

灌区降水量影响因子分析与集合降水预测*

邢贞相 郭皓 付强

(东北农业大学水利与建筑学院, 哈尔滨 150030)

摘要: 针对降水随机性较强、影响因素复杂、单一模型预测精度低的特点,采用集对分析法,研究降水量与气象影响因子的关系。将基于密度参数的径向基函数人工神经网络模型与灰色模型相结合,利用信息熵权重法计算2个单一模型的权重,构建基于信息熵的集合模型(Combing model based on information entropy, IE-CM),用于三江平原友谊农场的降水量预测。研究表明,与单一模型相比,IE-CM模型预测结果的决定系数、平均相对误差及均方根误差较单一模型均有所提高,分别为0.99、10.655%和3.03 mm,预测结果的合格率为83.3%,均满足水文预测要求。

关键词: 集对分析法 集合降水预测 径向基神经网络 灰色模型

中图分类号: P338 **文献标识码:** A **文章编号:** 1000-1298(2015)08-0097-07

Analysis of Influencing Factors of Rainfall in Irrigation Area and Combining Rainfall Forecasting

Xing Zhenxiang Guo Hao Fu Qiang

(School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin 150030, China)

Abstract: Sanjiang Plain is located in the east of Heilongjiang Province, which belongs to the humid climate area. In recent years, Sanjiang Plain's natural state has changed due to several factors, such as the warming climate and human activity. Precipitation is one of the major sources of agricultural irrigation in the irrigation area. Due to the strongly stochastic characteristic of precipitation which was influenced by many factors and the lower accuracy of single forecasting model, set pair analysis was introduced which could discuss the relation between rainfall and meteorological factors. In order to improve the training speed of the radial basis function neural network, the K-means algorithm based on density parameter was applied. In this way, the sensitivity of conventional K-means algorithm to initial clustering center was also removed. A combining model based on information entropy (IE-CM) was built, which combined the radial basis function artificial neural network based on density parameter with the grey model, and the weight of each single model was calculated by using the information entropy weight method. The constructed model was applied to forecast the rainfall over the Youyi Farm in Sanjiang Plain. The case study showed that the determination coefficient, average relative error and root mean square error of IE-CM were better than those of single models, which were demonstrated to be 0.99, 10.655% and 3.03 mm, respectively. The qualification rate of the forecasted result was 83.3%, which satisfied the requirements of hydrologic prediction. In conclusion, the built combining model could provide a new method for forecasting precipitation.

Key words: Set pair analysis Combining rainfall forecasting Radial basis function neural network Grey model

收稿日期: 2014-09-24 修回日期: 2014-10-19

* 国家自然科学基金资助项目(51109036, 51179032)、教育部高等学校博士学科点专项科研基金资助项目(20112325120009)、水利部公益性行业科研专项经费资助项目(201301096)、黑龙江省省级领军人才梯队后备带头人资助项目(500001)、黑龙江省博士后启动基金资助项目(LBH-Q12147)和黑龙江省自然科学基金资助项目(E2015024)

作者简介: 邢贞相, 副教授, 博士, 主要从事水文不确定性及概率预测研究, E-mail: zxxneau@hotmail.com

通讯作者: 付强, 教授, 博士生导师, 主要从事水文与水资源系统分析及水土资源高效利用研究, E-mail: fuqiang0629@126.com

引言

三江平原位于黑龙江省东部,属于气候湿润地区,是“千亿斤”粮食产能基地。近年来,在气候变暖和人类活动频繁干扰等多种因素的影响下,三江平原许多灌区的自然平衡状态正在发生改变。降水作为灌区农业灌溉的主要来源,同时也是自然界水循环的重要环节之一,其与气温、蒸发、湿度等气象因子之间存在着较强的随机性联系。因此,运用不确定性方法分析降水与各气象因子的关系,研究降水的变异特性和预测一直广受关注^[1]。目前,不确定性分析方法主要有随机分析法^[2]、灰色分析法^[3]和集对分析法^[4]等。其中,集对分析法(Set pair analysis, SPA)原理简单,计算简捷,成果合理,已在水资源领域得到了广泛应用。

由于水文系统的复杂性和水文要素的不确定性,单一的降水量预测方法具有一定的局限性。1969年,Bates等首次提出集合模拟和预测^[5],综合利用多种方法所提供的信息,构建集合模型^[6],进而提高模拟和预测的精度。目前为止,主要集合方法有:遗传算法与神经网络集合模型^[7]、免疫算法与粒子群算法集合模型^[8]、灰色关联度与GM(1,1)相结合的预测模型^[9]等。本文将基于密度参数的径向基函数人工神经网络模型(Radial basis function artificial neural network based on density parameter, RBF-ANN)与灰色模型(Grey model, GM)相结合,并利用信息熵权重法计算各单一模型的权重,构建基于信息熵的集合模型(IE-CM),对友谊农场降水量进行集合预测,可克服单一模型的缺陷,从而提高降水模型的精度。

