

DOI:10.13476/j.cnki.nsbdqk.2021.0116

姚思源,李国芳,钱睿智,等.南水北调东线工程源头长江潮位预报[J].南水北调与水利科技(中英文),2021,19(6):1136-1146. YAO S Y, LI G F, QIAN R Z, et al. Tidal level forecast of Yangtze River at the source of East Route of South-to-North Water Transfer Project[J]. South-to-North Water Transfers and Water Science & Technology, 2021, 19(6): 1136-1146. (in Chinese)

南水北调东线工程源头长江潮位预报

姚思源¹,李国芳¹,钱睿智²,王永东²,马放¹

(1 河海大学水文水资源学院,南京 210098;2 江苏省水文水资源勘测局扬州分局,江苏 扬州 225002)

摘要:南水北调东线工程通过江都水利枢纽抽引长江水,准确及时的引水口水位预报对提高水利工程调度和水资源配置的科学性至关重要。选择引水口附近的三江营潮位站为代表,由于三江营地处潮区界以内,所以在预报其潮位时需要综合考虑潮汐引力、上游来水量和区间降水量等影响因素。首先采用潮汐调和分析法计算得到天文潮位,再采用二次校正法、多元线性回归法、支持向量机(support vector machine, SVM)法和随机森林法,对三江营的日均潮位和逐时潮位进行预报,并比较4种方法的预报精度。结果表明,4种方法得到的年日均潮位预报结果均能达到乙级以上精度,预报效果较好,其中:二次校正法的预报精度最高且最稳定,日均潮位和逐时潮位预报均能达到甲级精度,适合用于水利工程精细化调度;SVM法的预报精度次之,多元线性回归法和随机森林法相对较差,但后3种方法的预见期长于二次校正法,适合用于对预见期要求较高的潮位预警。

关键词:潮位预报;二次校正法;多元线性回归法;支持向量机法;随机森林法

中图分类号:TV213;P338 文献标志码:A 开放科学(资源服务)标志码(OSID):



南水北调工程实现精细化调度的要点之一在于其引水口水位预报。南水北调东线工程的源头位于江苏省扬州市江都水利枢纽,三江营是其主要引水口。三江营位于长江大通站以下,属于潮区界以内^[1-2],其潮位不仅取决于长江大通以上径流量,也受到潮汐引力的作用^[3],同时还受大通至三江营区间来水的影响。由于影响因素众多,所以潮位预报难度较大,传统的水文预报方法难以满足要求。

长江下游感潮河段的常用潮位预报方法包括潮汐调和分析法^[4-5]、结合回归分析法和机器学习法等。李国芳等^[6]、钱睿智等^[7]、宋正逵等^[8]在预报长江下游感潮河段站点潮位时发现,仅采用潮汐调和分析法的预报精度较低,进一步研究上游径流和长江下游感潮河段的日均增水(潮位站的实测潮位与调和预报的天文潮位的差值)后发现,用上游径

流增量实时校正调和预报的潮位可以大幅提高潮位预报精度。王文才等^[9]在研究长江下游感潮河段水位相关关系时,也同样发现长江下游感潮河段站点潮位与上游径流量呈现出相同的变化趋势。近年来,随着计算机技术的发展,机器学习越来越多地应用于潮位预报。在长江下游的秦淮河流域的潮位预报中,张轩等^[10]运用BP神经网络对秦淮河流域东山站的水位进行预报,提高了预报精度;吴美玲等^[11]用KNN算法和GA算法对BP神经网络进行优化,对比传统潮位预报结果后发现改进后的KG-BP模型的预报精度更高。

综合考虑潮汐引力、上游来水量和区间降水量等影响因素,采用潮汐调和分析的潮位预报结果结合二次校正法、多元线性回归法、支持向量机(support vector machine, SVM)法和随机森林法^[12-14]等4种方法

收稿日期:2021-03-17 修回日期:2021-07-19 网络出版时间:2021-07-30

网络出版地址:https://kns.cnki.net/kcms/detail/13.1430.TV.20210729.1721.006.html

基金项目:国家重点研发计划项目(2018YFC1508200);江苏省水利科技项目(2019035);江苏高校“青蓝工程”;中国电力建设股份有限公司重大科技专项项目(DJ-ZDZX-2016-02-02)

作者简介:姚思源(1997—),女,江苏南京人,主要从事水文学及水资源研究。E-mail:517095924@qq.com

通信作者:李国芳(1971—),女,浙江杭州人,教授,博士生导师,主要从事水文学及水资源研究。E-mail:liguofang@hhu.edu.cn

对三江营的日均潮位进行预报,并比较这4种方法的预报精度。为满足工程的实际需要,使调水过程更加合理,在以上4种方法得到的日均潮位预报的基础上增加预报逐时潮位,并比较4种逐时潮位预报方法的预报精度,选择更适合三江营地区的潮位预报方法。

1 研究区概况

三江营是南水北调东线工程的主要引水口,地处长江干流和淮河入江水道夹江的交汇地,见图1。南水北调东线工程由三江营自长江干流引水,经夹江和芒稻河引江水至江都抽水站,再由江都抽水站抽水后一路向北送水。



图1 研究区水系及测站分布

Fig. 1 Water system and station distribution in the study area

由于三江营位于潮区界以内,其潮位受到海洋潮汐引力和长江上游来水的共同影响,为非正规半日潮型,水动力条件复杂。此外,三江营地区的气候类型为亚热带季风气候:夏季气温较高,降水量大,夏季降水量占到全年的60%以上,雨热同期;冬季温暖干燥,最冷月平均气温高于0℃,降水量较小。由于降水的年内分配不均匀,所以区间来水对三江营地区潮位变化的影响也不容忽视。

2 研究数据及方法

2.1 水文数据

所选用的水文资料包括:三江营站2006—2020年实测逐时潮位;上游来水控制站大通1964—2020年实测日均流量;长江大通站—三江营站区间童埠、荻港、汤沟、三山、芜湖、和县、马鞍山、江浦、南京和镇江等10个代表站2006—2020年的实测逐日降水量。

由于洪水从大通站到三江营站的传播时间约为2d^[7],因此在计算预报日的潮位时取预报日前2日的大通流量作为预报日当天的代表流量。区间来水考虑产汇流及洪水传播时间,同样取预报日前2日沿程雨量站实测降水量作为预报日当天的代表区间来水量。

2.2 研究方法

整体思路为采用潮汐调和分析方法揭示潮汐引力驱动下三江营站的水位涨落变化规律,以潮汐调和分析预报的潮位(文中称为“天文潮位”)为基础,

采用二次校正法、多元线性回归法、SVM法和随机森林法等4种方法对三江营站的潮位进行预报。具体方法如下。

(1)潮汐调和分析法^[15]是将实测任意周期潮位转化为富氏积分的数学方法,从而剥离各分潮并计算其调和常数,在预报天文潮位时将各分潮潮高进行叠加,计算公式见式(1)。首先对三江营站2006—2019年实测逐时潮位进行潮汐调和分析计算,分别对2006—2019年每年的潮位进行调和分析计算,得到每年各分潮的调和常数,再取各分潮的平均调和常数计算各分潮潮高,将各分潮潮高叠加预报每年的逐时天文潮位,由逐时天文潮位统计得到日均天文潮位。天文潮位反映了排除径流流量和区间来水等因素影响下由潮汐引力作用引起的潮位变化规律。

$$h_i = A_0 + \sum_{i=1}^p f_i H_i \cos[\sigma_i t + (V_0 + u)_i - g_i] \quad (1)$$

