

基于 ELM 模型的浅层地下水位埋深时空分布预测

喻黎明¹ 严为光² 龚道枝^{3,4} 李沅媛^{3,4} 冯禹^{3,4} 姜丹曦⁵

(1. 昆明理工大学现代农业工程学院, 昆明 650500; 2. 长沙理工大学水利工程学院, 长沙 410114;

3. 中国农业科学院农业环境与可持续发展研究所, 北京 100081;

4. 作物高效用水与抗灾减损国家工程实验室, 北京 100081;

5. 宁乡县水利水电勘测设计院, 长沙 410004)

摘要: 选用石家庄平原区补排因子的多种组合为输入参数, 利用 28 眼水井的实测资料作为预测目标值, 首次建立基于极限学习机 (Extreme learning machine, ELM) 的地下水位埋深时空分布预测模型, 讨论补排因子在不同缺失情况下对模型精度的影响; 利用 ArcGIS 分析误差空间分布趋势, 并与常用的三隐层 BP 神经网络模型进行对比。结果表明: 基于水均衡理论的 ELM 地下水位埋深模拟模型能够准确反映人类和自然双重影响下地下水系统的非线性关系, 模型输入因子中缺失降水量或开采量的模拟结果均方根误差 (RMSE) 比缺失其余因子的 RMSE 高 2.00 倍及以上, 同时模型有效系数 (E_{ns}) 和决定系数 (R^2) 进一步降低; 与 BP 模型相比, ELM 模型可使 RMSE 减小 43.6%, 误差区间降低 46.4%, E_{ns} 和 R^2 提高至 0.99, 且 RMSE 在空间相同区域上均明显呈现出 ELM 模型小于 BP 模型; ELM 模型在南部高误差区的移植精度 (RMSE 低于 1.82 m/a, E_{ns} 高于 0.95) 高于 BP 模型 (RMSE 超过 3.00 m/a, E_{ns} 低于 0.85); 因此, 影响地下水位埋深的主导因素是降水量和开采量, 且 ELM 模型在精度、稳定性和空间均匀性上较优, 移植预测效果较好, 可利用已知资料推求区域空间内其余未知水井的浅层地下水位埋深; 该模型可作为水文地质参数及补排资料缺乏条件下浅层地下水位埋深预测的推荐模型。

关键词: 地下水位; 时空分布; 神经网络; 极限学习机; 预测

中图分类号: P641 文献标识码: A 文章编号: 1000-1298(2017)02-0215-09

Temporal and Spatial Distribution Prediction of Shallow Groundwater Level Based on ELM Model

YU Liming¹ YAN Weiguang² GONG Daozhi^{3,4} LI Yuanyuan^{3,4} FENG Yu^{3,4} JIANG Danxi⁵

(1. Faculty of Modern Agricultural Engineering, Kunming University of Science and Technology, Kunming 650500, China

2. School of Hydraulic Engineering, Changsha University of Science and Technology, Changsha 410114, China

3. Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China

4. State Key Engineering Laboratory of Crops Efficient Water Use and Drought Mitigation, Beijing 100081, China

5. Ningxiang Hydro and Power Design Institute, Changsha 410004, China)

Abstract: In order to achieve high-precision prediction of temporal and spatial distribution of the groundwater level in shallow groundwater cones region, a model was constructed firstly based on extreme learning machine (ELM). By choosing different combination factors of groundwater recharge and discharge as the input parameters of model and observing data of 28 wells as predicted target in Shijiazhuang plain, the error of spatial distribution trend was analyzed by using ArcGIS software. The results showed that the ELM model based on the water balance theory could accurately reflect the non-linear relationship of groundwater system under the influence of human and nature activity. The root mean square error (RMSE) of model under the condition without exploitation or precipitation as input factor was two times higher than that under the condition without other factors, and the coefficient of efficiency

收稿日期: 2016-10-24 修回日期: 2016-12-13

基金项目: 中国农业科学院“华北节水保粮协同创新行动”项目 (CAAS-XTCX2016019)、国家自然科学基金项目 (51379024)、中央高校基本科研业务费项目 (51679243) 和“十二五”国家科技支撑计划项目 (2012BAD09B01、2015BAD24B01)

作者简介: 喻黎明 (1976—), 男, 副教授, 博士, 主要从事节水灌溉理论研究, E-mail: liming16900@163.com

通信作者: 龚道枝 (1976—), 男, 副研究员, 博士, 主要从事农业节水研究, E-mail: gongdaozi@caas.cn

(E_{ns}) and coefficient of determination (R^2) were further reduced. Compared with the BP model, the RMSE of ELM model was reduced by 43.6%, and the scope of error was reduced by 46.4%. E_{ns} and R^2 were improved to 0.99. The tendency of error distribution showed that it was decreased from the south and southeast to the central. The RMSE of ELM model was obviously lower than that of BP model in all the regions. The accuracy of ELM model (RMSE was less than 1.82 m, E_{ns} was higher than 0.95) was higher than that of BP model (RMSE was more than 3.00 m, E_{ns} was less than 0.85) in southern high error region. Therefore, exploitation and precipitation were the main impact factors on the groundwater dynamic in the model. Meanwhile, the stability, accuracy and space uniformity of ELM model were better than those of BP model. And the transplantation results of ELM model were more satisfactory. The model could be used to forecast groundwater level of other unknown wells based on given data. Therefore, the ELM model could be used as a recommended model for predicting groundwater level under conditions of missing hydrogeological and groundwater recharge data.

