

多重组合神经网络模型在年径流预测中的应用

崔东文

(云南省文山州水务局,云南 文山 663000)

摘要:针对线性组合预测模型预测精度不高、单一预测模型权重较难确定和非线性组合预测模型组合函数难以构造等问题,为最大限度地挖掘输入向量间的有用信息以及充分发挥神经网络模型的高度非线性映射能力,提出一种基于BP、Elman、RBF、GRNN这4种神经网络算法原理的多重组合年径流预测模型。以4种单一预测模型的预测结果作为一次组合预测模型的输入向量,实测流量作为输出向量,构建4输入1输出的一次组合预测模型;再以一次组合预测模型预测结果作为二次组合预测模型的输入向量,实测流量作为输出向量,构建4输入1输出的二次组合预测模型;依次类推,构建12种多重组合预测模型。以新疆伊犁河雅马渡站年径流预测为例,将预测结果与4种单一预测模型及IEA-BP模型的预测结果进行比较,结果表明:多重组合预测模型的预测精度和泛化能力较单一预测模型均有较大提高,随着模型组合重数的增加,预测精度呈提高趋势,是提高预测精度的有效方法。

关键词:径流预测;组合模型;BP神经网络;Elman神经网络;RBF神经网络;GRNN神经网络

中图分类号:TV121 **文献标志码:**A **文章编号:**1006-7647(2014)02-0059-05

Application of multiple combined neural network model in annual runoff prediction//CUI Dongwen (Wenshan Water Conservancy Bureau of Yunnan Province, Wenshan 663000, China)

Abstract: In light of the problems that the prediction accuracy of a linear combination model is not high, the weight of a single prediction model is hard to determine, and the function of a nonlinear combined prediction model is hard to construct, a multiple combined annual runoff prediction model was put forward based on the principle of neural networks such as BP, Elman, RBF, GRNN, so as to maximize the useful information between the input vectors, and give full play to the characteristics of neural network models, such as their nonlinear mapping ability. Using the results of four single prediction models as the input vectors of one combined prediction model, and the observed flow data as the output vector, one combined prediction model with four inputs and one output was constructed. Then, using the result of the one combined prediction model as the input vector for two combined prediction models, and the observed flow data as the output vector, two combined prediction models with four inputs and one output were constructed. Using the same method, multiple combined prediction models with 12 kinds of construction schemes were constructed. Taking the Yili River Yamadu hydrological station in Xinjiang as an example, its annual runoff prediction results were compared with those of four kinds of single BP models and IEA-BP models. The results show that the prediction accuracy and generalization ability of multiple combined prediction models are improved as compared with the single prediction model, and with the increase of the combined number of model, the prediction accuracy tends to be improved. Multiple combined prediction models can improve the prediction accuracy.

Key words: runoff prediction; combined model; back-propagation neural network; Elman neural network; radial basis function neural network; generalized regression neural network

1 研究背景

河川径流受多种因素的影响和制约,表现出复杂、随机、多维等特性,探寻能够提高预测精度的模型对河川径流预测预报具有重要的现实意义和应用价值。人工神经网络中常见的BP神经网络、Elman神经网络、RBF神经网络和GRNN神经网络均广泛

运用于径流预测^[1-5]。实践表明,不同的预测方法往往有着不同的预测结果,不同的预测方法挖掘不同的有用信息,不同模型的预测结果通常具有互补性,其预测精度也各不相同^[6],任何一种预测方法都有其适用性和局限性,应依据实际情况选择适当的模型与方法^[7]。组合预测是Bates等^[8]在1969年提出的,其目的是为了有效地利用各种模型的优点,将

不同的预测方法进行适当组合,综合利用各种方法所提供的信息,从而尽可能提高预测精度。这种预测模型构建的关键是寻找一个最优的组合方法,目前普遍采用线性组合预测,如加权平均法、回归法等。然而,线性组合预测存在两个关键性问题:一是如何客观确定各单一预测模型的权重;二是由于线性组合方法是不同预测方法的值之间的一种凸组合,对于一个复杂系统的预测对象,简单的线性相加显然不能满足预测精度要求,而建立非线性组合预测模型又存在难以构造组合函数等困难^[9]。从组合预测模型构成上讲,目前常见的模型组合形式主要是以移动平均、线性回归、自回归滑动平均等统计预测算法与神经网络模型相结合为主,而完全基于神经网络的组合预测模型并不多见,尤其是用于年径流预测的神经网络组合模型更是少见。