式中: A_0 表示平均海平面高度,m; p 表示总分潮数; f_i 表示分潮交点因子; H_i 表示分潮振幅,m; σ_i 表示角速度,rad/s; t 表示时间,s; $(V_0 + u)_i$ 表示天文初相角,rad; g_i 表示分潮迟角,rad。其中 H_i 和 g_i 为要求的调和常数。

(2)二次校正法分别考虑除潮汐以外的其他影响因素,对调和分析预报的天文潮位进行两次校正。第一次校正考虑上游相应来水量对潮位的影响,建立大通流量增量(预报日前2日的大通流量与2006—2019年同日大通平均流量之差)与日均增水(日均实测潮位与日均天文潮位之差)的相关关系,由于长江年内各月的水文特性差异较大,分月建立相关关系。根据预报日前2日的大通流量增量对天文潮位进行第一次校正。第二次校正假设连续3日的区间来水较稳定,建立第一次校正后残余误差的自回归模型,计算公式见式(2),将预报日前1日和前2日的日均潮位误差作为影响因子,对潮位进行第二次校正。将两次校正后的潮位作为最终预报的潮位。由于二次校正法包含一个二阶自回归模型,因此此种方法的日均潮位预报的预见期为1d。

$$\Delta Z_t = a_0 \Delta Z_{(t-1)} + a_1 \Delta Z_{(t-2)} \quad (2)$$

式中: ΔZ_t 表示预报日当天的残差,m; $\Delta Z_{(t-1)}$ 表示预报日前1日的残差,m; $\Delta Z_{(t-2)}$ 表示预报日前2日的残差,m; a_0 、 a_1 为自回归模型参数。

(3)多元线性回归法^[16]是一种统计模型,研究单因变量与多自变量的线性关系,并用最小二乘法估计模型参数,模型目标公式见式(3)。应用多元线性回归法时,把潮汐调和分析预报的天文潮位作为影响潮位的一个因素,与其他影响因素一起参与预

报,将预报日的天文潮位、预报日前 2 日的大通流量增量和预报日前 2 日的大通—三江营区间降水量作为自变量,将预报日实测潮位作为因变量,分月建立多元线性相关关系,再根据相关关系对潮位进行预报。由模型原理可知,多元线性回归法的日均潮位预报的预见期为 2 d。

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 \quad (3)$$

式中: y 表示因变量; x_1 、 x_2 、 x_3 表示自变量(x_1 为天文潮位, m ; x_2 为流量增量, m^3/s ; x_3 为区间降水量, mm); a_0 、 a_1 、 a_2 、 a_3 表示回归系数。

(4)支持向量机(SVM)法^[17-18]是一种监督式学习模型,常被运用于洪水预报、水位预报等水文预报中^[19-20]。SVM 方法的基本理念是:基于训练集寻找超平面,将样本进行划分;通过非线性映射将属于低维特征空间的训练集样本映射到高维特征空间中,使无法在低维特征空间中解决的线性估计转至高维特征空间,详见参考文献^[17-18]。将预报日的天文潮位、预报日前 2 日的大通流量增量和预报日前 2 日的大通—三江营区间降水量作为 SVM 模型的输入,将预报日实测潮位作为模型的输出。分月对数据进行训练,输出预报潮位。由模型原理可知,SVM 法的日均潮位预报的预见期为 2 d。

(5)随机森林法目前在径流预报中使用较多^[21-23],而在潮位短期预报中的应用还很少。它是一种基于袋装法改进的随机树模型,模型的构建首先从训练集中随机有放回地抽取样本,再以此构建回归树。随机森林的预测结果取各回归树的预测均值,模型目标公式见式(4),详见参考文献^[12-13]。本文尝试将预报日的天文潮位、预报日前 2 日的大通流量增量和预报日前 2 日的大通—三江营区间降水量作为随机森林模型的输入,将预报日实测潮位作为模型的输出。分月对数据进行训练,通过决策树对输入因子进行分类和非线性拟合,输出预报潮位。由模型原理可知,随机森林法的日均潮位预报的预见期为 2 d。

$$H(x) = \frac{1}{T} \sum_{i=1}^T h_i(x) \quad (4)$$

式中: $H(x)$ 表示预测结果, m ; T 表示决策树的个数; $h_i(x)$ 表示各决策树的预测结果, m 。

逐时潮位预报建立在日均潮位预报的基础上,因为缺乏逐时流量以及逐时降水资料,因此以上的 4 种方法都不能直接用来预报逐时潮位。本文在预报逐时潮位时,考虑到上游来水量和区间入流量比较稳定,日内变幅不大,故假设同一日 24 h 内的增水是相同的,将以上 4 种方法预报的日均潮位与日

均天文潮位作差,再将此差值加在预报的逐时天文潮位上,得到由 4 种方法间接预报的逐时潮位。因此,各种方法的逐时潮位预报预见期与其对应的日均潮位预报预见期相同。

将 2006—2019 年作为各种方法模型参数的率定期,并分别采用所率定的模型对 2020 年的日均潮位和逐时潮位进行预报,将预报潮位与 2020 年实测潮位进行对比,以验证其预报精度。

2.3 精度评定

根据《水文情报预报规范》^[24],潮位预报误差在 ± 30 cm 以内即为合格,预报项目的精度按合格率分为 3 个等级,合格率 $QR \geq 85.0\%$ 为甲级, $85.0\% > QR \geq 70.0\%$ 为乙级, $70.0\% > QR \geq 60.0\%$ 为丙级。

对于某年的预报精度评定均以按年单独统计的合格率作为标准,合格率的计算方法为误差在允许误差范围内的样本数除以总样本数。

在计算预报合格率和划分预报精度等级的基础上,统计各年预报潮位的年平均绝对误差 E_a 、相关系数 R 、均方根误差 E_{RMS} 来对比各方法预报潮位的误差和与实测潮位的拟合程度。

3 研究结果

3.1 代表年选择

为使率定的模型能适用于不同的洪水条件,模型在参数率定时应包含丰、平、枯水年。选取大通站 1964—2019 年的年平均流量进行频率分析,选择 P-III 型概率密度函数,采用适线法估计概率密度函数的参数。根据《水文基本术语和符号标准》(GB/T 50095—2014)^[25],确定丰、平、枯水年的设计频率所对应的设计年平均流量,划分丰、平、枯水年的年平均流量范围,见表 1。2006—2019 年归属丰、平、枯水年的情况列于表 2。由表 1 和表 2 可知,用来率定模型参数的年份涵盖了丰、平、枯水年。选择 2012、2014、2018 年作为模型参数率定期丰、平、枯水年的代表年,用来比较不同方法在参数率定期的精度。