1 研究方法

1.1 集对分析原理

集对分析的核心思想是对不确定性系统中的两个关联的集合构造集对,对集对的特性做同一性、差异性、对立性分析,然后建立集对的同、异、反联系度,以联系度表征集对的关系,从而达到对研究对象相互联系(关系)发展过程的认识^[10]。所谓集对,指有一定联系的两个集合A和B组成的对子,记为 $H(A, B)$ 。描述 $H(A, B)$ 间关系的联系度定义为^[11]

$$\mu_{A \sim B} = \frac{S}{N} + \frac{F}{N}I + \frac{P}{N}J \quad (1)$$

式中 N ——集合特性的总数
 S ——同一特性的个数
 P ——对立特性的个数
 F ——差异特性的个数

I ——差异不确定性系数,在 $(-1, 1)$ 区间视不同情况取值,有时仅作差异标记

J ——对立度系数, $J = -1$,有时仅起对立标记作用

为简便,称 $a = \frac{S}{N}$ 为同一度, $b = \frac{F}{N}$ 为差异度,

$c = \frac{P}{N}$ 为对立度,则式(1)可改写为

$$\mu_{A \sim B} = a + bI + cJ \quad (2)$$

其中 a, b, c 非负,且满足 $a + b + c = 1$ 。

1.2 单一降水预测模型

1.2.1 RBF-ANN模型

RBF-ANN模型是一种收敛速度快、逼近能力强的神经网络模型,广泛应用于水资源领域^[11]。其中,基函数中心的选取是该模型能否成功实现的关键所在,而较为有效的选取方法是K-均值聚类法^[12]。从其原理^[12]不难看出K-均值算法对初始聚类中心很敏感,聚类结果随不同的初始输入而波动,这势必影响最终各样本组群的特征,从而对RBF-ANN模型的收敛结果造成影响。

因此,本文采用基于密度参数的K-均值算法,寻求RBF-ANN模型基函数的中心。将 k 个处于高密度区域的数据对象作为初始聚类中心,以消除传统K-均值算法对初始聚类中心的敏感性,具体步骤见文献^[13]。

RBF-ANN模型的创建采用Matlab的RBF工具箱来实现,具体的程序语言为

$$\text{net} = \text{newrb}(\mathbf{p}, \mathbf{t}, \text{goal}, \text{spread}, \text{mn}, \text{df}) \quad (3)$$

其中 $\text{spread} = bd_i \quad (4)$

式中 \mathbf{p}, \mathbf{t} ——样本的输入和输出向量

goal——均方误差,取 $\text{goal} = 0.001$,防止过度拟合

spread——径向基函数的宽度,通过隐层神经元各个中心的间距来确定

b ——重叠系数,一般取大于1的整数

mn——神经元的最大数目

df——2次显示之间添加的神经元数目

d_i ——基函数中心的最小间距

基函数中心由基于密度参数的K-均值算法求得。本文采用上述建立的RBF-ANN模型作为构建降水量预测的单一模型。

1.2.2 GM(1, N)模型

灰色模型是一种基于灰色系统理论的预测模型,通过对原始数据的处理来寻求其变化规律^[14]。目前大多数水文预测研究多集中于针对水文单变量数据序列进行建模,从而实现水文要素的模拟和预

测^[15]。实际上,降水过程是多种因素共同作用的结果,如气温、蒸发等,单变量的灰色模型很难真实地反映降水过程,故应采用多变量的 GM(1, N) 模型 (N 为因变量的个数) 进行降水预测,具体建模步骤见文献[16]。

1.3 基于信息熵的集合预测模型(IE-CM)

采用信息熵权重法^[17]确定集合预测模型中各单一模型的权重,构建基于信息熵的集合预测模型(IE-CM)。根据求得的各单一模型的权重系数,对 RBF-ANN 模型和 GM(1, N) 模型的预测结果进行集合,具体计算步骤如下:

(1) 计算各单一模型预测结果的相对误差为

$$\begin{cases} e_{\text{RBF-ANN}}(i) = \frac{F_{\text{RBF-ANN}}(i) - O_{\text{RBF-ANN}}(i)}{O_{\text{RBF-ANN}}(i)} \\ e_{\text{GM}(1,N)}(i) = \frac{F_{\text{GM}(1,N)}(i) - O_{\text{GM}(1,N)}(i)}{O_{\text{GM}(1,N)}(i)} \end{cases} \quad (5)$$

式中 i ——时间序列号,取 $i = 1, 2, \dots, n$

$F_{\text{RBF-ANN}}(i)$ 、 $F_{\text{GM}(1,N)}(i)$ ——各单一模型的降水量预测值

$O_{\text{RBF-ANN}}(i)$ 、 $O_{\text{GM}(1,N)}(i)$ ——各单一模型实测降水量序列值

各单一模型的相对误差权重为

$$\begin{cases} p_{\text{RBF-ANN}}(i) = \frac{|e_{\text{RBF-ANN}}(i)|}{\left| \sum_{i=1}^n e_{\text{RBF-ANN}}(i) \right|} \\ p_{\text{GM}(1,N)}(i) = \frac{|e_{\text{GM}(1,N)}(i)|}{\left| \sum_{i=1}^n e_{\text{GM}(1,N)}(i) \right|} \end{cases} \quad (6)$$

(2) 计算各单一模型相对误差的熵值和变异程度系数

$$\begin{cases} H_{\text{RBF-ANN}}(i) = -k \sum_{i=1}^n p_{\text{RBF-ANN}}(i) \ln p_{\text{RBF-ANN}}(i) \\ H_{\text{GM}(1,N)}(i) = -k \sum_{i=1}^n p_{\text{GM}(1,N)}(i) \ln p_{\text{GM}(1,N)}(i) \end{cases} \quad (7)$$

