

中国内地分区年用水总量预测模型研究

周宇茗¹, 陈恬玥¹, 郭青¹, 宋松柏^{1,2}

(1. 西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100; 2. 西北农林科技大学
旱区农业水土工程教育部重点实验室, 陕西 杨凌 712100)

摘要: 为了完善中国内地大尺度区域用水量预测模型的研究, 为水资源科学利用提供支撑, 应用 ARMA、灰色 GM(1,1) 及 BP 神经网络模型原理, 建立省级行政区划、流域、地理大区 3 种尺度的年用水总量预测模型, 对模型优选结果进行统计分析, 并应用最优模型对 2021–2025 年各省级行政区用水总量进行预测。结果表明: 对于内地省级行政区划尺度, 9 个行政区的年用水总量最优预测模型为 ARMA 模型, 6 个行政区的最优预测模型为灰色 GM(1,1) 模型, 16 个行政区的最优预测模型为 BP 神经网络模型; 对于流域尺度, 5 个流域的年用水总量最优预测模型为 ARMA 模型, 3 个流域的最优预测模型为灰色 GM(1,1) 模型, 长江流域的最优预测模型为 BP 神经网络模型; 对于地理大区尺度, 北方 6 区的年用水总量最优预测模型为 BP 神经网络模型, 南方 4 区的最优预测模型为灰色 GM(1,1) 模型; 2021–2025 年内陆地各省级行政区的年用水总量总体保持稳定。研究结果可为中国内地用水总量的管理提供依据。

关键词: 年用水总量预测; ARMA 模型; 灰色 GM(1,1) 模型; BP 神经网络模型; 中国内地

中图分类号: TV213.4

文献标识码: A

文章编号: 1672-643X(2022)06-0111-09

Prediction models of regional annual total water consumption in mainland China

ZHOU Yuming¹, CHEN Tianyue¹, GUO Qing¹, SONG Songbai^{1,2}

(1. College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling 712100, China;

2. Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of
Education, Northwest A&F University, Yangling 712100, China)

Abstract: In order to improve large-scale water consumption prediction models for mainland China and provide a technical support for the scientific utilization of water resources, the annual total water consumption prediction models of provincial administrative divisions, river basins and geographical regions were established based on the principles of ARMA, gray GM (1,1) and BP neural network model. The optimized results of the models were statistically analyzed, and the selected optimal models were used to predict the total water consumption from 2021 to 2025. The results show that for the provincial administrative division scale, the optimal prediction model of total annual water consumption is ARMA model in nine administrative divisions, gray GM (1,1) model in six administrative divisions, and neural network model in 16 administrative divisions. At the river basin scale, the optimal prediction model is ARMA model in five basins, gray GM (1,1) model in three basins, and neural network model in the Yangtze Basin. For the scale of large geographical regions, the optimal model of the six northern regions is the neural network model, and that of the four southern regions is the gray GM (1,1) model. The water consumption in the next five years will be generally stable. These findings are expected to provide some statistical reference for total water consumption management in China.

Key words: annual total water consumption prediction; ARMA model; gray GM (1,1) model; BP neural network model; mainland China

收稿日期: 2022-01-26; 修回日期: 2022-06-23

基金项目: 国家自然科学基金项目 (52079110); 西北农林科技大学大学生科技创新项目 (S202110712560)

作者简介: 周宇茗 (2001–), 女, 陕西西安人, 在读本科生, 研究方向为水文及水资源。

通讯作者: 陈恬玥 (2001–), 女, 陕西西安人, 在读本科生, 研究方向为水文及水资源。

1 研究背景

近年来,全球气候变暖影响不断加剧,各国都在为可持续发展而蓄力,我国为实现碳达峰、碳中和的“双碳”目标也必将向清洁能源高效利用转型。水资源作为重要的绿色能源,在过去经济快速发展过程中却普遍存在被污染和浪费的现象,尤其是在用水阶段,缺乏节水意识引导、定量用水制度规范以及用水现状实时监管等,因而亟需对用水总量进行较为精准的预测,以期为我国水资源规划和管理以及调整用水结构等决策提供支撑。

目前,我国对用水量的研究涉及水量统计调查的差异分析、用水实时监测方法修正、用水结构成因分析等^[1-7],而用水量预测一直是研究的热点,其中包括从用水量变化驱动因素的角度进行趋势规律定性分析^[8-9],例如,刘晶等^[10]采用 Mann-Kendall 法揭示了我国 10 个水资源一级区的水资源量及用水总量的变化规律;赵勇等^[11]对全国 358 个地级行政区的用水总量增长趋势进行了研究并得到用水极限峰值及发生时间。其次包括基于不同理论模型的定量估计,较为成熟的方法有定额法、时间序列法、回归分析法、灰色 GM(1,1)模型法、神经网络法及系统动力学法等,各种方法均取得了一定的研究成果^[12-13],例如,何靖等^[13]通过建立灰色系统,对陕西省农业、工业、生活和生态用水总量进行了预测。但目前更为常见的是为提高预测精度而对单一模型组合升级后进行用水量预测^[14-20],包括分数阶灰色幂次模型、灰色马尔科夫和非线性灰色贝努利模型等。由于大尺度范围内用水量数据汇总的工作量较大,在预测模型的优选方面,则较多为小区域范围的研究^[21-22],例如,白鹏等^[23]基于京津冀地区年用水总量数据分别采用自回归模型法、灰色 ANN 法和年增长率法 3 种预测方法进行了预测及对比;杨建强等^[24]通过建立 ARMA、GM(1,1)、ES-GM(1,1)和 ES-ARMA 模型对酒泉市用水总量进行了预测。国外对于用水量预测研究起步较早,早期研究内容多为中短期用水预测的基础算法分析和模型建构^[25-27],在此基础上随着技术的发展不断提出了新模型并进行对比,例如 Antunes 等^[28]基于葡萄牙两个水务公司的实测数据建立了 ANN、ARMA、支持向量机等模型进行预测和结果对比。当今依托发达的卫星遥感技术,各地区的水文模型及相关统计参数日臻完备,气象、农业、自然资源数据库系统也逐渐成熟,已有研究基于 BESS-STAIR 数据和应用卫星数据模型可完成对全球农业用水量的精准预测。