表 1 丰、平、枯水年划分依据

Tab. 1 Division basis of high flow year, median water year and low flow year

| 年份划分 | 概率 p | 年均流量 $Q/(m^3 \cdot s^{-1})$ |
|------|--------------------------|-----------------------------|
| 特丰水年 | $p \leq 12.5\%$ | $Q \geq 32\ 586$ |
| 偏丰水年 | $12.5\% < p \leq 37.5\%$ | $29\ 338 \leq Q < 32\ 586$ |
| 平水年 | $37.5\% < p \leq 62.5\%$ | $27\ 000 \leq Q < 29\ 338$ |
| 偏枯水年 | $62.5\% < p \leq 87.5\%$ | $24\ 141 \leq Q < 27\ 000$ |
| 特枯水年 | $p > 87.5\%$ | $Q < 24\ 141$ |

表2 2006—2019年丰、平、枯水年划分

Tab. 2 Division of high flow year, median water year and low flow year from 2006 to 2019

| 年份 | 年均流量 $Q/(\text{m}^3 \cdot \text{s}^{-1})$ | 年份划分 |
|------|-------------------------------------------|------|
| 2006 | 21 800 | 特枯 |
| 2007 | 24 400 | 偏枯 |
| 2008 | 26 200 | 偏枯 |
| 2009 | 24 800 | 偏枯 |
| 2010 | 32 400 | 偏丰 |
| 2011 | 21 200 | 特枯 |
| 2012 | 31 700 | 偏丰 |
| 2013 | 25 000 | 偏枯 |
| 2014 | 28 300 | 平 |
| 2015 | 29 000 | 平 |
| 2016 | 33 100 | 特丰 |
| 2017 | 29 700 | 偏丰 |
| 2018 | 25 500 | 偏枯 |
| 2019 | 29 300 | 平 |

3.2 潮位预报及精度评定

首先对三江营站 2006—2019 年的实测潮位进行潮汐调和计算,取平均调和常数,对 2012、2014 和 2018 年 3 个代表年以及 2020 年的天文潮位进行预报,得到的日均天文潮位和逐时天文潮位

预报结果的合格率见表 3。

表3 调和计算的天文潮位合格率

Tab. 3 The qualified rate of astronomical tide calculated by harmonic analysis

| 年份 | 预报误差/cm | 日均天文潮位/% | 逐时天文潮位/% |
|------|------------|----------|----------|
| 2012 | $< \pm 30$ | 62.84 | 60.22 |
| | $< \pm 15$ | 37.16 | 33.89 |
| 2014 | $< \pm 30$ | 75.07 | 71.80 |
| | $< \pm 15$ | 44.93 | 38.63 |
| 2018 | $< \pm 30$ | 68.77 | 63.99 |
| | $< \pm 15$ | 37.26 | 33.91 |
| 2020 | $< \pm 30$ | 43.17 | 40.03 |
| | $< \pm 15$ | 22.95 | 20.24 |

在潮汐调和计算得到的天文潮位基础上,采用 4 种不同的潮位预报方法,对各模型的参数进行率定,然后分别对 3 个代表年和 2020 年的日均潮位进行预报,得到的日均潮位预报结果见图 2~5,日均潮位预报合格率统计结果见表 4,其余精度指标的统计结果见表 5。根据 4 种方法预报的日均潮位、调和计算预报的日均天文潮位和逐时天文潮位,得到 4 种方法预报的逐时潮位,逐时潮位预报合格率见表 6,其余精度指标的统计结果见表 7。

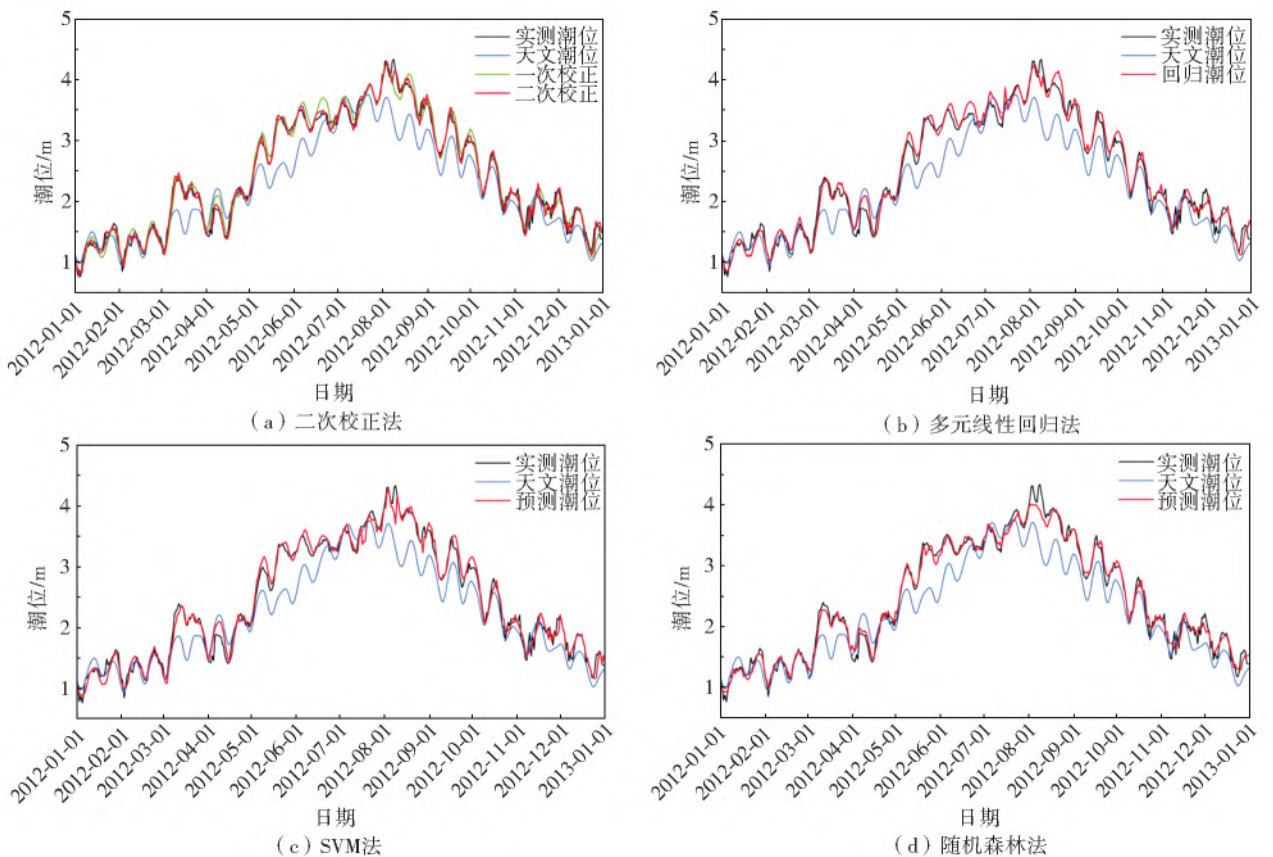


图2 4种方法预报2012年的日均潮位与实测日均潮位对比

Fig. 2 Comparison chart of forecasting and measured mean daily tidal level by four methods of 2012