Key words: groundwater level; temporal and spatial distribution; neural network; extreme learning machine; prediction

引言

由于多年来不合理的地下水开采和管理导致许多地区含水层开始疏干,地下水流动场发生异变,形成地下水漏斗,并衍生了严重的生态环境问题。准确可靠的地下水水位预测是规划和实施流域地下水和地表水资源科学管理的重要组成部分^[1]。相比于深层地下水,浅层地下水(潜水或微承压水)对自然和人为影响更为敏感,径、补、排条件更加复杂,使得浅层地下水动态的预测变得极其困难。

人工神经网络(Artificial neural network, ANN)因其在解决复杂非线性系统问题上的良好表现使其广泛用于水文预测^[2-5],且在美国土木工程师协会(ASCE)专家委员会报告中有大量讨论^[6-7]。LALLAHEM等^[8]、SREEKANTH等^[9]和霍再林等^[10]分别在不同地区将ANN用于地下水位的预测,验证了不同ANN模型模拟地下水位的可靠性。COPPOLA等^[11]和MOHAMMADI^[12]研究表明ANN模型对地下水位埋深的预测精度优于数值模拟。YANG等^[13]使用BP人工神经网络模型(Back-propagation artificial neural network)模拟吉林地下水位埋深,验证了ANN优于综合时间序列模型(ITS)。

随着ANN模型的发展,BP模型因其较强的自学习和自适应能力,使其在地下水位埋深预测中成为运用最为广泛的模型^[10,13-15],但其存在训练速度慢、初始权值和阈值选择敏感程度高、很难求得全局最优解等固有缺点。极限学习机算法(Extreme learning machine, ELM)作为一种新型的单隐层前馈神经网络学习算法,在算法执行过程中不需要调整网络的输入权值以及隐元的偏置,并可获得全局最优解,且具有很强的泛化性^[16-17],较好地克服了BP模型的缺点。另外,输入因子的选择对模型模拟结果有着重大影响^[14],而目前地下水位埋深的

ANN模型研究多是直接选取以目标测井前几个时间段的地下水位埋深,或者以人为和自然影响的某一特定因素作为模型的输入因子,并没有基于水均衡原理对各因子的影响程度进行较为全面的讨论。

石家庄平原区作为我国地下水超采最严重的典型浅层地下水漏斗区,同时也是我国重要的粮食高产区,地下水水位变化受农业影响较大,具有复杂的人为影响效应,有着较强的代表性。因此,本文建立基于水均衡理论的石家庄平原区浅层地下水位埋深ELM时空预报模型,研究各因子对模型精度的影响程度,并与三隐层BP模型进行比较,提出水文地质参数及补排资料缺乏条件下浅层地下水位埋深的高精度模拟模型,以期对地下水管理提供科学依据。

1 研究区概况与神经网络模型

1.1 研究区域概况与数据资料来源

石家庄平原区(图1)为滹沱河流域,属太行山前冲洪积平原,位于东经 $114^{\circ}18'$ ~ $115^{\circ}30'$ 、北纬 $37^{\circ}30'$ ~ $38^{\circ}40'$ 之间,总面积为 $6\,976.4\text{ km}^2$,是华北平原中形成最早发展最快的浅层地下水漏斗区。年平均气温 13.4°C ,多年平均蒸发量 $1\,616.6\text{ mm}$,多年平均降水量 534.5 mm ,其中6—9月份降水量约占全年降水量的75.0%,且空间分布极其不均。区内分布有沙河、磁河、槐河和滹沱河,主要渠道为石津渠,水利工程中影响最大的有岗南水库和黄壁庄水库。研究区自上而下可划分为4个含水组,其第1和第2含水组之间无连续隔水层,加之多年混合开采将其视为统一含水层,统称为浅层地下水,地下水水力性质属潜水—微承压水类型。浅层地下水系统底板埋深在 $40\sim 60\text{ m}$ 之间,表层多为亚砂土、黏土,下部岩性较粗,含水层岩性以卵石、卵砾石、粗砂、中砂为主^[18]。

本研究中地下水位埋深、滹沱河流量、石津渠流

量、黄壁庄水库水位资料均由河北省水文水资源勘测局提供，地下水开采量和补排量数据来自文献[19]，灌溉水量来源于《河北省水资源公报》，降水资料来源于中国气象科学数据共享服务网。

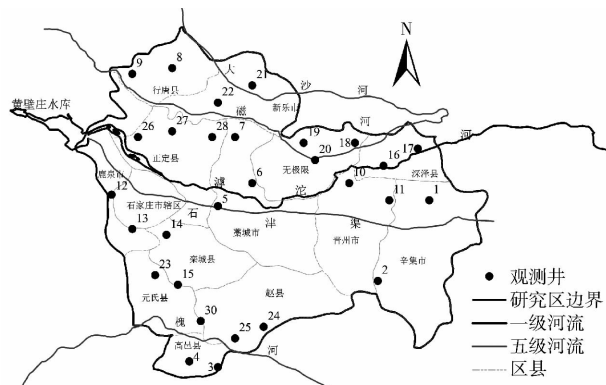


图 1 研究区域与观测井位置

Fig. 1 Study area and location of observation wells

1.2 地下水补排条件

石家庄平原区地下水补给项包括降水、河渠、灌溉及侧向补给，其中降水入渗为该区主要补给源超过总补给量的 50.0%；区域内滹沱河为最主要也是

最大的河流，河床岩性结构简单，为连续厚层砂、卵砾石，局部夹薄层或透镜体状粘性土，垂向连通性好，是重要的河道渗漏补给来源，除 1996 年供水期外河道行洪量小或基本断流，2006—2010 年河道补给量仅为 0.27 亿 m³/a；石津渠横穿石家庄市、正定县和藁城县等 4 县 1 市，全长 134.00 km，为石家庄最大渠道，输水能力达 100 m³/s，除市区段外没有进行防渗措施，为主要渠系渗漏补给源，随着石津渠流量由 1976 年之前的 12.37 亿 m³/a 降至 2000 年以来的 4.47 亿 m³/a，渠系渗漏补给量一直维持在较低水平。黄壁庄水库的渗漏量为研究区重要的侧向补给来源，2001 年黄壁庄水库副坝完成加固防渗工程后，造成水库渗漏补给量减少 57.5%，是近年来侧向补给减少的主要原因。根据前人研究结果，在地下水位大于 4 m 条件下，潜水蒸发基本为零，而本研究区自 1975 年以来地下水位埋深小于 4 m 的分区基本已经全部消失^[19]，加之多年侧向流出量只有人工开采量的 6.0%，可忽略不计，故主要排泄项为人工开采。1991—2010 年各项补排情况见表 1。