目前我国在大尺度用水量预测研究方面,既缺少定量预估,又缺乏不同模型的对比分析。本文基于 ARMA、GM(1,1)、BP 神经网络 3 种模型,分别从省级行政区划、流域、地理大区 3 种尺度建立用水总量预测模型,得到预测值并进行模型优选,再统计出中国内地各分区用水量最优预测模型种类的数量,以期找出其随地理空间分布变化是否具有区域普遍适用性的规律,为各地区不同用水预测模型的组合创新提供参考,以提高用水预测精度,同时也为我国用水量数据的统一规范和智慧水务互联管理提供依据,减少由各地区用水统计差异而造成的误差,从而更加精准地制定用水制度,有利于节水型城市建设和国家能源结构转型。

2 数据来源与研究方法

2.1 研究区概况与数据来源

我国水资源开发利用工程多样,供用水结构复杂,水管理部门繁多,在进行水资源用水评价时往往从不同角度出发进行分区。本文将在中国内地范围按照省级行政区、水文地理(流域)大区和自然地理大区 3 种尺度划分,分别建立年用水总量预测模型。

全国用水总量数据主要来源为《中国水资源公报》(不包括香港、澳门地区和台湾省的相关数据),具有一定的科学性和权威性。分别选用 1998-2017 年用水总量数据为训练期,2018-2020 年用水总量数据为验证期,建立省级行政区尺度年用水总量预测模型;流域和地理大区尺度年用水总量预测模型则选择 1997-2017 年用水总量数据为训练期,2018-2020 年用水总量数据为验证期。在建立预测模型前,通过序列组分识别和提取处理,满足序列平稳性的要求。

2.2 研究方法

2.2.1 自回归滑动平均模型 ARMA 自回归滑动平均(autoregressive moving average, ARMA)模型^[18]是描述平稳随机水文序列的一类最主要的线性平稳模型,可根据前后数据相关关系来建立回归方程。该模型具有两个优点:(1)能够揭示平稳随机水文序列的一些主要统计特性(如变差系数、自相关系数等)。(2)模型参数具有一定的物理意义。ARMA 模型的一般形式为:

$$x_t = u + \varphi_1(x_{t-1} - u) + \varphi_2(x_{t-2} - u) + \cdots + \varphi_p(x_{t-p} - u) + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} \quad (1)$$

式中: x_t 为平稳、正态、零均值的时间序列; u 为常数项; $\varphi_1, \varphi_2, \cdots, \varphi_p$ 为自回归系数; p 为自回归阶数;

$\theta_1, \theta_2, \dots, \theta_q$ 为滑动平均系数; q 为滑动平均阶数; ε_t 为独立的误差项。

ARMA 模型的建立一般分为以下 6 个步骤: (1) 组分识别。采用相应的方法对序列的趋势、跳跃及周期成分进行识别并排除, 获得平稳随机序列。(2) 选择模型类型。绘制序列的自相关、偏相关函数图, 进行截尾性判定分析。(3) 识别模型形式。给定 p, q 的上限值, 分别计算 $AIC(p, q)$ 或 $BIC(p, q)$, 确定模型阶数。(4) 估计模型参数。依据 Yule - Walker 方程、矩法等方法计算模型参数。(5) 模型检验。采用综合自相关系数法进行检验, 分析残差随机变量的独立性。(6) 模型适用性分析。采用长序列、短序列法分析所建立模型是否能保持实测序列的主要统计特性。

2.2.2 灰色 GM(1,1) 模型 灰色预测^[10]是利用各因素之间的关联性估计其趋势的相异程度, 其关键在于对原始数据做相加、累减或平均形成近似的指数规律后再建立灰色 GM(1,1) 模型。灰色 GM(1,1) 模型是指应用于中短序列且仅包含一个变数的一阶变量微分方程模型, 在水文预报中作为基于长时间序列的中期或短期预报。一般选择后验差检验法测试模型的可靠性, 并利用最小误差概率与后验差的比值 C 判断模型的正确性, 分别以 $C = 0.35, 0.50, 0.65$ 为分界点将精度划分为优、良、合格及不合格 4 个等级。

2.2.3 BP 神经网络模型 人工神经网络 (artificial neural network, ANN) 模型是模仿动物神经反馈机制

进行信息处理, 能够自学习、自适应的一种数学模型。其工作原理是基于人工神经元, 将输入数据经过激活函数处理得到输出数据, 常用的激活函数包括线性函数、S 形函数、双极 S 形函数等, 训练方法包括梯度下降法、量化共轭梯度法、拟牛顿算法和一步正割算法等。根据输出值的检验精度调整激活函数类型, 激活函数的选择耗时长、影响深, 是神经网络预测过程中的重点。

神经网络由大量神经元构成, 其结构包括前馈、反馈神经网络和自组织网络 3 种, 工作模式为有、无导师学习两大类, BP (back propagation) 神经网络属于前馈有导师学习神经网络系统。前馈网络反馈信号在训练中产生, 进而在分类时向前传递到输出层。有导师学习算法是将样本原值与真实输出值之差作为调整标准, 不断重复操作直到该误差处于满足条件的范围内。这种能学习和存贮大量的激励 - 响应模式映射关系并不断学习调整连接权的多层前馈网络, 是目前普遍应用的神经网络模型之一。

