

南天, 曹文庚, 任印国, 等. 京津冀平原浅层地下水漏斗演变规律与影响因素[J]. 南水北调与水利科技(中英文), 2024, 22(1): 110-121. NAN T, CAO W G, REN Y G, et al. Evolution and influence factors of shallow groundwater depression cone in Beijing-Tianjin-Hebei Plain[J]. South-to-North Water Transfers and Water Science & Technology, 2024, 22(1): 110-121. (in Chinese)

京津冀平原浅层地下水漏斗演变规律与影响因素

南天^{1,2}, 曹文庚^{1,2}, 任印国³, 孙龙⁴, 高媛媛⁵

(1. 中国地质科学院水文地质环境地质研究所, 石家庄 050061; 2. 河北沧州平原区地下水与地面沉降国家野外科学观测研究站, 石家庄 050061; 3. 河北省水文勘测研究中心, 石家庄 050061; 4. 水利部信息中心, 北京 100053; 5. 水利部南水北调规划设计管理局, 北京 100038)

摘要: 为研究华北地区河湖生态补水对地下水漏斗演变的影响, 以京津冀平原浅层地下水漏斗 2003 年至 2022 年的相对变化作为识别目标, 从气象因素、地形因素、人为因素和含水层水力学特性 4 个方面进行考虑, 选取 8 个具体指标构建特征变量数据集, 使用逻辑回归(logistic regression, LR)、支持向量机(support vector machine, SVM)和随机森林(random forest, RF)方法建立漏斗演变识别模型, 并利用敏感度、特异度和决定系数 R^2 对拟合效果进行对比评价, 结果显示随机森林为最优模型。进而利用模型分析研究区地下水漏斗演变规律, 阐明具体因素对漏斗演变的影响作用。研究表明: 京津冀平原区浅层地下水漏斗在 2010 年之前整体呈扩张趋势, 之后在部分地区呈现缩减和消失的态势。河湖补水前, 地下水漏斗发展主要受开采影响, 其重要度约 50%; 2018 年后河湖补水对抑制漏斗扩张发挥了较为明显的作用, 重要度达 16%。从发展过程来看, 地下水开采依然是控制京津冀平原浅层地下水漏斗变化最重要的因素。对比宁柏隆和高蠡清两个典型浅层地下水漏斗的发展变化可知, 河道生态补水对宁柏隆漏斗变化的贡献率接近 10%, 而对高蠡清漏斗变化影响的重要度仅为 1%, 因此持续的河流生态补水对宁柏隆漏斗水位恢复会产生积极影响, 而对于高蠡清漏斗则需要以水源置换压减农业灌溉地下水量为关键手段实现水位恢复。

关键词: 京津冀平原; 地下水降落漏斗; 多源数据驱动模型; 机器学习; 演化机制

中图分类号: TV211 **文献标志码:** A **DOI:** 10.13476/j.cnki.nsbdqk.2024.0013

地下水是支撑京津冀地区社会、经济和农业发展的重要水源^[1-2]。自 20 世纪 70 年代以来, 京津冀地区地下水开采强度由每年约 200 亿 m^3 增加至 2017 年的 363 亿 m^3 。在天然补给量不足以满足开采的情况下, 主要含水层水位持续下降, 在京津冀地区形成了多处大小不一的地下水漏斗^[3-4]。由此引发诸如地面沉降、海水入侵和河湖衰竭等一系列环境地质问题^[5-7]。近年来, 国家和地方政府采取了包括南水北调、生态补水和地下水压采等措施旨在修复京津冀地区地下水漏斗。2019 年京津冀平原山前河道补水达 13.2 亿 m^3 , 2021 年底河北省累积压减地下水开采量 52.3 亿 m^3 , 部分地区水位已有所回升, 治理效果初步显现^[8-11]。2022 年 6 月份, 通过多渠道调水措施, 京杭大运河亦实现了全线贯通。

在此形势下, 需要总结京津冀平原地下水漏斗变化规律, 查明控制漏斗变化的主要因素, 提出漏斗治理的科学建议。

京津冀地区浅层含水层和深层含水层均存在地下水漏斗, 而由于开采条件、补给条件和水力学特性差异, 地下水漏斗呈现不同的分布规律, 浅层漏斗主要分布在山前平原区, 而深层漏斗主要分布在东部平原^[12]。浅层地下水漏斗形成时间长, 分布范围广, 变化经历复杂, 将作为本次研究的主要对象。众多学者均对京津冀平原浅层地下水漏斗展开过调查研究。邵景力课题组^[13-17]自 20 世纪初开始利用数值模拟方法系统地研究了京津冀平原浅层地下水漏斗区水资源量和水位变化关系, 指出了漏斗产生和发展主要受地下水开采影响, 并预测了不同

收稿日期: 2023-02-22 修回日期: 2023-07-09 网络出版时间: 2023-10-09

网络出版地址: <https://link.cnki.net/urlid/13.1430.tv.20231007.1622.015>

基金项目: 国家自然科学基金项目(41972262); 河北省自然科学基金优秀青年科学基金项目(D2020504032); 地下水位降落漏斗成因及关键控制因子研究(1300022P00329410101J)

作者简介: 南天(1986—), 男, 河北石家庄人, 助理研究员, 主要从事水文地质和地下水模拟研究。E-mail: 419416433@qq.com

通信作者: 曹文庚(1985—), 男, 山东禹城人, 副研究员, 主要从事水文地球化学和水文地质研究。E-mail: 281084632@qq.com

开采方式对地下水位的影响作用。目前针对京津冀地区河湖补水前地下水漏斗变化的研究较为详尽,但对漏斗整体发展与个体发展的差异性和对影响因素作用的量化研究尚未深入展开。对漏斗发展差异性的研究能够更加具体地分析漏斗演变机制,为漏斗区的地下水水位恢复提出有针对性的建议。赵伟等^[18]利用回归模型分析了京津冀平原宁柏隆漏斗水位变化影响因素,认为漏斗中心水位变化主要受开采和调水影响。靳博文等^[19]利用历史回归和差分方法从时空角度分析了京津冀平原高蠡清漏斗地下水位变化的主要驱动力,认为自然降水是驱动水位变化的重要驱动力,而人工开采则是地下水位变化的主导驱动力。可以看出,学者们已经逐渐重视对漏斗差异性的研究工作。

数据分析和数值模拟是以往对地下水漏斗研究的主要方法。两种方法模拟精度通常由建立模型时使用的源汇项和参数数据短板所决定,存在较大的不确定性,容易产生过拟合问题^[20]。此外,在大区域高精度剖分情况下,由于计算单元繁多,计算耗时冗长^[4,21-23]。随着大数据时代的到来,数据驱动的机器学习模拟方法进入了地下水领域研究者的视野。经典模型包括多元线性回归(multivariable linear regression, MLR)、支持向量机算法(support vector machine, SVM)和贝叶斯模型等^[24-25]。近年来随着算法的迭代更新,模拟方法则以集成学习和深度学习等高阶复杂算法为主^[26],如Wang等^[27]2018年利用随机森林(random forest, RF)方法评估了青岛大沽河流域地下水位,Jiang等^[28]在2022年利用集成学习算法预测了华北地区地下水水位变化。其中随机森林算法具有参数调节灵活、预测精度高、泛化能力强不易产生过拟合现象、对数据质量要求较低及对特征变量选择具有较好的鲁棒性,目前常用于进行分类识别^[29]。该方法可用于分析与地下水漏斗发展变化有关的各类驱动因素,指出各因素和目标变量之间的关系。针对漏斗发展变化的高度空间不均一性,可基于二元目标变量建模,分析特征变量对漏斗分布的影响作用,量化特征重要度,进而阐述漏斗演变机制,为漏斗治理提供科学依据。

