

基于 WNN-SVM 的水文时间序列预测方法研究

余 洋, 万定生

(河海大学 计算机与信息学院, 江苏 南京 211100)

摘 要: 在面对具有随机性、突变性的复杂时间序列数据(如流量等水文时间序列数据)时,传统单一的模型的预测精度不尽人意,对单一模型的优化不能完全克服其局限性。因此,文中提出一种基于 WNN-SVM 组合的水文时间序列预测模型。首先对水文时间序列数据作均值归一化处理,然后对预处理后的水文时间序列进行小波分解,将分解后的子序列通过相空间重构的方法使其从低维时间序列向高维转换;根据其分解后的特点,对尺度变换序列采用支持向量机(SVM)进行建模预测,小波变换序列采用小波神经网络(WNN)进行建模预测,再将两者的预测结果进行小波重构,得到原始序列预测值。随后采用屯溪流域 1980 年至 2007 年 43 996 个小时流量数据进行实验验证,结果表明该模型的预测精度高于单一模型,证明了该模型的有效性。

关键词: 均值归一化;组合模型;时间序列预测;小波神经网络;支持向量机

中图分类号: TP39

文献标识码: A

文章编号: 1673-629X(2019)09-0013-05

doi: 10.3969/j.issn.1673-629X.2019.09.003

Research on Hydrological Time Series Prediction Method Based on WNN-SVM

YU Yang, WAN Ding-sheng

(School of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract: When faced with hydrological time series data with randomness and inconsistency, such as flow, etc., the prediction accuracy of traditional single model is not satisfactory, and the optimization of single model cannot completely overcome its limitations. Therefore, we present a hydrological time series prediction model based on WNN-SVM combination. First of all, the hydrological time series data are normalized by means, and then the wavelet decomposition is performed on the pretreatment hydrological time series and the low-dimensional time series is transformed to high-dimensional by phase space reconstruction. According to the characteristics of decomposition, the scaling sequence is modeled and predicted by support vector machine (SVM), and the wavelet transform sequence is modeled and predicted by wavelet neural network (WNN), and the predicted value of the original sequence is obtained after the wavelet reconstruction of the predicted result. The experiment of 43 996 data of hourly flow in Tunxi basin from 1980 to 2007 indicates that the prediction accuracy of the model is higher than that of a single model, which proves the validity of the model.

Key words: mean normalization; combined model; time series prediction; WNN; SVM

0 引言

在大数据时代,数据已经成为重要的资源。面对海量的数据,对其进行有效的处理和分析变得非常重要^[1]。而时间序列数据挖掘(time series data mining, TSDM)已成为数据挖掘领域的重要分支,主要通过不同的数据挖掘技术^[2]分析时间序列的局部特征从而挖掘出时间序列内在的规律,这将有助于实现时间序列的异常检测、预测等目的^[3]。

而具有随机性、突变性的复杂时间序列如水文时

间序列的预测已成为国际热点。此类时间序列预测方法主要有三类:一是定性预测方法,根据预测对象的特性、状态,采用类推对比、专家打分等方法来预测序列;二是因果关系预测法,对目标时间序列和影响因子序列基于因果关系进行统计预测;三是时间序列预测法,根据时间序列自身变化规律进行预测,常见的有自回归模型(AR)、自回归滑动模型(ARMA)^[4]、人工神经网络(ANN)、相空间模型、灰色理论和支持向量机(SVM)等。Kottegoda^[5]提出在线性随机模型中

收稿日期:2018-10-15

修回日期:2019-02-18

网络出版时间:2019-04-24

基金项目:国家科技支撑计划课题(2018YFC0407900);公益性行业科研专项(201501022)

作者简介:余 洋(1995-),男,硕士研究生,研究方向为数据挖掘;万定生,教授,CCF 会员(08015S),研究方向为信息处理与信息系统。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190424.1047.032.html>

ARMA 对 Vanern 河流量进行了预测,但其在非平稳序列中的表现不够精确。Wei Shouke 等^[6]将小波神经网络应用于江河月径流量拟合和预测上,得出 WNN 相对于后者能显著增加预测精度的结论,但其对多变化的流量精度仍有缺陷。

由于这些模型的适用范围不同,而此类复杂时间序列既包含线性成分又含有非线性成分,对单一模型的优化不能完全克服其局限性^[7]。针对此问题,文中提出了基于 WNN-SVM 组合的水文时间序列预测模型。