3 结果与分析

3.1 模型预测结果

3.1.1 ARMA 模型预测结果 依据 ARMA 自回归滑动平均模型原理, 借助 MATLAB 进行建模, 分别得到省级行政区划、流域、地理大区 3 种尺度下的 ARMA 模型系数, 见表 1、2。3 种不同尺度下预测模型结果表明, ARMA(p, q) 模型占比较大, 只有少数省市为 ARMA(p) 模型。

表 1 部分省级行政区 ARMA 模型参数

省份	模型类别	自回归系数			滑动平均系数			常数项
		φ_1	φ_2	φ_3	θ_1	θ_2	θ_3	
山东省	ARMA(1,0)	0.863						29.595
云南省	ARMA(3,3)	0.776	-0.601	-0.171	-0.042	-0.042	1.000	248.600
河南省	ARMA(1,1)	0.616			-0.338			85.017
青海省	ARMA(1,0)	0.794			-1.000			6.568
天津市	ARMA(1,2)	1.000			-0.609	-0.391		0.194
内蒙古自治区	ARMA(3,3)	0.871	-0.768	-0.421	0.031	0.989	0.335	197.390
浙江省	ARMA(2,2)	1.975	-0.976		-2.000	0.999		
江西省	ARMA(1,0)	0.816						43.126
安徽省	ARMA(1,2)	0.974			-0.887	1.000		11.944
广东省	ARMA(1,0)	0.749						112.950
湖北省	ARMA(2,1)	0.066	0.796		1.000			36.662
海南省	ARMA(3,1)	0.999	-0.388	-0.150	-1.000			24.358

表2 各流域及地理大区 ARMA 模型系数

流域	模型类别	自回归系数			滑动平均系数			常数项
		φ_1	φ_2	φ_3	θ_1	θ_2	θ_3	
松辽	ARMA(1,2)	0.887			0.220	0.529		71.981
海河	ARMA(1,2)	0.875			-1.000			45.313
黄河	ARMA(3,0)	0.254	-0.073	0.389				165.380
淮河	ARMA(2,3)	-0.192	-0.109		0.341	0.208	-0.219	727.020
长江	ARMA(1,1)	1.000			-1.000			14.872
东南诸河	ARMA(3,1)	-0.471	0.579	0.149				238.810
珠江	ARMA(1,0)	0.746						217.210
西南诸河	ARMA(3,3)	1.763	-1.412	0.553	-1.661	1.484	-0.377	9.9630
西北诸河	ARMA(3,2)	0.547	-0.425	0.878	0.703	1.000		-0.684
北方6区	ARMA(2,1)	0.202	0.646		0.539			360.950
南方4区	ARMA(2,2)	1.662	-0.667		-1.125	0.125		24.704

3.1.2 灰色 GM(1,1) 模型预测结果 根据灰色 GM(1,1) 原理,经计算分别得到省级行政区划、流域、地理大区 3 种尺度下的灰色 GM(1,1) 模型系数 a 、 b 及后验差比值 C , 见表 3、4。由表 3、4 可知,在全国内地 31 个省级行政区中,仅河北省年用水总量的拟合精度达到了优,后验差比值恰为 0.35;内蒙古、安徽、山西、陕西和新疆 5 个省级行政区模拟精度为良,后验差比值分别为 0.36,0.43,0.46,0.36,0.38;山东、江西、湖北、湖南等 10 个省级行政区模拟精度为合格;云南、河南等 15 个省级行政区模拟精度均为不合格。在九大流域中,长江、西北诸河流域的模拟精度为良,后验差比值分别为 0.46、0.36;松辽流域后验差比值为 0.59,模拟精度为合格;其他流域模拟精度均为不合格。南方 4 区后验差比值为 0.41,模拟精度为良;北方 6 区后验差比值为 0.63,模拟精度为合格。上述结果表明,虽然灰色 GM(1,1) 模型能够对一些地区用水总量序列进行预测,但并不具有普遍适用性,仍然存在大量精度不符合标准的情况,这是由于灰色 GM(1,1) 模型的适用性和用水总量序列自身规律难以拟合造成的。

3.1.3 BP 神经网络模型预测结果 依据 BP 神经网络模型原理,借助 MATLAB 软件,先通过 mapminmax 函数将输入值特征化归一以便识别处理,再应用 newff 函数建立用水总量预测 BP 神经网络模型,可分别得到省级行政区划、流域、地理大区 3 种尺度下神经网络模型的预测评价结果,如表 5 所示。由表 5 可以看出,所有尺度下的预测平均相对误差均在 30% 以内;流域尺度下珠江和东南诸河的预测误差相对较大,