为能最大限度地挖掘输入变量间的有用信息以及充分发挥神经网络模型具有的高度非线性映射能力的特性,本文提出一种基于 BP、Elman、RBF、GRNN 这 4 种神经网络算法原理的多重组合年径流预测模型,构建了 12 种多重组合的 BP、Elman、RBF 和 GRNN 神经网络年径流组合预测模型,并与 BP、Elman、RBF、GRNN 这 4 种单一预测模型及文献[10]提出的 IEA-BP 模型的预测结果进行比较。鉴于网络在训练过程中出现“欠拟合”和“过拟合”而降低网络泛化能力的问题,提出了网络拟合度的概念和经验公式,用于网络“欠拟合”和“过拟合”的控制和评价。为验证该方法的有效性,以新疆伊犁河

雅马渡站年径流预测为例进行验证,结果表明所提出的多重组合神经网络年径流预测模型能够以较高的精度进行预测。

2 多重组合神经网络年径流预测模型

BP、Elman、RBF 和 GRNN 神经网络是人工神经网络中较为常见的学习型神经网络,在函数逼近、模式识别、分类和数据压缩等领域有着广泛的应用。其网络的拓扑结构及算法可参阅文献[11-14]。

多重组合神经网络年径流预测模型预测原理是:将影响年径流预测的相关因子作为 BP、Elman、RBF 和 GRNN 神经网络模型的输入向量,实测流量作为输出向量,构建 BP、Elman、RBF 和 GRNN 神经网络单一预测模型;以 BP、Elman、RBF 和 GRNN 神经网络单一预测模型预测结果作为一次组合 BP、Elman、RBF 和 GRNN 神经网络预测模型的输入向量,实测流量作为输出向量,从而构建 4 输入 1 输出的一次组合 BP、Elman、RBF 和 GRNN 神经网络年径流预测模型(模型 1~4);再以一次组合预测模型预测结果作为二次组合 BP、Elman、RBF 和 GRNN 神经网络预测模型的输入向量,实测流量作为输出向量,构建 4 输入 1 输出的二次组合 BP、Elman、RBF 和 GRNN 神经网络年径流预测模型(模型 5~8);依次类推,构建多重(本文以“三重”为例)组合神经网络年径流预测模型(模型 9~12)。多重组合神经网络年径流预测模型的算法流程见图 1。