综上,本次研究将2003年和2022年的京津冀平原浅层地下水漏斗分布数据形成样本数据集,在分析研究区浅层地下水漏斗演变规律的基础上,利用逻辑回归、支持向量机和随机森林算法,以气象

因素、地形因素、人为因素和含水层水力学特性作为驱动因子,对漏斗空间分布变化建模,识别漏斗扩张区,进而评价分析特征变量对漏斗发展的控制行为和作用程度,阐明京津冀浅层地下水漏斗现阶段演变机制,对比分析典型漏斗——宁柏隆漏斗和高蠡清漏斗演变控制因素的差异,提出有针对性的漏斗修复建议。

1 研究区概况

1.1 自然地理

京津冀平原位于我国东部,东临渤海,北起燕山,西抵太行山,是华北平原的主要组成部分。本次研究范围包括北京、天津和河北3省(直辖市)的全部平原区,总面积约8.85万km²。研究区属于暖温带半干旱季风型气候,平均气温10~15℃,多年平均降水量约为560mm。年降雨时空分布不均,多集中在7—9月份,占全年降水量的70%~80%。年均蒸发量在1100~2000mm^[30]。区内主要分布海河和滦河水系,干流主要包括发源于背风山区的漳河、滹沱河、永定河和潮白河,以及发源于迎风坡的蓟运河、北运河、大清河和滏阳河。

1.2 水文地质条件

京津冀平原第四系含水介质是几何形态复杂、多种类型叠加的含水层组结构(图1)。通常按岩性特征、沉积规律、含水层分布形式和水动力条件分为4个含水层组对应着全新世到更新世地层(Q₄-Q₁)^[31],而由于第I、II含水层组之间存在较好的水力联系,一般将I、II含水层组概化为浅层含水层,第I层厚度基本为60m,第II含水层底界深度一般在120~170m,而将III、IV含水层组概化为深层含水层,目前第II和III含水层为主要开采层^[32]。地下水漏斗随着地下水开采的进行而持续发生变化。本次研究工作主要以浅层含水层作为研究对象。

研究区地下水补给主要是大气降水补给,其次是山前侧向补给。近年来河湖补水渗漏也成为区域补给的重要来源之一。其次补给项还包括渠系渗漏补给、田间灌溉入渗补给、井灌回归补给以及山前边缘的隐伏碳酸盐岩溶水的顶托补给。20世纪50年代以前,研究区地下水基本处于天然状态,由山前流向渤海或海岸带洼地。地下水开发初期,山前平原区浅部地下水大体近南东和东方向流动,中东部平原及滨海平原区地下水近北东方向流动。目前地下水由冲洪积扇流向中部平原的漏斗区,在

中部平原靠滨海平原一侧,仍可见流向渤海的趋势。消耗。
研究区地下水主要排泄方式为人工开采、潜水蒸发

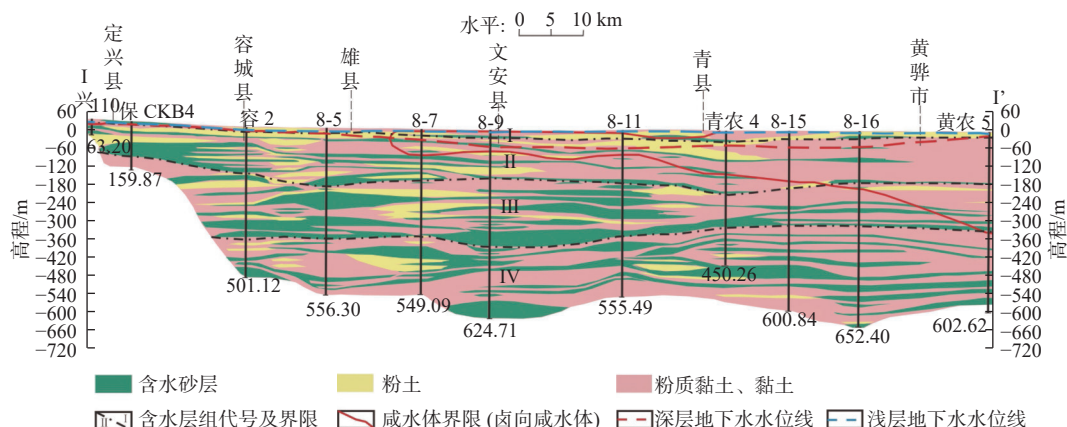


图 1 研究区定兴县-文安县-黄骅市水文地质剖面

Fig. 1 Hydrogeological profile of Dingxing County - Wen'an County - Huanghua City study area

研究区地下水动态由 20 世纪 50 年代之前的天然状态,经历了伴随工农业和城镇发展的地下水开采增加引起的局部地下水下降阶段,发展至 20 世纪末至 21 世纪初的区域地下水位快速下降阶段。如今,随着南水北调工程的运行、地下水超采综合治理和河湖补水的实施,地下水开采量得到控制,地下水位整体下降趋势减缓,部分地区地下水已呈现回升态势。

京津冀平原区沉积环境变化和相应的物理气候过程,塑造了平原区由山前至滨海的地下水循环系统的基本面貌。浅层含水层以淡水亚系统为主,自西向东山前及中东部平原为以溶滤淡化作用为主的淡水亚系统,过渡到中部平原以大陆盐化作用为主的咸水亚系统,至滨海地区则是以海水入侵影响为主的滨海咸水亚系统。山前平原第四系地下水矿化度一般小于 1 g/L。浅层含水层水化学类型以 $\text{HCO}_3\text{-Ca}\cdot\text{Mg}$ 型为主。中部平原通常为矿化度大于 2 g/L 的咸水体,水化学类型在浅部以 $\text{SO}_4\cdot\text{HCO}_3\text{-Ca}\cdot\text{Mg}$ 型为主。滨海平原存在“咸-淡”或“全咸”两种水质结构,咸水体化学类型一般为 Cl (或 $\text{Cl}\cdot\text{HCO}_3$)- Na ($\text{Na}\cdot\text{Mg}$)型,矿化度最大可达 35 g/L;下伏淡水水化学类型一般为单一的 $\text{HCO}_3\text{-Na}$ 型或 $\text{HCO}_3\cdot\text{SO}_4\text{-Na}$ 型,矿化度通常小于 1 g/L。

1.3 地下水漏斗现状

京津冀地区地下水开采量从 20 世纪 70 年代末起迅速增加,地下水位相应产生大幅下降,至 80 年代在平原区已有漏斗产生,2000 年以后随着地下水开采量的持续增加,地下水位不断下降,漏斗在平原区广泛发展。根据 2020 年监测统计数据计算,在

京津冀平原浅层含水层分布的地下水漏斗总面积接近 1.5 万 km^2 。浅层地下水位埋深在山前地区普遍超过 20 m,中部平原埋深在 10~20 m,东部滨海地区普遍小于 10 m,部分地区水位埋深在 5 m 之内。随着华北地下水超采综合治理工作的开展,至 2019 年河北平原河湖生态补水超过 10 亿 m^3 ,累积地下水压采量超过 50 亿 m^3 [33-35]。虽然漏斗整体处于扩张状态,但是在局部地区显示出一定程度的缩减。

2 研究方法和数据

本次研究整体流程见图 2。首先在分析研究区地下水漏斗演变规律的前提下,选取特征变量并进行数据清洗,继而与漏斗分类标签形成样本数据,然后利用逻辑回归、支持向量机和随机森林模型对漏斗演变情况进行识别验证,最后通过对模拟结果的分析,解释地下水漏斗变化机制,评价特征因素对漏斗演变的影响作用。

2.1 数据集构建

2.1.1 标签数据集

利用河北省水文局 2003 年和 2022 年监测孔水位统测数据绘制地下水流场,其中,2003 年流场与邵景力课题组在 2013 年研究中选用的流场保持了高度一致,2022 年流场与中国地质科学院水文地质研究所 2022 年根据监测数据绘制的流场基本相符[36]。根据京津冀三地漏斗发展变化特点,北京和天津地区以 0 m 等水位线作为浅层地下水漏斗边界,河北省地区则以 -10 m 等水位线作为浅层地下水漏斗边界。利用两期流场进行变差计算,进而对漏斗变化