表3 各省级行政区灰色 GM(1,1)模型构建结果

省份	发展系数 a	灰色作用量 b	后验差比值 C
山东省	0.0087	246.19	0.60
云南省	-0.0021	145.73	0.78
河南省	-0.0051	208.90	0.89
青海省	-0.0016	30.59	1.00
天津市	-0.0135	20.31	0.73
内蒙古自治区	-0.0059	168.34	0.36
浙江省	0.0063	214.29	0.69
江西省	-0.0149	195.41	0.61
安徽省	-0.0287	181.50	0.43
广东省	0.0010	456.84	0.96
湖北省	-0.0094	248.00	0.62
海南省	-0.0002	45.00	0.94
湖南省	-0.0028	314.89	0.56
四川省	-0.0132	190.19	0.91
重庆市	-0.0197	59.95	0.57
江苏省	-0.0130	457.57	0.63
广西壮族自治区	-0.0016	293.92	0.95
甘肃省	0.0017	123.40	0.75
上海市	0.0000	114.30	1.00
贵州省	-0.0065	88.09	0.77
辽宁省	-0.0025	134.53	0.94
吉林省	-0.0167	96.36	0.59
河北省	0.0089	217.77	0.35
山西省	-0.0207	50.67	0.46
福建省	-0.0071	178.62	0.62
北京市	0.0001	36.70	0.93
陕西省	-0.0113	74.21	0.36
黑龙江	-0.0188	253.25	0.56
新疆维吾尔 自治区	-0.0115	467.81	0.38
宁夏回族自 治区	0.0150	87.32	0.55
西藏自治区	-0.0079	28.66	0.66

分别为 11.46%、29.23%,其他流域的预测平均相对误差则在 10% 以下;省级行政区划尺度下,广东、四川、江苏和吉林 4 个省级行政区的预测误差相对较大,分别为 9.16%、7.38%、10.33%、8.16%,其他省级行政区的预测平均相对误差基本在 5% 左右。可见 BP 神经网络模型对用水总量预测的适用性较好,能够应用于中长期用水量预测。

表 4 流域及地理大区灰色 GM(1,1)模型构建结果

流域	发展系数 a	灰色作用量 b	后验差比值 C
松辽	-0.0122	556.20	0.59
海河	0.0056	404.32	0.68
黄河	-0.0034	368.15	0.93
淮河	-0.0051	569.73	0.92
长江	-0.0106	1695.10	0.46
东南诸河	-0.0039	311.90	0.87
珠江	-0.0012	845.05	0.92
西南诸河	-0.0071	94.26	0.69
西北诸河	-0.0120	549.69	0.36
北方 6 区	-0.0067	2442.60	0.63
南方 4 区	-0.0077	2920.40	0.41

3.2 3 种模型优选

对于上述 3 种模型,选择 2018 - 2020 年用水总量数据进行省级行政区划、流域、地理大区 3 种尺度的年用水总量预测结果验证,用平均绝对百分比误差 (mean absolute percentage error, $MAPE$) 判断预测精度,择出最优模型。图 1 为 3 个尺度中较为典型的区域的年用水总量不同模型预测结果与实际值对比,从图 1 中可以看出,不同尺度下的最优模型均与实际用水总量的年际波动变化一致;灰色 GM(1,1)模型与 ARMA 模型的拟合差异不明显,在用水总量变化为单一线性分布时,灰色 GM(1,1)模型优势较为明显,而当变化为非线性时,ARMA 模型优势更显著;BP 神经网络模型的拟合程度不太稳定,这可能与其激活函数的选择有关。模型在训练期的拟合度较好,而在验证期则会表现出显著差异,以云南省为例(图 1(b)),训练期 BP 神经网络和 ARMA 模型的预测效果均明显优于灰色 GM(1,1)模型,验证期 ARMA 模型优势依旧,BP 神经网络模型则表现出偏离。因而训练期拟合结果不能反映出模型的真实预测水平,应以验证期相对误差作为模型优选标准。

表 5 省级行政区划、流域及地理大区 BP 神经网络模型预测结果平均相对误差 %

流域、大区	平均相对误差	省份	平均相对误差	省份	平均相对误差	省份	平均相对误差
南方 4 区	3.97	山东省	2.72	广西壮族自治区	4.34	海南省	1.73
北方 6 区	2.03	云南省	3.21	甘肃省	4.86	湖南省	2.87
松辽	8.79	河南省	2.11	上海市	5.45	四川省	7.38
黄河	2.81	青海省	4.01	贵州省	3.50	重庆市	1.98
淮河	2.27	天津市	6.02	辽宁省	1.59	陕西省	2.83
长江	1.05	内蒙古自治区	3.37	吉林省	8.16	黑龙江省	1.81
海河	4.61	浙江省	7.01	河北省	0.77	新疆维吾尔自治区	4.89
珠江	11.46	江西省	2.80	山西省	2.44	宁夏回族自治区	2.79
西南诸河	5.26	安徽省	4.69	福建省	1.96	西藏自治区	5.59
西北诸河	2.80	广东省	9.16	湖北省	4.32		
东南诸河	29.23	江苏省	10.33	北京市	4.46		

