

中文核心期刊 CSCD核心期刊 中科双效期刊 中国科技核心期刊 Caj-cd规范获奖期刊

地下水机器学习方法研究

齐永强,李文鹏,郑跃军,李 慧,王成见

Application of machine learning to aquifer analysesLocating hydrogeological boundaries with water table monitoring data QI Yongqiang, LI Wenpeng, ZHENG Yuejun, LI Hui, and WANG Chengjian

在线阅读 View online: https://doi.org/10.16030/j.cnki.issn.1000-3665.202104030

您可能感兴趣的其他文章

Articles you may be interested in

运用地下水对潮汐的响应识别压力传导系数

Identification of hydraulic conductivity using the response of groundwater levels to oceanic tide 阳玲, 杜金月, 王同科, 赵志学, 郝永红 水文地质工程地质. 2019, 46(6): 26-30

区域地下水污染调查取样点布设量化分配方法

Quantitative sample distribution in regional groundwater contamination investigation 李颖智, 蔡五田, 耿婷婷, 张涛, 张智印 水文地质工程地质. 2019, 46(5): 24-30

西北地区地下水水量--水位双控指标确定研究——以民勤盆地为例

A study of the determination of indicators of dual control of groundwater abstraction amount and water table in northwest China: a case study of the Minqin Basin

王晓玮, 邵景力, 王卓然, 崔亚莉, 张秋兰 水文地质工程地质. 2020, 47(2): 17-24

江汉平原东北部地区高铁锰地下水成因与分布规律

Distribution and genesis of high Fe and Mn groundwater in the northeast of the Jianghan Plain 蔡玲, 胡成, 陈植华, 王清, 王宁涛, 常威, 黄琨 水文地质工程地质. 2019, 46(4): 18-25

伊犁河支流大西沟河水与地下水转化关系研究

A study of the conversion between the Daxigou river and groundwater 张琛, 段磊, 刘明明, 李瑛, 宋浩 水文地质工程地质. 2019, 46(3): 18-18

赣江新干航电枢纽左库岸地下水浸没控制效果研究

Research on the control effects of groundwater immersion on the left reservoir bank of Xingan Navigation and Power Junction in Ganjiang River

张宇,任国澄,杨蕴,周志芳,郭生根,熊鸿强 水文地质工程地质. 2020, 47(3): 147-155



关注微信公众号,获得更多资讯信息

DOI: 10.16030/j.cnki.issn.1000-3665.202104030

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.

QI Yongqiang, LI Wenpeng, ZHENG Yuejun, *et al.* Application of machine learning to aquifer analyses: Locating hydrogeological boundaries with water table monitoring data[J]. Hydrogeology & Engineering Geology, 2022, 49(1): 1-11.

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

齐永强¹,李文鹏²,郑跃军²,李 慧¹,王成见³ (1. 北京北水国际科技有限公司,北京 102206;2. 中国地质环境监测院,北京 100081; 3. 青岛市水文局,山东 青岛 266071)

摘要:随着我国地下水监测工作的高速发展,高频率高密度水位监测数据的出现催生了对其进行深入信息挖掘的需求。 在传统地下水模型研究中,地下水水位监测值常位于模型构建过程的下游,当水位监测的时空密度逐渐增大时,新增信息 无法有效传导至模型的规划阶段并指导概念模型的修订。文章提出了一种地下水系统补排边界的识别方法,在不建立地 下水数值模型的前提下,以监测井空间位置为节点,按照德劳内原则建立三角网格。在此网格系统中,首先定义一个水力 梯度变换函数 grad*F*,以求取网格中任意位置的水力梯度;借鉴机器学习领域的优化算法,使用水力梯度场驱动含水层中随 机分布质点的运行轨迹,并以此推断和识别区域内地下水补给和排泄边界。在环境地学计算平台 EnviFusion-CGS 中实现, 并构建了详细工作流程。以山东省青岛市大沽河中下游含水层为示范区,对含水系统的补给区和排泄区的空间分布及其 动态变化进行了分析,取得了良好效果。本研究为构建和修订已有含水层概念模型提供了新思路。 关键词:高密度;水位监测;概念模型;地下水;补给边界;排泄边界

中图分类号: P641.2 文献标志码: A 文章编号: 1000-3665(2022)01-0001-11

Application of machine learning to aquifer analyses: Locating hydrogeological boundaries with water table monitoring data

QI Yongqiang¹, LI Wenpeng², ZHENG Yuejun², LI Hui¹, WANG Chengjian³

(1. Beijing Water International Co. Ltd., Beijing 102206, China; 2. China Institute of Geo-Environment Monitoring, Beijing 100081, China; 3. Hydrological Bureau of Qingdao, Qingdao, Shandong 266071, China)

Abstract: Groundwater level fluctuations in China are being monitored with unprecedented frequency and density, which drives the need for mining such types of data. In a typical aquifer analysis project, groundwater level data is generally applied after the completion of the aquifer conceptual framework. When the temporal and spatial density of groundwater level data gradually increases, the information gain needs to be effectively transformed into conceptual knowledge of the model. In this study, we propose a method to identify hydrological boundaries based on the groundwater level monitoring data. In this method we discretize space into a triangular mesh using monitoring wells as the initial nodes, and a transformation function gradF is defined to calculate the hydraulic gradient at any given location on the mesh. The hydraulic gradient field is subsequently use to drive an array of

randomly scattered particles to obtain the streamline representation of the flow field, which will in turn serve as the basis for deducing and refining the recharge and discharge boundaries of a hydrogeological domain. This method is implemented into the geo-environmental scientific computation platform (EnviFusion-CGS), and a detailed work flow is developed to facilitate the development of the aquifer conceptual model. This method is applied to the hydrogeological investigation of the Dagu aquifer located in Qingdao of Shandong Province, where the spatial distributions and dynamic fluctuations of the hydrogeological boundaries are identified.