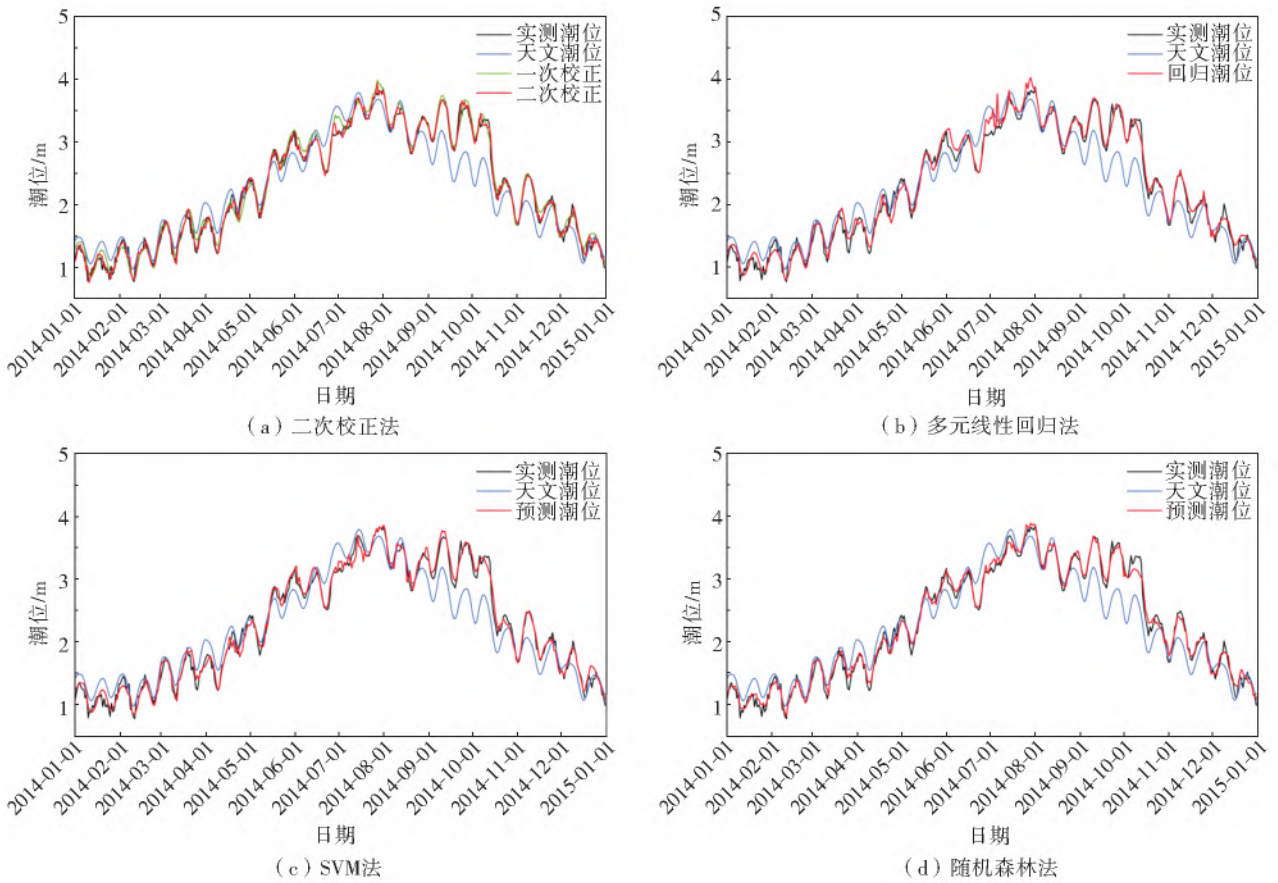


图 3 4 种方法预报 2014 年的日均潮位与实测日均潮位对比

Fig. 3 Comparison chart of forecasting and measured mean daily tidal level by four methods of 2014

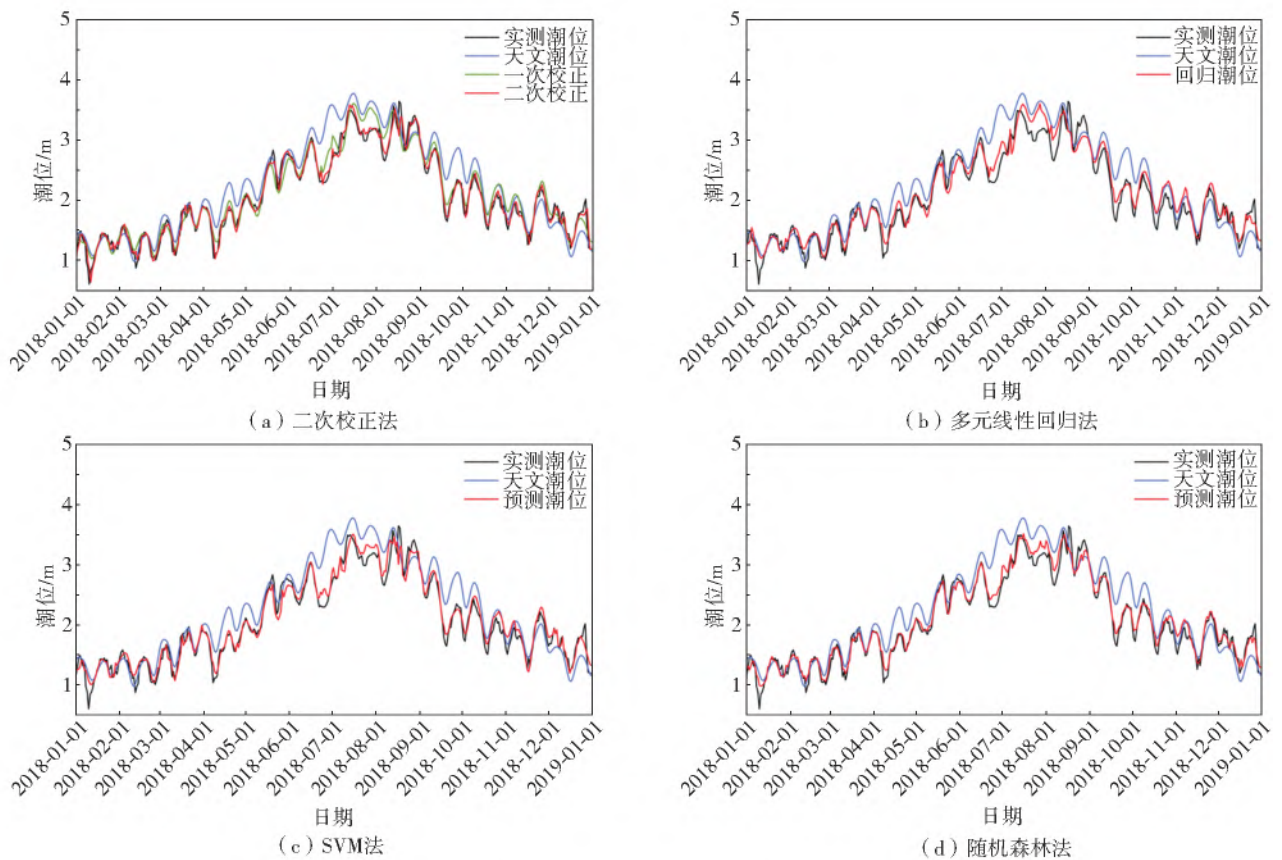


图 4 4 种方法预报 2018 年的日均潮位与实测日均潮位对比

Fig. 4 Comparison chart of forecasting and measured mean daily tidal level by four methods of 2018

