

基于径向基函数神经网络的地下水数值模拟模型的替代模型研究

伊燕平^{1,2}, 卢文喜², 张 耘¹, 芦贵君³, 王大中³, 洪德法⁴

(1. 吉林省宏利水土保持咨询有限公司, 长春 130033; 2. 吉林大学 环境与资源学院, 长春 130026; 3. 吉林省水土保持科学研究院, 长春 130033; 4. 长春市城乡规划设计研究院, 长春 130021)

摘 要:近年来提出的替代模型方法是一种连接数值模拟模型与优化模型的有效途径,替代模型质量的好坏取决于采样方法和替代模型种类。以金泉工业园区地下水水源地为研究区,基于拉丁超立方抽样方法,结合研究区地下水数值模拟模型,获取输入(抽水量)输出(水位降深)数据集,运用人工神经网络方法,建立径向基函数神经网络模型,作为地下水数值模拟模型的近似替代模型。经验证,径向基函数神经网络模型输出得到的水位降深均值与模拟模型计算结果的拟合平均相对误差为 0.038;水位降深剩余标准差的拟合平均相对误差为 0.042。拟合平均相对误差较小,表明径向基函数神经网络模型能够有效地替代地下水数值模拟模型,为日后替代模型的深入研究提供了科学依据。

关键词:替代模型; 径向基函数神经网络; 拉丁超立方抽样; 金泉工业园区

中图分类号:P641.8

文献标识码:A

文章编号:1005-3409(2012)04-0265-05

Study on the Surrogate Model of Groundwater Numerical Simulation Model Based on Radial Basis Function Neural Network

YI Yan-ping^{1,2}, LU Wen-xi², ZHANG Yun¹, LU Gui-jun³, WANG Da-zhong³, HONG De-fa⁴

(1. Hongli Soil and Water Conservation Consulting Co., Ltd. of Jilin Province, Changchun 130033, China; 2. College of Environment and Resources, Jilin University, Changchun 130026, China; 3. Jilin Institute of Soil and Water Conservation, Changchun 130033, China; 4. Changchun Institute of Urban Planning & Designing, Changchun 130021, China)

Abstract: In recent years, the surrogate model has become an effective way to connect the numerical simulation model and optimization model. The quality of surrogate model depends on sampling method and the type of surrogate model. In this paper, Jinqian Industrial Park groundwater sources was selected as the study area, the radial basis function neural network model was established as the surrogate model of groundwater numerical simulation model by combining with the groundwater numerical simulation model of the study area, accessing to input (pumping) output (water level drawdown) data set, and using artificial neural network based on Latin Hypercube Sampling techniques. It was proven that the mean water level drawdown output of RBF neural network model and simulation model results fit the average relative error was 0.038; water level drawdown of the remaining average relative standard deviation of the fitting error was 0.042. Fitting the average relative error was small, showing that the radial basis function neural network model can effectively replace groundwater numerical simulation model. The result of research provides a scientific basis for the future in-depth study of surrogate model.

Key words: surrogate model; radial basis function neural network; Latin Hypercube Sampling; Jinqian Industrial Park

理论研究和工程实践证明,模拟优化才能解决地下水修复过程中决策方案的优选问题,因此基于数值

收稿日期:2011-07-11

修回日期:2011-10-05

资助项目:国家自然科学基金资助项目(41072171)

作者简介:伊燕平(1984—),女,吉林长春人,硕士,助理工程师,主要从事水土保持规划设计工作。E-mail:yianping-101@163.com

通信作者:卢文喜(1956—),男,吉林德惠人,教授,主要从事地下水数值模拟与优化管理等研究。E-mail:luwenxi@jlu.edu.cn

模拟的优化模型日益受到人们的重视^[1-2]。但实现模拟模型与优化模型耦合,以往运用的嵌入法、响应矩阵法、状态转移方程法都有自身的局限性^[3],如嵌入法,它将状态变量(水位)和可控输入变量(抽水量)同时作为优化模型中的决策变量,因此对于多时段的不稳定流问题,就会产生“维数灾难”;响应矩阵法只适用于线性系统的模型耦合;状态转移方程法只适用于动态规划。为了克服以往耦合技术方法的局限性,近年来,提出了替代模型法^[4],它在功能上逼近模拟模型,在优化模型迭代求解过程中可直接调用替代模型,大幅度地减少计算负荷,节省运算时间,是一种连接模拟模型与优化模型的有效途径。