Keywords: high density; groundwater level monitoring; conceptual model; groundwater; recharge boundary; discharge boundary

近年来我国地下水监测工作取得了长足的进步, 获得了高频率高密度的水位监测数据,是极佳的地下 水系统信息来源。但目前监测数据未得到充分利 用^[1-2]。在传统的地下水模型应用中,水位监测数据 处于从属地位,与区域概化所得的水文地质边界以及 水文地质参数共同决定模型的最终表现。随着水位 监测的密度、频率和精度不断提升,监测数据中蕴含 的信息可以更多更好地指导水文地质边界的识别和 水文地质参数的确定^[3]。本文应用了一种机器学习方 法,用于提取区域地下水补排边界。

地下水工作者常借助地下水流模型系统整合已 有的含水层信息,认识含水层行为,测试溶质运移假 说^[4-5]。地下水流模型模拟是使用解析方法或数值方 法对饱和带水分运移控制方程求解的过程。水流模 型的计算结果是在给定的初始条件和边界条件下,水 头在空间各处和时间各点上的分布¹⁶。一般使用数值 方法对控制方程求解,因为解析方法求解时对初始条 件和边界条件有严格的限制,而客观世界中的情况极 为复杂,只有数值方法能够处理与之相对应复杂性的 模拟问题^[7]。由于地下水数据存在长期而普遍的稀缺 性,模型工作者需要对模型结构和参数进行插值和调 整。对这一过程的校准通常通过试错法进行,即反复 调整模型中指定的输入参数,直到水头的模拟值与监 测值达到一致。在数据和参数量较大的情况下,也可 以使用专门为参数识别而设计的计算程序实现自动 反演,如李家兴^[8]利用 GAN 实现了水头场与底层参 数场间的双向预测,李竞生等¹⁹¹利用遗传算法尝试识 别含水层参数。

在传统地下水模型研究中,地下水水位监测值常 位于模型构建过程的下游,其作用主要体现在模型校 正过程中对模型参数的调整。这主要是由于之前地 下水水位监测数据无论从时间密度或空间密度均存 在严重短缺^[5],其中蕴含的信息不足以勾画含水层的 全貌。相应地,研究人员在实际利用地下水水位监测数据之前,一般需要针对系统补给和排泄边界和含水 层参数分布作出一系列假设,即构建概念模型。但是 假设中不符合实际情况的错误信息,将在模型中不断 传导和放大,造成模型的多解性^[10]。单纯使用地下水 水位的实测数据修正和微调地下水模型不仅效率较 低,而且容易陷入循环论证的困局。本文旨在于建立 一套可行的技术方法,能够直接使用高密度、高频率 的地下水水位监测数据对区域内地下水补给和排泄 边界进行识别,帮助构建和修正含水层概念模型。

机器学习是人工智能的一个分支,其核心目的是 从大量数据中自动分析获得规律^[11],近30年已发展 成为一门多领域交叉学科,涉及概率论、统计学^[12]、 计算机图形学、逼近论、凸分析^[13]、计算复杂性理论 等多门学科。机器学习通常基于神经网络架构,以认 知科学、神经科学等为依托,设计类脑智能的学习模 型^[14-15]。机器学习常分为监督学习和无监督学习。 从高密水位数据中提取地下水补给和排泄边界属于 模式识别,是一种形式的无监督学习。无监督学习指 从现有的训练集出发,经过训练得出一定结果。与监 督学习不同,无监督学习不需要人为标注,仅着重于 捕捉和描述数据中存在的模式和规律^[16]。机器学习 与基于物理逻辑的传统模型不同,是以统计与算法为 基础,将逻辑搭建交由神经网络自行生成,旨在挖掘 数据中无法人为识别的信息与信号^[17]。

本文介绍了一种地下水系统补排边界的识别方法,其核心是在三维空间中构建研究区域,针对区域上的任意点判断其落入补给区或排泄区的概率,引入机器学习处理优化问题中的梯度下降^[18-19]。梯度下降法由法国数学家 Cauchy^[20]发明,由 Robbins 等^[21]重新提出,后由 Rumelhart 等^[22]发扬并运用于机器学习领域。近十几年来,机器学习受到越来越多的学者与业界关注^[23-24],梯度下降算法也得到了广泛的研究与长

足的发展^[25-26], 被应用于约束优化^[27]、正则优化^[28-29] 和函数复合^[30]等不同的场景中。机器学习实践中, 常 用损失函数描述模拟结果和实测数据的差异, 在指定 数据集上, 损失函数越小, 模型预测越准确。为了确 定何种参数组合下的损失函数最小, 可通过参数在损 失函数的"场"中向损失函数值减小的方向移动。其 中损失函数值下降最快的方向称为负梯度方向, 使用 的算法称为梯度下降法, 是机器学习中最常见的优化 算法。

梯度下降法与地下水的流动具有天然的相似 性。在地下水流场中的任意质点,如果允许其沿地下 水流动,在下一时刻这一质点位于流场排泄区的几率 增大;反之如果质点向地下水上游流动,下一时刻位 于流场补给区的几率增大。这一理念在地下水动力 学领域已有应用,最为典型的是美国地质调查局 (USGS)开发的 MODPATH 质点追踪程序包^[31]。然而 MODPATH 必须应用在已经求解的 MODFLOW 流场 中,对于仅有水位监测信息的情形并不适用。本文描 述的方法将尝试解决此问题。