1 支持向量机和小波分析

1.1 支持向量机

支持向量机 (support vector machine, SVM) 是一种基于统计学习理论构建的典型神经网络,它的基本思想是使用一个线性模型解决非线性的问题,将一些非线性的输入空间映射到高维的特征空间。利用最佳的超平面构造边界,使决策类之间实现最大分隔,在新的空间构建一个线性模型表示原来空间的非线性决策边界,最后转化为一个凸二次规划问题进行求解。在线性可分的情况下,二元决策类分离表示为:

$$\mathbf{y} = \omega_0 + \omega_1 x_1 + \omega_2 x_2 \quad (1)$$

其中, \mathbf{y} 为输出向量, x_i 为特征值, ω_i 为平面的权重值。

支持向量机将输入转换成高维特征空间,以构造一个线性模型来实现非线性类边界。在非线性可分的情况下,高维的最大超平面边界方程可以表示为:

$$\mathbf{y} = b + \sum_{i=0} a_i y_i K(x(i), x) \quad (2)$$

其中, $K(x(i), x)$ 为核函数。输入空间中产生的不同类型的支持向量机有不同的核函数,常见的核函数有多种,例如多项式核和高斯径向基核等。

1.2 小波分析

1.2.1 小波变换

小波 (wavelet) 是一种特殊的长度有限,平均值为 0 的波形。它的“小”体现在时域具有紧支性或近似紧支性,它的“波”体现为正负交替的波动性,即直流分量为零。小波变换 (WT) 利用信号在一簇基函数构成的函数空间的投影表征该信号。在低频部分具有较低的时间分辨率和较高的频率分辨率,在高频部分具有较高的时间分辨率和较低的频率分辨率,能实现信号的多分辨率分析,得到信号的等价描述^[8]。小波变换可分为连续小波变换 (CWT) 和离散小波变换 (DWT)^[9]。

定义 1: 设 $\Psi(t) \in L^2(R)$, $L^2(R)$ 表示平方可积的实数空间 (能量有限的信号空间), 其傅里叶变换为

$\hat{\Psi}(w)$ 。若 $\hat{\Psi}(w)$ 满足容许性条件:

$$\mathbb{C}_{\Psi} = \int_0^{+\infty} \frac{|\hat{\Psi}(w)|^2}{w} dw < +\infty \quad (3)$$

则称 $\hat{\Psi}(w)$ 为母小波, 对 $\hat{\Psi}(w)$ 进行伸缩、平移, 得到小波函数族:

$$\{\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \hat{\Psi}\left(\frac{t-b}{a}\right), a > 0, b \in R\} \quad (4)$$

其中, a 是尺度因子; b 是位移因子。

定义 2: 任意 $L^2(R)$ 空间中的一维信号 $f(t)$ 作离散小波变换的表达式为:

$$W_f(j, k) \leq f(t) \\ \Psi_{j,k}(t) \geq \int_{-\infty}^{\infty} f(t) \overline{\Psi_{j,k}(t)} dt \quad (5)$$

1.2.2 紧致型小波神经网络

小波神经网络 (WNN) 从结构上分为两大类^[10], 一类是小波变换与神经网络松散耦合形成的松散型小波神经网络。整个系统由小波变换和神经网络以串联形式构成。小波分解作为神经网络的前置处理手段^[11]。另一类是小波变换与神经网络紧密融合形成的紧致型小波神经网络^[12]。文中采用的是第二种, 其基本思想是以小波分析中的小波基函数代替 BP 神经网络中的 Sigmoid 族激活函数, 并以小波函数的尺度因子置换输入到隐含层的权值, 以位移因子置换隐含层阈值。紧致型 WNN 由输入层、输出层和隐含层组成, 信号前向传播, 误差反向传播。其表达式为:

$$f(x) = \sum_{i=1}^N w_i \psi[D_i R_i(x - t_i)] + \bar{f} \quad (6)$$

其中, 输入向量 $\mathbf{x} \in R^n$; t_i 为平移因子构成的矢量; D_i 为尺度参数构成的对角阵; R 为旋转阵; \bar{f} 为 f 的均值估计; w_i 为权值。