表 1 石家庄平原区 1991—2010 年地下水补排状况

Tab. 1 Groundwater recharge and drainage in Shijiazhuang plain from 1991 to 2010

类别	时段累积量/亿 m ³				总计/ 亿 m ³	占多年总补给、 排泄量比率/%	
	1991—1995 年	1996—2000 年	2001—2005 年	2006—2010 年			
补给项	降水入渗	37.90	42.40	40.37	41.10	161.77	51.0
	侧向流入	18.37	16.19	10.26	9.91	54.73	17.2
	河道渗漏	6.40	17.11	2.48	1.36	26.99	8.5
	渠系渗漏	5.73	7.75	2.08	3.07	18.63	5.9
	灌溉回归	12.44	16.38	13.50	13.01	55.33	17.4
排泄项	侧向流出	6.02	3.67	5.58	10.14	25.41	5.8
	人工开采	108.60	108.70	102.60	90.38	410.28	94.2

1.3 极限学习机模型

ELM 为典型的单隐层前馈神经网络，其在训练之前可随机产生输入层和隐含层间的连接权值和隐含层神经元间的阈值，使之在训练过程中无需再进行调整，只需设置网络的隐含层节点个数及隐含层神经元的激活函数即可产生唯一的最优解，有效地克服了大多数梯度下降法训练速度慢、学习率选择敏感和易陷入局部极小等缺点。ELM 的结构如图 2 所示，其中输入层、隐含层和输出层神经元个数分别为 n 、 l 和 m 。设 x_n 为输入样本，则 t_m 为对应的输出样本， ω_{ln} 为输入层与隐含层的连接权值， β_{lm} 为隐含层与输出层连接权值， b_l 为隐含层神经元阈值， $g(x)$ 为激活函数，单隐层神经网络 (SLFN) 的数学模型为

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix}_{m \times 1} = \begin{bmatrix} \sum_{i=1}^l \beta_{i1} g(\omega_i x_j + b_i) \\ \sum_{i=1}^l \beta_{i2} g(\omega_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^l \beta_{im} g(\omega_i x_j + b_i) \end{bmatrix}_{m \times 1} \quad (j = 1, 2, \dots, D) \quad (1)$$

将式(1)转换为

$$H\beta = T' \quad (2)$$

其中 T' 为 T 的转置， $T = [t_1, t_2, \dots, t_n]_{m \times D}$ ； H 为隐含层输出矩阵，其具体形式为

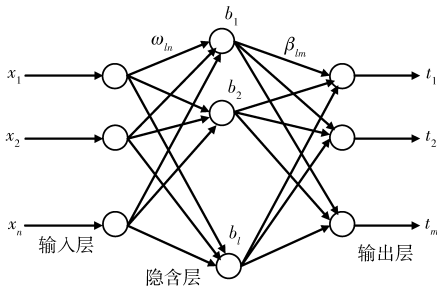


图2 极限学习机网络结构图

Fig.2 Topological structure diagram of extreme learning machine

learning machine

$$\mathbf{H} =$$

$$\begin{bmatrix} g(\omega_1 x_1 + b_1) & g(\omega_2 x_1 + b_2) & \cdots & g(\omega_l x_1 + b_l) \\ g(\omega_1 x_2 + b_1) & g(\omega_2 x_2 + b_2) & \cdots & g(\omega_l x_2 + b_l) \\ \vdots & \vdots & & \vdots \\ g(\omega_1 x_D + b_1) & g(\omega_2 x_D + b_2) & \cdots & g(\omega_l x_D + b_l) \end{bmatrix}_{D \times l}$$

(3)

根据 HUANG 等^[16]提出的定理:当隐含层神经元个数 l 小于训练集样本个数 D 时,SLFN 的训练误差可逼近任意 $\varepsilon > 0$ 。因此,当激活函数 $g(x)$ 无限可微时, ω_{in} 和 b_i 在训练前可随机选择,并保持不变。而隐含层与输出层的连接权值 β_{in} 可通过以下方程组的最小二乘解获得

$$\min_{\beta} \|\mathbf{H}\beta - \mathbf{T}'\| \quad (4)$$

其解为

$$\hat{\beta} = \mathbf{H}^+ \mathbf{T}' \quad (5)$$

其中 \mathbf{H}^+ 为隐含层输出矩阵 \mathbf{H} 的 Moore-Penrose 广义逆。

1.4 BP 神经网络模型

BP 神经网络与 ELM 的不同在于 BP 方法采用梯度下降法重复迭代计算调整权值和阈值,使网络输出值与实际样本值的误差平方和不断减小直至达到预期目标。主要步骤为:

(1)网络初始化。确定网络输入层、隐含层和输出层节点数,并初始化连接权值,初始化隐含层阈值和输出层阈值,给定学习速率和神经元激励函数。

(2)隐含层输出计算。根据输入变量 \mathbf{X} ,输入各权值和阈值,计算隐含层输出矩阵。

BP 神经网络的详细原理与实现过程可参照文献[20]。

2 预测模型建立

2.1 输入因子选择

基于水均衡方程可推导出任意水井的地下水位埋深为

$$h(x, y|t) = \frac{\hat{Q}(t) - Q(t)}{\mu\alpha F} + h(x, y|(t-1)) \quad (6)$$

式中 $h(x, y)$ ——经、纬度为 x, y 处水井地下水位埋深

$\hat{Q}(t)$ —— Δt 时段补给量, m^3

$Q(t)$ —— Δt 时段排泄量, m^3

μ ——各透水介质综合弹性释水系数

α ——区域埋深变幅的差异系数

F ——均衡区面积, m^2

简化后可得

$$h(x, y|t) = f(\hat{Q}(t), Q(t), h(x, y|(t-1)), x, y) \quad (7)$$

根据式(6)结合前文关于研究区补排条件的分析,本研究选取以年为 Δt 时段:研究区年平均总降水量 $p(t)$ 、滹沱河年平均流量 $q_1(t)$ 、石津渠年平均流量 $q_2(t)$ 、黄壁庄水库年平均水位与研究区浅层地下水位埋深差值 $w(t)$ 、研究区年总灌溉水量 $I(t)$,分别反映各项补给的物理量以作为补给项因子;研究区年开采量 $Q(t)$ 代表排泄项因子;水井的经、纬度 x, y 代表方位因子,加上水井上一年地下水位埋深 $h(t-1)$,共 9 个时间序列为输入因子。其中,因侧向补给主要受水力梯度与渗透系数影响,根据达西定律可知,当渗透系数改变时可将这一变化系数反映到水力梯度上,得到渗流速度的物理量不变。因此,为反映黄壁庄水库于 2001 年防渗加固后使水库渗漏补给量减少 57.5% 这一突变影响,将 2001 年后的 $w(t)$ 值进行 42.5% 的折算以反映物理变量。

2.2 训练样本及模型结构选取

采用研究区 28 眼浅层地下水观测井 1990—2015 年浅层地下水位埋深资料,其位置和编号见图 1。即选择 1~28 号水井 1991—2010 年资料为训练样本,2011—2015 年资料为检测样本,输入向量为 $\mathbf{X} = (p(t), q_1(t), q_2(t), w(t), I(t), Q(t), x, y, h(t-1))$,输出样本为对应地下水位埋深 $\mathbf{Y} = [h(t)]$ 。文献[21]表明三隐层 BP 神经网络能够逼近任何有理函数,使用 Matlab 2013a 分别建立 ELM 和 BP 神经网络模型,采用试错法分别确定 BP 神经网络结构为 9-10-10-5-1,激励函数为 tansig,学习率为 0.01;ELM 模型的神经元个数为 50,输入层和输出层节点数与 BP 模型一致,激励函数为 sig。

2.3 模型精度评估指标及其计算方法

采用决定系数 (R^2)、均方根误差 (Root mean square error, RMSE) 和模型有效系数 (E_{ns}) 评价各模型计算结果与实际观测值之间的误差及拟合程度,计算公式分别为

$$R^2 = \frac{\left[\sum_{i=1}^m (X_i - \bar{X})(Y_i - \bar{Y}) \right]^2}{\sum_{i=1}^m (X_i - \bar{X})^2 \sum_{i=1}^m (Y_i - \bar{Y})^2} \quad (8)$$

$$R_{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (Y_i - X_i)^2} \quad (9)$$

$$E_{ns} = 1 - \frac{\sum_{i=1}^m (Y_i - X_i)^2}{\sum_{i=1}^m (X_i - \bar{X})^2} \quad (10)$$

式中 m ——数据样本数

X_i ——实际观测的地下水位埋深

Y_i ——对应模型模拟值

\bar{Y} —— Y_i 的平均值

\bar{X} —— X_i 的平均值

R_{RMSE} ——均方根误差

R^2 越接近于 1,表明模型吻合程度越高; E_{ns} 越接近于 1,表明模型效率越高;RMSE 越小即模型偏差越小。

3 结果与分析

3.1 补排资料缺失条件下 ELM 精度分析

在 2.2 节确定的模型结构基础上,调整不同组合的补排因子作为输入参数建立 8 个对应的 ELM 模型,计算时不输入各模型对应的缺失因子(此处 x 、 y 和 $h(t-1)$ 视为自身因子,不参与讨论),模拟结果如表 2 所示。当输入参数为全部 6 个补排因子时,ELM1 模拟的拟合 RMSE 为 1.07 m/a, R^2 和 E_{ns} 达到 0.99 ($P < 0.01$),达极显著水平。因此 ELM1

模型对地下水位埋深的模拟可靠性较高,能够反映浅层地下水与人为、自然因素之间的非线性关系。当缺乏 I 和 w 情况下模型 (ELM2 和 ELM3) 的 RMSE 比输入全部 6 个补排因子时 (ELM1) 分别增加 0.44 m/a 和 0.62 m/a,当缺乏河 (q_1)、渠 (q_2) 补给资料时 (ELM6) 模型模拟的 RMSE 增加幅度达到 0.83 m/a,与 ELM1 相比,这 3 种情况下的 E_{ns} 和 R^2 均略有降低。然而,当模型输入因子中缺乏 Q 和 p 时 (ELM4、ELM5),模型 RMSE 的增幅较之 ELM2、ELM3、ELM6 明显加大,其幅度达到 2 倍及以上,同时 E_{ns} 和 R^2 有进一步的降低,且 ELM5 的 E_{ns} 未达到显著水平 ($P > 0.05$),说明降水量和开采量这 2 个因子的缺失使模型精度下降的幅度是所有补排因子中最大的。

为进一步探究降水量和开采量对模型精度的影响,在输入补给因子仅有降水量和开采量的 ELM7 中,模型 RMSE 比由 4 个输入因子时的 ELM6 模型低 0.07 m/a,尤其是决定系数 R^2 达到 0.99 ($P < 0.01$),达极显著水平。说明增加补给因子的输入不一定提高模拟精度,而输入参数对模拟精度的影响主要取决于输入参数对输出项的影响水平与贡献率^[22]。此外,当输入参数仅为自身因子时,ELM8 模拟精度过低,而 ELM7 可使 RMSE 减小 77.1%, E_{ns} 提高 55.4%,大幅提升了模拟精度。

综上所述,降水量和开采量是影响模型模拟地下水位埋深的主导因子,且降水量的贡献率要大于开采量,因此准确收集区域内这 2 个因子资料能大幅提高预测结果的精度。

表 2 不同补排因子组合下 ELM 模型的浅层地下水位埋深模拟精度

Tab. 2 Statistic performance evaluation criteria of ELM model with different input factors