其中 $k = 1/\ln(\cdot)$

式中 $\ln(\cdot)$ ——自然对数

根据式(7)求得的熵值,利用式(8)可得到各模型误差序列的变异程度系数 D 。

$$\begin{cases} D_{\text{RBF-ANN}}(i) = 1 - H_{\text{RBF-ANN}}(i) \\ D_{\text{GM}(1,N)}(i) = 1 - H_{\text{GM}(1,N)}(i) \end{cases} \quad (8)$$

在集合模型中,单一模型误差的信息熵越小,则变异程度越大,不确定程度越大,则该模型的权重越小^[18]。权重系数 W 计算式为

$$\begin{cases} W_{\text{RBF-ANN}}(i) = \frac{1}{M-1} \left(1 - \frac{D_{\text{RBF-ANN}}(i)}{D_{\text{RBF-ANN}}(i) + D_{\text{GM}(1,N)}(i)} \right) \\ W_{\text{GM}(1,N)}(i) = \frac{1}{M-1} \left(1 - \frac{D_{\text{GM}(1,N)}(i)}{D_{\text{RBF-ANN}}(i) + D_{\text{GM}(1,N)}(i)} \right) \end{cases} \quad (9)$$

式中 M ——单一模型的个数,取 $M = 2$

(3) 构建集合预测模型

将各单一模型的预测结果根据权重系数加权求和,即可得集合预测模型的结果,即

$$F_{\text{IE-CM}}(i) = W_{\text{RBF-ANN}}(i) F_{\text{RBF-ANN}}(i) + W_{\text{GM}(1,N)}(i) F_{\text{GM}(1,N)}(i) \quad (10)$$

2 实例分析

2.1 基础数据

为验证所建集合预测模型的有效性和准确性,以三江平原友谊农场为研究区域。友谊农场位于黑龙江省东北部,幅员 18.88 万 km^2 ,耕地面积 10.67 万 hm^2 。境内中型灌区较多,是我国商品粮的主产地。基础数据包括友谊农场 1981—2010 年 360 个月逐月降水量 (y) 和逐月气象数据 (x_1 蒸发量、 x_2 平均气温、 x_3 平均相对湿度、 x_4 平均风速)。

2.2 降水量主要影响因子识别与相关分析

基于集对分析的降水量与气象因子间联系度的确定步骤如下:

(1) 构造集对:1981—2010 年的逐月降水量为因变量 ($Y = y_1, y_2, \dots, y_n, n = 12$) 记作集合 A ,各影响因子为自变量 ($X_i = x_{i1}, x_{i2}, x_{i3}, x_{i4}$) 记作集合 $B_i (i = 1, 2, 3, 4)$,则 A 与 B_i 构成集对 (A, B_i) ,即 (A, B_1) 、 (A, B_2) 、 (A, B_3) 、 (A, B_4) 。

(2) 制定分类标准:通过特定的分类方法,将集合 A 和 B_i 中的元素进行分级量化处理。分级量化标准的关键在于如何确定分类标准,而均值标准差是一种常见的分类方法。本文利用该方法将数据分成 3 级 (I、II、III),分别对应区间为 $[0, \bar{x} + k_1 s)$ 、 $[\bar{x} + k_1 s, \bar{x} + k_2 s)$ 和 $[\bar{x} + k_2 s, \infty)$,其中 \bar{x} 、 s 分别为集合 A 或 B_i 中各元素的均值和均方差, k_1 、 k_2 分别为 0.5、-0.5。集合 A 和 B_i 中各元素的分类区间与分级标准如表 1 所示。

(3) 同一度 a 、差异度 b 、对立度 c 的计算:根据表 1 所示的分类区间,将 A 和 B_i 中各元素进行分类。以 1981 年为例 (其他年份同理),表 2 中给出 A 和 B_i 中各元素分类结果。其中,级数相同的,定义为同,其个数记为 S ;级数相差一级的,定义为异,其个数记为 F ;级数相差两级的,定义为反,其个数记为 P 。由此得到各因子同一度、差异度、对立度计算结果如表 3 所示。

(4)联系度的计算:根据式(1)和表3计算各集对的联系度

$$\mu_{A-B_1} = \frac{204}{360} + \frac{140}{360}I + \frac{16}{306}J = 0.567 + 0.389I + 0.044J \quad (11)$$

$$\mu_{A-B_2} = \frac{238}{360} + \frac{114}{360}I + \frac{8}{306}J = 0.661 + 0.371I + 0.022J \quad (12)$$

$$\mu_{A-B_3} = \frac{163}{360} + \frac{153}{360}I + \frac{44}{306}J = 0.453 + 0.425I + 0.122J \quad (13)$$

$$\mu_{A-B_4} = \frac{81}{360} + \frac{187}{360}I + \frac{92}{306}J = 0.225 + 0.519I + 0.256J \quad (14)$$

(5)主要影响因子识别:联系度 μ 反映了降水量和各影响因子之间的相关程度。 a 表示正相关度; b 表示既非正又非负的不定相关度; c 表示负相关度。如果 $a > c$,表示两个水文变量存在正的相关关系;如果 $a < c$,表示两个水文变量存在负的相关关系;如果 b 较大时,表示两个水文变量间的不定相关关系较强^[18]。因此,能根据 a 、 b 、 c 的大小判断相关性的性质。