2 WNN-SVM 组合预测模型

具有随机性、突变性的复杂时间序列如水文时间序列的确定性成分、周期性成分和非确定性因素, 使得单一模型及其改进对预测精度的提升空间十分有限。传统时间序列模型对序列平稳性要求较高且预测效果不尽人意, 对此提出一种基于小波变换的组合预测模型。与传统的以权系数最优为目标的组合预测方法不同, 通过引入小波变换, 使得序列的高低频成分分离, 针对高低频序列特征, 分别采用不同的预测算法。以水文时间序列为例, 其建模流程如图 1 所示。

2.1 均值归一化

复杂时间序列普遍波动频繁且差值较大, 因此首先有必要对时间序列进行归一化处理以减小模型误差。处理数据的归一化方法通常为线性函数归一化

(Min-Max scaling)和0均值标准化(Z-score standardization),但由于水文时间序列数据在个别处峰值极高,采用线性函数归一化方法会导致大多数数据处于一个极低的值,损害模型的训练,甚至导致梯度消失或爆炸;而0均值标准归一化方法则要求原始数据的分布近似为高斯分布,否则归一化的效果会变得更糟糕。

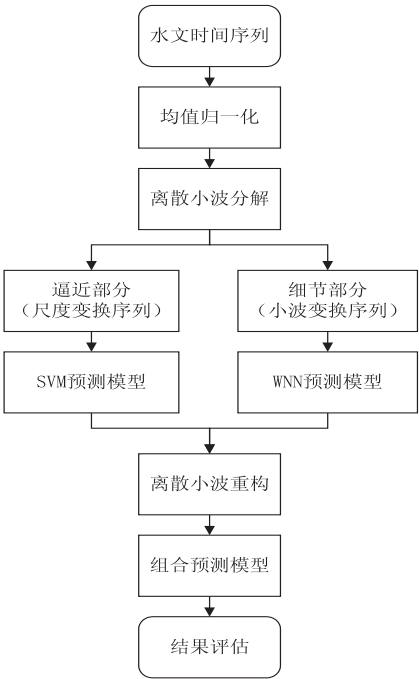


图1 组合预测模型建模流程

因此,提出一种基于均值的归一化方法,经验证其有效规避了梯度消失和梯度爆炸,具有较好的训练效果。归一化公式如下:

$$n(t)=\frac{x(t)}{x(t)}$$

(7)

其中, $x(t)$ 为样本数据; $n(t)$ 为输入样本; $\overline{x(t)}$ 为样本数据的均值。

2.2 组合预测模型建模

算法思想:在对时间序列进行预处理后,通过相空间重构的方法将低维时间序列向高维转换。随后对时间序列信号进行小波分解,有助于发挥小波的多分辨率分析优势,同时也是组合模型并行预测的必要步骤。结合子序列的波形特征和波动频率,针对波动性很高的高频小波变换序列,用 WNN 模型对其进行拟合预测,能有效提高预测的准确性。针对包含趋势项和周期项的低频尺度变换序列,利用 SVM 模型预测其未来值,同时能降低计算量。最后将子模型输出的预测值序列进行小波重构,得到预测值。

算法步骤:

- (1)对原始复杂时间序列进行去噪、均值归一化等预处理。
- (2)对清洗后的时间序列数据进行小波分解,得

到不同频段的小波。小波分解函数通常按照自相似原则和小波支撑集长度进行选择。大部分应用选择支撑集长度为 5~9 的小波,因为支集太长会产生边界问题,支集太短消失矩太低,不利于信号能量的集中。朱跃龙等^[13]采用 db10 小波作为流量预测时的分解小波,因 Daubechies(db)小波适合提取时间序列跃变特征,而实验流域流量序列细节波动与 db 小波非常相似,因此选择 db10 小波作为原始时间序列分解的小波基,分解层级为 6。分解算法见算法 1,重构算法见算法 2。

算法 1:Signal_Wavedec(db,sourceData)。

输入:db:过滤器类型;sourceData:归一化时间序列。

输出:cA[n]:分解后的近似部分序列-低频部分;cD[n]:

分解后的细节部分序列-高频部分。

```
for(n=0 to decLen) do
    cA[n]=0,cB[n]=0;//遍历小波变换结果序列长度,初始化分解序列
    for(k=0 to filterLen) do
        p=2 * n-k;
        if((p<0)&&(p>=-filterLen+1))
            tmp=sourceData[-p-1];
        else if((p>dataLen-1)&&(p<=-dataLen+filterLen-2))
            tmp=sourceData[2 * dataLen-p-1];
        else if((p>=0)&&(p<dataLen-1))
            tmp=sourceData[p];
        else tmp=0;//信号边沿对称延拓
        cA[n]+=m_db.lowFilterDec[k] * tmp;
        cD[n]+=m_db.highFilterDec[k] * tmp;
    end for
end for
```

算法 2:Signal_Waverec(cA,cD,db)。

输入:db:过滤器类型;cA,cD:分解后的低频高频序列。

输出:recData:重构后输出的数据。

```
for(n=0 to recLen) do
    recData[n]=0//遍历小波变换结果序列长度,初始化重组序列
    for(k=0 to cAlength) do
        p=n-2 * k+filterLen-1;
        if((p>=0)&&(p<filterLen))
            recData[n]+=m_db.lowFilterRec[p] * cA[k]+m_db.highFilterRec[p] * cD[k];//信号重构
        end for
    end for
```

(3)在时间序列数据分解后对各子序列进行相空间重构^[14],由此可以将低维时间序列向高维转换,尽可能挖掘到序列间的相关信息^[15]。设时间序列为 $x(t)$, $t=1,2,\cdots,N$,嵌入维为 m ,时间延迟为 τ ,则重构相空间为:

$$y(t)=\{x(t),x(t+\tau),\cdots,x[t+(m-1)\tau]\},t=$$

1,2,⋯,M (8)

其中, $M = N - (m - 1)$; τ 为相空间中的相点数。实际应用中,若时间延迟 τ 太小,则相空间矢量在数值上过于接近,无法提供两个独立的坐标分量;若时间延迟 τ 太大,则两坐标分量的相关关系又趋于独立,根据自相关函数法分析,选定 $\tau = 1, m = 25$ 。

(4) 针对分解后的低频尺度变换序列构建支持向量机模型进行预测。普遍认为,基于高斯核函数的支持向量机具有较好的模拟效果^[12],实验选取高斯核函数(RBF)作为支持向量机模型的核函数,定义为:

$$K(x,xc) = \exp(-g \|x - xc\|^2) \tag{9}$$

模型中惩罚参数 C 设置为 10.0,不敏感损失函数 ε 设置为 0.1。将相空间重构后的二维数据输入至模型中进行训练,并对测试数据进行预测。

(5) 对小波分解后的高频小波变换序列应用小波神经网络进行预测。小波神经网络输入维数和小波函数是影响网络性能的重要因素^[16]。通过对小波系数进行相关性评价,发现嵌入维 $m = 25$,时间延迟 $\tau = 1$ 时,具有较高的相关性。经测试选取小波函数为 Morlet 小波,其基函数和导函数分别为:

$$\Psi(x) = \exp(-x^2/2) * \cos 1.75x \tag{10}$$

$$\Psi'(x) = -(x\cos 1.75x + 1.75\sin 1.75x) * \exp(-x^2/2) \tag{11}$$

随后设计小波神经网络的结构。根据相关性分析设置输入层节点数为 128,输入维度为 25,隐含层数为 1,输出层节点数为 1,优化器为 Adam 函数,损失函数为 MSE,验证数据集为训练数据集的 10%,最大迭代次数为 100。进行网络迭代训练得到 WNN 模型。具体过程如算法 3。

算法 3:WNN_premodel(train,test)。

输入:train:小波变换序列训练集;test:小波变换序列测试集。

输出:测试集预测值 test_simu。

初始化参数:输入节点数 m ,输出节点数 N ,隐含节点数 n ,训练样本数 S ,迭代步长 etha,最大迭代步数 StepMax;

```
G<-normrnd(0,1,n,m+1), W<-normrnd(0,1,N,n+1);
//权值系数矩阵随机初始化
while(flag==0) do //未达到训练要求或迭代步数
WaveNetIn<- G * [X;ones(1,S)];
V<-psi(x); //小波函数基函数
Y<-tansig(W * [V;ones(1,S)]); //计算 WNN 输出
if(Step>=StepMax) flag<-1; //退出
else
dtansigWV<-4 * exp(2W[V;ones(1,S)])/(exp(2W[V;ones(1,S)]+1)^2);
W<- W - etha * (-(dtansigWV * EA) * [V;ones(1,S)]'); //输出层系数矩阵更新
DL<-psi'(x); //小波函数导函数
```

```
dEdG<- - DL * (W(:,1:n))' * Gamma) * [X;ones(1,S)]';
G<- G -etha * dEdG; //输入-隐含层权值系数矩阵更新
end if
end while
```