ELM 模型	输入的补排因子个数	输入补排因子	缺失的补排因子	RMSE/(m·a ⁻¹)	E_{ns}	R^2
ELM1	6	q_1, q_2, w, p, Q, I		1.07	0.99	0.99
ELM2	5	q_1, q_2, w, p, Q	I	1.51	0.98	0.98
ELM3	5	I, q_1, q_2, p, Q	w	1.69	0.97	0.98
ELM4	5	I, q_1, q_2, w, p	Q	2.95	0.92	0.94
ELM5	5	I, q_1, q_2, w, Q	p	3.49	0.89	0.93
ELM6	4	I, w, p, Q	q_1, q_2	1.90	0.97	0.97
ELM7	2	p, Q	I, q_1, q_2, w	1.83	0.97	0.99
ELM8	0		I, q_1, q_2, w, p, Q	7.98	0.43	0.80

3.2 ELM、BP 模型模拟值与测井实测值对比分析

采用反距离加权插值法,使用 ArcGIS 10.2 软件,绘制 ELM、BP 模型模拟值与 28 眼测井实测值之间的 RMSE 空间分布图(图 3),RMSE 空间变化呈现出南部较高、自中部逐渐降低的趋势。由于降水是地下水位埋深动态变化的主导因素,根据韩军彩

等^[23]基于 17 个测站 36a 资料研究表明石家庄市东南部(即上述南部、东南部)年降水天数较中西部明显减少,而暴雨天数增加,两模型却使用平均降水量来反映降水渗透补给并作为模型的输入因子,从而导致两模型 RMSE 呈此趋势。由于 ELM 模型本身具有的优点克服了 BP 模型采用梯度下降法使之容

易陷入局部极小的缺点,使得图3显示相同区域上明显呈现出ELM模型的RMSE小于BP模型,表明

空间上ELM模型稳定性较好,且模拟精度高于BP模型。

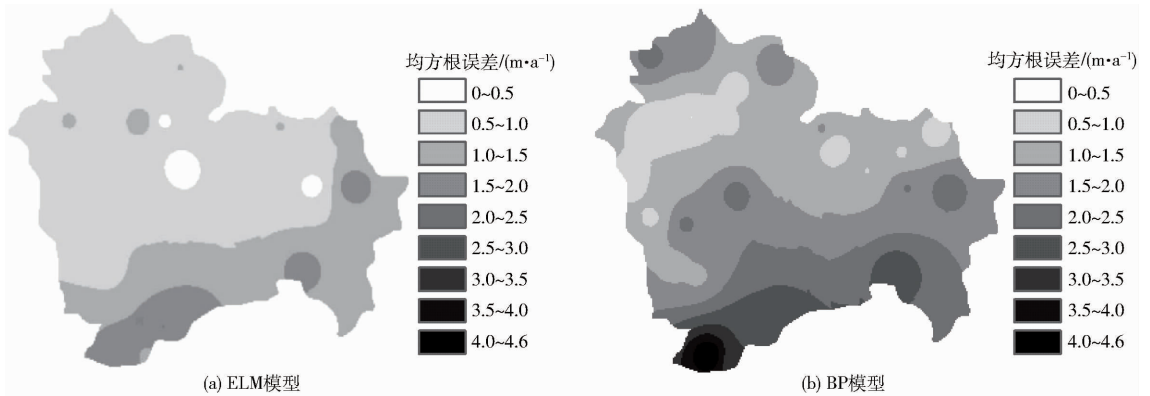


图3 ELM和BP模型的RMSE空间分布

Fig.3 Spatial distributions of root mean square error of ELM and BP models

ELM和BP模型模拟28个测井结果与实测结果整体分析对照如表3所示,相比BP模型,ELM模型可使RMSE减小43.6%,而 R^2 和 E_{ns} 提高至0.99($P < 0.01$),达到极显著水平。比较ELM、BP模型的 e_{max} 、 e_{min} (最大、最小绝对误差)可知,两模型的 e_{min} 比较接近,而 e_{max} 差别较大,后者是前者的1.87倍,由此可知ELM模型的误差区间(最大和最小绝对误差的差值)为3.63,BP模型为6.78,即前者的误差区间较之后者减小了46.4%。同样,两模型的 δ_{min} (最大相对误差,其中相对误差 δ 是指绝对误差

值与其对应井的多年(1990—2015年)埋深变幅的比值)接近,而 δ_{max} (最大相对误差)差距较大,前者为29.6%,远低于后者的83.9%,说明ELM模型稳定性要优于BP模型。参照文献[24],ELM和BP模型模拟结果都能满足要求,但ELM模型有90.0%的样本 $\delta < 10.0%$,63.6%的样本 $\delta < 5.0%$;而BP模型 $\delta < 10.0%$ 的样本只有70.0%, $\delta < 5.0%$ 的样本仅占45.7%,ELM模型的高精度样本数($\delta < 5.0%$ 的样本数)较BP模型多17.9%,很明显,ELM模型模拟精度显著高于BP模型。

表3 ELM和BP模型模拟结果对比分析

Tab.3 Comparison of statistic performance evaluation criteria of ELM and BP models

模型	e_{max}/m	e_{min}/m	$\delta_{max}/\%$	$\delta_{min}/\%$	RMSE/($m \cdot a^{-1}$)	E_{ns}	R^2
ELM	3.63	≈ 0	29.59	≈ 0	1.07	0.99	0.99
BP	6.78	≈ 0	83.93	≈ 0	1.90	0.97	0.97

3.3 典型测井移植案例分析

选用误差分布较大的南部和东南部的2、4、24、25号4眼观测井1991—2015年序列资料共100组为移植样本对模型精度进行检测,且均不参与模型训练。其余所有观测井的序列共600组为训练样本,得到模型的移植结果见图4。由图4可知,BP模型对2号井的模拟结果与实测值比较接近,而对4、24、25号井的模拟误差较大,特别是对24、25号井(高误差区)模拟的最大绝对误差均超过10.00 m, RMSE超过3.00 m/a, E_{ns} 均低于0.85($P > 0.05$)远未达到显著水平,模拟的动态曲线波动剧烈,且整体上明显低估了实际埋深。ELM模型对4眼井的模拟误差较小, RMSE均低于1.82 m/a, E_{ns} 均超过0.95($P < 0.05$),达显著水平,且与地下水埋深的实际动态拟合程度较高。可见,ELM模型计算过程中无需使用梯度下降法对权值和阈值进行调整,有

