

# 基于稠密连接网络的地下水污染替代模型研究

江思珉<sup>1</sup>, 孔维铭<sup>1</sup>, 吴延浩<sup>1</sup>, 刘金炳<sup>2</sup>, 张春秋<sup>1</sup>, 夏学敏<sup>3</sup>

(1. 同济大学 土木工程学院, 上海 200092; 2. 长江生态环保集团有限公司, 江苏南京 210019;

3. 上海理工大学 环境与建筑学院, 上海 200093)

**摘要:** 地下水污染溯源和含水层参数反演问题等地下水逆问题通常面临计算负荷量的制约, 考虑使用替代模型作为解决方法, 从而降低地下水反演问题的计算成本。借鉴卷积神经网络的图像识别过程, 考虑将地下水流运动和污染物运移问题转化为输入场图像(渗透系数场、污染源信息等)与输出场图像(水头场、浓度场等)之间函数关系的图像回归问题, 利用基于稠密连接网络的 AR-Net-WL 构建地下水流运动和污染物运移的替代模型。算例研究表明, 针对替代模型的过拟合现象, 尽可能选择较大的训练样本可获得约 10% 的精度提升; 当没有条件增加训练样本时, 采用最优正则项系数的 AR-Net-WL 在训练样本较少的情形下(训练样本 500)也能够取得良好的性能, 能够精确预测地下水流运动和污染物运移。

**关键词:** 地下水污染; 替代模型; 地下水逆问题; 深度学习; 稠密连接网络

中图分类号: P641.2; X523

文献标志码: A

## A DenseNet Network Based Surrogate Model for Simulating Contaminant Transport in Groundwater Systems

JIANG Simin<sup>1</sup>, KONG Weiming<sup>1</sup>, WU Yanhao<sup>1</sup>, LIU Jinbing<sup>2</sup>, ZHANG Chunqiu<sup>1</sup>, XIA Xuemin<sup>3</sup>

(1. College of Civil Engineering, Tongji University, Shanghai 200092, China; 2. Yangtze Ecology and Environment Co., Ltd., Nanjing 210019, China; 3. School of Environment and Architecture, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** Groundwater inverse problems such as groundwater contaminant identification and aquifer parameter inversion problems are usually restricted by the computational load. In order to reduce the computational cost of groundwater inversion, the surrogate method is a feasible solution. In this paper, imitating the image

recognition process of the convolutional neural network, the groundwater flow movement and contaminant transport problem is transformed into an image regression problem of the functional relationship between input image (hydraulic conductivity field, pollution source information, etc.) and output image (groundwater level, contaminant concentration, etc.). The surrogate model of groundwater flow movement and contaminant transport is constructed by using AR-Net-WL based on the DenseNet network. The case study shows that, for the overfitting of the surrogate model, a 10% improvement in accuracy can be obtained by selecting training samples as large as possible. When there are no conditions to increase the training sample, AR-Net-WL with an optimal regularization coefficient can also achieve a good performance with fewer training samples (500 training samples) and can accurately predict the groundwater flow movement and contaminant transport.

**Key words:** groundwater contamination; surrogate model; groundwater inverse problem; deep learning; DenseNet

地下水系统一旦受到污染则很难恢复, 并将长期影响地下水水质、地下水系统生态平衡和人体健康。与地表水相比, 地下水一旦遭受污染其治理和修复难度更大, 往往需要更长的时间(常常持续 30 年以上), 费用非常昂贵<sup>[1]</sup>。通常来说, 要对地下水进行有效的治理和修复, 需要大致弄清污染源信息(地下水污染溯源)以及能够对含水层性质(渗透系数等)进行准确表征(模拟参数反演), 这是地下水数值模拟领域亟待解决的重要逆问题。地下水污染溯源和渗透系数场参数反演问题等除了逆问题具有不

收稿日期: 2021-09-26

基金项目: 国家自然科学基金(42077176); 南京水利科学院水文水资源与水利工程科学国家重点实验室重点基金(2019nkzd01)

通信作者: 江思珉(1980—), 男, 副教授, 工学博士, 主要研究方向为地下水科学与工程。

E-mail: jiangsimin@tongji.edu.cn



论文  
拓展  
介绍

适定性的挑战,还面临计算负荷量的制约<sup>[2]</sup>。为降低地下水反演问题的计算成本,已有研究考虑提出效率更高的反问题求解算法<sup>[3]</sup>、采用性能更高的计算设备进行并行计算等<sup>[4]</sup>。作为另一种可行的解决方法,替代模型方法在近年受到越来越多的关注。

替代模型依据其构建方式一般可分为两大类,第一类是数据驱动模型<sup>[5-6]</sup>,第二类是过程简化的模型<sup>[7-8]</sup>。但是,当前的应用主要集中于相对低维的问题,在刻画高维问题时替代模型方法的性能和效率急剧下降<sup>[9-11]</sup>。近年来,随着计算机性能的大幅提高,尤其是图形处理器(GPU)的出现,神经网络凭借其优秀的鲁棒性和泛化能力成为计算机视觉、空间降维计算、替代模型等领域的热点研究方法。Zhu 等<sup>[12]</sup>提出了用无监督学习(unsupervised learning)的方法来构造替代模型。Zhu 等<sup>[13]</sup>将贝叶斯理论引入卷积神经网络(convolutional neural networks, CNN),提出了一种新的 CNN 架构—BNN(Bayesian neural network),在训练数据有限的情况能够构建出较高精度的替代模型。Peterson 等<sup>[14]</sup>提出基于时间序列的地下水自回归(auto-regressive, AR)替代模型,预测不同气候情景下的地下水位。Mo 等<sup>[15]</sup>针对污染源附近区域非线性浓度梯度较大的问题,提出了在网络损失函数中对污染源及邻近网络加大权重(weighted loss, WL)的自回归替代模型。

现有研究在利用各类神经网络构建地下水污染替代模型时,并未结合具体的应用场景对神经网络的函数选择、参数敏感性等进行充分考量。本文借鉴卷积神经网络的图像识别过程,考虑将地下水流运动和污染物运移问题转化为输入场图像(渗透系数场、污染源信息等)与输出场图像(水头场、浓度场等)之间函数关系的图像回归问题,通过数值算例研究稠密连接网络替代模型的有效性以及模型性能的改进策略,主要包括:针对实际应用中样本数量较少的情形提出在神经网络中采用最优正则项系数的改进策略,针对污染源附近区域存在非线性浓度梯度问题采用改进权重的网络损失函数以及神经网络激活函数的选择。