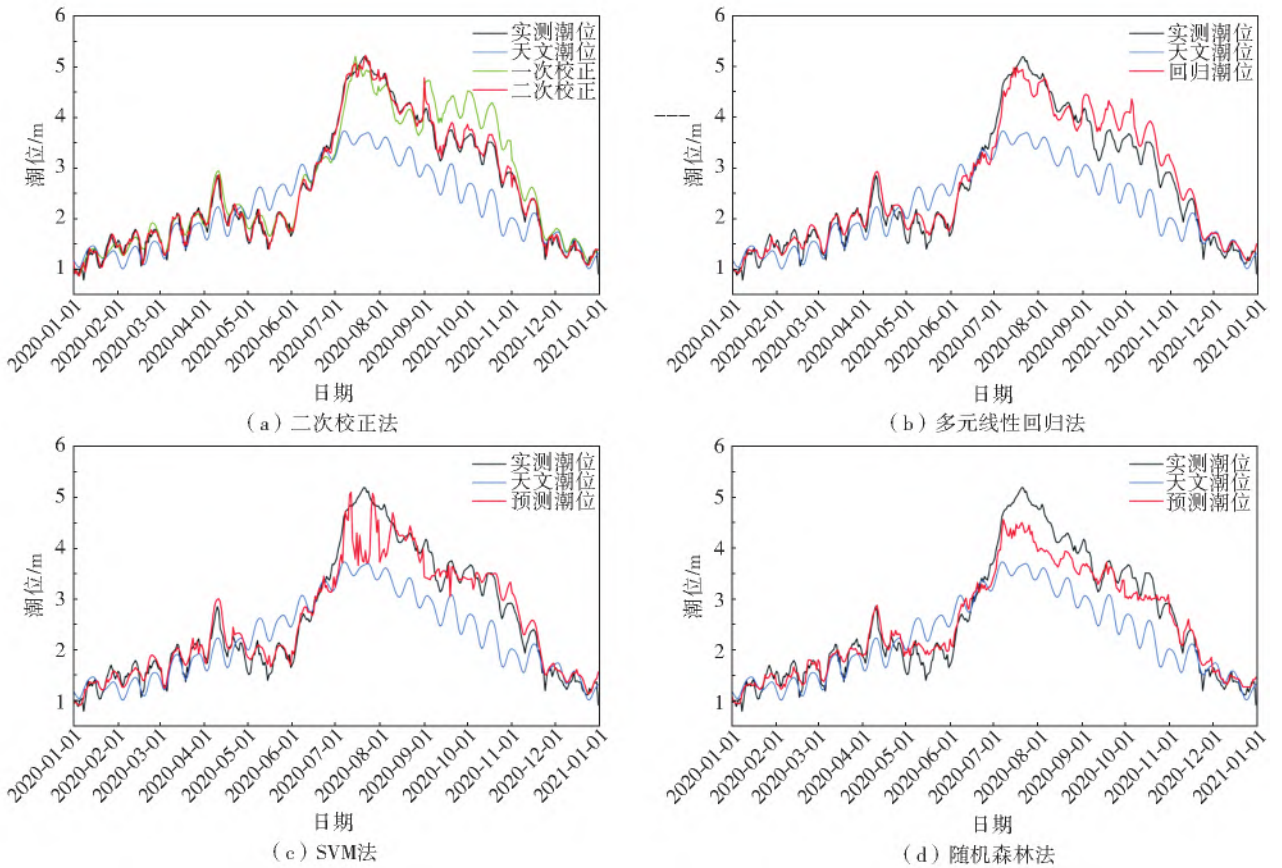


图5 4种方法预报2020年的日均潮位与实测日均潮位对比

Fig. 5 Comparison chart of forecasting and measured mean daily tidal level by four methods of 2020

表4 4种方法预报的日均潮位的合格率

Tab. 4 The qualified rate of daily mean tide level predicted by four methods

| 年份 | 预报误差/cm | 二次校正法/% | 多元线性回归法/% | SVM法/% | 随机森林法/% |
|------|------------|---------|-----------|--------|---------|
| 2012 | $< \pm 30$ | 99.45 | 98.36 | 98.36 | 98.91 |
| | $< \pm 15$ | 89.07 | 74.86 | 80.05 | 88.25 |
| 2014 | $< \pm 30$ | 99.73 | 98.36 | 99.18 | 99.73 |
| | $< \pm 15$ | 92.05 | 84.11 | 84.66 | 90.41 |
| 2018 | $< \pm 30$ | 99.18 | 90.41 | 96.71 | 98.63 |
| | $< \pm 15$ | 88.49 | 64.66 | 77.53 | 83.01 |
| 2020 | $< \pm 30$ | 98.63 | 76.50 | 84.15 | 71.86 |
| | $< \pm 15$ | 83.33 | 48.63 | 51.64 | 42.35 |

表5 4种方法预报的日均潮位精度

Tab. 5 The precision of daily mean tide level predicted by four methods

| 年份 | 评价指标 | 二次校正法 | 多元线性回归法 | SVM法 | 随机森林法 |
|------|-----------|---------|---------|---------|---------|
| 2012 | E_a | 0.074 0 | 0.103 2 | 0.088 0 | 0.079 0 |
| | R | 0.994 5 | 0.990 8 | 0.991 4 | 0.993 7 |
| | E_{RMS} | 0.096 2 | 0.129 6 | 0.120 4 | 0.103 4 |
| 2014 | E_a | 0.065 7 | 0.085 7 | 0.077 7 | 0.070 0 |
| | R | 0.995 3 | 0.992 3 | 0.993 2 | 0.994 9 |
| | E_{RMS} | 0.085 4 | 0.113 5 | 0.103 9 | 0.089 5 |
| 2018 | E_a | 0.072 0 | 0.134 2 | 0.095 0 | 0.084 2 |
| | R | 0.989 8 | 0.971 5 | 0.981 9 | 0.987 4 |
| | E_{RMS} | 0.097 9 | 0.172 1 | 0.132 0 | 0.113 5 |
| 2020 | E_a | 0.085 6 | 0.197 2 | 0.210 4 | 0.233 1 |
| | R | 0.995 4 | 0.980 1 | 0.966 2 | 0.981 1 |
| | E_{RMS} | 0.115 3 | 0.252 6 | 0.322 8 | 0.304 0 |