替代模型质量的好坏取决于采样方法和替代模型种类的选定。替代模型种类的选取已有初步研究,Qin 等运用回归分析方法建立了双响应面模型作为替代模型^[5-6];Johnso 等应用人工神经网络方法建立误差逆传播(BP)神经网络模型作为替代模型^[7-9]。本文以金泉工业园区地下水水源地为研究区,在已有研究区地下水数值模拟模型的基础上,采用拉丁超立方抽样方法,应用人工神经网络方法,建立了径向基函数神经网络模型,作为地下水数值模拟模型的替代模型,并对替代模型的有效性进行了验证,旨在为日后替代模型的深入研究提供科学的理论依据。

1 拉丁超立方抽样

输入(抽水量)—输出(水位降深)数据集是建立替代模型的基本前提。本文应用拉丁超立方抽样的方法,在抽水井的抽水量可行范围内进行采样,选取有代表性的抽水量作为地下水数值模拟模型输入数据集(表 1)。

表 1 实际抽水量数据资料 m^3/d

数值	井号				
	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅
最小值	1	1	1	1	1
最大值	2011	1981	1831	1681	1741

1.1 基本原理

拉丁超立方抽样(Latin Hypercube Sampling,简称 LHS)属于分层抽样,是一种用采样值反映随机变量整体分布的方法^[10]。拉丁超立方抽样可以避免直接抽样法数据点集中而导致的仿真循环重复问题,有效地避免了大量反复的抽样工作^[11-12]。同时,它强制抽样过程中采样点必须离散分布于整个抽样空间,使得抽样的覆盖度得到了很好的保证,抽取的样品具有一定代表性,如图 1 所示。

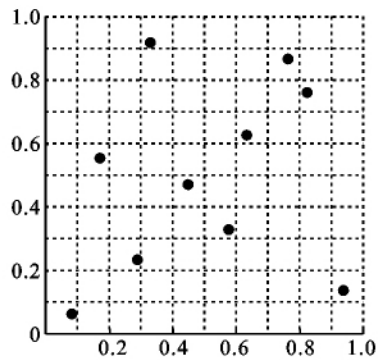


图 1 二维变量 LHS 抽样示意图

1.2 抽样过程

设随机变量的维数为 N , $x^i, x_u^i, i=1, 2, \dots, n$; 其中 x^i 为第 i 维变量, x_l^i 与 x_u^i 分别为第 i 维变量的上限和下限。首先,确定抽样规模 H ;其次,将每维变量 x^i 的定义域区间 $[x_l^i, x_u^i]$ 划分成 H 个相等的小区间: $x_l^i = x_0^i < x_1^i < x_2^i < \dots < x_j^i < x_{j+1}^i < \dots < x_H^i = x_u^i$,将原来的一个超立方体划分成 H^n 个小超立方体;然后,将产生一个 $H \times n$ 的矩阵 A ,矩阵 A 的每列都是数列 $\{1, 2, \dots, H\}$ 的一个随机全排列组合。在拉丁超立方抽样过程中, N 维变量分成 M 层,满足要求的小超立方体的最大组合数可以通过公式计算: $M^{n-1} (M-n)^{N-1}$;最后,矩阵 A 的每行就对应一个被选中的小超立方体,在每个被选中的小超立方体内随机产生一个样本,这样就选出 H 个样本。

1.3 抽样结果

基于拉丁超立方抽样的思想,确定每口抽水井的抽样数目, $H=30$ 。通过调用 VB 程序中的 Randomize() 函数实现抽样,结果见表 2。

2 径向基函数神经网络

径向基函数(Radial Basis Function,简称 RBF)神经网络是一种局部逼近的前馈式神经网络。它由输入层、隐含层和输出层组成,网络拓扑结构如图 2 所示。与误差逆传播(Error Back—Propagation,简称 BP)神经网络相比,RBF 神经网络收敛速度快,能够快速找到全局极小值。