## 1 研究方法

地下水模拟模型是研究地下水流和污染物运移的关键,而地下水模型的输出是通过求解地下水流和溶质运移的偏微分方程获得的,这往往需要较长

的计算时间。为充分刻画地下空间的异质性、考虑更多的物理过程、提高数值精确度、扩大模拟区域尺度等,都会导致计算量增大和模型运行时间的增加。单个模型运行时间的增加还会限制其在需要大量调用模拟模型的实际问题中的应用,如地下水逆问题、不确定性分析、敏感性分析等。为了提高计算效率,可以在模型中使用替代模型来替换原始模型<sup>[16]</sup>。

### 1.1 地下水流与溶质运移模型

目前常用的污染物迁移模拟程序有 MOC3D、MT3DMS、RT3D、FEMWATER、FEFLOW 等。其中,MT3DMS 是应用最为广泛的污染物迁移模拟程序。本文采用 MT3DMS 程序进行地下水污染物迁移模拟,MT3DMS 本身不包括地下水流模拟程序,在模拟计算时,MT3DMS 需和 MODFLOW 一起使用<sup>[17-18]</sup>。

地下水流方程如式(1):

$$\frac{\partial}{\partial x}\left(K_x\left|\frac{\partial h}{\partial x}\right|\right)+\frac{\partial}{\partial y}\left(K_y\left|\frac{\partial h}{\partial y}\right|\right)+\frac{\partial}{\partial z}\left(K_z\left|\frac{\partial h}{\partial z}\right|\right)+W=S_s\frac{\partial h}{\partial t} \quad (1)$$

式中:假定渗透系数的主轴方向与坐标轴方向一致, $K_x$ 、 $K_y$ 、 $K_z$ 分别为渗透系数在 $x$ 、 $y$ 、 $z$ 方向上的分量, $\text{m}\cdot\text{d}^{-1}$ ;  $h$ 为水头, $\text{m}$ ;  $W$ 为单位时间从单位体积含水层流入或流出的水量, $\text{d}^{-1}$ ;  $S_s$ 为孔隙介质的贮水率, $\text{m}^{-1}$ ;  $t$ 为时间, $\text{d}$ 。

地下水溶质运移方程如式(2):

$$\frac{\partial(\theta|c_k)}{\partial t}=\frac{\partial}{\partial x_u}\left(\theta|D_{uw}\left|\frac{\partial c_k}{\partial x_w}\right|\right)-\frac{\partial}{\partial x_u}(\theta|\mathbf{v}_u|c_k)+q_sc_{ks}+\sum R_n \quad (2)$$

式中: $\theta$ 为有效孔隙度; $c_k$ 为溶质组分 $k$ 的浓度, $\text{g}\cdot\text{m}^{-3}$ ;  $D_{uw}$ 为水动力弥散系数张量, $\text{m}^2\cdot\text{d}^{-1}$ ,  $\mathbf{u}$ 为纵向弥散方向,  $\mathbf{w}$ 为横向弥散方向;  $\mathbf{v}_u$ 为含水介质中的实际水流速度,  $\text{m}\cdot\text{d}^{-1}$ , 与 Darcy 流速  $q_u$  的关系为  $\mathbf{v}_u=\mathbf{q}_u/\theta$ ;  $q_s$ 为单位体积含水层的源汇项流量, $\text{d}^{-1}$ ;  $c_{ks}$ 为源汇项中溶质组分 $k$ 的浓度, $\text{g}\cdot\text{m}^{-3}$ ;  $\sum R_n$ 化学反应项总和, $\text{g}\cdot\text{m}^{-3}\cdot\text{d}$ 。

### 1.2 基于 DenseNet 网络结构的替代模型框架

深度卷积神经网络可以有效处理高维输入-输出映射,其适用于处理具有空间结构的图像数据<sup>[19]</sup>。具体地,地下水流与溶质运移模型中输入(渗透系数等)和输出(水头场和浓度场)都具有显著空间相关性,通过图像化处理后可以利用卷积神经网络建立高维输入-输出映射,该映射关系即为替代模型。模

型的输入与输出的关系表示如式(3):

$$f: I_{d_x \times H \times W} \xrightarrow{\text{CNN}} O_{d_y \times H \times W} \quad (3)$$

式中:  $H \times W$  为模型的空间剖分网格数(对应于1.1节MODFLOW有限差分网格);  $d_x$  和  $d_y$  分别表示模型输入和输出的个数。对于二维模型,输入参数包含渗透系数场、污染源信息等,输出状态量包括污染浓度场和水头场。在卷积神经网络中,将输入参数和输出状态量以近似于图像的数据结构表示(高维图像),通过网络训练建立输入场图像与输出场图像之间函数关系的图像回归问题。

在图像回归问题中,全卷积神经网络是用编码网络(encoder)对高维输入图像进行特征提取,然后在解码网络(decoder)根据特征重建输出图像。随着设计越来越深的网络,深刻理解“新添加的层如何提升神经网络的性能”变得至关重要。针对这一问题,He等<sup>[20]</sup>提出的残差网络(ResNet)在一定程度上避免深度网络的梯度消失、梯度爆炸问题和网络退化

问题。稠密连接网络(DenseNet)在某种程度上可以看成是ResNet的逻辑扩展,它采用一种密集连接结构将不相邻层进行连接,从而减少网络传输过程的信息损失,同时避免深度增加引起的网络参数爆炸<sup>[21]</sup>。

DenseNet由稠密块(dense block)与过渡层(transition layer)2个重要部分组成。具体地,稠密块的层与层之间的信息传递关系表示如式(4):

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (4)$$

式中:  $[x_0, x_1, \dots, x_{l-1}]$  表示0到  $l-1$  层的输出特征图的串联操作;  $H_l(\cdot)$  包括批归一化(batch normalization, BN)、ReLU激活函数和卷积操作(convolution, Conv)的顺序组合等。稠密块主要由2个参数决定:结构内的网络层数  $L$  和每一层的输出特征面数量  $K'$ 。通过整个稠密块的特征图数量为  $N_0 + L \times K'$ , 其中  $N_0$  为输入特征图的数量。图1为稠密块结构示意图。