进行二分类处理,漏斗扩张区为正样本标记为 1,反之为负样本标记为 0。

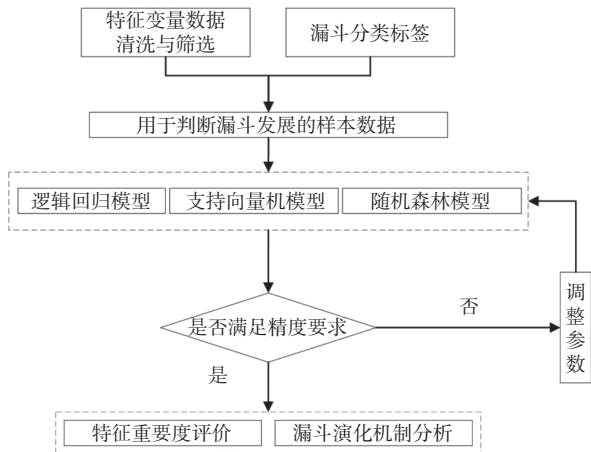


图 2 研究流程
Fig. 2 Flowchart

2.1.2 特征变量数据集

本次建模考虑了气象、地形、人为和含水层水力学特性 4 类因素共计 8 个具体指标用作建模的特征变量,见表 1。变量的选取基于以往研究对漏斗发展机制的认识^[37-38]。2003 年至 2022 年的降水和蒸发量来自国家气象监测站逐月数据。地下水开采量数据来自河北省、北京市和天津市 2003 至 2022 年水资源公报。河湖补水数据来自水利部海河水利委员会 2018 年至 2022 年调查统计资料。高程数据和坡度数据来源于 NASA 卫星遥感解译数据,解译精度为 30 m×30 m。给水度和渗透系数由 2009 年中国地图出版社出版的华北平原地下水可持续利用图集获取。

表 1 特征变量数据集组成
Tab. 1 List of the feature variable data set

类别	漏斗特征变量数据集		
	指标	来源	变化特点
气象因素	降水量	统计数据	累加
	蒸发强度	统计数据	累加
地形因素	地形坡度	遥感数据	分布式
	地表高程	遥感数据	分布式
人为因素	河流补水量	统计数据	累加
	开采量	统计数据	累加
含水层水力学特性	渗透系数	数值模型率定	分布式
	给水度	数值模型率定	分布式

各类型特征变量通过克里金插值方式离散至 4 000 km×4 000 km 的单元格,其中,开采量在对应行政区平均分配至计算单元格,河湖补水量首先按

补水河道段进行划分,然后再平均分配到各段对应的网格中,水文地质参数则根据分布情况矢量化处理后赋值至各计算单元作为模型输入数据。本次研究共使用 88 433 个输入样本。

2.1.3 数据归一化处理

为了消除特征变量之间的量纲差异,对特征变量进行归一化处理,即用数据集中各数据减去均值后再除以标准差,可表示为

$$\bar{X}_i = \frac{X_i - E(X)}{S_d(X)} \quad (1)$$

式中: X_i 为特征变量; \bar{X}_i 为标准化处理后的特征变量; $E(X)$ 为该特征变量的期望; $S_d(X)$ 为该特征变量的标准差。

2.2 模型搭建

2.2.1 模型选择

对比逻辑回归(LR)、支持向量机(SVM)和随机森林(RF)模型,其中随机森林方法在对比中显示出了最好的模拟性能。

逻辑回归(LR)。逻辑回归从广义上属于线性回归分析模型,其推导过程与计算方式类似于回归问题,但主要用于解决二分类问题,属于机器学习中的有监督学习。通过对样本数据进行训练,对给定的测试数据可进行分类。其核心 Sigmoid 函数可表述为

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

对于输入模型的每一组训练样本 $x_{(i)}$,均能通过该函数得到一个 0~1 的值,当函数值大于 0.5 时认为属于分类 1,否则为分类 0。通过对样本进行训练可以在训练集获取准确的模拟值,进而用于测试和预测。其损失函数通常使用交叉熵函数。

以是否是漏斗区扩张作为识别标准,其中,正样本类别为漏斗扩张区,反之为负样本类别。

支持向量机(SVM)。支持向量机是常用的二分类分类器,其决策边界是对学习样本求解的最大边距超平面,同样属于一类有监督的机器学习方法,原理见图 3。具体来讲:针对线性可分问题时,在原空间寻找分类超平面的最优解;在线性不可分时,加入松弛变量使用非线性映射将低维空间样本映射到高维空间从而转化为线性可分问题,对其进行分类超平面最优解的求取。其常用方法为核方法,重点是构造用于求解超平面的核函数,超平面可表

示为

$$\omega^T \phi(X) + b = 0 \quad (3)$$

其中 ϕ 为映射函数。由于其难以直接计算内积, 因此以定义核函数的方法回避内积计算, 核函数可表示为

$$k(X_1, X_2) = \phi(X_1)^T \phi(X_2) \quad (4)$$

利用核函数找到超平面后即可根据超平面和预测值位置关系确定预测值类别。

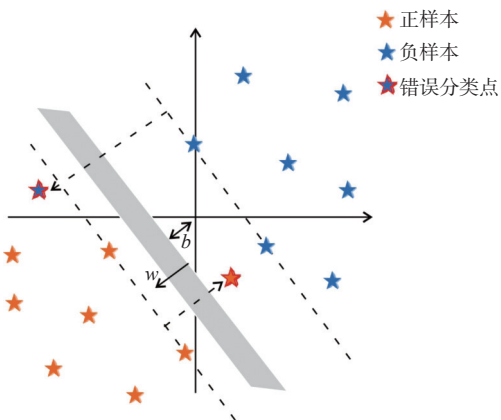


图 3 支持向量机方法原理

Fig. 3 Schematic diagram of support vector machine

为与 LR 模型进行对比, SVM 模型中同样以 Sigmoid 函数作为核函数将数据样本进行维度转化后搜寻特征数据的超平面, 对漏斗扩张区进行识别。

随机森林(RF)。随机森林是一个高度灵活的机器学习方法, 可对数据进行回归和分类, 具备处理大数据的特性, 属于机器学习算法中的 Bagging 集成学习方法^[39]。其工作原理是生成多个分类器, 各自独立学习和预测, 最后结合成单预测, 因此优于任何一个单分类器进行的预测。RF 模型算法过程见图 4。RF 由一系列决策树组成, 通过自助法重采样技术, 从原始训练样本集中有放回地重复随机抽取 m 个样本, 生成新的训练样本集合, 然后根据自助样本集生 k 个分类树组成随机森林, 新数据的分类结果按分类投票多少而定。其实质是将多棵决策树合并在一起, 每棵树的建立依赖于一个独立抽取的样本, 森林中每棵树具有相同的分布, 分类误差取决于每棵树的分类能力和他们之间的相关性。

模型需要设置的主要参数为决策树分类器的数量和树节点划分时随机选取的预测变量数。一般决策树越多, 拟合精度越高, 而树节点划分所保留的预测变量数据较为敏感, 通常可取预测变量的平方根。由于研究区范围较大, 漏斗区和非漏斗区存在数量的偏差, 建模时将分类权重设置为平衡条件用以保证模拟的准确性。