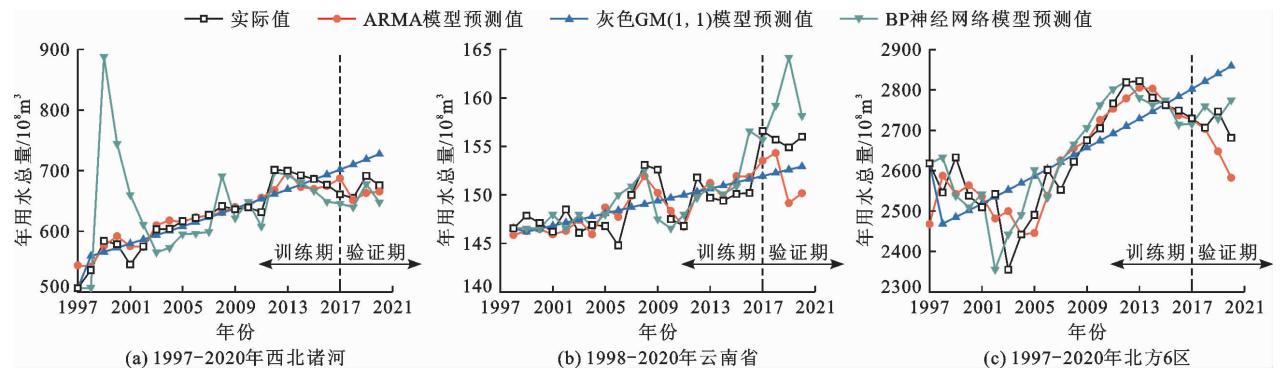


图 1 3 种尺度中典型区域年用水总量的不同模型预测结果与实际值对比

因数据较多,本文选取部分省份、流域及地理大区 2018 - 2020 年 3 种模型的年用水总量预测结果列于表 6,并计算出预测值与实际值的相对误差 (RE) 及平均绝对百分比误差 ($MAPE$) 进行对比。由表 6 可以看出,省级行政区划尺度下,ARMA、灰色 GM(1,1) 和 BP 神经网络 3 种模型对辽宁省年用水总量预测值的 $MAPE$ 分别为 0.52%、9.31%、1.59%,优选模型为 ARMA 模型;对河北省年用水总量预测值的 $MAPE$ 分别为 0.95%、1.42%、0.77%,优选模型为 BP 神经网络模型;对云南省年用水总量预测值的 $MAPE$ 分别为 2.78%、1.88%、3.21%,优选模型为灰色 GM(1,1) 模型。流域尺度下,3 种模型对西北诸河年用水总量预测

值的 $MAPE$ 分别为 2.06%、4.36%、2.80%,优选模型为 ARMA 模型;对长江流域年用水总量预测值的 $MAPE$ 分别为 3.17%、2.47%、1.05%,优选模型为 BP 神经网络模型;对黄河流域年用水总量预测值的 $MAPE$ 分别为 1.30%、0.73%、2.81%,优选模型为灰色 GM(1,1) 模型。地理大区尺度下,对北方 6 区年用水总量预测值的 $MAPE$ 分别为 2.46%、2.69%、2.03%,优选模型为 BP 神经网络模型;对南方 4 区年用水总量预测值的 $MAPE$ 分别为 5.61%、3.63%、3.97%,优选模型为灰色 GM(1,1) 模型。上述结果表明,3 种模型在同一尺度下对不同区域的年用水总量预测误差均有所不同。

表 6 3 种模型对部分省份、流域及地理大区 2018 - 2020 年的年用水总量预测结果对比

类别	年份	实际值/ 10^8 m^3	ARMA 模型			灰色 GM(1,1) 模型			BP 神经网络模型			优选模型
			预测值/ 10^8 m^3	$RE /$ %	$MAPE /$ %	预测值/ 10^8 m^3	$RE /$ %	$MAPE /$ %	预测值/ 10^8 m^3	$RE /$ %	$MAPE /$ %	
辽宁省	2018	130.3	129.72	-0.45		141.71	8.75		131.59	0.99		ARMA
	2019	130.3	129.02	-0.98	0.52	142.07	9.03	9.31	131.60	1.00	1.59	
	2020	129.3	129.48	0.14		142.42	10.15		125.71	-2.77		
河北省	2018	182.4	182.26	-0.08		181.51	-0.49		181.66	-0.40		BP 神经网络
	2019	182.3	181.25	-0.57	0.95	179.91	-1.31	1.42	181.04	-0.69	0.77	
	2020	182.8	178.80	-2.19		178.32	-2.45		180.56	-1.23		
云南省	2018	155.7	154.33	-0.88		152.28	-2.20		159.21	2.25		ARMA
	2019	154.9	149.16	-3.71	2.78	152.60	-1.48	1.88	164.18	5.99	3.21	
	2020	156.0	150.17	-3.74		152.93	1.95		158.15	5.43		
西北诸河	2018	654.9	651.32	-0.55		710.30	8.46		639.58	-2.34		ARMA
	2019	691.4	663.02	-4.10	2.06	718.86	3.97	4.36	678.17	-1.91	2.80	
	2020	675.7	665.45	-1.52		727.51	7.67		647.71	-4.14		
长江	2018	2071.6	2079.20	0.37		2129.10	2.78		2056.60	-0.72		BP 神经网络
	2019	2064.5	2094.00	1.43	3.17	2151.80	4.23	2.47	2027.10	-1.81	1.05	
	2020	1957.6	2108.90	7.73		2174.70	11.09		1945.80	-0.60		
黄河	2018	391.7	391.46	-0.06		396.44	1.21		404.40	3.24		灰色 GM(1,1)
	2019	400.2	388.05	-3.04	1.30	397.80	-0.60	0.73	391.32	-2.22	2.81	
	2020	392.7	389.50	-0.81		399.17	1.65		404.40	2.98		
北方 6 区	2018	2706.4	2703.40	-0.11		2821.10	4.24		2759.70	1.97		BP 神经网络
	2019	2746.5	2647.70	-3.60	2.46	2840.00	3.40	2.69	2728.20	-0.67	2.03	
	2020	2681.3	2582.50	-3.68		2859.00	6.63		2774.30	3.47		
南方 4 区	2018	3309.1	3384.60	2.28		3449.30	4.24		3348.60	1.19		灰色 GM(1,1)
	2019	3274.7	3418.90	4.40	5.61	3476.10	6.15	3.63	3570.00	9.02	3.97	
	2020	3131.6	3449.70	10.16		3503.10	11.86		3078.20	-1.71		