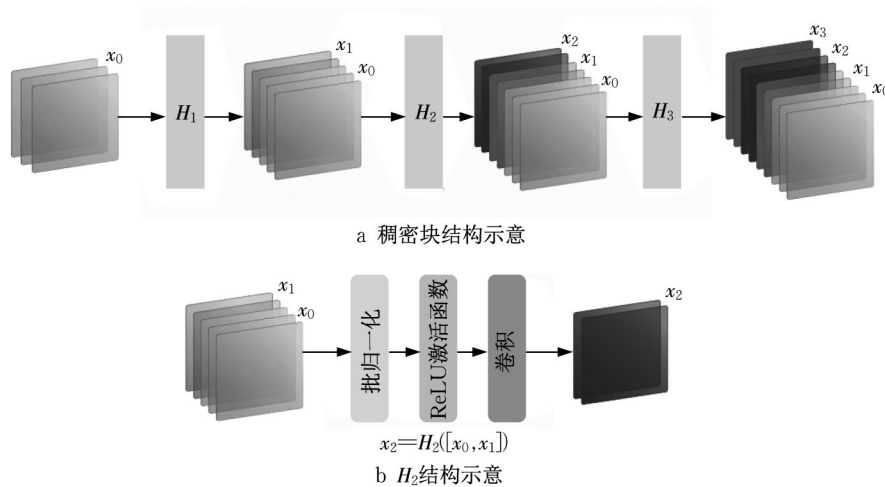


图1 DenseNet的稠密块结构示意图<sup>[20]</sup>

Fig.1 Dense block structure of DenseNet<sup>[20]</sup>

图2为图像回归问题的DenseNet网络结构示意图,其中第1层为初始卷积层,随后为稠密块和过渡层(编码层和解码层)的组合。由于每个稠密块都会带来特征图数量的增加,使用过多则会过于复杂化模型,因此需要过渡层来连接2个相邻的稠密块,通过批归一化、ReLU激活函数和  $1 \times 1$  卷积等组合操作减小特征图数量<sup>[21]</sup>。此外,在图像回归问题中,过渡层在编码过程中被称为编码层(encoding layer),利用卷积操作完成特征图尺寸改变的任务(downsampling),从而实现特征提取;过渡层在解码过程中被称为解码层(decoding layer),通过反卷积

操作恢复特征图尺寸(upsampling)。

## 2 算例研究

### 2.1 问题概述

图3考虑非均质各向异性承压含水层,含水层水流运动为稳定流。研究区大小为  $12.5\text{ m} \times 17.5\text{ m}$ ,离散为  $51 \times 71$  的均匀网格(网格尺寸为  $0.25\text{ m} \times 0.25\text{ m}$ )。含水层厚度为  $1\text{ m}$ 。南北边界为二类隔水边界,东西边界为定水头边界,水头  $H$  分别为  $6\text{ m}$  和  $5\text{ m}$  (图4)。场地东侧存在一口抽水井



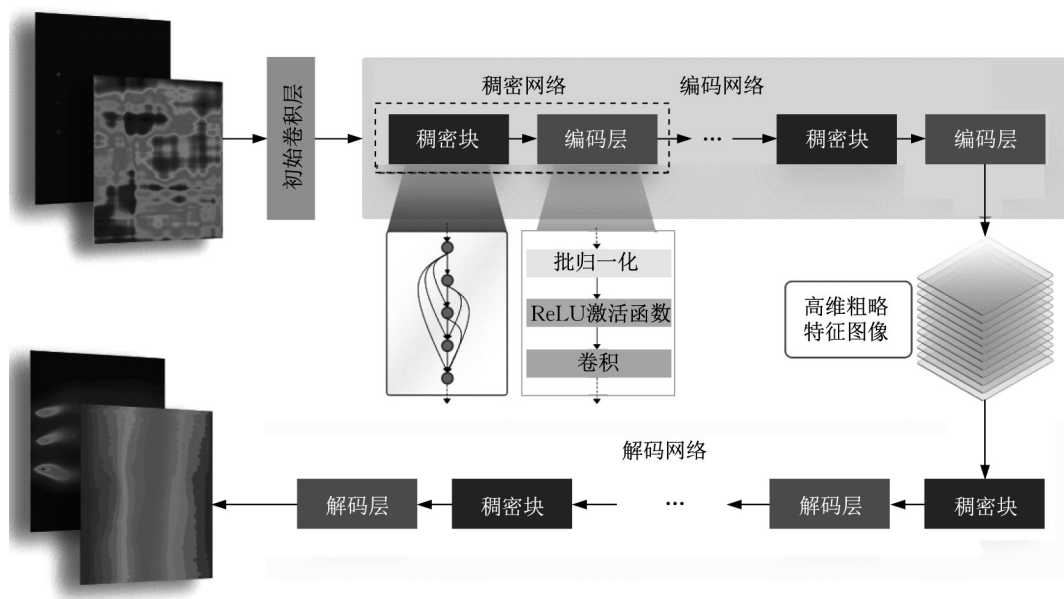


图 2 图像回归问题的DenseNet网络结构示意图

Fig.2 DenseNet network structure for image regression problem

(W1, 流量  $2 \text{ m}^3 \cdot \text{d}^{-1}$ ), 降雨入渗强度为  $0.005 \text{ m} \cdot \text{d}^{-1}$ 。初始时刻, 含水层中无污染物。含水介质孔隙度为 0.30, 纵向弥散度为 1.50 m, 水平横向弥散度为 0.15 m。含水层渗透系数  $K$  满足对数正态分布, 其均值 ( $\ln K$ ) 与方差 ( $\sigma_{\ln K}^2$ ) 分别为 2.0 和 0.5, 在  $x$ 、 $y$  方向的相关长度分别为 4 m 和 2 m, 变差函数为指数型。模拟总时长为 160 d, 共 10 个应力期 (每个应力期 16 d)。