由式(11)~(14)可以看出, A 和 B_1 的正相关度为0.567,负相关度为0.044,不定相关度为0.389,综合起来表现为正相关。计算 A 和 B_1 的相关系数为0.458,可见是吻合的。 A 和 B_2 的正相关

度为0.661,负相关度为0.022,不定相关度为0.371,综合起来表现为正相关。计算 A 和 B_2 的相关系数为0.678。 A 和 B_3 的正相关度为0.453,负相关度为0.122,不定相关度为0.425,综合起来表现为正相关。计算 A 和 B_3 的相关系数为0.406。 A 和 B_4 的正相关度为0.225,负相关度为0.256,不定相关度为0.519,综合起来表现为负相关,但不定的相关度占50%以上。计算 A 和 B_4 的相关系数为-0.238。

综上所述,联系度能从微观层面上揭示 A 和 B_i 的正负相关和不定相关情况,无疑是对从宏观层面分析总体相关系数的补充和完善。同时,揭示出在某种情况下可能并不存在理想的相关关系。如 A 和 B_4 。

根据传统的分析方法, x_1 、 x_2 、 x_3 、 x_4 均通过 $\alpha = 0.05$ 的显著性检验,说明 x_1 、 x_2 、 x_3 、 x_4 均可作为降水量的主要影响因子。应用集对原理对其进行探讨。根据式(11)~(14), A 和 B_i 的不定相关度分别为0.389、0.371、0.425和0.519。一般而言,不定相关度越大,可确定的相关度所占的比例就越小,表示该因子作为主要影响因子的可能性就越小;而不定相关度大于0.5时,可确定的相关度小于0.5,即可考虑排除该因子^[18]。由此可得, x_4 可不作为 y 的主要影响因子,即影响 y 的主要因子为 x_1 、 x_2 、 x_3 ,分别为蒸发量、平均气温和平均相对湿度。

表1 集合 A 和 B_i 中各元素的分类区间

Tab.1 Value range of elements in set A and B_i

指标	B_1/mm	$B_2/^\circ\text{C}$	$B_3/\%$	$B_4/(\text{m}\cdot\text{s}^{-1})$	A/mm
\bar{x}	99.1	3.8	67.7	3.1	42.6
s	73.8	14.4	9	0.9	49.3
I	[0, 62.2)	[0, -3.5)	[0, 63.2)	[0, 2.6)	[0, 18.0)
II	[62.2, 136.1)	[-3.5, 11.0)	[63.2, 72.2)	[2.6, 3.6)	[18.0, 67.3)
III	[136.1, ∞)	[11.0, ∞)	[72.2, ∞)	[3.6, ∞)	[67.3, ∞)

表2 A 和 B_i 中各元素分类结果(1981年)

Tab.2 Classification of elements in set A and B_i

月份	B_1/mm	B_1 分类	$B_2/^\circ\text{C}$	B_2 分类	$B_3/\%$	B_3 分类	$B_4/(\text{m}\cdot\text{s}^{-1})$	B_4 分类	A/mm	A 分类
1	14.7	I	-17.3	I	70	II	4.2	III	0.1	I
2	20.3	I	-14.2	I	69	II	2.8	II	2.2	I
3	58.7	I	-5.7	I	66	II	2.9	II	19.0	II
4	165.9	III	7.3	II	57	I	3.3	II	20.4	II
5	200.9	III	12.7	III	61	I	3.0	II	68.0	III
6	177.8	III	17.7	III	77	III	2.2	I	119.9	III
7	198.1	III	22.5	III	81	III	2.2	I	136.3	III
8	133.0	II	19.0	III	86	III	2.3	I	186.8	III
9	125.1	II	13.1	III	75	III	2.4	I	83.4	III
10	77.7	II	5.3	II	74	III	3.1	II	42.5	II
11	29.4	I	-8.7	I	68	II	3.9	III	4.3	I
12	15.3	I	-12.6	I	76	III	2.6	I	3.5	I

表3 各因子同一度、差异度及对立度

Tab.3 Identical degree, difference degree and opposite degree of factors

影响因素	同一度 a	差异度 b	对立度 c
B_1/mm	0.567	0.389	0.044
$B_2/^\circ\text{C}$	0.661	0.371	0.022
$B_3/\%$	0.453	0.425	0.122
$B_4/(\text{m}\cdot\text{s}^{-1})$	0.225	0.519	0.256

2.3 单一降水预测模型预测结果

2.3.1 基于密度参数的 RBF-ANN 模型的模拟和预测

本文选取 1981—2004 年为 RBF-ANN 模型的率定期, 2005—2007 年为模型的验证期, 2008—2010 年为模型预测期。由于蒸发量、平均气温和平均相对湿度为降水量的 3 个显著影响因子, 因此 RBF-ANN 网络的输入节点个数为 3 个。输出层为当月降水量, 节点数为 1 个。根据经验公式和试算, 利用率定期降水数据对模型进行训练, 确定隐含层节点数为 5 个, 至此, 网络拓扑结构确定为 3:5:1。