1 研究方法

1.1 地下水监测网空间剖分

地下水水位监测网络的核心是三维空间中的一 系列散点类水头值,位于每口监测井筛管的中心位 置。若要通过内插法得到模型范围内任意点的水头 值,需首先为散点建立空间网格,描述它们之间的拓 扑关系。由于在地下水区域监测网络中,水平方向的 长度(*X*、*Y*方向)一般远大于垂直方向(*Z*方向),将三 维散点沿*Z*轴投影到大地平面上,使用德劳内三角化 规则为二维平面上的点集建立三角网格。德劳内三 角化是计算几何中常见的三角网格建立规程,其核心 为不应有任何一个顶点在三角形网的任一三角形外 接圆内部,从而避免产生狭长三角形(图1)。



注:(a)不符合德劳内的三角形剖分,右侧顶点出现在左侧三点的外接圆之内;(b)重新选择剖分对角线后,符合德劳内规则;(c)按照德劳内规则在 XY平面上生成的三角网格;(d)将网格投影至三维空间后的最终网格。

1.2 求取水力梯度

设三角网格 M上的点集合为 V, 三角集合为 F, 函数 h代表三角网格节点(监测井位)上的水位值, 即 $h \in \mathbb{R}^{|V|}$ 。当集合 V中的 h值均已知时, 可以在集合 F上求取函数 h的梯度。定义一个针对三角集合 F的 水力梯度变换算子 gradF, 用来将集合 V上的水位值 映射为集合 F上的水力梯度值:

$$\operatorname{grad} F: \mathbb{R}^{|V|} \to \mathbb{R}^{3|F|} \tag{1}$$

 $\operatorname{grad} F \in \mathbb{R}^{3|F| \times |V|}$

式中:gradF——水力梯度变换算子。

对于F集合中的一个三角形j, gradF变换为:

$$(\operatorname{grad} Fh)(j) = \frac{1}{2A(j)} \sum_{i=1}^{3} h_i T \overrightarrow{e_{ji}}$$
(2)

 h_i ——三角形 j 的第 i 个顶点上的水位值;

ē_{ji} ── 第*i* 个顶点的对边,即除顶点*i* 之外其余 两个顶点沿逆时针方向组成的向量;

T*e*^{*j*}_{*j*}^{*j*} → *e*^{*j*}_{*j*}向量沿三角形平面旋转 90°后的向量, 可由*e*^{*j*}_{*j*}向量与三角形*j*的单位法向量做 内积所得。

$$\boldsymbol{T}\overrightarrow{\boldsymbol{e}_{ji}} = \begin{bmatrix} (\boldsymbol{T}\boldsymbol{e}_{ji})_1 \\ (\boldsymbol{T}\boldsymbol{e}_{ji})_2 \\ (\boldsymbol{T}\boldsymbol{e}_{ji})_3 \end{bmatrix} = \overrightarrow{\boldsymbol{n}_j} \times \overrightarrow{\boldsymbol{e}_{ji}}$$
(3)

$$\vec{n}_{j} = \frac{\vec{e}_{j1} \times \vec{e}_{j2}}{\|\vec{e}_{j1} \times \vec{e}_{j2}\|}$$
(4)

式中: **n**_j — 三角形 j 的单位法向量, 其长度为 1, 方向 与三角形 j 垂直;

 $\vec{e_{\mu}}$ 和 $\vec{e_{\mu}}$ ——三角形 *j* 的任意两边。

综上,可以将 gradF 变换的矩阵形式写为:

$$\begin{bmatrix} (\operatorname{grad}Fh)(j)_{1} \\ (\operatorname{grad}Fh)(j)_{2} \\ (\operatorname{grad}Fh)(j)_{3} \end{bmatrix} =$$

$$\frac{1}{2A(j)} \begin{bmatrix} (\mathbf{T}e_{j_{1}})_{1} & (\mathbf{T}e_{j_{2}})_{1} & (\mathbf{T}e_{j_{3}})_{1} \\ (\mathbf{T}e_{j_{1}})_{2} & (\mathbf{T}e_{j_{2}})_{2} & (\mathbf{T}e_{j_{3}})_{2} \\ (\mathbf{T}e_{j_{1}})_{3} & (\mathbf{T}e_{j_{2}})_{3} & (\mathbf{T}e_{j_{3}})_{3} \end{bmatrix} \begin{bmatrix} h_{1} \\ h_{2} \\ h_{3} \end{bmatrix}$$

$$(5)$$

假设网格 *M* 中以某点 *N* 为顶点的三角形有 *k* 个, 顶点 *N* 处的水力梯度为此 *k* 个三角形中水力梯度最 大值。定义一个针对点集合 *V* 的水力梯度变换算子 grad *V*,用来为顶点 *V* 集合上的任一顶点 *N* 生成唯一的 水力梯度值:

$$(\operatorname{grad}Vh)(N) = \max_{i \in \mathcal{A}} (\operatorname{grad}Fh)(j_i)$$
 (6)

式中: gradV——针对网格顶点 N的水力梯度变换;

 j_i ——经过顶点 N 的所有三角形。

1.3 补给区和排泄区的定义与识别

地下水补给区和排泄区做如下定义:在同一含水 层内部指定随机分布的一系列质点,质点沿水位场梯 度向水位较低的方向运动,在给定时间后,质点密度 较高的区域称为排泄区;反之,质点沿梯度向水位较 高的方向运动指定时间后,质点密度较高的区域称为 补给区。