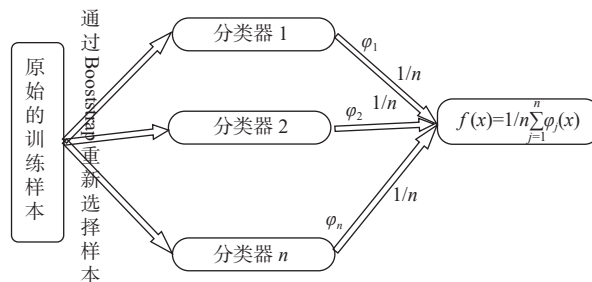


图 4 随机森林模型算法过程

Fig. 4 Algorithmic process of random forest

RF 模型通过不断地进行迭代分类, 以不同顺序的特征变量作为节点进行计算, 投票出整个研究空间训练样本所属分类, 迭代至损失函数保持稳定后最终可以确定样本点在空间上是否属于漏斗扩张区。

2.2.2 数据划分

利用漏斗分布标签数据集和特征变量数据集, 将 88 433 个样本数据随机划分为 80% 的训练集和 20% 的测试集。在此基础上, 利用 Python 语言编译机器学习模型, 进而应用模型对训练集和测试集数据进行拟合调参, 通过统计数据评估模型性能, 以评价模型对漏斗分布判别的准确性。

2.2.3 模型评价

选用敏感度 (δ_{sc})、特异度 (δ_{sp}) 和决定系数 R^2 这 3 个指标对模拟结果进行评价。

决定系数 R^2 可以通俗地理解为使用均值作为误差基准, 判断预测误差是否大于或者小于均值基准误差。 R^2 越接近 1 模型精度越高, 其计算公式为

$$R^2 = 1 - S_{sc}/S_{st} \quad (5)$$

其中: S_{sc} 为残差平方和即估计值与真实值的误差, 反映模型拟合程度; S_{st} 为总离差平方和即平均值与真实值的误差, 反映与数学期望的偏离程度。

敏感度 (δ_{sc}) 表示正样本被正确预测的比例, 敏感度越高表示模型识别正样本能力越强, 以此指标来衡量模型能够正确识别漏斗区的准确性, 其计算公式可表示为

$$\delta_{sc} = N_{TP}/(N_{TP} + N_{FN}) \quad (6)$$

式中: N_{TP} 为模型预测正确的正类样本数; N_{FN} 为模型预测错误的负类样本数; N_{TP} 与 N_{FN} 之和为正确正类样本的总数。

特异度 (δ_{sp}) 则表示负样本被正确预测的比例, 特异度值越高表示模型能够识别负样本的能力越强, 以此指标来评价模型能够正确识别非漏斗区的准确性, 其计算公式可表达为

$$\delta_{sp} = N_{TN}/(N_{TN} + N_{FP}) \quad (7)$$

式中: N_{TN} 为模型预测正确的负类样本数; N_{FP} 为模型预测错误的正类样本数; N_{TN} 与 N_{FP} 之和为正确负类样本的总数。 N_{TP} 、 N_{FN} 、 N_{TN} 和 N_{FP} 通过对模型预测结果进行统计获取。

特征变量对漏斗演变的影响作用主要用重要度特征和 SHAP 值进行评价。重要度分布可以用于评估特征变量对漏斗扩张的影响程度,通常通过基尼杂质减少方法进行统计。SHAP 值是根据特征组合次序计算全部特征组合及组合中每个特征变量的边际贡献,进而衡量各特征变量对输出结果的影响,SHAP 值越大表明特征变量对漏斗扩张起促进作用,反之则表示抑制作用。

在整体分析京津冀浅层漏斗变化主控因素的基础上,对比宁柏隆漏斗和高蠡清漏斗演变驱动因素的差异性,提出有针对性的建议。

3 结果与讨论

3.1 模型性能评价

对比模拟过程的关键要素(表 2)可以发现,LR 模型整体误差较大达到 42%,SVM 模型误差为 20%,RF 模型误差最小仅为 12%。从迭代拐点可以看出,SVM 模型在迭代 75 次后结果趋于平稳且精度最高,

而 LR 模型达到最高精度时的迭代次数约为 SVM 模型的 1/2,RF 模型在 20 次迭代后即可达到最佳精度。在同时迭代 100 次条件下,LR 模型运行时间为 2 min,SVM 模型略长为 3 min,RF 模型最短为 1 min。从可调节性角度看,在 TensorFlow 框架下,LR 和 SVM 模型可调节参数均为 15 个,RF 模型略多为 18 个,说明 RF 模型相对更加灵活。

表 2 模型性能对比
Tab. 2 Features of models

模型	误差/%	出现拐点的迭代次数/次	运行时间/min	可调节参数数量/个
逻辑回归	42	38	2	15
支持向量机	20	75	3	15
随机森林	12	20	1	18

各模型最终模拟效果见表 3。从表 3 可以看出:LR 模型模拟精度较差,其预测正分类样本即判断漏斗扩张区性能较好,敏感度可达 75% 以上,但不能准确地判断出非漏斗区,特异性仅为 42%;SVM 模型和 RF 模型预测性能相对更强;而 RF 模型判断漏斗扩张区的性能更好,敏感度可达 94%。在判断非漏斗区方面 SVM 模型和 RF 模型二者接近,RF 模型特异性稍高为 77.9%。

表 3 模型拟合结果对比

Tab. 3 Shallow model results for LR, SVM and RF models

模型	R^2			敏感度			特异度		
	训练集	测试集	整体	训练集	测试集	整体	训练集	测试集	整体
逻辑回归	0.589	0.582	0.587	0.759	0.754	0.757	0.428	0.421	0.426
支持向量机	0.800	0.800	0.800	0.830	0.830	0.830	0.733	0.732	0.733
随机森林	0.878	0.878	0.878	0.940	0.940	0.940	0.779	0.778	0.779

3.2 京津冀平原漏斗演变规律分析

根据京津冀平原地下水漏斗 2003 年至 2022 年的变化(图 5)可以看出:浅层地下水漏斗主要以山前大中型城市为中心,以受长期大量开采地下水影响作用为主,漏斗主要沿南北向发展。2003 年京津冀平原区浅层地下水漏斗面积 6 000 km²,至 2017 年虽然地下水漏斗整体赋存面积与 2003 年相差不大,但由于北京市率先进行了地下水压采,北京顺义-燕郊漏斗明显减小,漏斗整体由北京中心区移至东部燕郊地区,面积由 330 km² 减小至 250 km²,北京地区漏斗基本消失。

近年来,京津冀地区地下水大规模压采和河湖补水成效显著,地下水漏斗呈明显减小趋势,从 2003 年到 2022 年,地下水漏斗由 5 107.38 km² 下降

到了 976.71 km²,下降幅度达 80.88%,仅宁柏隆漏斗和高蠡清漏斗依然处于封闭状态,但封闭面积也呈现出明显缩减的趋势。其中:宁柏隆漏斗面积由 2003 年的 1 021 km² 下降至了 2022 年的 257 km²,下降幅度达 74.8%;高蠡清漏斗面积由 2003 年的 1 753 km² 左右下降至了 2022 年的 692 km²,下降幅度达 60.5%。由此可见,在持续的地下水压采和河湖补水的情况下,地下水漏斗得到了有效遏制,京津冀平原有望实现地下水位的全面回升。