图 1 显示验证期内的降水量模拟值和实际值拟合情况较好, 经计算, 验证期拟合结果的平均相对误差和决定系数分别为 12.055% 和 0.97, 达到水文模拟的相关精度要求。通过以上验证, 可利用已构建的 RBF-ANN 模型对 2008、2009、2010 年逐月降水量进行预测, 预测结果和实测值的拟合情况如图 2 所示, 平均相对误差、决定系数和均方根误差分别为 13.840%、0.97 和 3.24 mm (表 4), 预测合格率为 77.8% (以相对误差介于 -20% ~ 20% 为合格标准)。综上, 预测期内降水量预测精度一般, 平均相对误差较大, 精度有待提高, 但合格率较高。

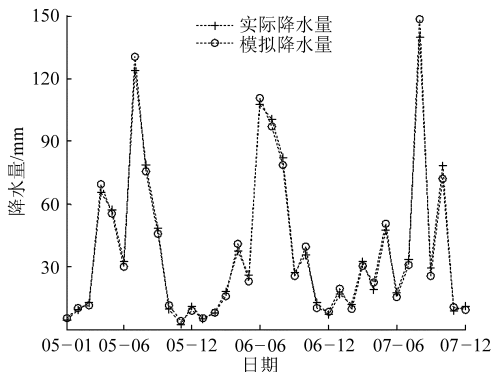


图1 RBF-ANN 模型拟合值与实际值对比

Fig.1 Comparison between simulated and observed values of RBF-ANN model

2.3.2 GM(1,4)模型的模拟和预测

采用 1981—2004 年逐月降水量及各主要影响因子的数据资料, 建立 GM(1,4) 模型, 模拟降水量与各影响因子的关系, 用 2005—2007 年的数据对模

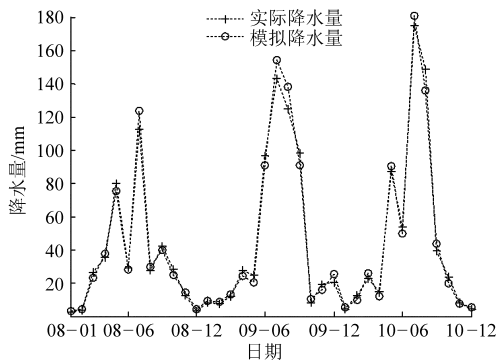


图2 2008—2010年逐月降水量预测值与实际值对比
Fig.2 Comparison between forecasted and observed monthly rainfall from 2008 to 2010

表4 RBF-ANN 模型逐月降水量模拟和预测评价
Tab.4 Simulation and forecasted evaluation of monthly rainfall by RBF-ANN model

时段	决定系数	平均相对误差/%	均方根误差/mm
验证期(2005—2007年)	0.97	12.055	3.58
预测期(2008—2010年)	0.97	13.840	3.24

型进行验证, 并对 2008—2010 年逐月降水量进行预测。根据参数估计, 得到 GM(1,4) 模型的参数列 $\hat{m} = [m, n_2, n_3, n_4]^T = [0.084 \quad 0.072 \quad 0.322 \quad -0.070]^T$, 即 $m = 0.084, n_2 = 0.072, n_3 = 0.322, n_4 = -0.070$ 。因此, GM(1,4) 模型的白热方程为

$$\frac{dg_1^{(1)}(t)}{dt} + mg_1^{(1)}(t) = n_2g_2^{(1)}(t) + n_3g_3^{(1)}(t) + n_4g_4^{(1)}(t) \quad (15)$$

其中, $g_1^{(1)}(t), g_2^{(1)}(t), g_3^{(1)}(t)$ 和 $g_4^{(1)}(t)$ 分别为降水量、蒸发量、平均气温、平均相对湿度的依次累加序列。据此, 可得 GM(1,4) 模型的近似时间响应函数为

$$\hat{g}_1^{(1)}(k+1) = \left(g_1^{(1)}(0) - \frac{1}{m} \sum_{i=2}^4 n_i g_i^{(1)}(k+1) \right) e^{-mk} + \frac{1}{m} \sum_{i=2}^4 n_i g_i^{(1)}(k+1) \quad (16)$$

将式(16)作一次累减生成, 便可得到还原序列 $\hat{g}_1^{(0)}(k)$, 并与原始序列 $g_1^{(0)}$ 对比。

对于验证期, 给出模型的模拟值与实际值的对比结果, 如图 3 所示。经计算可得, 平均相对误差和决定系数分别为 10.386% 和 0.98。通过以上验证, 可利用已构建的 GM(1,4) 模型对 2008、2009、2010 年逐月降水量进行预测, 预测结果和实测值的对比如图 4 所示, 平均相对误差、决定系数和均方根误差分别为 12.980%、0.98 和 3.58 mm (表 5), 预测合格率为 83.3%。与 RBF-ANN 预测结果比较, 降水量预测情况优于 RBF-ANN 模型, 平均相对误差小

于 RBF-ANN 模型的误差,且合格率更高,但均方根误差较大,精度还有待提高。

时,表7分别给出3种模型对预测年间月降水量最大值、最小值的误差分析。

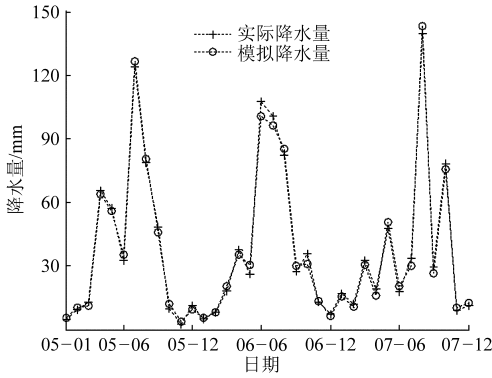


图3 GM(1,4)模型拟合值与实际值对比
Fig.3 Comparison between simulated and observed values of GM(1,4) model