考虑到实际地下水污染问题中, 根据前期场地调查往往能大致确定污染源位置, 而对应污染源的释放历史是未知的。本算例中考虑 3 个污染源, 为  $S_1$ 、 $S_2$  和  $S_3$ 。污染源参数由 30 个参数进行量化 (3 个污染源在 10 个应力期的污染释放强度), 即  $S_i = \{S_{ij}\}$ , 其中  $i=1, 2, 3; j=1, 2, \dots, 10$ 。非均质渗透系数场的反演估计通常会导致高维问题, 高维参数受制于维度灾难问题, 通常需要对渗透系数场进行概化和降维<sup>[22]</sup>。在本算例中, 渗透系数场的维度决定于网格剖分程度 ( $51 \times 71 = 3621$ ), 为了在可接受的计算时间内解决高维参数反演问题, 采用 KLE 方法进行降维操作, 保留前 400 个 KL 主要展开项 (即可保留该渗透系数场近 95% 的非均质特征)<sup>[23]</sup>。假设表征污染源释放强度的 30 个参数的先验分布是均匀分布, 其先验区间和真实值见表 1。

在构建替代模型前, 将正演模型的输入和输出项以图像的形式表示。输入包括渗透系数场和图像化处理后的污染源参数。输出为利用原始模型

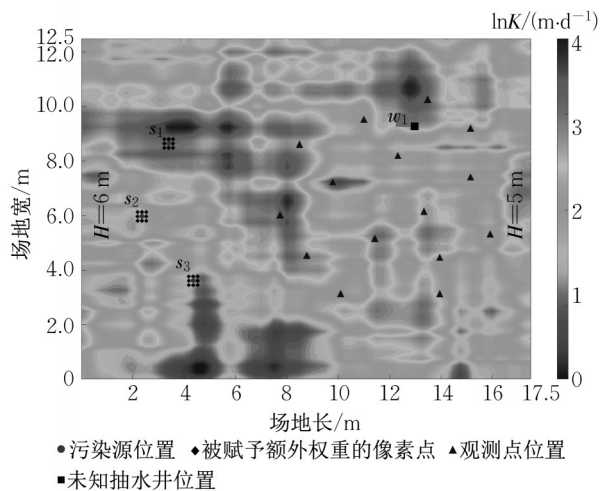


图 3 含水层结构平面

Fig.3 Plan view of aquifer structure

表 1 污染源释放强度参数的真实值及先验区间

Tab.1 True value and prior interval of release intensity parameters of pollution source

应力期	先验区间	污染源浓度真实值/( $\text{g} \cdot \text{d}^{-1}$ )		
		$S_1$	$S_2$	$S_3$
SP1	$u[0, 8]$	5.62	1.04	5.38
SP2	$u[0, 8]$	4.34	4.55	3.26
SP3	$u[0, 8]$	1.79	3.76	5.69
SP4	$u[0, 8]$	5.21	0.10	4.00
SP5	$u[0, 8]$	2.54	2.70	1.68
SP6	$u[0, 8]$	4.38	1.30	4.21
SP7	$u[0, 8]$	5.56	6.36	5.38
SP8	$u[0, 8]$	4.69	2.49	6.23
SP9	$u[0, 8]$	2.32	4.23	2.38
SP10	$u[0, 8]$	5.32	1.33	5.26

(MODFLOW 和 MT3DMS)得到的水头和多个时刻污染物浓度场。模型输入的污染源强度( $S$ )随时间变化,模型输出的污染物浓度除了受输入参数(渗透系数 $K$ )的影响外,还受随时间变化的污染源源强的影响。换言之,任意某一目标时刻污染物浓度场受到该目标时刻 $j$ 及更早时刻( $1, 2, \dots, j-1$ )的污染源源强的共同影响。利用自回归模型方法,可以表示为式(5):

$$c_j = f(K, S, c_{j-1}) \quad (5)$$

式中: $c_j$ 为时刻 $j$ 的污染浓度场; $K$ 、 $S$ 分别为当前时刻的渗透系数场和污染源参数; $c_{j-1}$ 为上一时刻的污染分布。

考虑到地下水污染点源会对释放区域周围产生较大非线性浓度梯度,为了提高替代模型在污染源附近的预测精度,在网络损失函数中分别赋予污染源及周围8个像素点额外权重 $w_c$ (见图4)。加权损失函数表示为式(6):

$$L'(\theta) = L(\theta) + L_c(\theta) = L(\theta) + w_c \sum_{i=1}^{N_t'} |\hat{c}_{si} - c_{si}| \quad (6)$$

式中: $L(\theta)$ 为网络训练的损失函数; $n_t'$ 为有污染物释放的时间段数, $n_t' = 10$ ;  $c_s$ 为污染点和周边像素点的9个浓度向量; $\hat{c}_s$ 为 $c_s$ 的神经网络预测值; $w_c$ 为权重, $w_c = 5.0$  [15]。

## 2.2 DenseNet 网络结构设计

算例所用的网络结构为 AR-Net-WL,即考虑式(6)定义的加权损失函数的自回归网络。AR-Net-WL 采用 DenseNet 网络结构,输入样本考虑自回归策略。网络输入图片为3个,即 $\tilde{x} = (K, |S|, |c_{i-1}|) \in R_{3 \times 51 \times 71}$  ( $i=1, \dots, 10$ ,  $c_0$ 为初始浓度场),输出图片为2个,即 $y = (h, |c_i|) \in R_{2 \times 51 \times 71}$ 。AR-Net 和 AR-Net-WL 的 DenseNet 网络如表2所示,包括初始卷积层、3个稠密块和3个过渡层。表中, $k'$ 为卷积核大小, $s$ 为步长, $p$ 为特征图填充宽度, $W_i$ 为输出特征面数量, $N_i \times H_i$ 为特征面大小, $L$ 和 $K'$ 分别为稠密块中的卷积层数和每层输出的特征面数量。其中,3个稠密块的内部层数 $L = [5, 10, 5]$ 以增加网络深度,每一层输出的特征面为 $K' = 40$ ;此外,每个过渡层包括2个卷积层( $1 \times 1$ 卷积);因此,DenseNet 网络由27个卷积层组成,并且不含全连接层,为全卷积网络。