对研究范围内3种尺度下的最优预测模型进行数学统计可知:在省级行政区划尺度下,陕西省、北京市、广东省等9个区域的年用水总量最优预测模型为ARMA模型;云南省、河南省、天津市等6个区域的最优预测模型为灰色GM(1,1)模型;山东省、青海省、安徽省等16个区域的最优预测模型为BP神经网络模型。在流域尺度下,松辽、海河等5个流域的年用水总量最优预测模型为ARMA模型;黄河、淮河和珠江3个流域的最优预测模型为灰色GM(1,1)模型;长江流域的最优预测模型为BP神经网络模型。在地理大区尺度下,北方6区的年用水总量最优预测模型为BP神经网络模型;南方4区的最优预测模型为灰色GM(1,1)模型。

3.3 中国内地分区年用水总量预测

从省级行政区划尺度出发,根据选取的各行政区年用水总量最优预测模型,分别对2021–2025年的年用水总量进行预测,结果如图2所示。

由图2可见,最优模型为ARMA的9个省级行政区中(图2(a)),2021–2025年浙江省年用水总量的减少趋势相对明显,年用水总量平均减少速率

为 $5.88 \times 10^8 \text{ m}^3/\text{a}$;辽宁省年用水总量呈小幅增加趋势,平均增加速率为 $1.87 \times 10^8 \text{ m}^3/\text{a}$;其余各省(自治区)年用水规模基本持平。最优模型为灰色GM(1,1)模型的6个省级行政区中(图2(b)),江苏省的年用水总量增加幅度相对较大,平均增大速率为 $8.27 \times 10^8 \text{ m}^3/\text{a}$,增速较快,该省2025年年用水总量预测值为 $653.3 \times 10^8 \text{ m}^3$ 。最优模型为BP神经网络模型的16个省级行政区中(图2(c)、2(d)),2021–2025年甘肃、安徽、湖南、福建、山西、海南和贵州7个省份的年用水总量基本不变;广西壮族自治区、宁夏回族自治区和黑龙江省3个省份(自治区)年用水总量均呈现先减少后增大的趋势,由减少到增大变化的转折点年份分别为2023、2022和2024年,特别注意到,黑龙江省2025年的用水总量增幅较大,比2024年增加了 $62.95 \times 10^8 \text{ m}^3$;新疆维吾尔自治区年用水总量在2021和2022年基本不变,2023和2024年连续减少,2025年与2024年又基本持平,2022–2025年的年用水总量平均减少速率为 $28.11 \times 10^8 \text{ m}^3/\text{a}$;青海省年用水总量呈连续小幅波动变化状态。

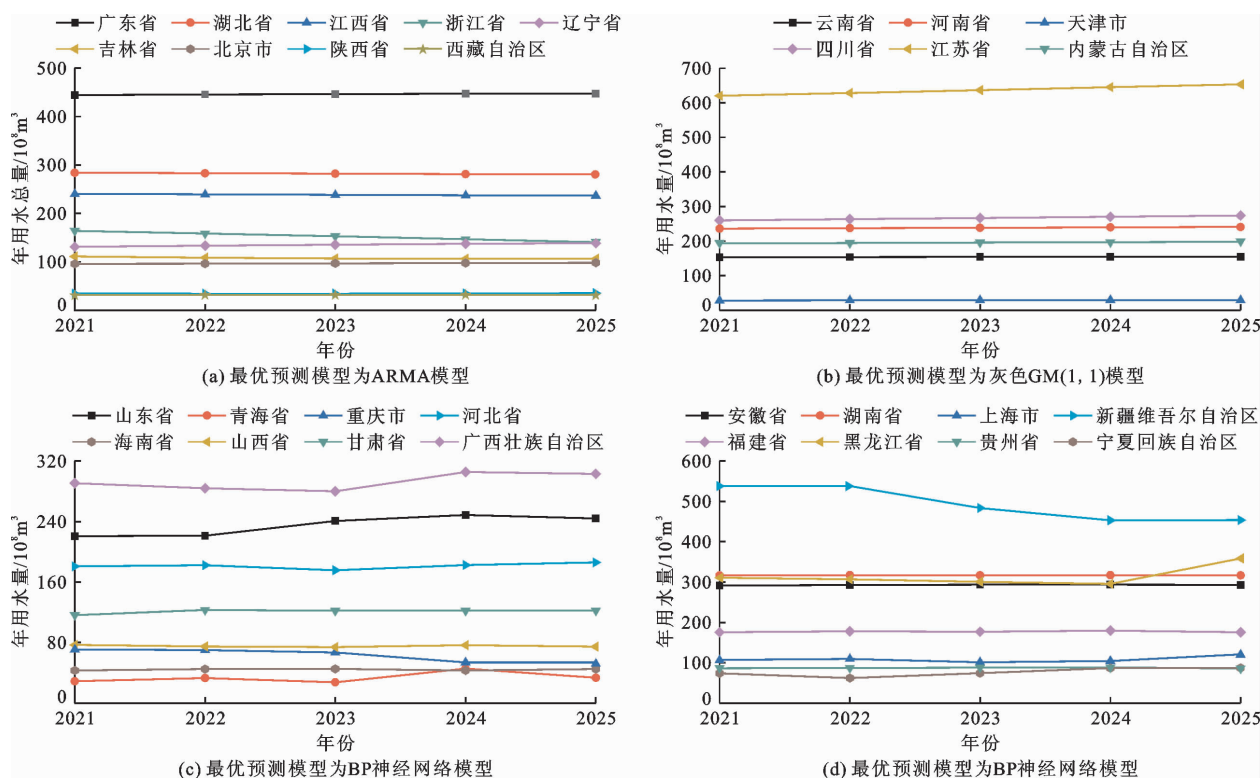


图2 中国内地各省级行政区2021–2025年年用水总量预测结果(以最优预测模型分类)