表 6 4 种方法预报的逐时潮位的合格率

Tab. 6 The qualified rate of hourly tide level predicted by four methods

| 年份 | 预报误差/cm | 二次校正法/% | 多元线性回归法/% | SVM 法/% | 随机森林法/% |
|------|---------|---------|-----------|---------|---------|
| 2012 | <±30 | 97.01 | 94.73 | 95.49 | 96.45 |
| | <±15 | 78.90 | 67.35 | 72.51 | 76.65 |
| 2014 | <±30 | 97.40 | 95.65 | 96.62 | 97.82 |
| | <±15 | 81.64 | 73.32 | 74.93 | 78.38 |
| 2018 | <±30 | 94.25 | 84.36 | 90.99 | 92.69 |
| | <±15 | 70.02 | 52.07 | 61.46 | 64.33 |
| 2020 | <±30 | 92.45 | 69.24 | 75.13 | 65.55 |
| | <±15 | 64.20 | 35.51 | 40.31 | 35.99 |

表 7 4 种方法预报的逐时潮位精度

Tab. 7 The precision of hourly mean tide level predicted by four methods

| 年份 | 评价指标 | 二次校正法 | 多元线性回归法 | SVM 法 | 随机森林法 |
|------|-----------|---------|---------|---------|---------|
| 2012 | E_a | 0.097 6 | 0.123 2 | 0.113 0 | 0.103 3 |
| | R | 0.991 8 | 0.988 8 | 0.989 3 | 0.991 0 |
| | E_{RMS} | 0.130 2 | 0.156 5 | 0.149 0 | 0.135 6 |
| 2014 | E_a | 0.092 8 | 0.110 4 | 0.104 6 | 0.097 7 |
| | R | 0.992 3 | 0.990 0 | 0.990 7 | 0.992 1 |
| | E_{RMS} | 0.123 0 | 0.143 9 | 0.136 5 | 0.125 9 |
| 2018 | E_a | 0.121 2 | 0.168 9 | 0.141 7 | 0.132 2 |
| | R | 0.982 9 | 0.970 8 | 0.978 0 | 0.981 6 |
| | E_{RMS} | 0.158 6 | 0.212 6 | 0.181 6 | 0.168 7 |
| 2020 | E_a | 0.133 2 | 0.232 6 | 0.244 8 | 0.260 5 |
| | R | 0.991 1 | 0.978 0 | 0.962 2 | 0.978 2 |
| | E_{RMS} | 0.171 4 | 0.282 7 | 0.346 8 | 0.329 4 |

为进一步比较 4 种潮位预报方法在不同时间的适用性,分别计算各方法预报的日均潮位于汛期

(5—10 月)和非汛期(11 月—次年 4 月)的合格率,统计结果见表 8、表 9。

表 8 4 种方法预报的日均潮位于汛期的合格率

Tab. 8 The qualified rate of the daily average tide predicted by the four methods in flood season

| 年份 | 预报误差/cm | 二次校正法/% | 多元线性回归法/% | SVM 法/% | 随机森林法/% |
|------|---------|---------|-----------|---------|---------|
| 2012 | <±30 | 99.46 | 98.91 | 98.37 | 97.83 |
| | <±15 | 92.93 | 75.00 | 81.52 | 89.13 |
| 2014 | <±30 | 99.46 | 98.37 | 100.00 | 99.46 |
| | <±15 | 94.57 | 86.41 | 83.15 | 90.22 |
| 2018 | <±30 | 99.46 | 85.33 | 95.65 | 97.83 |
| | <±15 | 89.13 | 52.17 | 70.11 | 78.80 |
| 2020 | <±30 | 98.91 | 60.33 | 72.83 | 50.00 |
| | <±15 | 82.07 | 33.15 | 45.65 | 23.91 |

由表 3 可以看出,仅使用传统的潮汐调和和分析法对潮位进行预报的合格率较低,达不到预期精度。特别在验证期,调和预报的天文潮位合格率相对于率定期大幅下降,因此需要改进预报方法。对比表 4 和表 6 可知,4 种预报方法的预报精度都在使用平均调和常数预报潮位的基础上大幅提升。

4 种方法中,二次校正法的预报合格率最高,对率定期代表年和验证期的日均潮位预报和逐时潮位预报均达到甲级精度。多元线性回归法、SVM 法和随机森林法对率定期代表年的日均潮位预报达到甲级精度,对验证期的日均潮位预报达到乙级精度,其中,以 SVM 法效果较好,多元线性回归法和随机森林

法的效果较差。在对率定期代表年的逐时潮位预报中,多元线性回归法、SVM法和随机森林法潮位预报达到甲级精度;在对验证期的逐时潮位预报中,

SVM法达到乙级精度,而多元线性回归法和随机森林法为丙级精度。

表9 4种方法预报的日均潮位于非汛期的合格率

Tab.9 The qualified rate of the daily average tide predicted by the four methods in non-flood season

| 年份 | 预报误差/cm | 二次校正法/% | 多元线性回归法/% | SVM法/% | 随机森林法/% |
|------|---------|---------|-----------|--------|---------|
| 2012 | <±30 | 99.45 | 97.80 | 98.35 | 100.0 |
| | <±15 | 85.16 | 74.73 | 78.57 | 87.36 |
| 2014 | <±30 | 100.00 | 98.34 | 98.34 | 100.0 |
| | <±15 | 89.50 | 81.77 | 86.19 | 90.61 |
| 2018 | <±30 | 98.90 | 95.58 | 97.79 | 99.45 |
| | <±15 | 87.85 | 77.35 | 85.08 | 87.29 |
| 2020 | <±30 | 98.35 | 92.86 | 95.60 | 93.96 |
| | <±15 | 84.61 | 64.29 | 57.69 | 60.99 |

为进一步比较4种方法的预报精度,将±15 cm作为新的潮位预报允许误差,再计算预报合格率,结果发现减小预报允许误差范围后仍然是二次校正法的预报精度最高。对比原预报允许误差范围(±30 cm),二次校正法的预报合格率降低幅度最小,其他3种预报方法的合格率降幅较大,可以初步认为二次校正法的预报效果更稳定,预报误差更小。通过表5和表7直观比较了4种方法预报潮位的误差和与实测潮位的拟合程度,再次说明二次校正法的平均绝对误差最小,均方根误差最小最稳定,与实测潮位的拟合程度最高,在模型参数率定期和验证期都保证了较高的精度。

总的来说,二次校正法在±30 cm允许误差范围和±15 cm允许误差范围的预报合格率均为最高,预报误差最小且稳定,与实测潮位的拟合度最高。但是这种预报方法的局限性在于预见期仅为1 d。多元线性回归法、SVM法和随机森林法对上游来水和区间降水变化的反应较为敏感,在上游流量或区间降水激增或骤减时预报效果不稳定,预报精度不如二次校正法高,但这3种方法仅需用到预报日前2日的流量和降雨资料,预见期至少有2 d,并且还可以随降水预报和大通流量预报预见期的延长而延长。

此外,在使用多元线性回归法预报日均潮位时,发现各影响因子与实测潮位的相关系数随月份而发生变化。在3种影响因子中,天文潮位和大通流量与实测潮位的相关系数较高,且此消彼长,如:5月至9月,天文潮位与实测潮位的相关系数较低,而大通流量与实测潮位的相关系数较高;10月至次年4月,天文潮位与实测潮位的相关系数较高,而大通流量与实测潮位的相关系数较低。这与实际相吻合:

在长江汛期,由于上游来水量大,因此潮汐引力对于潮位的影响减弱,流量对于潮位的影响增强;在非汛期,上游来水量小,因此潮汐引力对于潮位的影响增强,流量对于潮位的影响减弱。

由表8和表9可以看出,二次校正法在率定期代表年和验证期的汛期和非汛期潮位预报中均表现出较高的精度和稳定性。多元线性回归法、SVM法和随机森林法在率定期代表年的汛期和非汛期潮位预报中的表现均较为理想。在对验证期的非汛期潮位预报中,3种方法的精度也较高。而对于验证期的汛期潮位预报,尤其是像2020年汛期流量大和降水量大导致区间来水多的年份,后3种方法的预报精度降低较大,其中,SVM法达到乙级精度,多元线性回归法为丙级精度,随机森林法更差。但对于汛期持续高潮位的预报,SVM法出现了预报偏低较大的情况。

综上所述,在模型率定中,4种方法对率定期代表年的潮位预报结果均较理想,认为4种方法均可用于部分缺资料年份的补遗。在如2020年的特丰水年潮位预报中,二次校正法的预报效果最好,SVM法次之,多元线性回归法和随机森林法对于汛期的预报效果不佳。

4 结 论

(1)根据《水文情报预报规范》中潮位的预报精度要求,4种预报方法对率定期代表年的日均潮位预报和逐时潮位预报结果均较为理想,可用于缺短序列潮位资料的补遗;在验证期的日均潮位和逐时潮位预报中,二次校正法的精度最高且最稳定,SVM法次之,多元线性回归法和随机森林法相对较差。

(2)4 种方法在非汛期潮位预报中,表现均较为理想。在特丰水年汛期潮位预报中,二次校正法达到甲级精度,SVM 法达到乙级精度,而多元线性回归法和随机森林法精度较低。

(3)4 种方法中,二次校正法的预报精度最高且表现最为稳定,但其缺点在于潮位预见期较短,仅为 1 d。相比之下,多元线性回归法、SVM 法和随机森林法预报精度要低些,但预见期较长。

在实际工程应用中,可根据工程需要选择不同的预报方法。在对预见期要求较高时,汛期可以采用 SVM 法预报潮位,非汛期可以采用 SVM 法、多元线性回归法和随机森林法预报潮位,为相关水利工程调度计划制定和潮位预警提供参考;随着时间的临近和实时资料的获取,采用二次校正法预报潮位,以取得更高精度的预报潮位,为水利工程精细化调度提供科学依据。

参考文献(References):

- [1] 陈吉余. 21 世纪的长江河口初探[M]. 北京:海洋出版社,2009. (CHEN J Y. The Yangtze River estuary in the 21st century [M]. Beijing: China Ocean Press, 2009. (in Chinese))
- [2] 徐汉兴,樊连法,顾明杰. 对长江潮区界与潮流界的研究[J]. 水运工程,2012(6): 15-20. (XU H X, FAN L F, GU M J. Study on tidal mark and tidal current mark in the Yangtze River [J]. Port & Waterway Engineering, 2012(6): 15-20. (in Chinese)) DOI: 10. 16233/j. cnki. issn1002-4972. 2012. 06. 005.
- [3] 方新. 长江感潮河段水位过程预报模型研究[D]. 南京:南京师范大学,2014. (FANG X. Study on prediction model of water level process in tidal reach of Yangtze River[D]. Nanjing: Nanjing Normal University, 2014. (in Chinese))
- [4] 石景元,路川藤. 潮汐调和与分析与应用研究[J]. 海洋技术学报,2019,38(6): 46-50. (SHI J Y, LU C T. Harmonic analysis and application study of tides [J]. Journal of Ocean Technology, 2019, 38(6): 46-50. (in Chinese)) DOI: 10. 3969/j. issn. 1003-2029. 2019. 06. 007.
- [5] 江海东,费岳军,应岳. 使用 MATLAB 拟合工具箱进行潮汐调和与分析[J]. 海洋技术,2012,31(3): 33-36. (JIANG H D, FEI Y J, YING Y. Tidal harmonic analysis using the MATLAB fitting toolbox [J]. Ocean Technology, 2012, 31(3): 33-36. (in Chinese)) DOI: 10. 3969/j. issn. 1003-2029. 2012. 03. 009.
- [6] 李国芳,谭亚,张秀菊. 感潮河段上游流量对潮位预报的影响[J]. 河海大学学报(自然科学版),2006,34(2): 144-147. (LI G F, TAN Y, ZHANG X J. Influence of upstream discharge in tidal level prediction for tidal reaches [J]. Journal of Hohai University(Natural Sciences), 2006, 34(2): 144-147. (in Chinese)) DOI: 10. 3321/j. issn: 1000-1980. 2006. 02. 007.
- [7] 钱睿智,李国芳,王永东. 南水北调东线源头潮汐预报模型研究[J]. 人民长江,2018,49(20): 35-39. (QIAN R Z, LI G F, WANG Y D. Research on tidal forecasting model for source of eastern route project of South-to-North Water Diversion [J]. Yangtze River, 2018, 49(20): 35-39. (in Chinese)) DOI: 10. 16232/j. cnki. 1001-4179. 2018. 20. 007.
- [8] 宋正逵,黄丽亚,仇娟娟. 长江上游径流对天生港站潮位影响浅析[J]. 水文,2010,30(4): 56-58. (SONG Z K, HUANG L Y, QIU J J. Influence of runoff from upper reaches of Yangtze River on tidal level at Tianshenggang station [J]. Journal of China Hydrology, 2010, 30(4): 56-58. (in Chinese)) DOI: 10. 3969/j. issn. 1000-0852. 2010. 04. 013.
- [9] 王文才,李一平,杜薇,等. 长江感潮河段潮汐变化特征[J]. 水资源保护,2017,33(6): 121-124, 132. (WANG W C, LI Y P, DU W, et al. Tidal variation features of tidal reach of Changjiang River [J]. Water Resources Protection, 2017, 33(6): 121-124, 132. (in Chinese)) DOI: 10. 3880/j. issn. 1004-6933. 2017. 06. 19.
- [10] 张轩,张行南,江唯佳,等. 秦淮河流域东山站水位预报研究[J]. 水资源保护,2020,36(2): 41-46, 52. (ZHANG X, ZHANG X N, JIANG W J, et al. Study on water level forecast of Dongshan station in Qinhuai River basin [J]. Water Resources Protection, 2020, 36(2): 41-46, 52. (in Chinese)) DOI: 10. 3880/j. issn. 1004-6933. 2020. 02. 007.
- [11] 吴美玲,杨侃,张铖铖. 基于 KG-BP 神经网络在秦淮河洪水水位预测中的应用[J]. 水电能源科学,2019,37(2): 74-77, 81. (WU M L, YANG K, ZHANG C C. Application of KG-BP neural network in flood forecasting of Qinhuai River [J]. Water Resources and Power, 2019, 37(2): 74-77, 81. (in Chinese))
- [12] BREIMAN L. Bagging predictors[J]. Machine Learning, 1996, 24(2): 123-140. DOI: 10. 1023/A: 1018054314350
- [13] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5-32. DOI: 10. 1023/A: 1010933404324.
- [14] PETERS J. Random forests as a tool for ecohydrological distribution modeling[J]. Ecological Modelling, 2007, 207(2/4): 304-318. DOI: 10. 1016/j. ecolmodel. 2007. 05. 011.
- [15] 童章龙. 潮汐调和和分析的方法和与应用研究[D]. 南