3 实验及结果分析

3.1 实验结果

实验平台为 PyCharm2017.3.4,采用的数据集为屯河流域从 1981 年 5 月 1 日至 2007 年 3 月 25 日共计 43 996 个小时流量数据。数据集被分为两部分,一部分是训练数据共 39 998 条,用于训练调整网络权值和系数;另一部分是测试数据共 3 998 条,用于测试网络性能。

对归一化流量序列进行离散小波变换,采用 Mallat 算法,小波的低通滤波长度不能太长,同时能提取序列跃变特征,采用 db10 小波作为小波分解函数,分解层级为 6 层,分解后的低频尺度变换序列为 a6,高频小波变换序列为 d1,d2,d3,d4,d5,d6。因为随着分解层级的提高,每次循环抽样都会减少低频近似分量的数据量,因此对分解后序列应进行单支重构,便于预测。测试数据分解序列如图 2 所示,训练数据与其相似。

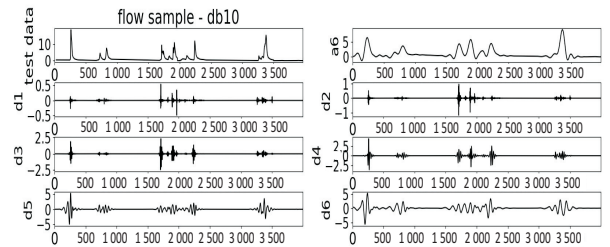


图 2 测试数据流量时间序列小波分解

利用 SVM 模型对尺度变换序列 a6 进行仿真预测,得到 SVM 模型对测试数据 a6 序列的拟合及预测结果如图 3 所示。

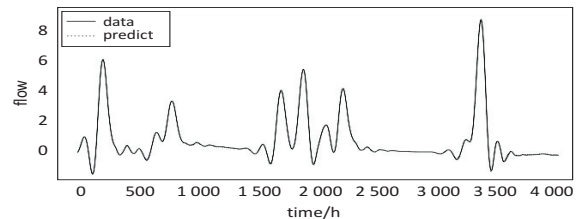


图 3 SVM 模型预测 a6 序列效果

依次导入训练数据的 d1,d2,d3,d4,d5,d6 序列,对 WNN 进行网络迭代训练,在迭代次数不到 50 次时损失函数值已趋于稳定,说明 WNN 的收敛速度良好;利用训练好的 WNN 对测试数据集的 3 998 个样本进行测试,其中对 d6 小波变换序列的预测效果如图 4 所示。

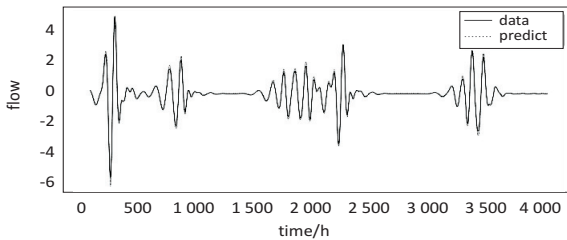


图4 WNN 模型预测 d6 序列效果

将 SVM 子模块和 WNN 子模块得出的预测结果进行重构,即可得到组合模型的预测结果。将尺度变换序列 a6 和小波变换序列 d1,d2,d3,d4,d5,d6 的预测值进行重构得到组合模型的输出,如图 5 所示。