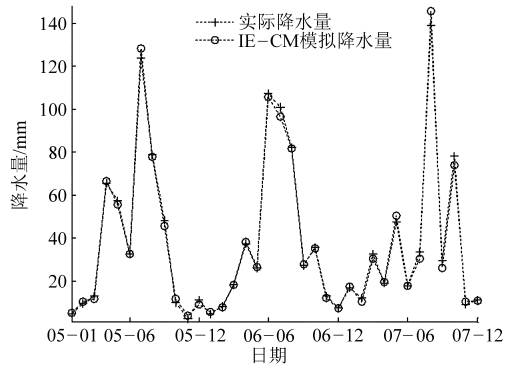


图5 IE-CM模型拟合值与实际值对比
Fig.5 Comparison between simulated value and observed value of IE-CM model

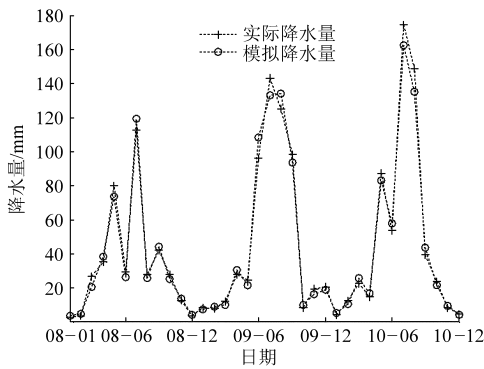


图4 2008—2010年逐月降水量预测值与实际值对比
Fig.4 Comparison between forecasted and observed monthly rainfall from 2008 to 2010

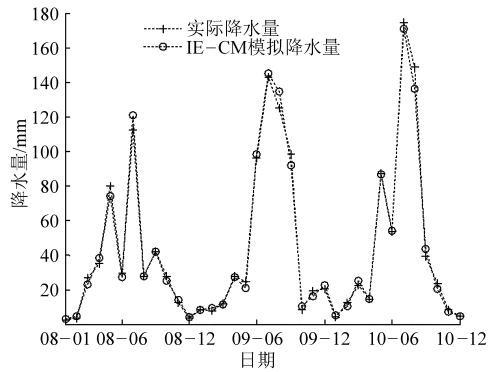


图6 2008—2010年逐月降水量预测值与实际值对比
Fig.6 Comparison between forecasted and observed monthly rainfall from 2008 to 2010

表5 GM(1,4)模型逐月降水量模拟和预测评价
Tab.5 Simulation and forecasted evaluation of monthly rainfall by GM(1,4) model

时段	决定系数	平均相对误差/%	均方根误差/mm
验证期(2005—2007年)	0.98	10.386	3.72
预测期(2008—2010年)	0.98	12.980	3.58

表6 IE-CM模型逐月降水量模拟和预测评价
Tab.6 Simulation and forecasted evaluation of monthly rainfall by IE-CM model

时段	决定系数	平均相对误差/%	均方根误差/mm
模拟期(2005—2007年)	0.99	7.606	2.80
预测期(2008—2010年)	0.99	10.655	3.03

2.4 IE-CM模型预测结果

根据2.3节的信息熵权重法得到模拟期RBF-ANN模型和GM(1,4)模型在集合模型中的权重系数分别为 $W_1 = 0.5014$, $W_2 = 0.4986$ 。再根据权重系数建立IE-CM模型,并对2005—2007年逐月降水量进行模拟,模拟结果如图5所示,平均相对误差和决定系数分别为7.606%和0.99。而对于2008—2010年来说,RBF-ANN模型和GM模型在集合模型中的权重系数分别为 $W_1 = 0.5094$, $W_2 = 0.4901$,建立IE-CM模型对2008—2010年逐月降水量进行预测,预测值与实际值的对比如图6所示,平均相对误差、决定系数和均方根误差分别为10.655%、0.99和3.03mm(表6),预测合格率为83.3%。同

表7 3种模型对2008—2010年间月降水量最大值、最小值的误差分析

Tab.7 Error analysis of maximum and minimum monthly rainfall by three models from 2008 to 2010

模型	%					
	最大值相对误差			最小值相对误差		
	2008年	2009年	2010年	2008年	2009年	2010年
RBF-ANN	9.786	7.983	3.668	26.087	18.919	28.571
GM(1,4)	5.516	5.112	8.028	39.130	24.324	21.429
IE-CM	7.692	1.559	2.067	32.486	21.571	25.067

综合分析可知:①RBF-ANN模型和GM(1,4)模型预测结果的决定系数相当,GM(1,4)模型的平均相对误差较低,但均方根误差明显高于RBF-ANN模型。而集合预测模型的决定系数、平均相对