在三角网格 M 内定义任意质点,其空间位置属性 p 随时间 t 的变化可以表示为:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = f(t,p) \tag{7}$$

$$p(t_0) = p_0 \tag{8}$$

式中:p---质点的空间坐标;

t---质点时间。

此处时间为计算质点轨迹过程中使用的时间项, 应与地下水监测网的实际监测时间进行区别。

在计算质点轨迹时,本方法假定流场为稳定态,即在计算期间不发生变化。在初始质点时间 t_0 ,质 点位置为 p_0 。对于给定质点时间步长 $\Delta t > 0$,Runge-Kutta^[32-33]给出了质点在未来时间位置的近似值:

$$p_{n+1} = p_n + \frac{1}{6}\Delta t \left(k_1 + 2k_2 + 2k_3 + k_4\right) \tag{9}$$

$$t_{n+1} = t_n + \Delta t \tag{10}$$

$$k_1 = f(t_n, p_n) \tag{11}$$

$$k_2 = f\left(t_n + \frac{\Delta t}{2}, p_n + \Delta t \frac{k_1}{2}\right) \tag{12}$$

$$k_3 = f\left(t_n + \frac{\Delta t}{2}, p_n + \Delta t \frac{k_2}{2}\right) \tag{13}$$

$$k_4 = f(t_n + \Delta t, p_n + \Delta t k_3) \tag{14}$$

如果 *p*(*t*_{n+1}) 是质点在下一质点时间步的位置, *p*_{n+1} 即为 *p*(*t*_{n+1}) 的四阶近似值。*p*_{n+1} 完全由当前质点 位置 *p*_n 和其位置函数 *f* 共同决定。

地下水中沿流场运行的质点,其位置函数由达西 定律确定:

$$f(t,p) = \frac{\mathrm{d}p}{\mathrm{d}t} = v = \frac{q}{n_{\mathrm{e}}} = \frac{KI}{n_{\mathrm{e}}} \tag{15}$$

式中:v——质点流速;

q——达西流速;

- ne---介质有效孔隙度;
- K——含水层渗透系数;
- *I*——本含水层的水力梯度。

在饱和渗流介质理论体系中,K值和n。值均为不随时间变化的固定值。在含水层的任一瞬时,一个质 点的运动轨迹仅与其轨迹上的I值有关。假设一个流场的 I值在给定时刻为已知,可以根据以上推导求取质点 迹线,只要使用的Δt值足够小,这条迹线的形态即为 固定值,与K值和n。值无关。若要求取地下水流场中 任意质点的迹线形态,理论上仅需流场的水力梯度(I) 的分布函数,不需要含水层的渗透系数和孔隙度参数。 1.4 应用实现

对于同一含水层中任意给定的一组观测井,可以 使用观测井筛管的中点生成一组点集合 V,在 XY 平面 上为点集 V 构造符合德劳内规则的三角网格,将网格 重新投影至三维空间中即形成三角网格曲面 M, M 由 点集 V 和三角集合 F 组成。在给定的时间切片上,点 集 V 上的水头 h 值均为已知。在曲面 M 上任意点的 水力梯度 I,均可以通过 gradF(三角形内)或 gradV(三 角形顶点处)计算得到。根据水力梯度在 M 上的分布 函数,可以计算曲面上任意质点的近似运动迹线,监 测井的空间分布越密集,质点运动的质点时间步长越 短,此近似运动迹线就越接近真实迹线。并将计算逻 辑在环境地学计算平台 EnviFusion-CGS 中实现。

EnviFusion-CGS 是中国地质环境监测院和北水国际联合开发的环境地学可视化计算平台,用于环境地 学领域基于数据的融合建模和知识发现。可针对多 种环境要素(土壤、地下水、地表水、大气等)的不同 维度数据(现场分析、仪器测试、数值模拟、地球物理 等)进行融合分析。EnviFusion-CGS 将本学科多源异 构数据统一到四维时空的多组分属性数据体中,帮助 研究人员以所见即所得的形式对海量数据进行可视 化分析和统计挖掘。

2 示范区概况

大沽河是山东半岛最大的河流。向南流经烟台 市的3个区县和青岛市的5个区县,长约179km,最 后注入山东半岛南部沿海岸的胶州湾。大沽河含水层 位于大沽河流域的中下游,距离青岛市城区约50km, 面积约420km²。大沽河含水层由大沽河流域冲积的 松散沉积地层形成。在大沽河流域,农业灌溉是用水 量最大的项目,同时也是区域经济可持续性发展的重 要因素^[34]。农业和轻工业对于大沽河含水层的经济 2.1 水文地质条件

大沽河流域位于山东半岛东部地区,地势较低, 地表向南倾斜延伸至海岸线,古岘镇是该流域的分水 岭。古岘镇北部是由花岗岩和变质岩构成的低矮山 丘,海拔高度为50~200 m;大泽山海拔高度达736 m, 为该地区最高的山。古岘镇的南部是胶莱盆地平原 地区,该地区的平均海拔高度为20~50 m。大沽河附 近的平原曾是一片广阔的山谷,山谷宽度达6 km,后 经古代冲积物填充。山谷的底部由白垩系砂岩和页 岩基岩组成,为承压含水层底板。山谷东西边界被冲 积物覆盖,目前很难鉴定地貌。山谷中充满了砂质冲 积物,是大沽河含水层的主要组成部分[35-36]。

潜水含水层由 2 个含水层组成:上部含水层(由黏 质砂土和砂质黏土组成,渗透系数较低)和下部含水 层(由渗透系数和孔隙度比较大的砂卵砾石组成)。 地下水的总体流向为由东北向西南。大沽河含水层 的上中下游分布有一系列的横断面,见图 2。包气 带厚度大小不一,为1~6m。包气带由粉质黏土,黏 质粉土以及砂土组成。南村水文监测站的北部地 区(该监测站位于大沽河含水层的中游)包气带为较 薄的砂层,是降雨入渗的天窗,入渗面积达 9.2 km²。 大沽河含水层渗透系数为 2.3~159.3 m/d,给水度为 0.05~0.18^[37-38]。