设输入向量为 n 维向量 X ,输出向量为 l 维向量 Y ,则 RBF 网络隐含层第 i 个节点的输出为: $q_i = R(PX - c_i P)$,式中: c_i 是隐含层第 i 个节点的中心, $i=1, 2, \dots, m$, m 是隐含层节点数; $PX - c_i P$ 是向量 $X - c_i$ 的范数,代表输入向量离开该径向基函数中心 c_i 的程度; $R(\cdot)$ 为径向基函数,它在 c_i 处有唯一的一个最大值,随着 $PX - c_i P$ 的增大, $R(\cdot)$ 迅速衰减到零,对于给定的输入变量,只有一小部分靠近中心的被激活。

网络输出层第 k 个节点的输出为隐含层节点输出的线性组合:

$$y_k = a_i w_{ik} q_i - q_k$$

式中: w_{ik} ——第 i 个隐含层节点到第 k 个输出层节点的连接权值; q_k ——第 k 个输出层节点的阈值。

表 2 LHS 抽样结果

方案	井号				
	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅
1	45	30	304	165	625
2	150	130	1089	19	400
3	100	195	119	110	985
4	430	260	28	213	23
5	370	300	1219	390	1100
6	265	393	492	278	108
7	525	461	1437	336	1254
8	385	530	185	441	170
9	600	595	1219	1106	1432
10	735	657	365	500	215
11	1335	726	981	997	1319
12	800	790	1674	554	278
13	665	842	1325	629	1608
14	1407	918	743	1288	336
15	866	990	1561	1505	1384
16	1069	1055	230	599	456
17	800	1120	1620	1546	1537
18	938	1188	425	1210	1653
19	1742	1255	1507	720	518
20	1000	1319	552	1455	1198
21	1809	1386	620	778	579
22	1135	1452	1365	1391	1496
23	1900	1500	1277	884	684
24	1671	1569	1035	835	729
25	1268	1649	928	946	1132
26	1742	1715	679	1669	806
27	2000	1778	1175	1037	1021
28	1532	1846	800	1153	869
29	1875	1905	1158	1325	934
30	1600	1979	865	1615	916

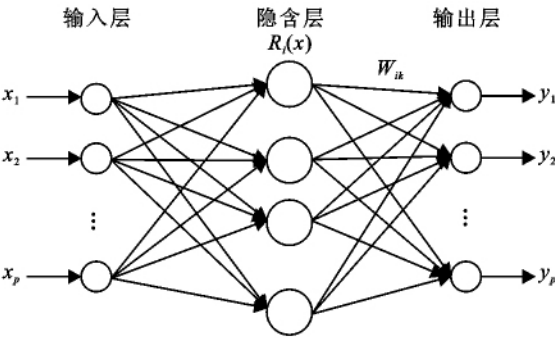


图 2 RBF 神经网络的拓扑结构

在学习阶段通过用大量的样本进行训练,不断修正神经网络的各参数,使神经网络对样本具有模式识

别能力,这样把训练样本数据集输入到 RBF 神经网络中就能够得到相应的目标输出数据集。基于 RBF 神经网络模型具有识别模式特征的能力,建立 RBF 神经网络模型作为地下水数值模拟模型的近似替代模型^[13-15]。

3 替代模型的建立

将抽样得到的数据代入到地下水数值模拟模型中运行,计算得到观测井地下水水位降深数据集,即输出数据集。

3.1 数据处理

为了模拟计算区内 5 口抽水井同时抽水对地下水水位降深的影响,利用 RBF 神经网络构建抽水强度与观测井水位降深之间的非线性函数关系。将拉丁超立方抽样得到的数据作为训练样本,并任意选取 5 组数据作为验证替代模型有效性的测试样本,如表 3 所示。