在神经网络的激活函数中,sigmoid 函数存在计算量大、梯度消失等局限性,ReLU 函数具有在0点处梯度不连续、在 $x < 0$ 时不传播等特点;softplus 能够有效克服 ReLU 和 sigmoid 函数的上述不足之

处 [24]。在地下水污染问题中,输出的浓度值和水头值均大于等于零,在网络的最后一层采用 softplus 激活函数,能够保证网络的浓度和水头预测值为非负数。Adam 优化算法的初始学习率设定为0.005,网络训练周期数(epoch)设定为200,若损失函数在10个epochs没有变化,则将学习率降至当前学习率的10%,选择 $L_1$ 损失函数(lasso),其正则项系数为 $\omega_d = 5 \times 10^{-5}$ 。

表2 AR-Net-WL 的网络结构

Tab. 2 Network structure of AR-Net-WL

结构	卷积核	$N_i \times H_i \times W_i$
输入层		$3 \times 51 \times 71$
初始卷积层	$k'7s2p3$	$48 \times 26 \times 36$
稠密块 1 ( $K'40L5$ )	$k'3s1p1$	$248 \times 26 \times 36$
编码层	$k'1s1p0, k'3s2p1$	$124 \times 13 \times 18$
稠密块 2 ( $K'40L10$ )	$k'3s1p1$	$524 \times 13 \times 18$
解码层 1	$k'1s1p0, k'4s2p1$	$262 \times 26 \times 36$
稠密块 3 ( $K'40L5$ )	$k'3s1p1$	$462 \times 26 \times 36$
解码层 2	$k'1s1p0, k'6s2p2$	$2 \times 51 \times 71$

## 3 结果与讨论

研究 DenseNet 网络 (AR-Net-WL) 构建得到地下水流与溶质运移替代模型的效果,分别对样本数量、DenseNet 网络超参(权重损失函数、正则项系数)等影响进行分析。AR-Net-WL 在 NVIDIA GeForce RTX2080s GPU 上训练。在替代模型训练时,采用决定系数( $R^2$ )和均方根误差(root-mean-square error,  $R_{MSE}$ )评估深度卷积网络替代模型的精度。决定系数( $R^2$ )的值越大,均方根误差( $R_{MSE}$ )的值越小,表明替代模型的预测精度越高。

$$R^2 = 1 - \frac{\sum_{n=1}^{N_e} \|y_n - \tilde{y}_n\|_2^2}{\sum_{n=1}^{N_e} \|y_n - \bar{y}\|_2^2} \quad (7)$$

$$R_{MSE} = \sqrt{\frac{1}{N_e} \sum_{n=1}^{N_e} \|y_n - \tilde{y}_n\|_2^2} \quad (8)$$

式中: $y_n$ 为样本 $n$ 的原始模型输出结果; $\tilde{y}_n$ 为样本 $n$ 的替代模型预测结果; $\bar{y} = 1/N_e \sum_{n=1}^{N_e} y_n$ ,  $N_e$ 为样本总数。

### 3.1 训练样本数量的影响

为讨论训练样本数量对替代模型近似精度的影响,在 AR-Net-WL 网络结构下考虑了4个训练样本集合,包含训练样本数量分别500、1 000、2 000和3 000。测试样本集合由 $N_e = 500$ 个随机生成的

测试样本组成,包括地下水稳定水位  $h$  和 10 个时刻  $t = [16, 32, 48, 64, 80, 96, 112, 128, 144, 160]$  d 的污染物浓度观测数据  $\{c_j\}_{j=1}^{t_c=10}$ 。因此,网络输入图片为 3 个,即  $\tilde{x} = (K, |S|, |c_{i-1}|) \in R_{3 \times 51 \times 71}$  ( $i=1, \dots, 10$ ,  $c_0$  为初始浓度场),输出图片为 2 个,即  $y = (h, |c_i|) \in R_{2 \times 51 \times 71}$ 。

图 4 所示为不同训练样本数量条件下训练样本和测试样本的  $R_{MSE}$  值随训练周期数(epoch)的变化趋势。在 160 个训练周期后,所有情形下的训练样本集合与测试样本集合的  $R_{MSE}$  值都趋于稳定;完成训练周期数为 200 的训练后,随着训练样本数量的增加,  $R_{MSE}$  值呈现降低趋势,即替代模型的预测精度越高。具体地,当训练样本从 1 000 增至 2 000 时,训练集合的  $R_{MSE}$  从 0.067 6 下降到 0.052 5;当训练样本从 2 000 增至 3 000 时,训练集合的  $R_{MSE}$  从 0.052 5 下降到 0.033 3。

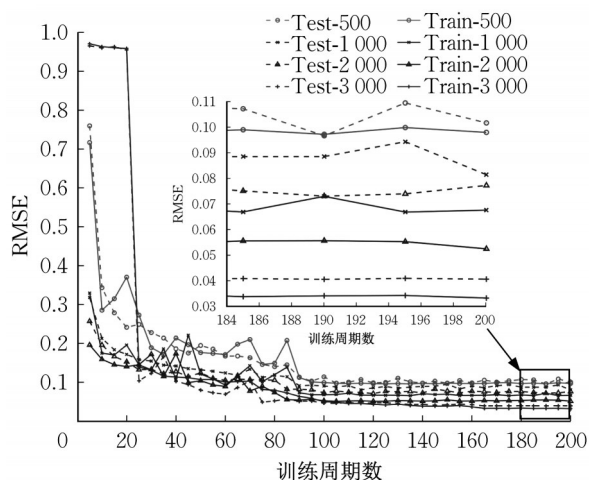


图 4 基于不同训练样本数量建立的替代模型的  $R_{MSE}$  精度比较

Fig. 4 Comparison of  $R_{MSE}$  accuracy of surrogate models established based on different numbers of training samples

图 5 比较了不同训练样本数量条件下替代模型的训练样本集合和测试样本集合的  $R^2$  值的变化趋势。训练样本为 500 时,替代模型的训练集合与测试集合的  $R^2$  值分别为 0.989 2 和 0.988 1;训练样本数量增至 2 000 时,训练集合与测试集合的  $R^2$  值分别增加至 0.997 9 和 0.996 0,替代模型具有较高的精度和较强的泛化性能;随着训练样本的继续增加,测试集合与训练集合的  $R^2$  值的提高幅度趋于平缓。具体地,训练样本数量从 2 000 增至 3 000 时,训练集合的  $R^2$  值增至 0.998 7 (增幅约 0.08%),测试集