2.2 地下水位波动

青岛市大沽河中下游地下水系统监测始于 1995 年。2008—2010年青岛水文局建立了地下水水位自 动监测网,在大沽河含水层范围内共建设 147 口自动 监测井,对于研究农业灌溉抽水与含水层储变量关系 至关重要。区域内春旱一般和春季灌溉同时发生,结 果导致地下水水位严重下降。在灌溉的峰值季节,地 下水水位呈现下降趋势,当 8—9 月降水丰富时,水位 逐渐回升。入秋之后,水位逐步下降。12 月至翌年 2 月农业灌溉较少,地下水水位相对稳定。全年地下 水水位波动一般在3m之内。

地下水资源的长期动态取决于气候变化周期和 人类活动。根据 1975—1995 年地下水位连续观测结 果,地下水资源经历了以下几个阶段:

(1)小幅下降期(1975-1980年)

该时期是地下水开采的初期,地下水水位开始逐 年下降,降幅约为 0.1 m/a。

(2)大幅下降期(1980-1984年)

该时期是地下水开采的高峰期,地下水水位大幅 下降,降幅约为0.5~1.0m/a。 (3)迅速补给期(1985年)

第九号台风登陆,青岛市连续4d暴雨,地下水水 位迅速上升,平均升幅达到了2.05m。

(4)大幅下降期(1986-1989年)

由于持续的干旱天气,这一时期的地下水水位又 开始急剧下降,降幅约为 0.5 m/a。

(5)补给和稳定期(1990-1995年)

1990年青岛市降水量达到了 900 mm,超出该地 区多年平均降水量的 50%。地下水也在这一年得到 了迅速补给。随着 1989年引黄济青工程正式通水,青 岛市对大沽河含水层的依赖明显减小。

2.3 地下水补给

大沽河含水层的补给来自降水、农业灌溉、河流 入渗以及少量的侧向补给。7—9月的降水量约占全 年降水量70%,因此这一时期也是地下水补给的高峰 期。贾立华^[36]研究了大沽河地下水污染脆弱性计算 得到了大沽河补给地下水的数值。利用 DRASTIC 绘 制的大沽河含水层显示约有10%的含水层在这一时 期的补给量达到约200 mm,部分地区高达400 mm。 该区有一个高入渗天窗,面积为9.2 km²,也是地下水 补给的重要来源之一^[36-38]。

2.4 地下水排泄

大沽河含水层是地下水埋深较浅的潜水含水 层。其砂质含水层的特性使得地下水和地表水相互 作用密切。在 20 世纪 80 年代,地下水被大规模开采 之前,由于淤泥层和黏土层的存在,局部地区可能是 半承压含水层。随着季节交替,河流补给地下水或地 下水排泄至河流。地下水开采时,含水层变成典型的 潜水含水层。随着地下水水位下降,河流下渗补给地 下水成为地下水和地表水之间交换的主导作用。从 此,大沽河含水层以垂向循环为主,通过降雨或者灌 溉补给;通过抽水和蒸散排泄。当地地下水动态存在 季节性和年际变化^[39-40]。

3 模型构建与讨论

大沽河含水层中地下水资源的补给和排泄受到 人类经济生产活动的深刻影响。区域内大范围的农 业种植区域依赖地下水进行灌溉,而且大沽河沿线存 在大量水利工程构筑物。到 2003 年,沿河道建有 8 个 中型水库,90 个小型水库和 1 223 座闸门,对地下水的 循环路径与模式产生了巨大影响。2013 年 1 月 1 日 - 12 月 31 日为研究期,在大沽河含水层范围内的 147 口监测井中,选取了 78 口具有连续逐日水位监测 数据的监测井作为研究对象(其余监测井均存在不同 程度的数据缺失)。仅使用监测井的空间位置信息和 连续监测水位数据,在不作任何概念模型假设的前提 下,使用机器学习算法对含水层的补给和排泄区域进 行识别。

3.1 网格构建

根据 78 口监测井的 X、Y坐标,基于德劳内规则,



构建大沽河含水层的三角网格,如图 3 所示。生成的 三角网格范围,见图 3(b),与传统水文地质勘察工作 所确定的"含水层边界",见图 3(a),略有不同。这是 因为本算法仅具有内插功能,对于没有落在三角网格 范围内的区域无法进行预测。当三角形网格过于粗 糙时,会导致质点运动迹线不够平滑难以收敛,对图 3 (b)中的三角网格进行加密,见图 3(c)。三角网格的 加密仅为增加空间分辨率,不会增删原始网格中所携 带的信息,其算法较为普遍,在此不再赘述。

3.2 水力梯度场计算

在图 3(c)所示的三角网格中,根据网格节点处的 水位数值,计算网格三角形内部和节点上的水力梯度 值。由于输入数据为逐日监测数据,所以分别计算 365 d 的水力梯度场,见图 3(c)。含水层北部靠近山 区,受地形坡度影响,水力梯度普遍较南部大。此外, 靠近大沽河河床的位置由于含水层透水性较好,水力 梯度普遍较小。