效地避免了模型陷入局部极小,使得模型可在一定程度上弱化因输入参数的精确度低所带来的误差,进而达到全局最优化。总之,ELM模型利用已知水井资料对区域内其余未知水井的浅层地下水埋深动态的模拟效果良好,其移植性能优于BP神经网络模型。

4 讨论

本研究在对石家庄平原区浅层地下水补排条件变化分析中得到,因石家庄平原区1980年来河道基本干涸,降水逐渐成为地下水唯一重要补给源^[25],且其补给量已经超过总补给量的50.0%,同时排泄项中人工开采量达到总排泄量的90.0%以上。区域内地下水降深与开采量和降水量的关联度分别达到0.71和0.60^[26]。降水量和开采量成为区域场流演变主导因素^[27-28],浅层地下水水位动态成为降水

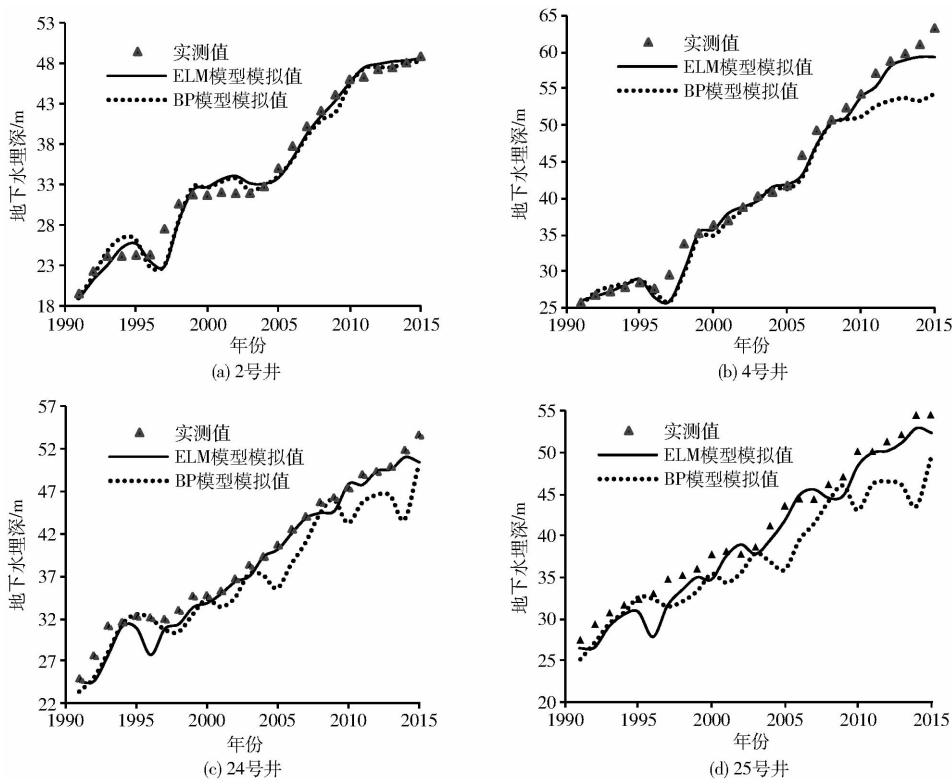


图 4 ELM 和 BP 模型移植模拟结果比较

Fig. 4 Comparisons of simulated results of ELM and BP models

入渗—开采排泄型^[19]。即该区地下水动力环境演化是开采量和降水量双重影响所致^[29]。本研究运用 ELM 对模型主导因子的识别结果符合研究区地下水动力环境演化特征,表明 ELM 模型通过根据调整不同输入参数组合的模拟结果对比,识别的主要驱动因子具有可靠性。

但本研究由 ELM 模型模拟结果判断出地下水埋深的主导因素中,降水量的贡献率要大于开采量,与大多数前人研究的开采量影响程度大于降水量这一观点不一致^[26-28]。分析原因:石家庄平原区是我国重要的粮食产区,地下水以农业开采为主,占开采量的 83.7%^[30],即开采量受农业用水需求的控制,而农业需水量与降水量相关。农业开采量随降水量增加以幂函数特征呈减小趋势^[31]。说明降水量与开采量之间存在有较强耦合作用,对地下水的影响强度则表现在降水量增加可使开采量的影响显著减弱。可能是由于这种耦合关系使得在本模型计算过程中降水量的动态能更好地反应开采量的变化,从而导致本模型判断降水量的影响程度要更大一些。

本研究在 ELM 和 BP 模型的精度对比中 BP 模型虽然有较好的表现,模拟结果符合文献[24]的基本要求,但 BP 模型对参数敏感度较高,使得在模型调试过程中为确定最优学习率及三隐含层的各个节点数,需要反复调整学习率和每层隐含层节点数进

行计算,该过程会花费大量时间。另外,由于 BP 神经网络使用梯度下降法反复迭代来修正权值和阈值容易使模型陷入局部极小,从而使得 BP 模型在本研究的运用中整体精度低、误差区间大、高精度样本数量少。同时,BP 模型对输入参数的精确度敏感,致使高误差区的模拟结果与实际地下水动态相差甚远,影响模型的移植效果。而 ELM 模型对参数敏感度低,且无需设置学习率,无需调整权值和阈值,有效地避免了局部极小问题。因此,ELM 模型参数设定更简单,训练速度更快,模拟结果更优^[32-33]。

由于区域内降水为最大的浅层地下水补给源,且为地下水埋深动态的主导因素,同时降水受太行山焚风、辐射影响,呈现南部和东南部降水天数少且暴雨时日多的分布特征。分析中认为,本研究以平均面降水量作为模型的输入因子,而实际上降水空间分布极其不均,是造成模型模拟的 RMSE 在南部、东南部明显高于其余区域的主要原因。但本研究所使用的资料有限,未能获得区域空间的开采强度分布,而无法确定开采强度对误差分布趋势的影响,因此在后续研究中应结合开采强度分布趋势,进一步分析模型精度与输入因子之间的关系。