3.2 网络的构建与训练

利用输入(抽水量)、输出(降深)数据集分别建立抽水井抽水量—观测井水位降深均值、抽水井抽水量—观测井水位降深剩余标准差两个 RBF 神经网络模型。以 Matlab 6.5 为平台,调用 newrb()函数构建 RBF 神经网络模型。首先建立抽水井抽水量—观测井水位降深均值的 RBF 神经网络模型。输入层代表抽水井的抽水量,神经元个数与抽水井数目一致,定为 5;输出层表示观测井水位降深均值,神经元个数定为 1。通过运行计算机程序,得到 RBF 神经网络模型的水位降深均值与模拟模型拟合的误差。同理,构建抽水井抽水量—观测井水位降深剩余标准差的 RBF 神经网络模型,输出层表示观测井水位降深剩余标准差,神经元个数定为 1。通过调用 Matlab 程序,得出 RBF 神经网络模型的水位降深剩余标准差及与模拟模型结果拟合的误差。

3.3 替代有效性验证

RBF 神经网络模型的输出结果与地下水数值模拟模型计算得到的数据拟合效果较好。为了检验 RBF 神经网络模型是否能够替代地下水数值模拟模型,将 RBF 神经网络模型输出的水位降深均值和水位降深剩余标准差,分别与地下水数值模拟模型得到的结果进行误差拟合计算,如表 4 所示。

经验证,RBF 神经网络模型计算得到的水位降深均值与模拟模型计算结果拟合的平均相对误差为 0.038;水位降深剩余标准差与模拟模型计算结果拟合的平均相对误差为 0.042。拟合误差较小,说明

RBF 神经网络模型在功能上逼近模拟模型,能够有效地替代地下水数值模拟模型。

表 3 RBF 神经网络训练样本与测试样本数据

样本号	抽水量/(m ³ ·d ⁻¹)					观测井水位 降深均值/m	观测井水位降深 剩余标准差/m	备注
	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅	\bar{y}	S	
1	45	30	304	165	625	0.496968	0.447849	训练 样本
2	150	130	1089	19	400	0.491058	0.443612	
3	100	195	119	110	985	0.490845	0.443836	
4	430	260	28	213	23	0.498168	0.434339	
5	370	300	1219	390	1100	0.537995	0.465451	
6	265	393	492	278	108	0.51251	0.447243	
7	525	461	1437	336	1254	0.543607	0.473492	
8	385	530	185	441	170	0.51537	0.430239	
9	600	595	1219	1106	1432	0.580342	0.44418	
10	735	657	365	500	215	0.528012	0.431454	
11	1335	726	981	997	1319	0.577758	0.433291	
12	800	790	1674	554	278	0.566497	0.470527	
13	665	842	1325	629	1608	0.56024	0.457654	
14	1407	918	743	1288	336	0.589012	0.421565	
15	866	990	1561	1505	1384	0.615923	0.448406	
16	1069	1055	230	599	456	0.536002	0.421353	
17	800	1120	1620	1546	1537	0.619618	0.450367	
18	938	1188	425	1210	1653	0.572725	0.415609	
19	1742	1255	1507	720	518	0.584738	0.453291	
20	1000	1319	552	1455	1198	0.590627	0.417635	
21	1809	1386	620	778	579	0.566305	0.423437	
22	1135	1452	1365	1391	1496	0.610618	0.439895	
23	1900	1500	1277	884	684	0.590645	0.440086	
24	1671	1569	1035	835	729	0.579652	0.435076	
25	1268	1670	928	946	1132	0.5785	0.43143	
26	1742	1715	679	1669	806	0.616283	0.415575	
27	2000	1778	1175	1037	1021	0.59905	0.432311	
28	1532	1846	800	1153	869	0.590392	0.421759	
29	1875	1905	1158	1325	934	0.613212	0.427985	
30	1600	1979	865	1615	916	0.618205	0.420699	
1	720	356	983	1067	562	0.512664	0.46123	测试 样本
2	348	1385	862	1468	289	0.593256	0.53165	
3	185	942	257	359	827	0.512534	0.39012	
4	529	487	246	1322	1479	0.532882	0.40175	
5	765	654	234	925	1632	0.563761	0.42518	

表 4 RBF 模型输出结果与地下水数值模拟模型的平均相对拟合误差

样本	观测井水位降 深均值/m		观测井水位降深剩余 标准差/m	
	RBF 模型	地下水数值 模拟模型	RBF 模型	地下水数值 模拟模型
1	0.5628	0.512664	0.4481	0.46123
2	0.5926	0.593256	0.4904	0.53165
3	0.5184	0.512534	0.4217	0.39012
4	0.5581	0.532882	0.4074	0.40175
5	0.5446	0.563761	0.4198	0.42518