3.3 迹点与迹线计算

在图 3(c)所示的三角网格中,首先为一定数量的 质点随机给定初始位置,按照追踪手段依次绘制质点 在下一个时间步长的位置并连成迹线,见图 4(a)。初 始质点的数量并没有具体的规定,但在计算能力允许 的前提下尽量提高质点密度有助于更好地利用机器 学习算法的优势。迹线的分布定性展示了流场的排 泄方向。为了排泄区域定量化,在*XY*二维空间中划



分了均匀方形网格,见图 4(b)。依次计算方形网格中 每一个网格周边迹点数量,并以此计算迹点的相对密度。 按照排泄迹点密度大小得到迹点密度图,见图 4(b)。

在普遍意义上,此密度分布并不能直接等同于含水层的排泄区,因为水力梯度I的分布函数仅反映潜在的渗流方向,而实际流速受渗透系数K和介质有效 孔隙度n。的共同约束。当含水层内存在强烈影响渗透 系数或有效孔隙度分布的地质异常(如阻水构造)时, 排泄迹点的分布不能用来推断排泄区的分布。反之, 在大沽河含水层此类含水介质分布较为稳定,渗流特 征较为均一的地区,此密度分布图可以看作含水层排 泄区域的近似表达。可以看出含水层的地下水排泄 基本以大沽河河道为主轴。偶有偏移的原因可能是: (1)傍河开采井的汇流作用;(2)古河道的地下渗流作 用;(3)监测井密度不足造成排泄区形态失真。

类似地,将水力梯度方向反转,可以求取地下水 的补给迹线和迹点密度分布图,见图 5。含水层内的 地下水补给作用明显集中在北侧的上游地区,含水层 西北侧古岘镇一高岚村一线存在明确的分水岭,此线 以西的地下水向西流动,以东的地下水向东流动。含 水层东部边界存在明显的地下水流入,最终被大沽河 吸纳。此外,含水层的南部底端存在局部的补给区 域,与棘洪滩水库位置重合。



3.4 地下水补排规律的时间变化 为考察本含水层补给区和排泄区随时间的变化 规律,使用 EnviFusion-CGS 中内置的自动化参数设置,计算 2013 年每月 1 日计算补给区和排泄区的分布 图,见图 6。2013 年含水层补给区与排泄区分布总体 变化不大,其显著特征(西北部分水岭、大沽河道排泄 带、南端水库补给等)均维持稳定状态。含水层与东 侧边界的沟通方式在雨季来临后(8月)发生一定变 化,其补给区与排泄区的分布位置有一定调整。这一 变化与大沽河沿线分布的橡胶坝(图 2)有关,雨季河 道蓄水后,扰动了地下水流场,从而改变了补给和排 泄的形态。

3.5 含水层边界识别与细化

图 2、图 3 和图 6 中的外侧轮廓线为以往研究人员为大沽河含水层划定的边界,此边界是融合了钻孔数据、地质信息、水位统测和模型研究的综合成果,长期以来指导了大沽河含水层的各项水资源研究,具有重要的历史意义。但本次研究进一步改进和细化:

(1)含水层西北部存在较为明确的分水岭,见图 5 (b),其形态显著且稳定。在 2013 年全年未见其东西



Fig. 6 Recharge and discharge areas of the study site (first days of each month, 2013)

两侧的含水系统存在沟通,结合更长期的地下水水文 监测数据判定后,考虑将含水层边界向东调整至分水 岭处。

(2)含水层东北侧边界在历史研究中常考虑为隔 水边界,但本研究显示其可能存在侧向径流补给。应 考虑在此区域收集更多监测井数据,首先排除由监测 点稀疏造成的潜在误判;若确有外源流入,应向东查 找稳定的地下水分水岭,扩展含水层东侧边界,或者 在模型研究中将其定义为侧向流入边界。

(3)大沽河上橡胶坝的蓄水状态对区域流场的扰动在雨季和旱季存在差别,后续研究中应对其进行重 点考察。

3.6 方法局限与适用条件

本方法的应用前提是空间上高密度的水位监测 点分布,但含水层空间异质程度与监测点密度需求之 间尚无经验,亟待大量的实证研究,进行深入的探索 和进一步细化。

本方法将地下水流场 f(t,p)简化为梯度场 I,并假 定质点沿梯度最大的方向前进。这一假设是建立在 K_{ne}的数值分布较为均匀的前提下,在渗透系数连续分 布的单一含水层中大多可以满足。但若含水层渗透 系数分布存在强烈的各向异性时(如裂隙含水层),这 一前提将不复存在。假设 2 个相邻监测点分别属于 2 个垂向分隔的含水层,当其间存在水头差时,虽然水 力梯度客观存在,但由于 2 点间渗透系数的分布存在 强烈的各向异性(水平渗透系数远大于垂向渗透系 数),其间不会存在真实的地下水流动。又如在隔水 断层两侧水力梯度固然存在,但质点不会沿梯度运行 到断层对侧,所以由此推断的补给区和排泄区有可能 存在偏差。

如果含水层中存在渗透系数强烈各向异性的情形,其必然结果是在这类边界附近的质点运行速率(水力梯度)远大于含水层的平均值。EnviFusion-CGS中记录了所有位置的水力梯度和全部质点的运行速率,可以据此首先对渗透系数突变进行筛查,如果确实存在此情形,可以将含水层拆分后分别进行分析,或者与传统水文地质勘察和模型研究配合进行进一步的信息迭代。

4 结论

(1)高频率高密度水位监测数据是极佳的地下水 系统信息来源。在不建立地下水数值模型的前提下, 以监测井空间位置为节点,通过建立德劳内三角网格、求取网格中任意位置的水力梯度数值、借鉴使用机器学习领域中的优化算法,推断和识别区域内地下水补给和排泄边界。