合的  $R^2$  值增至 0.998 1 (增幅约 0.20%)。

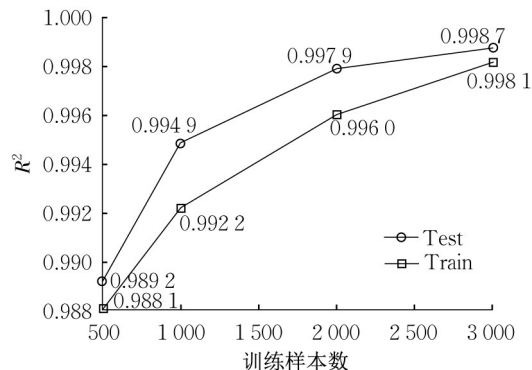


图 5 基于不同训练样本数量建立的替代模型的  $R^2$  精度比较

Fig. 5  $R^2$  accuracy comparison of surrogate models established based on different training sample numbers

### 3.2 权重损失函数的影响

为了量化 AR-Net-WL 网络中权重损失函数对替代模型精度的影响,分别选择 AR-Net-WL 和 AR-Net 网络构建替代模型(训练样本数量为 3 000),比较训练样本和测试样本的  $R^2$  与  $R_{MSE}$  指标。由表 3 可知,AR-Net 和 AR-Net-WL 替代模型都具有较高的精度,考虑权重损失函数的 AR-Net-WL 网络相较使用 AR-Net 网络在计算代价没有显著提高的前提下,计算精度( $R^2$  与  $R_{MSE}$ )有所提升。具体地,AR-Net-WL 相较于 AR-Net 网络,训练集合的  $R_{MSE}$  值由 0.038 8 降至 0.033 3,降幅约 14.2%。

表 3 训练样本 3 000 时 AR-Net 和 AR-Net-WL 网络的  $R^2$  与  $R_{MSE}$  指标

Tab. 3  $R^2$  and  $R_{MSE}$  indicators of AR-Net and AR-Net-WL networks (training samples = 3 000)

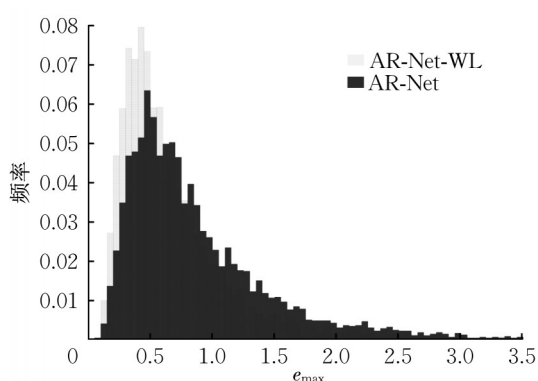
网络结构	时间/h	$R^2$		$R_{MSE}$	
		测试集	训练集	测试集	训练集
AR-Net	4.15	0.997 9	0.998 2	0.042 7	0.038 8
AR-Net-WL	4.27	0.998 1	0.998 7	0.040 6	0.033 3

为了进一步说明权重损失函数的作用,分别统计 AR-Net 和 AR-Net-WL 这 2 种替代模型的污染物浓度场预测误差绝对值的最大值  $e_{\max} = \max(|y| - |\hat{y}|)$ 。由图 6 可知,AR-Net-WL 网络的  $e_{\max}$  值相较 AR-Net 网络更小,再次验证了 AR-Net-WL 网络中添加的权重损失函数对提高污染源附近区域的预测精度的有效性,从而提高替代模型整体的预测精度。

### 3.3 正则项系数 $\omega_d$ 的影响

构建替代模型时,应该从训练样本中尽可能学



图6 AR-Net和AR-Net-WL的 $e_{\max}$ 频率Fig. 6  $e_{\max}$  frequency diagram of AR-Net and AR-Net-WL

出适合所有潜在样本的“普遍规律”<sup>[25]</sup>,但是如前所述在训练样本不是很充足的情况下,很容易出现过

拟合现象。针对过拟合现象,如果没有条件增加训练样本,可以考虑采用正则化方法(并不能完全解决)。下面讨论在样本数量较少的情形下,通过优化正则项系数提高替代模型的精度。为了找到最优正则化系数,算例中考虑5个数量级范围内的30个正则项系数值进行计算(训练样本为500)。具体地,设置 $\omega_d$ 在 $2 \sim 7$ 与 $10^{-7} \sim 10^{-3}$ 数量级组合,进行30次模型计算。

图7为30个 $\omega_d$ 对应替代模型的测试样本集合 $R_{\text{MSE}}$ 值,其中 $A$ 代表 $\omega_d$ 的整数部分, $B$ 代表 $\omega_d$ 的指数部分。在同一数量级下,随着 $\omega_d$ 值的变化, $R_{\text{MSE}}$ 值的变化规律不太明显;在 $10^{-7}$ 至 $10^{-5}$ 的数量级下 $R_{\text{MSE}}$ 值总体更小,对应替代模型精度更高;最小 $R_{\text{MSE}}$ 为0.063,明显优于未进行正则项系数优化的替代模型(训练样本为1 000对应的 $R_{\text{MSE}}=0.068$ )。