5 结论

(1) ELM 模型能够准确地反映人类和自然双重影响下地下水系统的非线性关系,能根据通过调整

不同输入参数组合的模拟结果对比识别出区域内影响地下水位埋深动态的主要驱动因子。补排资料缺失条件下 ELM 模型精度分析表明,与缺失补排因子的其余 ELM 模型相比 ELM4 和 ELM5 的模拟精度相对最低,表明石家庄平原区内开采量和降水量为地下水位埋深动态的主导因子。仅将开采量和降水量作为模型输入参数的 ELM7 也具有较高的精度 (RMSE、 E_{ns} 和 R^2 分别为 1.83 m/a、0.97 和 0.99), 可作为补排资料缺乏条件下石家庄平原区浅层地下水预测的推荐模型。

(2)与常用的三隐层 BP 模型相比,ELM 模型可使 RMSE 减小 43.6%, 误差区间降低 46.4%, 高精度样本增加 17.9%, R^2 和 E_{ns} 提高至 0.99 ($P <$

0.01), 达到极显著水平,且空间上的 RMSE 均明显呈现出 ELM 模型小于 BP 模型,显然 ELM 模型在精度、稳定性和空间均匀性上更优,可利用 ELM 模型较精确地检验未来各开采模式下的地下水响应趋势。

(3)模型可移植性分析发现,ELM 模型对高误差区的模拟精度远高于 BP 模型,表明 ELM 模型在利用石家庄平原区其余部分水井资料预测南部和东南部的地下水位埋深时可在一定程度上弱化因输入参数的精确度低所带来的误差,能够较好地表达高误差区地下水的非线性动态特征,并取得较高的模拟精度。因此,ELM 模型可在区域水井缺测情况下使用区域内相近站点资料进行较高精度的模拟预测。

参 考 文 献

- 1 MOHANTY S, MADAN K, JHA A, et al. Artificial neural network modeling for groundwater level forecasting in a river island of eastern India[J]. *Water Resources Management*, 2010, 24(9): 1845 - 1865.
- 2 DALIAKOPOULOS I N, COULIBALY P, TSANIS I K. Groundwater level forecasting using artificial neural networks[J]. *Journal of Hydrology*, 2005, 309(1 - 4): 229 - 240.
- 3 AQIL M, KITA I, YANO A, et al. Analysis and prediction of flow from local source in a river basin using a neuro-fuzzy modeling tool[J]. *Journal of Environmental Management*, 2007, 85(1): 215 - 223.
- 4 JAIN A, SRINIVASULU S. Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques[J]. *Water Resources Research*, 2004, 40(4): W04302.
- 5 COULIBALY P, ANCTIL F, ARAVENA R, et al. Artificial neural network modeling of water table depth fluctuations[J]. *JSAE Review*, 2001, 37(4): 885 - 896.
- 6 GOVINDARAJU R S. Artificial neural networks in hydrology. I: preliminary concepts[J]. *Journal of Hydrologic Engineering*, 2000, 5(2): 115 - 123.
- 7 GOVINDARAJU R S. Artificial neural networks in hydrology. II: hydrologic applications[J]. *Journal of Hydrologic Engineering*, 2000, 5(2): 124 - 137.
- 8 LALLAHEM S, MANIA J, HANI A, et al. On the use of neural networks to evaluate groundwater levels in fractured media[J]. *Journal of Hydrology*, 2005, 307(1 - 4): 92 - 111.
- 9 SREEKANTH P D, GEETHANJALI N, SREEDEVI P D, et al. Forecasting groundwater level using artificial neural networks[J]. *Current Science*, 2009, 96(7): 933 - 939.
- 10 霍再林, 冯绍元, 康绍忠, 等. 神经网络与地下水流动数值模型在干旱内陆区地下水位变化分析中的应用[J]. *水利学报*, 2009, 40(6): 724 - 728.
HUO Zailin, FENG Shaoyuan, KANG Shaozhong, et al. Application of ANN and FEFLOW model to simulate groundwater level variation in arid inland area[J]. *Journal of Hydraulic Engineering*, 2009, 40(6): 724 - 728. (in Chinese)
- 11 COPPOLA E, SZIDAROVSKY F, POULTON M, et al. Artificial neural network approach for predicting transient water levels in a multilayered groundwater system under variable state, pumping, and climate conditions[J]. *Journal of Hydrologic Engineering*, 2003, 8(6): 348 - 360.
- 12 MOHAMMADI K. Groundwater table estimation using MODFLOW and artificial neural networks[M]//ABRAHART R J, SEE L M, SOLOMATINE D P. *Practical hydroinformatics*. Springer Berlin Heidelberg, 2008: 127 - 138.
- 13 YANG Z P, LU W X, LONG Y Q, et al. Application and comparison of two prediction models for groundwater levels: a case study in Western Jilin Province, China[J]. *Journal of Arid Environments*, 2009, 73(4 - 5): 487 - 492.
- 14 NAYAK P C, RAO Y R S, SUDHEER K P. Groundwater level forecasting in a shallow aquifer using artificial neural network approach[J]. *Water Resources Management*, 2006, 20(1): 77 - 90.
- 15 张子贤, 袁李明, 李瑞森, 等. 承压水漏斗地区地下水位时空分布预报的 BP 网络模型[J]. *水利学报*, 2007, 38(7): 838 - 844.
ZHANG Zixian, YUAN Deming, LI Ruisen, et al. BP network model for forecasting temporal and spatial distribution of groundwater stage in confined water cones region[J]. *Journal of Hydraulic Engineering*, 2007, 38(7): 838 - 844. (in Chinese)
- 16 HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. *Neurocomputing*, 2006, 70(1 - 3): 489 - 501.
- 17 张淼, 孔盼, 李雁华, 等. 基于极限学习机的土壤硝态氮预测模型研究[J/OL]. *农业机械学报*, 2016, 47(6): 93 - 99.
http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20160612&flag=1. DOI: 10.6041/j.issn.1000-1298.