误差及均方根误差较单一模型均有所提高,决定系数分别提高2%、1%,平均相对误差分别降低22%、18%,均方根误差分别降低6%、15%。②从月降水量最大值的预测结果看出,GM(1,4)模型的相对误差总体上小于RBF-ANN模型,仅在2010年远大于RBF-ANN模型的误差。从月降水量最小值的预测结果看出,RBF-ANN模型的相对误差总体上小于GM(1,4)模型,仅在2010年大于RBF-ANN模型的误差。③应用集合模型对月降水量最大值预测的相对误差明显小于单一模型,仅在2008年介于2种单一模型之间,而3种模型对月降水量最小值的预测结果不佳,原因有待进一步研究。④GM(1,4)模型和集合模型的预测合格率均明显大于RBF-ANN模型。综上,集合建模具有一定的可行性,在不同程度上提高了降水量预测精度。

3 结论

(1)从2008—2010年逐月降水量的预测结果

可知,集合预测模型的决定系数、平均相对误差及均方根误差较单一模型均有所提高,分别为0.99、10.655%和3.03 mm,且预测结果的合格率较高。

(2)从2008—2010年月降水量最大值的预测结果可知,集合预测模型预测结果的相对误差明显小于单一模型,分别为7.692%、1.559%和2.067%。可见,集合预测模型对降水最大值具有较好的适应性。

(3)从2008—2010年月降水量最小值的预测结果可知,3种模型的预测结果均不佳。这说明最小值降水量具有较强的随机性,因此,对于降水量的最小值预测应进行针对性的研究,并结合降水发生概率,进行概率预测,是今后研究的重点方向。

(4)集合预测模型能够降低单一预测模型的误差,从整体上提高降水量的预测精度,但也受到各单一模型预测结果及其权重的影响,因此,可针对集合模型权重的确定进行研究,以提高集合模型的适应性。

参 考 文 献

- Lin G F, Jhong B C, Chang C C. Development of an effective data-driven model for hourly typhoon rainfall forecasting [J]. *Journal of Hydrology*, 2013, 495(12): 52-63.
- Pang L, Li L P, Chen G J, et al. Parameter uncertainty analysis of reservoir operating rules based on implicit stochastic [J]. *Journal of Hydrology*, 2014, 514(6): 102-113.
- 袁喆,杨志勇,史晓亮,等.灰色微分动态自记忆模型在径流模拟及预测中的应用[J].水利学报,2013,44(7):791-799. Yuan Zhe, Yang Zhiyong, Shi Xiaoliang, et al. Differential hydrological grey self-memory model for runoff simulation and prediction [J]. *Journal of Hydraulic Engineering*, 2013, 44(7):791-799. (in Chinese)
- Feng L H, Sang G S, Hong W H. Statistical prediction of changes in water resources trends based on set pair analysis [J]. *Water Resources Management*, 2014, 28(6):1703-1711.
- Bates J M, Granger C W. Combination of forecasts [J]. *Operational Research Quarterly*, 1969, 20(4): 451-468.
- 袁喆,严登华,杨志勇,等.集合建模在径流模拟和预测中的应用[J].水利学报,2014,45(3):351-359. Yuan Zhe, Yan Denghua, Yang Zhiyong, et al. Ensemble model and its application in runoff simulation and forecast [J]. *Journal of Hydraulic Engineering*, 2014, 45(3):351-359. (in Chinese)
- Kasisviswanathan K S, Cibin R, Sudheer K P, et al. Constructing prediction interval for artificial neural network rainfall runoff models based on ensemble simulations [J]. *Journal of Hydrology*, 2013, 499: 275-288.
- 吴静敏,左洪福,陈勇.基于免疫粒子群算法的组合预测方法[J].系统工程理论方法应用,2006,15(3):229-233. Wu Jingmin, Zuo Hongfu, Chen Yong. A combined forecasting method based on particle swarm optimization with immunity algorithms [J]. *Systems Engineering - Theory Methodology Applications*, 2006, 15(3): 229-233. (in Chinese)
- 曾波,刘思峰,方志耕,等.灰色组合预测模型及其应用[J].中国管理科学,2009,17(5):150-155. Zeng Bo, Liu Sifeng, Fang Zhigeng, et al. Grey combined forecast models and its application [J]. *Chinese Journal of Management Science*, 2009, 17(5):150-155. (in Chinese)
- Zhao K Q, Xuan A L. Set pair theory—a new theory method of non-define and its application [J]. *Systems Engineering*, 1996, 14(1):18-23.
- Muhammad S, Asaad Y S, Bruce W M. Comparative study of different wavelet based neural network models for rainfall-runoff modeling [J]. *Journal of Hydrology*, 2014, 515: 47-58.
- 张建辉. K-means 聚类算法研究及应用[D].武汉:武汉理工大学,2007. Zhang Jianhui. Research and application of K-means clustering algorithm[D]. Wuhan: Wuhan University of Technology, 2007. (in Chinese)
- 赖玉霞,刘建平. K-means 算法的初始聚类中心的优化[J].计算机工程与应用,2008,44(10):147-149. Lai Yuxia, Liu Jianping. Optimization study on initial center of K-means algorithm [J]. *Computer Engineering and Applications*, 2008, 44(10): 147-149. (in Chinese)