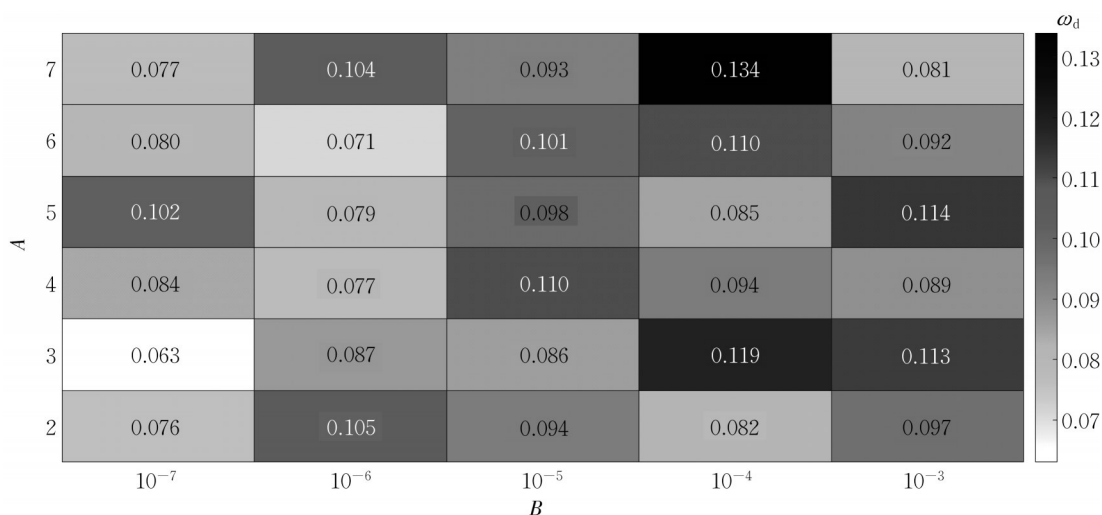
图7 不同 $\omega_d$ 对应替代模型的测试集合 $R_{\text{MSE}}$ 值(训练样本500)Fig.7  $R_{\text{MSE}}$  value for test set of surrogate model corresponding to different  $\omega_d$  values (training sample=500)

图8为训练样本数量为500得到的AR-Net-WL替代模型(选用最佳正则项系数 $\omega_d=3 \times 10^{-7}$ )对于某一随机选取的测试样本的预测结果。由图可知,替代模型的水头场和浓度场的预测结果与真实模型非常接近,这表明即使在训练样本较少的情形下(500个训练样本)选用较好的正则项系数得到的AR-Net-WL替代模型同样具有良好的性能,可以有效地预测水头场和不同时刻污染物浓度场。

## 4 结论

(1)借鉴卷积神经网络的图像识别过程,将地下水流运动和污染物运移问题转化为输入场图像(渗

透系数场、污染源信息等)与输出场图像(水头场、浓度场等)之间函数关系的图像回归问题。针对激活函数进行优化的AR-Net-WL建立的替代模型能够精确预测地下水流运动和污染物运移。

(2)针对替代模型的过拟合现象,在条件允许的情形下应尽可能选择较大的训练样本,算例研究表明训练样本从500增至3 000,替代模型的性能( $R^2$ 和 $R_{\text{MSE}}$ )明显得到提升;当没有条件增加训练样本时,采用最优正则项系数的AR-Net-WL在训练样本较少的情形下(训练样本500)也能够取得良好的性能。

(3)后续研究拟利用得到地下水污染替代模型(最优正则项系数的AR-Net-WL)进行地下水污染溯源问题的求解,将更进一步地研究影响替代模型

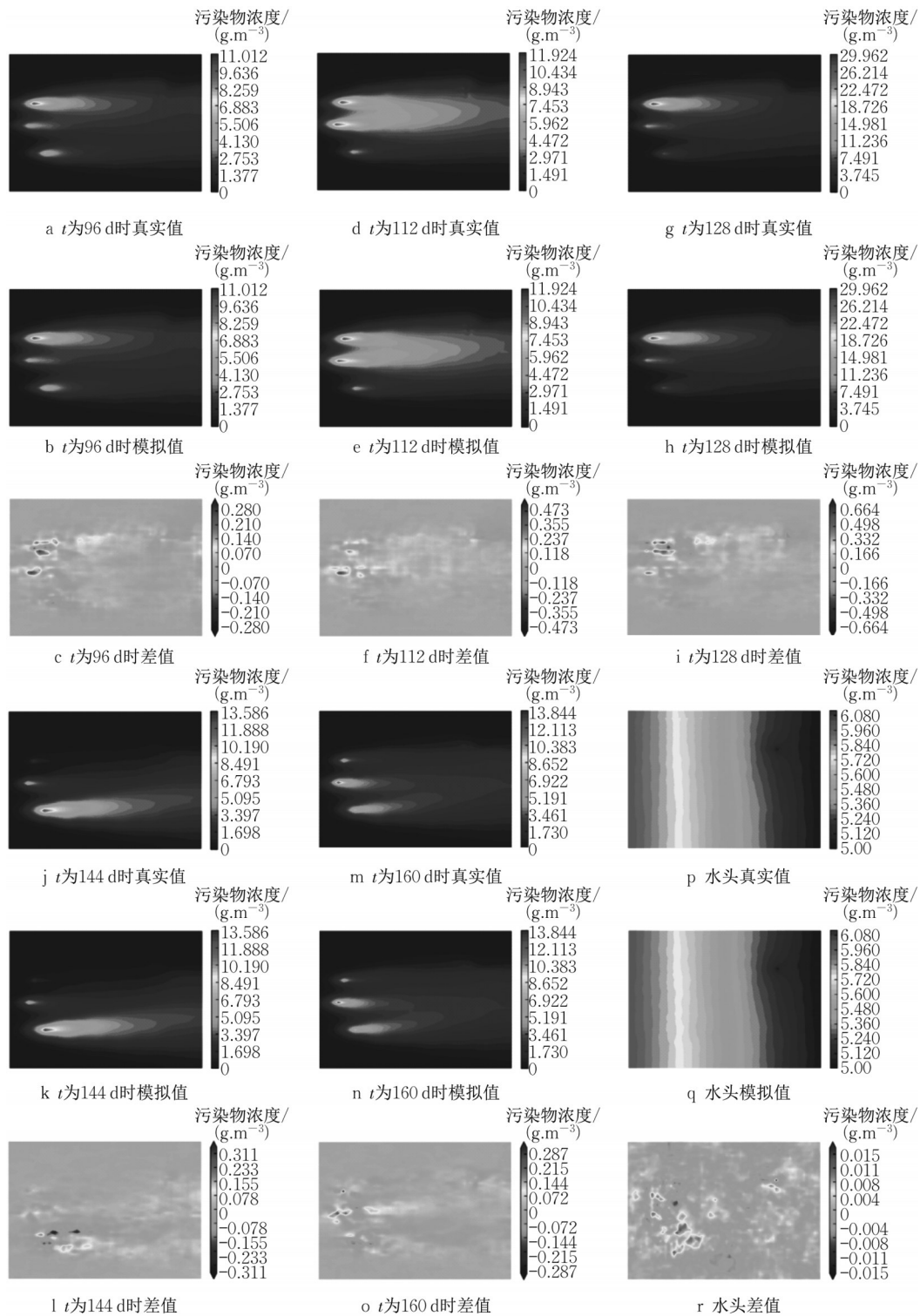


图 8 AR-Net-WL 替代模型的预测结果(随机测试样本)