4 结 论

结合地下水数值模拟模型,采用拉丁超立方抽样,获取建立替代模型所需的输入输出数据集,应用人工神经网络的方法,构建了 RBF 神经网络模型作为地下水数值模拟模型的替代模型。经验证,RBF 神经网络模型输出的结果与模拟模型的计算结果拟合误差较小,表明 RBF 神经网络模型在功能上逼近模拟模型,能够有效地替代地下水数值模拟模型,可以在优化模型迭代求解过程中直接调用替代模型解决多目标优化问题。

替代模型法能够克服以往耦合技术方法存在的局限性,大幅度地减少优化模型求解计算过程中直接调用模拟模型所造成的计算负荷,节省大量时间,是一种具有挖掘潜力和实用价值的解决问题的途径。但本文只验证了 RBF 神经网络模型本身的替代有效性,并未对模拟优化迭代求解过程中替代模型的实用性进行验证,因此建议将 RBF 神经网络模型应用到实际中,继续检验其作为替代模型的适用性,进一步完善替代模型理论,提高替代模型的准确性。

参考文献:

- [1] He L, Huang G H, Lu H W. A stochastic optimization model under modeling uncertainty and parameter certainty for groundwater remediation design-Part 1: model development[J]. *Journal of Hazardous Materials*, 2010, 176(1/3):521-526.
- [2] Qin X S, Huang G H, He L. Simulation and optimization technologies for petroleum waste management and remediation process control[J]. *Journal of Environmental Management*, 2009, 90(1):54-76.
- [3] 卢文喜. 地下水系统的模拟预测和优化管理[M]. 北京: 科学出版社, 1999.
- [4] Qin X S, Huang G H, Chakma A, et al. Simulation-based process optimization for surfactant-enhanced aquifer remediation at heterogeneous DNAPL-contaminated sites[J]. *Science of the Total Environment*, 2007, 381(1/3):17-37.
- [5] He L, Huang G H, Lu H W, et al. Optimization of surfactant-enhanced aquifer remediation for a laboratory BTEX system under parameter uncertainty[J]. *Environmental Science & Technology*, 2008, 42(6):2009-2014.
- [6] Qin X S, Huang G H, Chakma A, et al. Simulation-based process optimization for surfactant-enhanced aquifer remediation at heterogeneous DNAPL-contaminated sites[J]. *Science of the Total Environment*, 2007, 381(1/3):17-37.
- [7] Johnson V M, Rogers L L. Accuracy of neural network approximators in simulation-optimization[J]. *Journal of Water Resources Planning and Management (ASCE)*, 2000, 126(2):48-56.
- [8] Yan S Q, Minsker B. Optimal groundwater remediation design using an adaptive neural network genetic algorithm[J]. *Water Resources Research*, 2006, 42(5):407-420.
- [9] 薛禹群. 地下水动力学[M]. 北京: 地质出版社, 1997.
- [10] McKay M D, Beckman R J, Conover W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. *Technometrics*, 1979, 21(2):39-245.
- [11] 熊学玉, 顾炜. 基于改进 LHS 方法的预应力混凝土结构长期性能概率分析[J]. *工程力学*, 2010, 27(4):163-168.
- [12] 吴振君, 王水林, 葛修润. LHS 方法在边坡可靠度分析中的应用[J]. *岩土力学*, 31(4):1047-1054.
- [13] Asim R, Sandeep G, Raymond M. A neural-network learning theory and a polynomial time RBF algorithm[J]. *IEEE Transactions on Neural Network*, 1997, 18(6):1301-1313.
- [14] Lin Faajeng, Wai Rongjong, Duan Rouyong. Fuzzy neural networks for identification and control of ultrasonic motor drive with LLC resonant technique[J]. *IEEE Transactions on Industrial Electronics*, 1999, 46(5):999-1011.
- [15] 王洪斌, 杨香兰, 王洪瑞. 一种改进的 RBF 神经网络学习算法[J]. *系统工程与电子技术*, 2002, 24(6):103-105.