(2)使用环境地学计算平台 EnviFusion-CGS 构建 了识别研究区域补给区和排泄区空间信息的详细工 作流程,帮助地下水科学家构建和修订已有的含水层 概念模型。

(3)以山东省青岛市大沽河中下游含水层作为示范 区,使用本文提出的方法对含水系统补给区和排泄区 的空间分布及其动态变化进行分析,取得了良好效果。

参考文献(References):

- [1] 严宇红,周政辉.国家地下水监测工程站网布设成果综述[J].水文,2017,37(5):74-78. [YAN Yuhong, ZHOU Zhenghui. Introduction to network layout of national groundwater monitoring project[J]. Journal of China Hydrology, 2017, 37(5):74-78. (in Chinese with English abstract)]
- [2] 田志仁,李名升,夏新,等.我国地下水环境监测现状和工作建议[J].环境监控与预警,2020,12(6):1-6.
 [TIAN Zhiren, LI Mingsheng, XIA Xin, et al. Current situation and suggestions of groundwater monitoring work in China[J]. Environmental Monitoring and Forewarning, 2020, 12(6):1-6. (in Chinese with English abstract)]
- [3] 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.
- [4] 薛禹群,吴吉春.地下水数值模拟在我国——回顾与展望[J].水文地质工程地质,1997,24(4):21-24.
 [XUE Yuqun, WU Jichun. Numerical simulation of groundwater in retrospect and prospect in China[J]. Hydrogeology & Engineering Geology, 1997, 24(4):21-24. (in Chinese)]
- [5] 薛禹群.中国地下水数值模拟的现状与展望[J].高校 地质学报, 2010, 16(1): 1-6. [XUE Yuqun. Present situation and prospect of groundwater numerical simulation in China[J]. Geological Journal of China Universities, 2010, 16(1): 1-6. (in Chinese with English abstract)]
- [6] 中国地质调查局.水文地质手册[M].2版.北京:地质 出版社,2012. [China Geological Survey. Handbook of hydrogeology[M]. 2nd ed. Beijing: Geological Publishing House, 2012. (in Chinese)]

- [7] 易立新,徐鹤.地下水数值模拟[M].北京:化学工业出版社,2009. [YI Lixin, XU He. Numerical simulation of groundwater[M]. Beijing: Chemical Industry Press, 2009. (in Chinese)]
- [8] 李家兴.基于深度学习的含水层参数反演研究[D].北京:中国地质大学(北京), 2020. [LI Jiaxing. Research on inverse problem of aquifer parameters based on deep learning[D]. Beijing: China University of Geosciences, 2020. (in Chinese with English abstract)]
- [9] 李竞生,姚磊华. 含水层参数识别方法[M]. 北京: 地质 出版社, 2003. [LI Jingsheng, YAO Leihua. Identification method of aquifer parameter[M]. Beijing: Geological Publishing House, 2003. (in Chinese)]
- [10] 何清,李宁,罗文娟,等.大数据下的机器学习算法综述[J].模式识别与人工智能,2014,27(4):327-336.
 [HE Qing, LI Ning, LUO Wenjuan, et al. A survey of machine learning algorithms for big data[J]. Pattern Recognition and Artificial Intelligence, 2014, 27(4): 327-336. (in Chinese with English abstract)]
- [11] 刘光亚.水文地质数值模拟的拟合问题[J].水文地质 工程地质,1992(4):2-4. [LIU Guangya. The fitting problem of hydrogeological numerical simulation[J]. Hydrogeology & Engineering Geology, 1992(4):2-4. (in Chinese)]
- [12] 张双圣,强静,刘汉湖,等.基于贝叶斯公式的地下水 污染源识别[J].中国环境科学,2019,39(4):1568-1578. [ZHANG Shuangsheng, QIANG Jing, LIU Hanhu, et al. Identification of groundwater pollution sources based on Bayes' theorem[J]. China Environmental Science, 2019, 39(4): 1568 - 1578. (in Chinese with English abstract)]
- [13] NESTEROV Y. Introductory lectures on convex optimization: a basic course (applied optimization)[M].
 Boston: Springer Science & Business Media, 2004: 51–110.
- [14] 焦李成,杨淑媛,刘芳,等.神经网络七十年:回顾与展望[J].计算机学报,2016,39(8):1697-1716. [JIAO Licheng, YANG Shuyuan, LIU Fang, et al. Seventy years beyond neural networks: retrospect and prospect[J]. Chinese Journal of Computers, 2016, 39(8):1697-1716. (in Chinese with English abstract)]
- [15] 褚蕾蕾,陈绥阳,周梦.计算智能的数学基础[M].北京:科学出版社,2002. [CHU Leilei, CHEN Suiyang, ZHOU Meng. Mathematical basis of Computational Intelligence [M]. Beijing: Science Press, 2002. (in Chinese)]