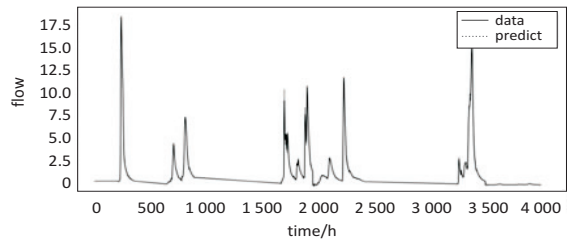


图5 WNN-SVM 组合模型预测效果

同时,应用单一的 WNN 模型及 SVM 模型对流量序列进行预测,得到的预测序列如图 6、图 7 所示。

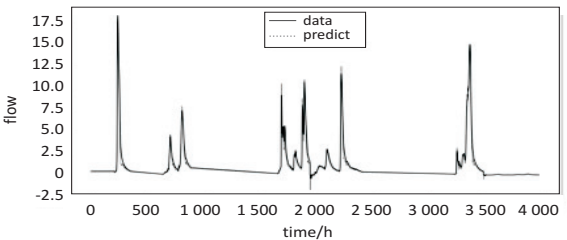


图6 WNN 模型预测效果

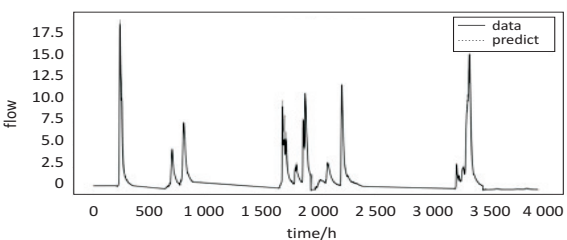


图7 SVM 模型预测效果

将基于小波变换的并行组合模型与单一的 WNN 模型以及 SVM 模型的预测值进行对比,同时选择均方误差(MSE)、均方根误差(RMSE)和决定系数(R^2)对上述模型的预测效果进行参数检验,得到的参数统计结果如表 1 所示。

表1 模型性能对比

模型	MSE	RMSE	R^2
并行组合模型	0.025 5	0.159 8	0.992 5
WNN	0.085 3	0.292 1	0.962 6
SVM	0.162 7	0.403 4	0.959 1

3.2 实验分析

通过实验验证,SVM 模型对尺度变换子序列的拟合程度与 WNN 模型对小波变换子序列的拟合程度都处于较高的水平。同时,文中提出的 WNN-SVM 组合模型对测试序列的拟合程度非常贴近原始序列真实值。而 SVM 模型和 WNN 模型在整体上虽然保持了一定的准确性,但在波动转折点具有一定的迟滞性,导致较大误差。在参数检验上,MSE、RMSE 越贴近 0,则预测效果越好, R^2 则越贴近于 1,表示拟合程度越高。可以看出,并行组合模型的 MSE 和 RMSE 明显低于 WNN 和 SVM 模型的 MSE 和 RMSE,而 R^2 值则比 WNN 和 SVM 模型更贴近于 1,验证了其预测能力,相对单一模型具有更好的性能表现。

4 结束语

针对单一模型对具有随机性、突变性复杂时间序列预测精度的不足,提出一种基于小波变换的组合预测模型,将 SVM 模型与 WNN 模型相结合的组合模型应用于测试屯溪流域的小时流量数据。为验证方法的有效性,将其与传统的单一 WNN 模型和 SVM 模型进行对比,经实验验证组合模型的预测精度高于单一模型,证明了该并行组合模型的有效性。但对于小波基函数及分解层次的选择,仍依赖于实验对比分析。同时,对复杂时间序列预测期的延长还有待后续研究。

参考文献:

[1] 唐黎哲,冯大为,李东升,等. 以 LDA 为例的大规模分布式机器学习系统分析[J]. 计算机应用,2017,37(3):628-634.

[2] HAN Jiawei,KAMBER M. Data mining:concepts and techniques[M]. 2nd ed. [s.l.]:Morgan Kaufmann Publishers, 2004.

[3] 周金玉,万定生,肖 艳. 基于嵌入式索引的水文时间序列预测模型[J]. 微电子学与计算机,2018,35(1):118-123.

[4] BOX G E P,JENKINS G M,REINSEL G C. Time series analysis:forecasting and control-4t[M]. [s.l.]:John Wiley & Sons,Inc.,2008.

[5] KOTTEGODA N T. Stochastic water resources technology [M]. London:Macmillan Press,1980.

[6] WEI Shouke,YANG Hong,SONG Jinxi,et al. A waterlet-neural network hybrid modeling approach for estimating and predicting river monthly flows [J]. Hydrological Sciences Journal,2013,58(2):374-389.

[7] 成 云,成孝刚,谈苗苗,等. 基于 ARIMA 和小波神经网络组合模型的交通流预测[J]. 计算机技术与发展,2017,27(1):169-172.

[8] SEO Y,KIM S,KISI O,et al. Daily water level forecasting