3.3 京津冀平原漏斗演变机制分析

3.3.1 漏斗整体演变机制

特征变量对漏斗整体演变的影响重要度分布见图 6,可以看出地下水开采量、累积降水量影响程度最高,其重要度分别达到 45% 和 19%。

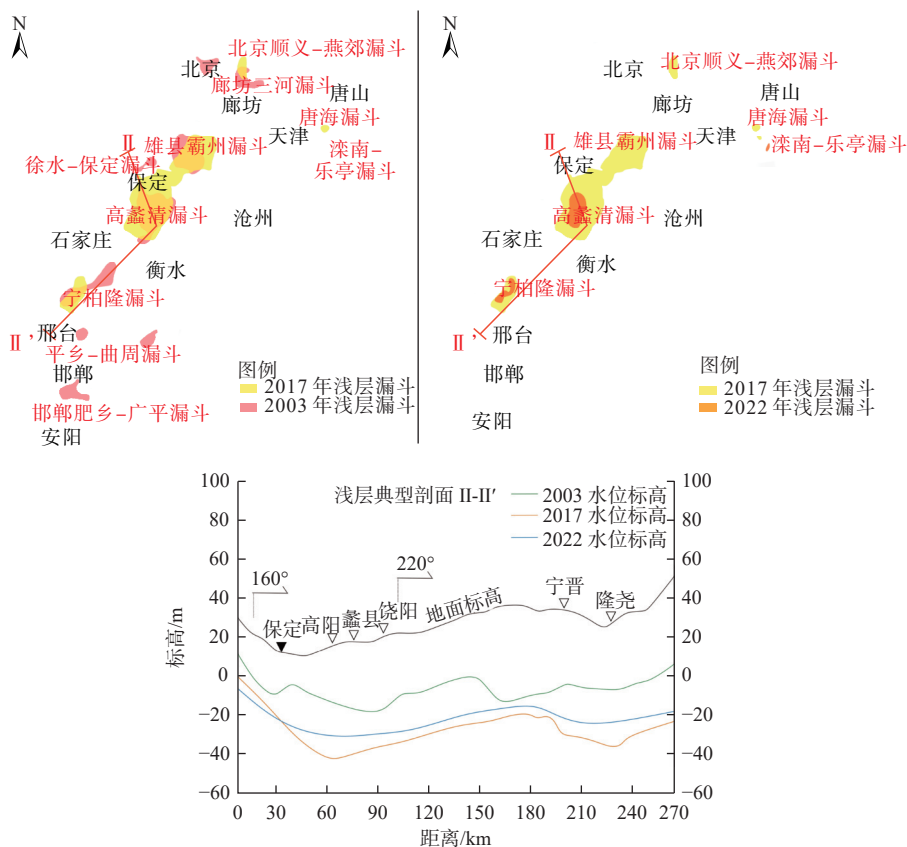


图 5 京津冀平原 2003—2022 年浅层地下水漏斗变化

Fig. 5 Variation of groundwater depression cone in Beijing-Tianjin-Hebei Plain from 2003 to 2022

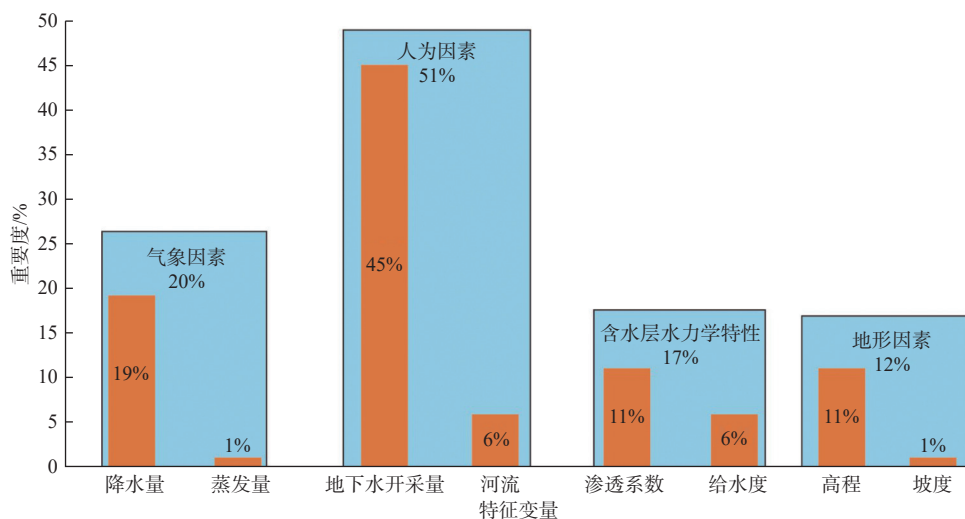


图 6 浅层漏斗演化特征重要度分布

Fig. 6 Distribution of feature importance for shallow groundwater depression cone

根据重要度排序可以看出地下水开采与漏斗演变显著相关,结合 SHAP 图(图 7)可知地下水开采量越大,漏斗越向扩张方向发展,体现了开采引起的水位下降作用^[40]。局部地下水位的快速持续下降是地下水漏斗产生的主要原因,特别是在太行山前主要城市区,随着城市发展地下水用量激增,漏斗也向扩大加深方向发展。

累积降水量对漏斗变化有着较高的影响^[41]。降水作为京津冀平原的主要地下水补给来源,降水量越大漏斗发展越得到抑制,而降水量减小意味着漏斗接受的补给将显著减小,从而漏斗产生扩张趋势。由于漏斗区地下水埋深普遍较大,因此蒸发对漏斗变化的影响十分微弱。

高程变化反映了侧向补排条件的差异,侧向补

给是研究区浅层含水层的重要补给来源^[42]。高程越大越接近山区,侧向补给充足而且补给速度快,漏斗不易形成或扩大,而高程越小相对离补给源越远,接受侧向补给困难,因此在强开采条件下容易形成漏斗或使漏斗区向周围扩张。侧向补给能力反映出京津冀平原浅层地下水漏斗西部边界的限制作用,而侧向排泄在整体排泄量中占比较小,对漏斗区变化的影响极为有限。

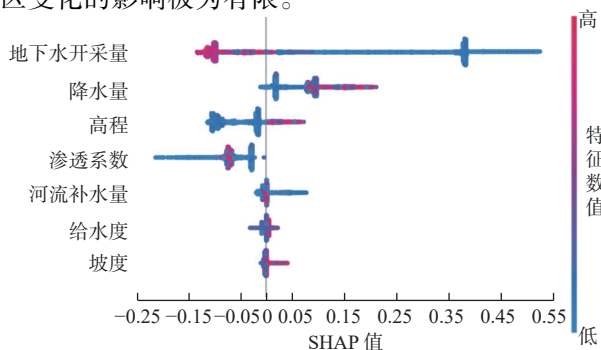


图7 浅层漏斗 SHAP 图

Fig. 7 SHAP value for shallow groundwater depression cone

渗透系数和给水度作为含水层水力学特性指标,反映着含水层接受侧向和垂向补给的性能^[43]。渗透系数和给水度的大小与组成含水层的沉积物岩性、沉积环境和构造特点相关。京津冀山前平原南北沿线含水层以粗颗粒沉积物为主,而在平原中东部、滨海地带含水层岩性以细颗粒沉积物为主^[44]。水力学特性的差异成为研究区浅层地下水漏斗向东部发展的限制条件。

随着近年河湖补水的持续增加,经由漏斗区的河道入渗量明显持续增加,河流周边含水层得到持续补给,漏斗扩张趋势得到抑制^[45]。综合来看,京津冀浅层地下水漏斗现阶段发展主要受人为因素控制,气象因素对漏斗发展产生抑制作用,而含水层水力学特性和地形因素共同影响着地下水漏斗的展布形式。

3.3.2 漏斗阶段性演变机制

按地下水开发利用情况和河湖补水条件将研究期进一步划分为3个阶段(表4),其中:2003—2010年为地下水大规模开采阶段;2011年至2017年为地下水开采减缓阶段;2018年至2022年为河湖补水阶段。各阶段地下水开采量、降水量和河湖补水量见表3,可以看出:2011年至2017年研究区地下水开采量较前一阶段已经下降近20%,漏斗在部分地区呈现缩减趋势;2017年以后在开采量持续减少,其次京津冀平原累积增加河湖补水240.44亿m³,促进了地下水水位的回升和地下水漏斗的进一步缩减。研究区降水在2003年至2010年阶段较小,