- [16] 陈海虹,黄彪,刘峰,等.机器学习原理及应用[M].成都:电子科技大学出版社,2017. [CHEN Haihong, HUANG Biao, LIU Feng, et al. Principle and application of machine learning[M]. Chengdu: University of Electronic Science and Technology of China Press, 2017. (in Chinese)]
- [17] WANG J, TAO Q. Machine learning: the state of the art[J]. IEEE Intelligent Systems, 2008, 23(6): 49 55.
- [18] BOTTOU L, CURTIS F E, NOCEDAL J. Optimization methods for large-scale machine learning[J]. SIAM Review, 2018, 60(2): 223 – 311.
- [19] POLYAK B T. Introduction to optimization[M]. New York: Optimization Software, 1987.
- [20] CAUCHY A L. Méthode générale pour la résolution des systemes d'équations simultanées[J]. Comptes rendus de l'Académie des Sciences, 1847: 536 – 538.
- [21] ROBBINS H, MONRO S. A stochastic approximation method[J]. The Annals of Mathematical Statistics, 1951, 22(3): 400 - 407.
- [22] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533 - 536.
- [23] 史加荣,王丹,尚凡华,等.随机梯度下降算法研究进展[J]. 自动化学报, 2021, 47(9):2103-2119. [SHI Jiarong, WANG Dan, SHANG Fanhua, et al. Research advances on stochastic gradient descent algorithms[J]. Acta Automatica Sinica, 2021, 47(9):2103-2119. (in Chinese with English abstract)]
- [24] ATTOUCH H, BOLTE J, SVAITER B F. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods[J]. Mathematical Programming, 2013, 137(1/2): 91 – 129.
- [25] 王惊晓,高乾坤,汪群山.一种具有最优收敛速度的正则化境面下降算法[J].计算机工程,2014,40(6): 148-153. [WANG Jingxiao, GAO Qiankun, WANG Qunshan. A regularization mirror descent algorithm with optimal convergence rate[J]. Computer Engineering, 2014,40(6):148-153. (in Chinese with English abstract)]
- [26] 刘梓懿. 多步梯度法的收敛性研究[D]. 武汉: 武汉科 技大学, 2020. [LIU Ziyi. On the convergence rate of the multi-step methods[D]. Wuhan: Wuhan University of Science and Technology, 2020. (in Chinese with English abstract)]
- [27] GHADIMI E, SHAMES I, JOHANSSON M. Multi-step

gradient methods for networked optimization[J]. IEEE Transactions on Signal Processing, 2013, 61(21): 5417 – 5429.

- [28] 孔康,汪群山,梁万路.L1正则化机器学习问题求解 分析[J]. 计算机工程, 2011, 37(17): 175 - 177.
 [KONG Kang, WANG Qunshan, LIANG Wanlu. Solution analysis of L1 regularized machine learning problem[J]. Computer Engineering, 2011, 37(17): 175 -177. (in Chinese with English abstract)]
- [29] TSENG P, YUN S. Incrementally updated gradient methods for constrained and regularized optimization[J].
 Journal of Optimization Theory and Applications, 2014, 160(3): 832 - 853.
- [30] YU N. Gradient methods for minimizing composite functions[J]. Mathematical Programming, 2013, 140(1): 125-161.
- [31] DAVID W P. User's guide for MODPATH/MODPATH-PLOT, Version 3: A particle tracking post-processing package for MODFLOW, the U. S. Geological Survey finite-difference ground-water flow model[R]. Reston: Geological Survey, 1994.
- [32] RUNGE C D T. Über die numerische Auflösung von Differentialgleichungen[J]. Mathematische Annalen, Springer, 1895, 46(2): 167 – 178.
- [33] KUTTA M. Beitrag zur n\u00e4herungsweisen Integration totaler Differentialgleichungen[J]. Zeitschrift f\u00fcr Mathematik und Physik, 1901, 46: 435 - 453.
- [34] 廖凯华.大沽河流域土壤水资源评价及农业节水灌溉模式研究[D].青岛:青岛大学,2009.[LIAO Kaihua. Study on evaluation of soil water resources and water saving irrigation modes in Dagu river basin[D]. Qingdao: Qingdao University, 2009. (in Chinese with English abstract)]
- [35] 尹子悦.大沽河流域地下水时空演变特征及数值模 拟[D].青岛:青岛大学,2019. [YIN Ziyue. Spatialtemporal variations of groundwater and its numerical

simulation in the Dagu river basin[D]. Qingdao: Qingdao University, 2019. (in Chinese with English abstract)]

- [36] 贾立华.大沽河地下水库脆弱性评价[D].青岛:中国 海洋大学, 2003. [JIA Lihua. Vulnerability assessment to the Dagu river groundwater reservoir[D]. Qingdao: Ocean University of China, 2003. (in Chinese with English abstract)]
- [37] 林国庆,郑西来,李海明.下水库人工补给的模型研究——以大沽河地下水库为例[J].中国海洋大学学报(自然科学版),2005,35(5):745-750.[LIN Guoqing, ZHENG Xilai, LI Haiming. Simulation of artificial recharge for a groundwater reservoir: a case study of the daguhe groundwater reservoir[J]. Journal of Ocean University of Qingdao, 2005, 35(5):745 750.(in Chinese with English abstract)]
- 【38】 张旭洋,林青,黄修东,等.大沽河流域土壤水-地下水流耦合模拟及补给量估算[J].土壤学报,2019,56(1):101 113. [ZHANG Xuyang, LIN Qing, HUANG Xiudong, et al. Numerical simulation coupling soil water/groundwater and estimation of groundwater recharge in dagu river basin[J]. Acta Pedologica Sinica, 2019, 56(1):101 113. (in Chinese with English abstract)]
- [39] 尹明泉,王治良,王通国,等.青岛市大沽河水源地地下水资源潜力分析[J].山东国土资源,2005,21(4):30-34. [YIN Mingquan, WANG Zhiliang, WANG Tongguo, et al. Potentiality analysis on underground water resource in daguhe water source area in Qingdao City[J]. Land and Resources in Shandong Province, 2005, 21(4):30-34. (in Chinese with English abstract)]
- [40] 杨丽娟. 大沽河平原区地下水资源评价[D]. 青岛: 中 国海洋大学, 2012. [YANG Lijuan. Evaluation of groundwater resources in dagu river[D]. Qingdao: Ocean University of China, 2012. (in Chinese with English abstract)]

编辑:张若琳