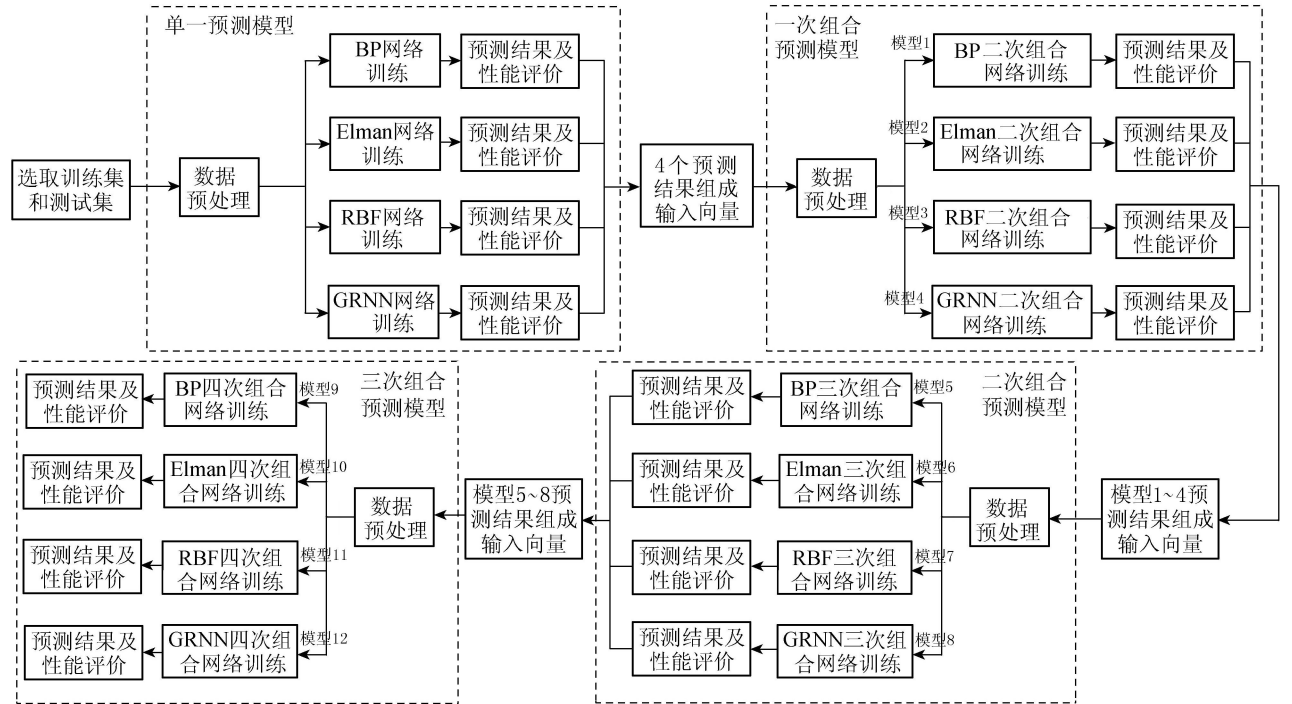


图 1 多重组合神经网络年径流预测模型预测流程

3 模型性能评价

网络在训练过程中出现的“欠拟合”和“过拟合”均会使网络的泛化能力下降。网络“欠拟合”表示模型无法完全探测到复杂数据集中的信号,使得训练达不到预期的评价或预测精度。网络“过拟合”表示模型会将信号连同噪声一起进行拟合,使得模型“记住”了训练样本的信息而使网络泛化能力降低,得不到理想的评价或预测精度^[15-16]。目前,理论上尚无较好解决网络训练过程中由于“欠拟合”和“过拟合”导致泛化能力降低的方法。笔者经过大量验证,提出以网络拟合度指标作为判定网络“欠拟合”与“过拟合”的依据。网络拟合度是指在预测样本平均相对误差足够小,且满足预测样本不少于样本总数的 1/4 条件下,拟合样本和预测样本平均相对误差之差的绝对值与平均相对误差之和的比值(以百分比表示),当网络拟合度小于 40% 时,认为网络处于较理想状态;小于 20% 时网络处于理想状态;大于 40% 时认为网络处于“过拟合”或“欠拟合”状态,即不理想状态。因此,本文选用的模型性能评价指标有平均相对误差 e_{MRE} 、最大相对误差 e_{maxRE} 、平均绝对误差 e_{MAE} 3 个统计学指标和本文提出的网络拟合度 e_{WLNHD} ,各评价指标越小,模型的性能越好。各指标的计算公式如下:

$$e_{MRE} = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \tag{1}$$

$$e_{maxRE} = \max_{1 \leq i \leq n} \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \tag{2}$$

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \tag{3}$$

$$e_{WLNHD} = \frac{|e'_{MRE} - e''_{MRE}|}{e'_{MRE} + e''_{MRE}} \times 100\% \tag{4}$$

式中: $\hat{y}_i$ 为第 i 个样本模拟值; y_i 为第 i 个样本实测值; n 为模拟样本数; e'_{MRE} 、 e''_{MRE} 分别为拟合样本和预测样本的平均相对误差。

4 实例验证

4.1 数据来源

以新疆伊犁河雅马渡站年径流预测为例进行实例分析,详细数据来源参见文献[10]。以文献[10]中的总降雨量等 4 个预报因子作为网络输入向量;以流量实测值作为输出向量,构建 4 输入 1 输出的预测模型。为便于与文献[10]比较,将前 16 年的资料作为各单一预测模型和组合模型的训练样本,后 7 年资料进行预测检验。为有利于网络训练,采用以下方法将各指标数据无量纲化到[0, 1]之间:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