- 14 赵骅,景峰. 阎家沟煤矿区弃土坡模拟降雨入渗规律研究[J]. 山西水土保持科技, 2008(3): 17-19.
Zhao Yi, Jing Feng. Study on simulation rainfall infiltration law in spoil from Yanjiagou mine area [J]. Soil and Water Conservation Science and Technology in Shanxi, 2008(3): 17-19. (in Chinese)
- 15 杨艳芬,王全九,曾辰,等. 土石混合介质水分入渗特性试验研究[J]. 水土保持学报, 2009, 23(5): 87-90, 132.
Yang Yanfen, Wang Quanjiu, Zeng Chen, et al. Experimental research on water infiltration characteristics of stony-soil medium [J]. Journal of Soil and Water Conservation, 2009, 23(5): 87-90, 132. (in Chinese)
- 16 刘晓丽,马理辉,杨荣慧,等. 黄土半干旱区枣林深层土壤水分消耗特征[J]. 农业机械学报, 2014, 45(12): 139-145.
Liu Xiaoli, Ma Lihui, Yang Ronghui, et al. Deep soil water depletion characteristic of jujube plantation in loess semiarid region [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12): 139-145. (in Chinese)
- 17 Hillel Daniel. Applications of soil physics[M]. San Diego, CA: Academic Press, 1998.
- 18 刘建国,王洪涛,聂永丰. 多孔介质非饱和导水率预测的分形模型[J]. 水科学进展, 2004, 15(3): 269-275.
Liu Jianguo, Wang Hongtao, Nie Yongfeng. Fractal model for predicting of unsaturated hydraulic conductivity of porous media [J]. Advance in Water Science, 2004, 15(3): 269-275. (in Chinese)
- 19 李卓,吴普特,冯浩,等. 容重对土壤水分入渗能力影响模拟试验[J]. 农业工程学报, 2009, 25(6): 40-45.
Li Zhuo, Wu Pute, Feng Hao, et al. Simulated experiment on effect of soil bulk density on soil infiltration capacity [J]. Transactions of the CSAE, 2009, 25(6): 40-45. (in Chinese)
- 20 吴军虎,张铁钢,赵伟,等. 容重对不同有机质含量土壤水分入渗特性的影响[J]. 水土保持学报, 2013, 27(3): 63-67, 268.
Wu Junhu, Zhang Tiegang, Zhao Wei, et al. Influence of soil bulk density on soil water infiltration characteristics under different soil organic matter contents [J]. Journal of Soil and Water Conservation, 2013, 27(3): 63-67, 268. (in Chinese)
- 21 刘春成,李毅,任鑫,等. 四种入渗模型对斥水土壤入渗规律的适用性[J]. 农业工程学报, 2011, 27(5): 62-67.
Liu Chuncheng, Li Yi, Ren Xin, et al. Applicability of four infiltration models to infiltration characteristics of water repellent soils [J]. Transactions of the CSAE, 2011, 27(5): 62-67. (in Chinese)
- 22 刘月梅,张兴昌. EN-1对黄土性固化土水分垂直入渗特征的影响[J]. 农业机械学报, 2012, 43(11): 65-73.
Liu Yuemei, Zhang Xingchang. Effects of EN-1 soil stabilizer on water vertical infiltration characteristics for loess stabilized soil [J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(11): 65-73. (in Chinese)
- 23 余冬立,郑加兴,刘营营,等. 围垦年限和土壤容重对海涂土壤水分运动参数的影响[J]. 农业机械学报, 2015, 46(2): 120-125.
She Dongli, Zheng Jiaying, Liu Yingying, et al. Effects of land reclamation time and bulk density on soil hydraulic parameters in the coastal reclamation region [J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(2): 120-125. (in Chinese)
- 24 邵明安,王全九,黄明斌. 土壤物理学[M]. 北京: 高等教育出版社, 2006.
- 25 党宏宇,陈洪松,邵明安. 喀斯特地区不同层次土石混合介质对土壤水分入渗过程的影响[J]. 农业工程学报, 2012, 28(8): 38-43.
Dang Hongyu, Chen Hongsong, Shao Ming'an. Effects of laminated rock fragments on soil infiltration processes in karst regions [J]. Transactions of the CSAE, 2012, 28(8): 38-43. (in Chinese)

(上接第 103 页)

- 14 Karmakar S, Mujumdar P P. Grey fuzzy optimization model for water quality management of a river system [J]. Advances in Water Resources, 2006, 29(7): 1088-1105.
- 15 解苗苗,王文圣,王红芳. 灰色自记忆模型在年径流预测中的应用[J]. 水电能源科学, 2007, 25(2): 8-10.
Xie Miaomiao, Wang Wensheng, Wang Hongfang. Comprehensive assessment of river health complex system based on Elman network model [J]. Water Resources and Power, 2007, 25(2): 8-10. (in Chinese)
- 16 王文圣,张翔,金菊良,等. 水文学不确定性分析方法[M]. 北京: 科学出版社, 2011: 261-263.
- 17 黎育红,陈玥,周建中,等. 熵权法并联组合建模在径流预测中的应用[J]. 水电能源科学, 2012, 30(2): 13-16.
Li Yuhong, Chen Yue, Zhou Jianzhong, et al. Entropy-based parallel combinational model and its application in runoff prediction [J]. Water Resources and Power, 2012, 30(2): 13-16. (in Chinese)
- 18 王文圣,李跃清,金菊良. 基于集对原理的水文相关分析[J]. 四川大学学报: 工程科学版, 2009, 41(2): 1-5.
Wang Wensheng, Li Yueqing, Jin Juliang. Relation analysis of hydrological variables based on set pair principle [J]. Journal of Sichuan University: Engineering Science Edition, 2009, 41(2): 1-5. (in Chinese)