2011年至2017年段有所增加,由于2021年为极端降水年,京津冀平原在2018年至2022年段平均降水量相对较大为572.56mm。

表4 重要特征数据阶段变化情况

Tab. 4 Variation of important feature variables

阶段	降水量/mm	地下水开采量/亿m ³	河湖补水量/亿m ³
2003—2010年	511.42	147.80	0
2011—2017年	542.23	119.02	0
2018—2022年	572.56	104.81	74.51

通过对各阶段影响漏斗变化的特征变量进行特征重要度分析(图8)可以看出,由于河湖补水在2017年之后实施,因此在前2个阶段未能影响研究区漏斗发展,而在2018—2020年阶段,其重要度可达16%,基本与水力学特性重要度相当,略小于降水的重要度21%。在此期间由于地下水开采的稳定减少,其重要度由前2个阶段的50%左右下降到33%。对比3个阶段地下水开采量的重要度变化可以看出,在漏斗发展阶段,主要是受地下水开采作用影响,随着开采量阶段性的减少,漏斗面积也相应有所减小。

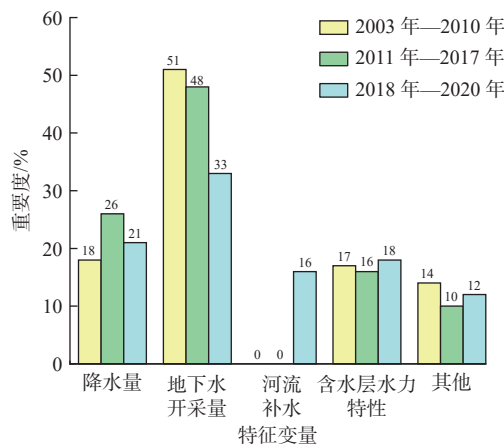


图8 各阶段浅层漏斗演化特征重要度

Fig. 8 Feature importance for shallow groundwater depression cone in different periods

3.3.3 典型漏斗演变机制对比

根据重要度排序(图9)和SHAP图(图10)可以看出,对宁柏隆漏斗演变影响最大的是降水量和开采量。宁柏隆漏斗一方面受开采量增大而扩张,另一方面在降水量相对丰沛的地区发展受到一定程度的抑制,地下水开采和降水对漏斗发展的影响十分接近,重要度分别为25%和24%。与宁柏隆漏斗不同,影响高蠡清漏斗发展变化的主要因素则是地下水开采,地下水漏斗的发展基本由开采导致,重要度高达85%。

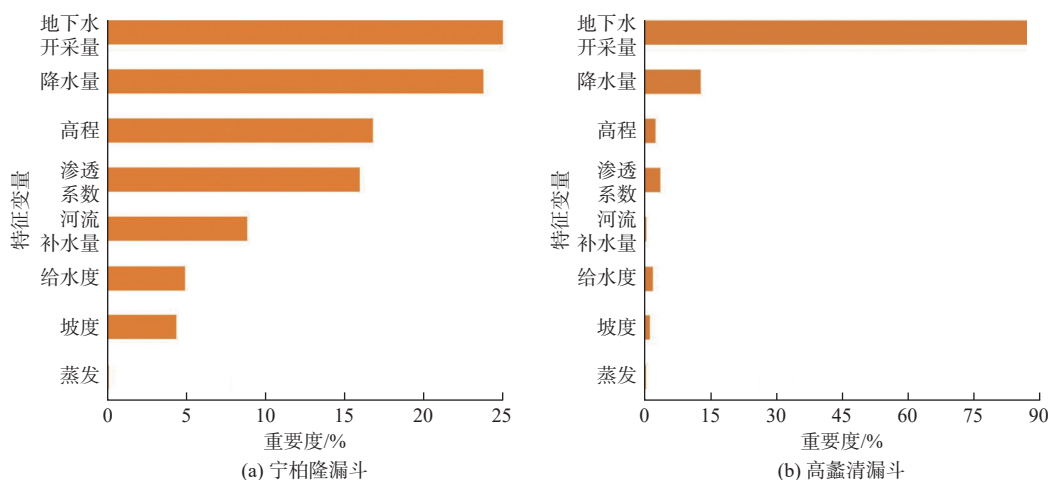


图 9 典型浅层漏斗演化特征重要度对比

Fig. 9 Comparison of feature variable importance distributions of typical shallow groundwater depression cone

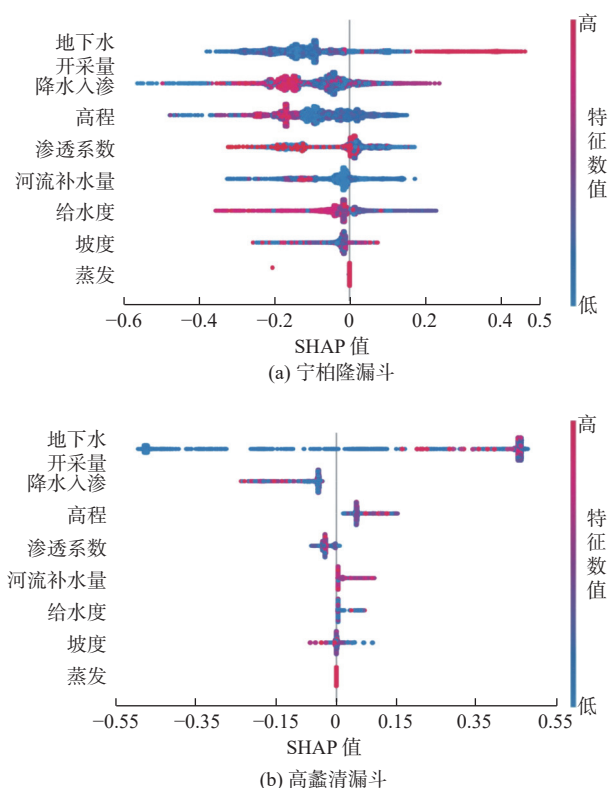


图 10 典型浅层漏斗特征变量 SHAP 图

Fig. 10 Comparison of SHAP value maps of typical shallow groundwater depression cone

由于宁柏隆漏斗区有滏阳河经过,其 2018 年以来利用南水北调中线、当地水库和再生水补水超过 4 亿 m^3 。河道的生态补水在一定程度上缓解了地下水漏斗的发展扩张,生态补水对漏斗演变的影响重要度接近 10%。虽然高蠡清漏斗区也有多条河流经过,但由于整体位于河流下游地区,含水层能够接受的河流渗漏补给十分有限,因此河流生态补水对漏斗的影响微弱。

4 结论

本次研究表明,京津冀平原地下水漏斗发展的主要控制因素为地下水开采,其影响重要度达 45%。地下水压采和河湖补水以来,漏斗区水位恢复效果显著,漏斗呈现缩减趋势。经计算,补水引起水位回升的重要度可达 16%。随着区域性地下水开采量的持续减少,开采对地下水漏斗扩张发展的驱动力有着明显的衰减,重要度由 2018 年之前的近 50% 下降至 2020 年的 33%。这说明地下水压采在抑制地下水漏斗扩张方面取得了明显成效。

对比宁柏隆漏斗和高蠡清漏斗可知:宁柏隆漏斗为开采-降水控制型漏斗,开采和降水对漏斗演变的重要度分别为 25% 和 24%,河道生态补水能够对漏斗发展起到一定的抑制作用,其重要度为 9%。高蠡清漏斗为开采控制型漏斗,开采对漏斗变化影响的重要度高达 85%,河道生态补水对漏斗发展影响微弱,重要度仅为 1%。因此,在宁柏隆漏斗区可适当增加滏阳河山前平原河道补水量,采取压补结合的方式恢复地下水水位,而在高蠡清漏斗区则建议在高阳西部、清苑区南部和蠡县北部继续合理配置水资源,通过控制地下水开采为主的方式对漏斗区进行进一步的修复。

参考文献:

- [1] 郑凌云,张永祥,贾瑞涛,等.基于GMS的北京市朝阳区地下水环境数值模拟与预测分析[J].水利水电技术(中英文),2022,53(1):114-123. DOI: 10.13928/j.cnki.wrahe.2022.01.012.
- [2] LIU J G, ZANG C F, TIAN S Y, et al. Water conservancy projects in China: Achievements, challenges

- and way forward[J]. *Global Environmental Change*, 2013, 23(3): 633-643. DOI: 10.1016/j.gloenvcha.2013.02.002.
- [3] LIU C M, YU J J, KENDY E. Groundwater exploitation and its impact on the environment in the North China Plain[J]. *Water International*, 2001, 26(2): 265-272. DOI: 10.1080/02508060108686913.
- [4] 曹文庚, 杨会峰, 高媛媛, 等. 南水北调中线受水区保定平原地下水质量演变预测研究[J]. *水利学报*, 2020, 51(8): 1-12. DOI: 10.13243/j.cnki.slxb.2020.0035.
- [5] 朱菊艳, 郭海朋, 李文鹏, 等. 华北平原地面沉降与深层地下水开采关系[J]. *南水北调与水利科技*, 2014, 12(3): 165-169. DOI: 10.13476/j.cnki.nsbdqk.2014.03.036.
- [6] 韩宇平, 刘存强, 赵雨婷, 等. 华北典型灌区气候变化条件下地下水响应研究[J]. *南水北调与水利科技*, 2019, 17(2): 107-115, 122. DOI: 10.13476/j.cnki.nsbdqk.2019.0040.
- [7] 何理, 焦蒙蒙, 王喻宣, 等. 基于组合模型的天津市地面沉降预测及危险性评价[J]. *水利水电技术(中英文)*, 2022, 53(1): 178-189. DOI: 10.13928/j.cnki.wrahe.2022.01.017.
- [8] 陈飞, 羊艳, 史文龙, 等. 河北省地下水超采综合治理农业措施压采效果与技术经济性分析[J]. *南水北调与水利科技(中英文)*, 2022, 20(5): 1019-1026. DOI: 10.13476/j.cnki.nsbdqk.2022.0101.
- [9] 陈飞, 丁跃元, 李原园, 等. 华北地区地下水超采治理实践与思考[J]. *南水北调与水利科技(中英文)*, 2020, 18(2): 191-198. DOI: 10.13476/j.cnki.nsbdqk.2020.0042.
- [10] ZHU L, GONG H L, CHEN Y, et al. Effects of water diversion project on groundwater system and land subsidence in Beijing, China[J]. *Engineering Geology*, 2020, 276: 105763. DOI: 10.1016/j.enggeo.2020.105763.
- [11] 李莎, 成建梅, 李敏敏, 等. 人类开采活动影响下的衡水地区地下水水质特征及演化[J]. *南水北调与水利科技*, 2016, 14(3): 55-61, 100. DOI: 10.13476/j.cnki.nsbdqk.2016.03.010.
- [12] 李文鹏, 王龙凤, 杨会峰, 等. 华北平原地下水超采状况与治理对策建议[J]. *中国水利*, 2020, 13: 26-30.
- [13] SHAO J L, LI L, CUI Y L, et al. 2013. Groundwater flow simulation and its application in groundwater resource evaluation in the North China Plain[J]. *Acta Geologica Sinica (English Edition)*, 2013, 87(1): 243-253. DOI: 10.1111/1755-6724.12045.
- [14] 邵景力, 赵宗壮, 崔亚莉, 等. 华北平原地下水流模拟及地下水资源评价[J]. *资源科学*, 2009, 31(3): 11-17. DOI: 10.3321/j.issn:1007-7588.2009.03.002.
- [15] 李玲, 邵景力, 崔亚莉, 等. 华北平原区地下水开采量估算研究[J]. *南水北调与水利科技*, 2013, 11(5): 82-85. DOI: 10.3724/SP.J.1201.2013.05082.
- [16] 李长青, 邵景力, 靳萍, 等. 平原地区水文地质结构条件模拟及其应用: 以华北平原为例[J]. *现代地质*, 2009, 23(1): 137-143. DOI: 10.3969/j.issn.1000-8527.2009.01.021.
- [17] 王仕琴, 邵景力, 宋献方, 等. 地下水模型MODFLOW和GIS在华北平原地下水资源评价中的应用[J]. *地理研究*, 2007, 5: 975-983. DOI: 10.3321/j.issn:1000-0585.2007.05.014.
- [18] 赵伟, 贾新颖, 武丽生, 等. 宁柏隆地下水漏斗水位变化影响因素分析[J]. *水利规划与设计*, 2020, 11: 48-52.
- [19] 靳博文, 王文科, 段磊, 等. 保定平原区地下水生态水位阈值的探讨[J]. *水文地质工程地质*, 2022, 49(5): 166-175. DOI: 10.16030/j.cnki.issn.1000-3665.202112020.
- [20] KITANIDIS P K. Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport[J]. *Water Resources Research*, 2015, 51, 5888-5904. DOI: 10.1002/2015WR017639.
- [21] 钟乐乐, 曾献奎, 吴吉春. 基于高斯过程回归的地下水模型结构不确定性分析与控制[J]. *水文地质工程地质*, 2019, 46(1): 1-10.
- [22] HU L T, XU Z X, HUANG W D. Development of a river-groundwater interaction model and its application to a catchment in northwestern China[J]. *Journal of Hydrology*, 2016, 543: 483-500. DOI: 10.1016/j.jhydrol.2016.10.028.
- [23] RAJABI M M, ATAIE-ASHTIANI B, SIMMONS C T. Model-data interaction in groundwater studies: Review of methods, applications and future directions[J]. *Journal of Hydrology*, 2018, 567: 457-477. DOI: 10.1016/j.jhydrol.2018.09.053.
- [24] 齐永强, 李文鹏, 郑跃军, 等. 地下水机器学习方法研究: 水位监测数据驱动的区域补排边界识别[J]. *水文地质工程地质*, 2022, 49(1): 1-11. DOI: 10.16030/j.cnki.issn.1000-3665.202104030.
- [25] HUSSAIN M M, BARI S H, TARIF M E, et al. Temporal and spatial variation of groundwater level in Mymensingh district, Bangladesh[J]. *International Journal of Hydrology Science & Technology*, 2016, 6(2): 188-197. DOI: 10.1504/ijhst.2016.075587.
- [26] ARABAMERI A, PAL S C, REZAIE F, et al. Modeling groundwater potential using novel GIS-based machine-learning ensemble techniques[J]. *Journal of Hydrology:Regional Studies*, 2021, 36: 100848. DOI: 10.1016/j.ejrh.2021.100848.