式中: $\hat{x}$ 为经过标准化处理的数据; x 为原始数据; $x_{\max}$ 和 $x_{\min}$ 分别为数据序列中的最大数和最小数。

4.2 网络训练

基于 MATLAB 环境,创建及训练 BP、Elman、RBF、GRNN 这 4 种单一预测模型及多重组合预测模型进行年径流预测。对于 BP、Elman 神经网络,由于隐含层节点数、期望误差、训练次数、传递函数和训练函数等参数目前并没有较理想的确定和选择方法,主要是凭经验确定和选取。本例的做法是^[14]:①依据经验给定期望误差和训练次数,根据 Kolmogorv 定理得出 1 个初始隐含层节点数,然后利用逐步增长或逐步修剪法确定最终隐含层节点数;②采用调整期望误差或训练次数,以及替换传递函数和训练函数的方法以提高网络预测精度;③最后采用网络拟合度评价网络的泛化能力,以获得最佳预测效果的相关参数。对于 RBF、GRNN 神经网络,人为调节的参数少,只有径向基函数扩展系数 SPREAD。本例的做法是:采用 MATLAB 软件编写循环算法程序,以网络拟合度评价网络的泛化能力,最终确定 RBF 和 GRNN 神经网络具有最佳预测效果时的 SPREAD 值。本例兼顾拟合和预测精度,经反复调试,在表 1 参数设置条件下,BP、Elman、RBF、GRNN 这 4 种单一预测模型及各多重组合预测模型具有较好的预测效果(除此之外的其他参数采用系统默认值)。

表 1 单一预测模型和组合预测模型最佳相关参数

预测模型	隐含层节点数	隐含层函数	输出层函数	训练函数	期望误差	训练次数	径向基函数系数
BP	4	logsig	purelin	traingdx	0.030	1 000	
Elman	4	logsig	purelin	traingdx	0.020	1 000	
RBF					0.020		3
GRNN							0.4
模型 1	6	tansig	purelin	traingdx	0.015	1 000	
模型 2	5	logsig	purelin	traingdx	0.010	1 000	
模型 3					0.015		20
模型 4							0.22
模型 5	3	logsig	purelin	traingdx	0.013	1 000	
模型 6	6	tansig	purelin	traingdx	0.010	1 000	
模型 7					0.015		2
模型 8							0.42
模型 9	3	logsig	purelin	traingdx	0.012	1 000	
模型 10	6	tansig	purelin	traingdx	0.010	1 000	
模型 11					0.015		20
模型 12							0.52

4.3 结果及分析

利用上述训练好的单一预测模型和多重组合预测模型(模型 1~12)对新疆伊犁河雅马渡站进行年径流拟合及预测,各统计指标值见表 2。

表 2 单一预测模型和组合预测模型拟合、预测统计指标值

预测模型	$e_{MRE}/\%$		$e_{maxRE}/\%$		e_{MAE}		$e_{WLNHD}/\%$
	拟合	预测	拟合	预测	拟合	预测	
BP	7.37	7.89	22.03	14.87	28.58	25.25	3.41
Elman	6.15	6.78	16.88	19.11	23.78	30.58	4.87
RBF	6.72	6.83	14.92	24.90	24.66	34.14	0.81
GRNN	7.91	7.53	19.16	20.78	29.31	34.84	2.46
模型 1	5.69	5.13	11.07	10.2	21.52	19.31	5.18
模型 2	4.81	4.79	10.31	11.11	17.57	17.26	0.21
模型 3	5.45	4.94	11.06	13.11	19.80	17.47	4.91
模型 4	5.24	5.97	14.82	12.47	19.96	22.43	6.51
模型 5	5.04	5.23	11.88	10.23	18.75	19.28	1.85
模型 6	4.45	4.46	10.60	9.74	16.28	16.21	0.11
模型 7	4.64	4.65	9.84	10.42	17.19	16.77	0.11
模型 8	5.25	5.43	11.92	10.48	19.56	20.13	1.69
模型 9	4.68	4.69	12.47	10.01	17.12	17.14	0.11
模型 10	4.47	4.49	11.39	9.97	16.33	16.37	0.22
模型 11	4.39	4.50	10.22	9.96	16.08	16.25	1.24
模型 12	5.35	5.82	13.81	11.31	19.81	21.60	4.21
IEA-BP ^[10]	6.62	5.14	22.19	10.66	25.10	17.49	12.59

