast range of possibilities for operational weather prediction.

Key words: genetic algorithm; artificial neural network; power basin; area rainfall; prediction; Guangxi