- [27] WANG X H, LIU T L, ZHENG X L, et al. Short-term prediction of groundwater level using improved random forest regression with a combination of random features[J]. *Applied Water Science*, 2018, 8: 125. DOI: [10.1007/s13201-018-0742-6](https://doi.org/10.1007/s13201-018-0742-6).
- [28] JIANG Z W, YANG S H, LIU Z Y, et al. Can ensemble machine learning be used to predict the groundwater level dynamics of farmland under future climate: A 10-year study on Huaibei Plain[J]. *Environmental Science and Pollution Research*, 2022, 29: 44653-44667. DOI: [10.1007/S11356-022-18809-8](https://doi.org/10.1007/S11356-022-18809-8).
- [29] CHEN Z Y, NIE Z L, ZHANG Z J, et al. Isotopes and sustainability of ground water resources, north China Plain[J]. *Ground Water*, 2005, 43(4): 485-493. DOI: [10.1111/j.1745-6584.2005.0038.x](https://doi.org/10.1111/j.1745-6584.2005.0038.x).
- [30] 刘敏, 聂振龙, 王金哲, 等. 华北平原地下水资源承载力评价[J]. *南水北调与水利科技*, 2017, 15(4): 13-18, 33. DOI: [10.13476/j.cnki.nsbdqk.2017.04.003](https://doi.org/10.13476/j.cnki.nsbdqk.2017.04.003).
- [31] 张兆吉, 雒国中, 王昭, 等. 华北平原地下水资源可持续利用研究[J]. *资源科学*, 2009, 31(3): 355-360. DOI: [10.3321/j.issn:1007-7588.2009.03.001](https://doi.org/10.3321/j.issn:1007-7588.2009.03.001).
- [32] 张宗祜. 华北平原地下水的历史和现状[J]. *自然杂志*, 2005, 27(6): 311-315. DOI: [10.3969/j.issn.0253-9608.2005.06.001](https://doi.org/10.3969/j.issn.0253-9608.2005.06.001).
- [33] 中国地质科学院水文地质环境地质研究所. 海河流域水文地质调查年度成果报告[R]. 2019.
- [34] 中国地质环境监测院. 地下水动态调查评价报告[R]. 2019.
- [35] 中国地质环境监测院. 国家地下水监测工程年度成果报告[R]. 2019.
- [36] 杨会峰, 曹文庚, 支传顺, 等. 近40年来华北平原地下水水位演变研究及其超采治理建议[J]. *中国地质*, 2021, 48(4): 1142-1155.
- [37] 狄胜同, 贾超, 张少鹏, 等. 华北平原鲁北地区地下水超采导致地面沉降区域特征及演化趋势预测[J]. *地质学报*, 2020, 94(5): 1638-1654. DOI: [10.19762/j.cnki.dizhixuebao.2020158](https://doi.org/10.19762/j.cnki.dizhixuebao.2020158).
- [38] 张宗祜, 沈照理, 薛禹群, 等. 人类活动影响下华北平原地下水环境的演化与发展[J]. *地球学报: 中国地质科学院院报*, 1997, 18(4): 337-344.
- [39] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32. DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).
- [40] 石建省, 李国敏, 梁杏, 等. 华北平原地下水演变机制与调控[J]. *地球学报*, 2014, 35(5): 527-534. DOI: [10.3975/cagsb.2014.05.01](https://doi.org/10.3975/cagsb.2014.05.01).
- [41] 卫磊, 杨桂莲, 鲁程鹏, 等. 华北平原超采区浅层地下水埋深变化及控制因素分析[J]. *水资源与水工程学报*, 2019, 30(6): 39-44. DOI: [10.11705/j.issn.1672-643X.2019.06.07](https://doi.org/10.11705/j.issn.1672-643X.2019.06.07).
- [42] 李文鹏, 邵新民, 祁晓凡, 等. 黑河中游盆地南部山区地下水对平原区侧向径流补给量的估算[J]. *水文地质工程地质*, 2022, 49(3): 1-10. DOI: [10.16030/j.cnki.issn.1000-3665.202202001](https://doi.org/10.16030/j.cnki.issn.1000-3665.202202001).
- [43] 裴源生, 李旭东, 赵勇, 等. 华北典型地下水大深埋区潜水层垂向补给特征及其给水度[J]. *南水北调与水利科技(中英文)*, 2020, 18(1): 176-193. DOI: [10.13476/j.cnki.nsbdqk.2020.0019](https://doi.org/10.13476/j.cnki.nsbdqk.2020.0019).
- [44] 朱静思, 王哲, 侯天宇, 等. 超采治理影响下的京津冀平原地下水动态演变特征[J]. *水文*, 2022, 42(5): 97-101. DOI: [10.19797/j.cnki.1000-0852.20210227](https://doi.org/10.19797/j.cnki.1000-0852.20210227).
- [45] 王晓光, 郭常来, 何海洋, 等. 地下水降落漏斗分类研究[J]. *地质与资源*, 2019, 28(5): 487-492. DOI: [10.13686/j.cnki.dzyzy.2019.05.012](https://doi.org/10.13686/j.cnki.dzyzy.2019.05.012).

Evolution and influence factors of shallow groundwater depression cone in Beijing-Tianjin-Hebei Plain

NAN Tian^{1,2}, CAO Wengeng^{1,2}, REN Yinguo³, SUN Long⁴, GAO Yuanyuan⁵

(1. *The Institute of Hydrogeology and Environmental Geology, CAGS, Shijiazhuang 050061, China*; 2. *Hebei Cangzhou Groundwater and Land Subsidence National Observation and Research Station, Shijiazhuang 050061, China*; 3. *Survey and Research Center of Hebei Province, Shijiazhuang 050061, China*; 4. *Information Center of the Ministry of Water Resources, Beijing 100053, China*; 5. *Bureau of South to North Water Transfer of Planning, Designing and Management, Ministry of Water Resources, Beijing 100038, China*)

Abstract: Since the 1960s, there is continuous groundwater exploitation in the North China Plain. With the rapid increase in water demand, groundwater overexploitation became an environmental geological problem. Recently, restrictions on groundwater exploitation and artificial groundwater recharge were developed to recover the groundwater level and remove the groundwater depression cone in Beijing-Tianjin-Hebei Plain. During the process of river ecological supplement, the recharge source of groundwater would be supplemented, and the water cycle mode could be changed. It is necessary to explain the groundwater depression cone evolution mechanism for accelerating the groundwater level recovery at this stage.

Numerical simulation is the traditional method to study the groundwater depression cone variation, but the model operation and construction are relatively complex. With the development of computer science, many machine-learning algorithms are proposed. Because of its simplicity and efficiency, machine learning models are widely used in the hydrogeological research field. Eight specified indicators have been selected to study the variation of groundwater depression cones, considering from natural factors, human activity factors, and hydrology factors. With these indicators, the feature variable data set is formed, and based on the feature variable data set, three typical machine learning models are developed to distinguish the variation of the groundwater depression cone. The logistic regression (LR) model and support vector machine (SVM) model are based on the traditional machine learning algorithm, and random forest (RF) model is a kind of ensemble algorithm based on the tree models. The established models were evaluated by sensitivity, specificity, and R^2 accuracy. The feature variable importance and shapely value were produced to quantify the contribution of each indicator to the groundwater depression cone and explain the behavior of each indicator.

The results showed that the RF model outperforms the LR and SVM models in terms of model performance. The sensitivity of the RF model was 0.94, the specificity was 0.78, and the R^2 accuracy was 0.88. It displayed that the RF model could be accurately identified both the groundwater depression cone area and the non-groundwater depression cone area. Model outputs suggested that the dominant influence indicator of the shallow groundwater depression cone was groundwater exploitation. Before 2018, the influence degree of groundwater pumping on the depression cone was about 50%. It played a positive role in the development of the groundwater depression cone. The river artificial recharge took 16% account for the variation of shallow groundwater depression cone development after 2018, and it had an obvious contribution to the groundwater level recovery. Two typical areas (Ningbailong area and Gaoliqing area) were selected to explore the evolution mechanism of groundwater depression cones in different regionals. The simulation results of the Ningbailong area and Gaoliqing area showed that the Ningbailong groundwater depression cone was governed by both precipitation and groundwater exploitation, the contribution rates for each indicator were 24% and 25%, respectively. Groundwater pumping dominated the development of the Gaoliqing groundwater depression cone, and it took 85% account for the evolution of the groundwater depression cone.

In summary, three different data-driven models were constructed to study the variation of shallow groundwater depression cones in the whole North China Plain and two typical areas. The RF model was the optimal model. It was suitable for identifying the groundwater depression cone. The main control factor of the shallow groundwater depression cone was groundwater artificial exploitation. The river's artificial recharge could take an obvious positive impact on the recovery of groundwater level in the Ningbailong area. But it had little effect in the Gaoliqing area. Therefore, restrained groundwater exploitation by replacing agricultural groundwater could be the crucial way to restore groundwater depression in the Gaoliqing area.

Key words: Beijing-Tianjin-Hebei Plain; groundwater depression cone; multisource data driven model; machine learning; evolution mechanism