分析表 2 可以得出以下结论：

a. 模型 1 ~ 12 的拟合和预测的各评价指标值均优于单一预测模型,12 种模型的拟合和预测精度较单一预测模型均有较大提高,表明将多重组合预测模型用于年径流预测是合理可行的,模型具有较好的预测精度和泛化能力,是提高预测精度的有效方法。

b. 比较而言,三次组合预测模型预测精度略优于二次组合预测模型,二次组合预测模型优于一次组合预测模型,表明随着模型组合重数的增加,预测精度呈提高趋势。但随着模型组合重数的增加,原始输入向量有用信息得到进一步挖掘,有用信息量逐渐减少,致使模型的预测精度随着组合重数的增加提高幅度不大。

c. 模型 2、3、6、7、9、10、11 拟合及预测的各评价指标值均优于文献[10]的 IEA-BP 模型,表明多重组合神经网络预测模型具有较高的预测精度和泛化能力。其中,模型 6、10、11 的拟合精度和泛化能力均达到了较理想状态,无论是拟合还是预测,平均相对误差均小于 4.5%;最大相对误差均小于 11.4%;平均绝对误差均小 17,网络拟合度均小于 1.5%,均大幅优于其他组合模型和文献[10]的 IEA-BP 模型。

5 结 语

目前,组合预测方法已成为预测领域中一个重要的研究方向,其线性组合预测模型中各单一预测模型权重的确定,或非线性组合预测模型中组合函数的构造是制约组合预测方法预测精度提高的难点

和关键。本文充分发挥了神经网络模型的高度非线性映射能力,提出了一种基于 BP、Elman、RBF、GRNN 这 4 种神经网络的多重组合神经网络年径流预测模型,以上一组合模型的预测结果作为下一组合模型的输入向量,构建多重组合神经网络年径流预测模型,巧妙地避开了线性组合预测模型预测精度不高、权重较难确定和非线性组合预测模型组合函数难以构造等实际问题,并提出网络拟合度指标,用于评价模型的泛化能力。以新疆伊犁河雅马渡站年径流预测为例进行验证,取得了较为理想的预测效果,为提高水文预测预报的精度提供了新的思路和方法。

参考文献：

[1] 蓝永超,康尔泗,徐中民,等. BP 神经网络在径流长期预测中的应用[J]. 中国沙漠,2001,21(1):97-100. (LAN Yongchao,KANG Ersi,XU Zhongmin,et al. Long-term runoff forecasting with BP neural network model[J]. Journal of Desert Research,2001,21(1):97-100. (in Chinese))

[2] 崔东文. 多隐层 BP 神经网络模型在径流预测中的应用[J]. 水文,2013,33(1):68-73. (CUI Dongwen. Application of hidden multilayer BP neural network model in runoff prediction[J]. Journal of China Hydrology,2013,33(1):68-73. (in Chinese))

[3] 崔东文. 改进 Elman 神经网络在径流预测中的应用[J]. 水利水运工程学报,2013(2):71-77. (CUI Dongwen. Application of improved Elman neural network in runoff prediction[J]. Hydro-Science and Engineering,2013(2):71-77. (in Chinese))

[4] 刘荻,周振民. RBF 神经网络在径流预报中的应用[J]. 华北水利水电学院学报,2007,28(2):12-14. (LIU Di,ZHOU Zhenmin. Application of RBF neural network in runoff prediction[J]. Journal of North China Institute of Water Conservancy and Hydroelectric Power,2007,28(2):12-14. (in Chinese))

[5] 陈仁升,康尔泗,张济世. 应用 GRNN 神经网络模型计算西北干旱区内陆河流域出山径流[J]. 水科学进展,2002,13(1):333-338. (CHEN Rensheng,KANG Ersi,ZHANG Jishi. Application of the generalized regression neural network to simulating runoff from the mountainous watersheds of inland river basins in the arid area of northwest China[J]. Advances in Water Science,2002,13(1):333-338. (in Chinese))

[6] 郭彦,金菊良,梁忠民. 基于集对分析的区域需水量组合预测模型[J]. 水利水电科技进展,2009,29(5):42-45. (GUO Yan,JIN Juliang,LIANG Zhongmin. Combined prediction model for regional water demand based on set pair analysis[J]. Advances in Science and Technology of