- 京:河海大学,2007. (TONG Z L. Methods and applications of tidal harmonic analysis. Nanjing: Nanjing Hohai University,2007. (in Chinese))
- [16] 黄振平. 水文统计学[M]. 南京:河海大学出版社,2003. (HUANG Z P, Statistical hydrology[M]. Nanjing: Hohai University Press,2003. (in Chinese))
- [17] CRISTIANINI N, SHAWA-TAYLOR J. An introduction of support vector machines and other kernel-based learning methods[M]. Cambridge: Cambridge University Press,2000.
- [18] BURGESS C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 127-167.
- [19] 章国稳,姬战生,孙映宏. 基于 SVM 的河道洪峰水位校正预报方法[J]. 水力发电, 2020, 46(4): 25-27, 40. (ZHANG G W, JI Z S, SUN Y H. Correction of the flood water level forecasting based on SVM[J]. Water Power, 2020, 46(4): 25-27, 40. (in Chinese)) DOI: 10. 3969/j. issn. 0559-9342. 2020. 04. 006.
- [20] 柳成,尹建川. 一种高精度的短期潮汐预报模型[J]. 上海海事大学学报, 2016, 37(3): 74-80. (LIU C, YIN J C. A high-accuracy short-term tide prediction model [J]. Journal of Shanghai Maritime University, 2016, 37(3): 74-80. (in Chinese)) DOI: 10. 13340/j. jsmu. 2016. 03. 014.
- [21] 万育生,王栋,黄朝君. 丹江口水库来水情势分析与径流预测[J]. 南水北调与水利科技(中英文), 2020. (WAN Y S, WANG D, HUANG C Z. Water inflow regime analysis and runoff prediction of Danjiangkou Reservoir[J]. South-to-North Water Transfers and Water Science & Technology, 2020. (in Chinese)) DOI: https://kns. cnki. net/kcms/detail/13. 1430. TV. 20201113. 1834. 002. html.
- [22] 赵钢铁,杨大文,蔡喜明,等. 基于随机森林模型的长江上游枯水期径流预报研究[J]. 水力发电学报, 2012, 31(3): 19-24, 38. (ZHAO T T G, YANG D W, CAI X M, et al. Predict seasonal low flows in the upper Yangtze River using random forest model [J]. Journal of Hydroelectric Engineering, 2012, 31(3): 19-24, 38. (in Chinese))
- [23] 李月玉,崔东文,高增稳. 基于多组群教学优化的随机森林预测模型及应用[J]. 人民长江, 2019, 50(7): 83-86, 91. (LI Y Y, CUI D W, GAO Z W. Random forest forecasting model and its application in hydrology based on optimization of multi-group teaching-learning [J]. Yangtze River, 2019, 50(7): 83-86, 91. (in Chinese)) DOI: 10. 16232/j. cnki. 1001-4179. 2019. 07. 013.
- [24] 中华人民共和国水利部. GB/T 22482—2008: 水文情报预报规范[S]. 北京: 中国标准出版社, 2008. (Ministry of Water Resources of the PRC. GB/T 22482-2008; Standard for hydrological information and hydrological forecasting [S]. Beijing: Standards Press of China, 2008. (in Chinese))
- [25] 中华人民共和国水利部. GB/T 50095—2014: 水文基本术语和符号标准[S]. 北京: 中国计划出版社, 2014. (Ministry of Water Resources of the PRC. GB/T 50095-2014; Standard for terms and symbols used in hydrography [S]. Beijing: China Planning Press, 2004. (in Chinese))

Tidal level forecast of Yangtze River at the source of East Route of South-to-North Water Transfer Project

YAO Siyuan¹, LI Guofang¹, QIAN Ruizhi², WANG Yongdong², MA Fang¹

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China;

2. Yangzhou Branch of Jiangsu Hydrology and Water Resources Survey Bureau, Yangzhou 225002, China)

Abstract: The East Route of South-to-North Water Transfer Project diverts water from the Yangtze River through Jiangdu Water Conservancy Project. Accurate and timely water-level forecast of the diversion project plays a key role in the scientific operation of water conservancy project and reasonable allocation of water resources. The Sanjiangying tide station near the diversion project was selected as a representative. Due to the Sanjiangying tide station locating in the tidal limit, the water level was affected by a combination of astronomical tide and upstream inflow. The interval precipitation should be taken into consideration in the tidal level forecast since the precipitation was unevenly distributed throughout the year because of the unique subtropical monsoon climate. Due to the numerous influencing factors, the tidal level forecast of Sanjiangying can be more complicated.

Tidal harmonic analysis method was used to forecast astronomical tide. The average daily tidal level and the hourly tidal level can be obtained by the tidal harmonic analysis method. Four methods, including secondary correction method, multiple linear regression method, support vector machine method and random forest method were applied to forecast the average daily tidal level of Sanjiangying station. The hourly tidal level forecast was based on the forecast of the average daily tidal level of the above

four methods and the tidal harmonic analysis method. The precision index, including qualified rate, absolute error and root mean square error were used to compare the accuracy of four methods for average daily tidal level and hourly tidal level. Moreover, a linearly dependent coefficient was used to compare the degree of fitting between simulated tidal level and measured tidal level. Finally, the optimum forecasting scheme was recognized.

The qualified rate of a simulated tidal level using tidal harmonic analysis method was too low to meet the actual engineering requirement and therefore other methods are needed to improve the forecast precision. The results of simulated average daily tidal level and simulated hourly tidal level showed that the accuracy of the simulated average daily tidal level determined the accuracy of simulated hourly tidal level. Better average daily tidal level forecast was followed by a better hourly tidal level forecast. In addition, the accuracy of the model training period was higher than the verification period, which was consistent with the assumption of the general case. The prediction accuracy of the average daily tidal level by four methods can reach class B or above in both the training period and verification period. Moreover, the precision index showed that secondary correction method owed the highest accuracy and stability with prediction accuracy reaching class A, followed by the support vector machine method, and the multiple linear regression method and random forest method had the relatively worst performance. The order was kept when forecasting the hourly tidal level, with secondary correction method also reaching class A. The result was also verified by linearly dependent coefficient; secondary correction method had the highest linearly dependent coefficient in the forecast of average daily tidal level and hourly tidal level, which means the simulated tidal level forecasted by secondary correction method was closer to the measured tidal level. Thus, it can be seen that secondary correction method was the best method to forecast tidal level in the four methods on account of its highest prediction accuracy and stability.

Out of comprehensive consideration, secondary correction method is suitable for applying in the fine scheduling of the water conservancy project in contrast to support vector machine, multiple linear regression and random forests, which are more appropriately used for tidal level warning owing to longer prediction period.

Key words: tidal level forecast; secondary correction method; multiple linear regression method; SVM method; random forest method

深入贯彻新发展理念 推进水资源集约安全利用