Fig. 8 Prediction results of AR-Net-WL surrogate model (in random test samples)

精度及泛化性能的重要因素(包括采样方法、深度神经网络结构等)。

#### 作者贡献声明:

江思珉:论文方向指定与指导。

孔维铭:论文主体内容撰写。

吴延浩:论文模型参数调整实验。

刘金炳:指导论文模型代码修改。

张春秋:论文图例设计与修改。

夏学敏:提供论文内容意见与协助修改。



## 参考文献:

- [1] 吴剑锋, 郑春苗. 地下水污染监测网的设计研究进展[J]. 地球科学进展, 2004, 19(3): 429.  
WU Jianfeng, ZHENG Chunmiao. Contaminant monitoring network design: Recent advances and future directions [J]. Advance in Earth Sciences, 2004, 19(3): 429.
- [2] LI L, ZHOU H, GÓMEZ-HERNÁNDEZ J J. A comparative study of three-dimensional hydraulic conductivity upscaling at the macro-dispersion experiment (MADE) site, Columbus Air Force Base, Mississippi (USA) [J]. Journal of Hydrology, 2011, 404(3/4): 278.
- [3] WHITE J T. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions [J]. Environmental Modelling & Software, 2018, 109: 191.
- [4] MIRGHANI B Y, MAHINTHAKUMAR K G, TRYBY M E, *et al.* A parallel evolutionary strategy based simulation - optimization approach for solving groundwater source identification problems [J]. Advances in Water Resources, 2009, 32(9): 1373.
- [5] YAN S, MINSKER B. Optimal groundwater remediation design using an adaptive neural network genetic algorithm: Optimal groundwater remediation design [J]. Water Resources Research, 2006, 42(5).
- [6] XIA X, ZHOU N, WANG L, *et al.* Identification of transient contaminant sources in aquifers through a surrogate model based on a modified self-organizing-maps algorithm [J]. Hydrogeology Journal, 2019, 27(7): 2535.
- [7] ZHENG Q, ZHANG J, XU W, *et al.* Adaptive multifidelity data assimilation for nonlinear subsurface flow problems [J]. Water Resources Research, 2019, 55(1): 203.
- [8] MAN J, ZHENG Q, WU L, *et al.* Adaptive multi-fidelity probabilistic collocation-based Kalman filter for subsurface flow data assimilation: Numerical modeling and real-world experiment [J]. Stochastic Environmental Research and Risk Assessment, 2020, 34(8): 1135.
- [9] LIN G, TARTAKOVSKY A M. An efficient, high-order probabilistic collocation method on sparse grids for three-dimensional flow and solute transport in randomly heterogeneous porous media [J]. Advances in Water Resources, 2009, 32(5): 712.
- [10] ASHER M J, CROKE B F W, JAKEMAN A J, *et al.* A review of surrogate models and their application to groundwater modeling: Surrogates of groundwater models [J]. Water Resources Research, 2015, 51(8): 5957.
- [11] LIAO Q, ZHANG D, TCHELEPI H. A two-stage adaptive stochastic collocation method on nested sparse grids for multiphase flow in randomly heterogeneous porous media [J]. Journal of Computational Physics, 2017, 330: 828.
- [12] ZHU Y, ZABARAS N, KOUTSOURELAKIS P-S, *et al.* Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data [J]. Journal of Computational Physics, 2019, 394: 56.
- [13] ZHU Y, ZABARAS N. Bayesian deep convolutional encoder - decoder networks for surrogate modeling and uncertainty quantification [J]. Journal of Computational Physics, 2018, 366: 415.
- [14] PETERSON T J, WESTERN A W. Nonlinear time-series modeling of unconfined groundwater head [J]. Water Resources Research, 2014, 50(10): 8330.
- [15] MO S, ZABARAS N, SHI X, *et al.* Deep autoregressive neural networks for high - dimensional inverse problems in groundwater contaminant source identification [J]. Water Resources Research, 2019, 55(5): 3856.
- [16] 张江江. 地下水污染源解析的贝叶斯监测设计与参数反演方法[D]. 杭州: 浙江大学, 2017.  
ZHANG Jiangjiang. Bayesian monitoring design and parameter inversion for groundwater contaminant source identification [D]. Hangzhou: Zhejiang University, 2017.
- [17] MGM M, AW H. MODFLOW, A modular three-dimensional finite difference ground-water flow model [M]. Denver: [S. n.], 1988.
- [18] ZHENG C, WANG P P. MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide [R]. Vicksburg: U.S. Army Engineer Research and Development Center, 1999.
- [19] JU L, ZHANG J, WU L, *et al.* Bayesian monitoring design for streambed heat tracing: Numerical simulation and sandbox experiments [J]. Groundwater, 2019, 57(4): 534.
- [20] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770 - 778.
- [21] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 2261 - 2269.
- [22] ZHANG J, LIN G, LI W, *et al.* An iterative local updating ensemble smoother for estimation and uncertainty assessment of hydrologic model parameters with multimodal distributions [J]. Water Resources Research, 2018, 54(3): 1716.
- [23] ZHANG D, LU Z. An efficient, high-order perturbation approach for flow in random porous media via Karhunen - Loève and polynomial expansions [J]. Journal of Computational Physics, 2004, 194(2): 773.
- [24] ZHENG H, YANG Z, LIU W, *et al.* Improving deep neural networks using softplus units [C]//2015 International Joint Conference on Neural Networks (IJCNN). Killarney: IEEE, 2015: 1 - 4.
- [25] 周志华. 机器学习 [M]. 北京: 清华大学出版社, 2016.  
ZHOU Zhihua. Machine learning [M]. Beijing: Tsinghua University Press, 2016.