4 讨论

本研究对中国内地不同尺度区域年用水总量最

优预测模型进行分析,可以得到在数量上的统计值。其中,省级行政区划尺度下最优预测模型中占比最大的为BP神经网络模型,流域尺度下占比最大的

为 ARMA 模型,地理大区尺度下北方 6 区最优预测模型为 BP 神经网络模型,南方 4 区最优预测模型为灰色 GM(1,1)模型,而最优预测模型在地理空间分布上没有显示出一定的规律性,可能是因为本文选取的 3 种预测模型均为时间序列模型,忽略了除用水量外的其他因素影响。在已有的其他小尺度区域用水量最优预测模型的研究中,同时考虑到了结构分析法,针对特定区域的用水量与其影响因素进行了相关分析,而大尺度区域的用水量预测多为农业用水量预测,是基于卫星遥感数据综合多种水文要素得到的。

虽然从本研究选用的用水量预测模型中没有发现一定的普适规律,但各尺度下模型预测值的相对误差均较为合理,具有一定推广价值。与以往研究相比,如 Meng 等^[30]利用灰色模型对我国内地 31 个省(自治区)的工业、农业、生活及总用水量进行预测,Meng 等^[31]考虑区域 GDP、人口与用水量关系建立 AGMC(1,N)模型对我国内地 31 个省(自治区)用水量进行预测,本研究还涉及了流域和地理大区尺度,并选用 3 种模型进行了优选。

综合以上方面,可考虑结合其他相关因素(如经济发展、人口增长等)选择适当模型,对 3 种尺度下用水量进行预测,以完善本研究结果。

5 结 论

(1)依据自回归滑动平均模型 ARMA、灰色 GM(1,1)模型及 BP 神经网络模型原理,利用各省级行政区划、流域、地理大区尺度 1997–2020 年的年用水量,得到各模型的相应参数。

(2)对 3 种尺度下 3 种用水量预测模型进行优选并得到数量统计结果:在省级行政区划尺度下,陕西省、北京市、广东省等 9 个区域的年用水量最优预测模型为 ARMA 模型,云南省、河南省、天津市等 6 个区域的最优预测模型为灰色 GM(1,1)模型,山东省、青海省、安徽省等 16 个区域的最优预测模型为 BP 神经网络模型;在流域尺度下,松辽、海河等 5 个流域的年用水量最优预测模型为 ARMA 模型,黄河、淮河和珠江 3 个流域的最优预测模型为灰色 GM(1,1)模型,长江流域的最优预测模型为 BP 神经网络模型;在地理大区尺度下,北方 6 区的年用水量最优预测模型为 BP 神经网络模型,南方 4 区的最优预测模型为灰色 GM(1,1)模型。

(3)应用选出的最优模型,预测出 2021–2025 年各省级行政区的用水量总量及其变化趋势,预测期

内各省级行政区用水量总量总体保持稳定。

参考文献:

- [1] 张绍强,沈莹莹.新形势下中国农业用水量统计调查工作现状及面临的问题[J].中国农村水利水电,2022(3):197–199.
- [2] 龙秋波,王绍兵.湖南省用水量统计口径分类与抽样调查方法研究[J].湖南水利水电,2019(2):27–30.
- [3] 龙秋波,贾绍凤,汪党献.中国用水数据统计差异分析[J].资源科学,2016,38(2):248–254.
- [4] VELASCO L C, GRANADOS A, ORTEGA J, et al. Performance analysis of artificial neural networks training algorithms and transfer functions for medium-term water consumption forecasting[J]. International Journal of Advanced Computer Science and Applications, 2018, 9(4): 109–116.
- [5] 宋丽娜,刘森,秦韬,等.基于 LOF 与 CEEMD 的城镇取用水监测数据异常值识别[J].水利信息化,2022(2):33–40.
- [6] 刘哲,张杨.2004–2018 年陕西省用水量及用水结构变化[J].农技服务,2021,38(5):65–66.
- [7] 张标,刘秀丽.我国用水量变动影响因素的结构分解分析[J].管理评论,2015,27(5):3–8+18.
- [8] 吕王勇,赵凌,陈东.基于主成分分析的四川省用水量预测[J].水资源与水工程学报,2009,20(6):84–87.
- [9] 王圃,王以知,张晋,等.基于用水量驱动因子的水量预测模型[J].排灌机械工程学报,2014,32(12):1051–1056.
- [10] 刘晶,鲍振鑫,刘翠善,等.近 20 年中国内地水资源及用水量变化规律与成因分析[J].水利水运工程学报,2019(4):31–41.
- [11] 赵勇,李海红,刘寒青,等.增长的规律:中国内地用水极值预测[J].水利学报,2021,52(2):129–141.
- [12] 刘真,吕同斌.灰色预测模型在合肥市年用水量预测中的应用[J].安徽水利水电职业技术学院学报,2019,19(1):18–20.
- [13] 何靖,王春林,董起广,等.基于灰色系统 GM(1,1)模型的陕西省 2021–2030 年需水量预测[J].农技服务,2021,38(1):118–121.
- [14] 刘献,袁旦,张小丽,等.基于残差灰色-马尔可夫链的生活用水量预测研究[J].人民珠江,2020,41(8):1–6.
- [15] ZHANG Xianqi, ZHAO Dong, WANG Tao, et al. Industrial water consumption forecasting based on combined CEEMD-ARMA model for Henan province, central chain: A case study[J]. Environmental Monitoring and Assessment, 2022, 194: 471.
- [16] 窦森,李金燕,崔岚博,等.相关性分析-神经网络模型在宁夏用水量预测中的应用[J].人民珠江,2022,43(8):71–77.
- [17] DU Hongyan, ZHAO Zhihua, XUE Huifeng. ARMA-