- 2016.06.012.
- ZHANG Miao, KONG Pan, LI Yanhua, et al. Prediction model of soil NO_3^- -N concentration based on extreme learning machine [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(6): 93–99. (in Chinese)
- 18 刘中培. 农业活动对区域地下水变化影响研究[D]. 北京: 中国地质科学院, 2010.
- LIU Zhongpei. Impact of agricultural activities on regional groundwater variation[D]. Beijing: Chinese Academy of Geological Sciences, 2010. (in Chinese)
- 19 冯慧敏. 石家庄平原区地下水流场演变特征与尺度效应[D]. 北京: 中国地质科学院, 2015.
- FENG Huiming. Groundwater flow field evolution characteristic and size effect of Shijiazhuang plain[D]. Beijing: Chinese Academy of Geological Sciences, 2015. (in Chinese)
- 20 王小川, 史峰, 郁磊, 等. MATLAB 神经网络 43 个案例分析[M]. 北京: 北京航空航天大学出版社, 2013: 11–19.
- 21 屈忠义, 陈亚新, 史海滨, 等. 地下水文预测中 BP 网络的模型结构及算法探讨[J]. 水利学报, 2004, 35(2): 88–93.
- QU Zhongyi, CHEN Yaxin, SHI Haibin, et al. Structure and algorithm of BP network for underground hydrology forecasting[J]. Journal of Hydraulic Engineering, 2004, 35(2): 88–93. (in Chinese)
- 22 冯禹, 崔宁博, 龚道枝, 等. 基于极限学习机的参考作物蒸散量预测模型[J]. 农业工程学报, 2015, 31(增刊1): 153–160.
- FENG Yu, CUI Ningbo, GONG Daozhi, et al. Prediction model of reference crop evapotranspiration based on extreme learning machine [J]. Transactions of the CSAE, 2015, 31(Supp.1): 153–160. (in Chinese)
- 23 韩军彩, 陈静, 岳艳霞, 等. 石家庄市不同等级降水日数的时空分布特征[J]. 气象与环境学报, 2009, 25(6): 34–38.
- HAN Juncai, CHEN Jing, YUE Yanxia, et al. Temporal-spatial distributions of different level precipitation days in Shijiazhuang, Hebei province [J]. Journal of Meteorology and Environment, 2009, 25(6): 34–38. (in Chinese)
- 24 GB/T 14497—1993 地下水资源管理模型工作要求[S]. 1993.
- 25 刘昌明, 魏忠义. 华北平原农业水文及水资源[M]. 北京: 科学出版社, 1989.
- 26 费宇红. 京津以南河北平原地下水演变与涵养研究[D]. 南京: 河海大学, 2006.
- FEI Yuhong. Evolution and conservation of groundwater in Hebei plain to the south of Beijing and Tianjin[D]. Nanjing: Hohai University, 2006. (in Chinese)
- 27 冯慧敏, 张光辉, 王电龙, 等. 近 50 年来石家庄地区地下水流场演变驱动力分析[J]. 水利学报, 2014, 45(2): 180–186.
- FENG Huimin, ZHANG Guanghui, WANG Dianlong, et al. Analysis on driving force for groundwater flow field evolution in Shijiazhuang area in recent 50 years[J]. Journal of Hydraulic Engineering, 2014, 45(2): 180–186. (in Chinese)
- 28 王电龙, 张光辉, 冯慧敏, 等. 降水和开采变化对石家庄地下水流场影响强度[J]. 水科学进展, 2014, 25(3): 420–427.
- WANG Dianling, ZHANG Guanghui, FENG Huimin, et al. Contribution degrees of precipitation and groundwater exploitation to the groundwater flow field in Shijiazhuang[J]. Advances in Water Science, 2014, 25(3): 420–427. (in Chinese)
- 29 朱延华, 刘淑芬, 郭永海. 河北平原地下水动力环境演化规律及影响因素[J]. 地球科学, 1995, 20(4): 433–437.
- ZHU Yanhua, LIU Shufen, GUO Yonghai. Groundwater dynamic environment evolution and its constraints in Hebei plain[J]. Earth Science, 1995, 20(4): 433–437. (in Chinese)
- 30 张光辉, 费宇红, 刘春华, 等. 华北平原灌溉用水强度与地下水承载力适应性状况[J]. 农业工程学报, 2013, 29(1): 1–10.
- ZHANG Guanghui, FEI Yuhong, LIU Chunhua, et al. Adaptation between irrigation intensity and groundwater carrying capacity in North China Plain[J]. Transactions of the CSAE, 2013, 29(1): 1–10. (in Chinese)
- 31 张光辉, 费宇红, 刘克岩, 等. 华北平原农田区地下水开采量对降水变化响应[J]. 水科学进展, 2006, 17(1): 43–48.
- ZHANG Guanghui, FEI Yuhong, LIU Keyan, et al. Regional groundwater pumpage for agriculture responding to precipitation in North China Plain[J]. Advances in Water Science, 2006, 17(1): 43–48. (in Chinese)
- 32 孙俊, 卫爱国, 毛罕平, 等. 基于高光谱图像及 ELM 的生菜叶片氮素水平定性分析[J/OL]. 农业机械学报, 2014, 45(7): 272–277. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20140742&flag=1. DOI: 10.6041/j.issn.1000-1298.2014.07.042.
- SUN Jun, WEI Aiguo, MAO Hanping, et al. Discrimination of lettuce leaves' nitrogen status based on hyperspectral imaging technology and ELM[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(7): 272–277. (in Chinese)
- 33 虞双吉, 苗春生, 王新. 极限学习机神经网络在短期降水预报中的应用[J]. 云南大学学报: 自然科学版, 2013, 35(4): 507–515.
- YU Shuangji, MIAO Chunsheng, WANG Xin. Application of extreme learning machine neural network to forecast of short-term precipitation[J]. Journal of Yunnan University: Natural Science Edition, 2013, 35(4): 507–515. (in Chinese)