- Water Resources, 2009, 29(5): 42-45. (in Chinese))
- [7] 谭满春,冯萃斌,徐建闽. 基于 ARIMA 与人工神经网络组合模型的交通流预测[J]. 中国公路学报, 2007, 20(4): 118-121. (TAN Manchun, FENG Luobin, XU Jianmin. Traffic flow prediction based on hybrid ARIMA and ANN model [J]. China Journal of Highway and Transport, 2007, 20(4): 118-121. (in Chinese))
- [8] BATES J M, GRANGER C W J. The combination of forecasts [J]. Operational Research Quarterly, 1969, 20(1): 451-468.
- [9] 张大斌,张景广,彭森. 基因表达式编程在组合预测建模中的应用[J]. 系统工程理论与实践, 2012, 32(3): 568-573. (ZHANG Dabin, ZHANG Jingguang, PENG Sen. Application of gene expression programming on combination forecasting modeling [J]. Systems Engineering-Theory & Practice, 2012, 32(3): 568-573. (in Chinese))
- [10] 郭淳,李祚泳,党媛. 基于免疫进化算法的 BP 网络模型在径流预测中的应用[J]. 水资源保护, 2009, 25(5): 1-4. (GUO Chun, LI Zuoyong, DANG Yuan. Runoff prediction application of BP neural network model based on immune evolutionary algorithm [J]. Water Resources Protection, 2009, 25(5): 1-4. (in Chinese))
- [11] 张德丰. MATLAB 神经网络应用设计[M]. 北京:机械工业出版社, 2009.
- [12] 丛爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 合肥:中国科学技术大学出版社, 2009.
- [13] MATLAB 中文论坛. MATLAB 神经网络 30 个案例分析[M]. 北京:北京航空航天大学出版社, 2010.
- [14] 崔东文. 几种神经网络模型在湖库富营养化程度评价中的应用[J]. 水资源保护, 2012, 28(6): 12-18. (CUI Dongwen. Applications of several neural network models to eutrophication evaluation of lakes and reservoirs [J]. Water Resources Protection, 2012, 28(6): 12-18. (in Chinese))
- [15] 史峰,王辉,郁磊,等. MATLAB 智能算法 30 个案例分析[M]. 北京:北京航空航天大学出版社, 2011.
- [16] 崔东文. 基于极限学习机的长江流域水资源开发利用综合评价[J]. 水利水电科技进展, 2013, 33(2): 14-19. (CUI Dongwen. Comprehensive evaluation of water resources development and utilization in Yangtze River Basin based on extreme learning machine [J]. Advances in Science and Technology of Water Resources, 2013, 33(2): 14-19. (in Chinese))

(收稿日期:2013-04-16 编辑:熊水斌)

(上接第 49 页)

参考文献:

- [1] SL 76—2009 小水电水能设计规程[S].
- [2] 何孝侠. 中国水利百科全书:水利规划分册[M]. 北京:中国水利水电出版社, 2004.
- [3] 全国勘察设计注册工程师水利水电工程专业管理委员会,中国水利水电勘测设计协会. 水利水电工程专业案例:工程规划与工程移民篇[M]. 郑州:黄河水利电力出版社, 2009.
- [4] 徐淑琴,刘小燕. 水利计算[M]. 北京:中国水利电力出版社, 2011.
- [5] 雒文生,宋星原. 工程水文及水利计算[M]. 北京:中国水利电力出版社, 2012.
- [6] 王文川. 水利水电规划[M]. 北京:中国水利电力出版社, 2013.
- [7] 钟琦,张勇传. 水电站径流调节和水库调度图的计算方法的研究[J]. 人民长江, 1987(12): 27-32. (ZHONG Qi, ZHANG Rongchuan. Calculation method research for the runoff adjustment and reservoir operation chart of hydropower station [J]. Yangtze River, 1987(12): 27-32. (in Chinese))
- [8] 王平. 灌溉水库兴利调度图长系列计算方法的改进[J]. 中国农村水利水电, 1997(7): 24-27. (WANG Ping. Improvement to long series calculation method of irrigation reservoir operation for promoting benefits [J]. China Rural Water and Hydropower, 1997(7): 24-27. (in Chinese))
- [9] 王平. 水库发电调度图常规计算方法的问题和改进建议[J]. 水利水电工程设计, 2007(3): 26-29. (WANG Ping. The problems and the improving suggestion for conventional calculation method of power generation dispatching drawings of reservoir [J]. Design of Water Resources & Hydroelectric Engineering, 2007(3): 26-29. (in Chinese))
- [10] 王平. 水库灌溉或供水调度图常规编制方法的改进[J]. 水电站设计, 2008, 24(4): 95-97. (WANG Ping. Reservoir irrigation or water supply chart conventional preparation improvement [J]. Design of Hydroelectric Power Station, 2008, 24(4): 95-97. (in Chinese))

(收稿日期:2013-04-09 编辑:熊水斌)