- M: A new model for daily water consumption prediction based on the autoregressive integrated moving average model and the Markov chain error correction[J]. *Water*, 2020, 12(3): 760.
- [18] WANG Zhaocai, WU Xian, WANG Huifang, et al. Prediction and analysis of domestic water consumption based on optimized grey and Markov model[J]. *Water Supply*, 2021, 21(7): 3887–3899.
- [19] YUAN Yanbin, ZHAO Hao, YUAN Xiaohui, et al. Application of fractional order-based grey power model in water consumption prediction[J]. *Environmental Earth Sciences*, 2019(8): 266.
- [20] YUAN Yanbin, LI Qian, YUAN Xiaohui, et al. A SAF-SA-and metabolism-based nonlinear grey bernoulli model for annual water consumption prediction[J]. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 2020, 44: 755–765.
- [21] 陈嘉彤,温立书,谭雅心. 基于灰色预测和 Elman 神经网络的全国用水量预测[J]. *江西科学*, 2018, 36(6): 961–967.
- [22] LIU Xingpo, ZHANG Yiqing, ZHANG Qichen. Comparison of EEMD – ARMA, EEMD – BP and EEMD – SVM algorithms for predicting the hourly urban water consumption[J]. *Journal of Hydroinformatics*, 2022, 24(3): 535–558.
- [23] 白鹏,龙秋波. 3种用水总量预测方法在京津冀地区的适用性比较[J]. *水资源保护*, 2017, 37(2): 102–107.
- [24] 杨建强,王启优,孙栋元,等. 基于不同模型的用水总量预测研究[J]. *水利规划与设计*, 2021(8): 29–34.
- [25] FARAH E, ABDALLAH A, SHAHROUR I. Prediction of water consumption using Artificial Neural Networks modeling (ANN)[C]// *MATEC Web of Conferences*, 2019.
- [26] VELASCO L C P, GRANADOS A R B, ORTEGA J M A, et al. Performance analysis of artificial neural networks training algorithms and transfer functions for medium-term water consumption forecasting[J]. *International Journal of Advanced Computer Science and Applications*, 2018, 9(4): 109–116.
- [27] GHARABAGHI S, STAHL E, BONAKDARI H. Integrated nonlinear daily water demand forecast model (case study: City of Guelph, Canada)[J]. *Journal of Hydrology*, 2019, 579: 124182.
- [28] ANTUNES A, ANDRADE – CAMPOS A, SARDINHA – LOURENCO A, et al. Short-term water demand forecasting using machine learning techniques [J]. *Journal of Hydroinformatics*, 2018, 20(6): 1343–1366.
- [29] 王文圣,金菊良,丁晶. 随机水文学(第3版)[M]. 北京:中国水利水电出版社, 2016.
- [30] MENG Xiangmei, WU Lifeng. Prediction of per capita water consumption for 31 regions in China[J]. *Environmental Science and Pollution Research*, 2021, 28: 29253–29264.
- [31] MENG Xiangmei, TU Leping, YAN Chen, et al. Forecast of annual water consumption in 31 regions of China considering GDP and population[J]. *Sustainable Production and Consumption*, 2021, 27: 713–736.
- [32] 杨舒媛,王军,魏保义,等. 城市需水预测方法综述及应用调研[J]. *北京规划建设*, 2016(4): 66–68.

(上接第110页)

- [16] 赵玲玲,夏军,杨芳,等. 粤港澳大湾区水生态修复及展望[J]. *生态学*, 2021, 41(12): 5054–5065.
- [17] 胡培,黄本胜,邱静,等. 粤港澳大湾区发展背景下珠海市水资源保障战略研究[J]. *广东水利水电*, 2020(7): 1–5+37.
- [18] 于洋,卢然超,李迎霞. 澳门水资源现状与展望[J]. *给水排水*, 2014, 50(2): 30–33.
- [19] 申子通,贺新春,郑久瑜. 珠澳供水安全现状及对策建议[J]. *广东水利水电*, 2015(11): 31–33.
- [20] 李骏飞,杨磊三,孟凡松. 粤港澳大湾区面临的水问题探析[J]. *中国给水排水*, 2019, 35(18): 10–14.
- [21] 王涛,李治军,姚蓉,等. 基于模糊层次分析法安庆市水资源承载力的评价[J]. *陕西水利*, 2021(8): 25–28+31.
- [22] 刘熙娟,于德珠,刘云. 基于 AHP 的阿拉尔市水资源系统评价研究[J]. *地下水*, 2021, 43(4): 185–188+245.
- [23] 何光洪,梁燕,邓炎倩,等. 基于层次分析法的泾阳南塬黄土滑坡影响因素综合研究[J]. *水利与建筑工程学报*, 2021, 19(2): 66–70+148.
- [24] 林龙圳,李达,林震. 基于熵权 – TOPSIS 模型的库布齐沙漠地区水资源承载力评价[J]. *华中师范大学学报(自然科学版)*, 2020, 54(4): 640–648.
- [25] 薛冰,宋新山,严登华. 基于系统动力学的天津市水资源模拟及预测[J]. *南水北调与水利科技*, 2011, 9(6): 43–47.
- [26] 王翠,杨广,何新林,等. 基于系统动力学的水资源承载力研究[J]. *中国农村水利水电*, 2016(9): 212–215+220.
- [27] 左其亭,张志卓,吴滨滨. 基于组合权重 TOPSIS 模型的黄河流域九省区水资源承载力评价[J]. *水资源保护*, 2020, 36(2): 1–7.