

NPSO - GRNN 在城市生活需水量预测中的应用

李万庆¹, 王静¹, 孟文清²

(1. 河北工程大学 经管学院, 河北 邯郸 056038; 2. 河北工程大学 土木学院, 河北 邯郸 056038)

摘要: 利用小生境粒子群算法(NPSO)的全局搜索能力,对GRNN的光滑因子进行优化,提高了广义回归神经网络(GRNN)的网络性能,建立了基于NPSO-GRNN的城市生活需水量预测模型。研究表明:利用该模型拟合和预测北京市1988-2012年生活需水量数据的平均相对误差绝对值分别为0.72%和0.36%,两者都比BP神经网络模型拟合预测的结果低;NPSO-GRNN模型能更好的拟合北京市城市生活需水量的变化趋势,预测精度更高,泛化能力更强。

关键词: GRNN; 小生境粒子群算法; 生活需水量; 预测模型

中图分类号: TU991

文献标识码: A

Application of NPSO - GRNN in Urban Domestic Water Demand Prediction

LI Wan-qing¹, WANG Jing¹, MENG Wen-qing²

(1. College of Economics and Management, Hebei University of Engineering, Hebei Handan 056038, China;

2. College of Civil Engineering, Hebei University of Engineering, Hebei Handan 056038, China)

Abstract: this paper used the niche particle swarm optimization (NPSO) global search ability to optimize the spread parameter of GRNN, which improved the general regression neural network (GRNN) performance; then a urban domestic water demand predictive model was established by basing on NPSO-GRNN. The results showed that NPSO-GRNN fitting and prediction average relative absolute error of Beijing domestic water demand data between 1988 and 2012 were 0.72% and 0.36%, respectively, the fitting and predicted results were lower than the result of BP neural network algorithm. NPSO-GRNN algorithm can be better fitting to the trend of urban domestic water demand in Beijing City, it has higher prediction accuracy and generalization ability.

Key words: GRNN; niche particle swarm optimization; domestic water demand; prediction model

随着我国城镇化的推进,居民对生活条件的要求不断提升,从而导致水资源的消耗量越来越大,特别是北方的大中型城市,因此,对城市生活需水量进行科学合理的分析预测,不但可以满足居民对水资源的需求,而且可以保证供水系统的正常运行。综合国内外研究现状,我们发现传统的预测模型中存在忽视影响需水量测算因素之间的动态关系,或样本数据量不足等诸多缺陷。例如,常淑玲等^[1]通过指数模型和移动平均数法来构建城市需水量综合动态模型,这种方法忽略了需水量各个影响因素之间的动态相互制约关系。张雅君等^[2]选用多元线性回归分析法,对影响北

京城市生活需水量的因素进行探讨,但该方法受到了显著性检验的制约,选定的自变量只有在通过显著性检验的前提下才能建立回归方程,计算量较大。随着科学和计算机技术的不断发展,新的人工智能方法,如BP神经网络算法,在需水量预测领域得到了广泛的应用^[3-5]。但该算法容易陷入局部极小值,收敛性差,给实际应用带来了一定的限制。

针对传统预测方法在需水量预测方面存在的不足,本文尝试将NPSO算法引入到GRNN神经网络中,建立NPSO-GRNN神经网络预测模型,希望能为未来城市生活需水量规划提供参考依据。

收稿日期: 2014-10-31

作者简介: 李万庆(1954-),男,河北临漳人,博士,教授,从事工程管理理论、技术及应用方面的研究。

1 需水量预测模型

1.1 GRNN 模型

GRNN 作为径向基神经网络的一个分支,在 1991 年,由美国学者 Donald F. Specht 提出^[6-7]。

GRNN 网络结构由四层构成,如图 1 所示。

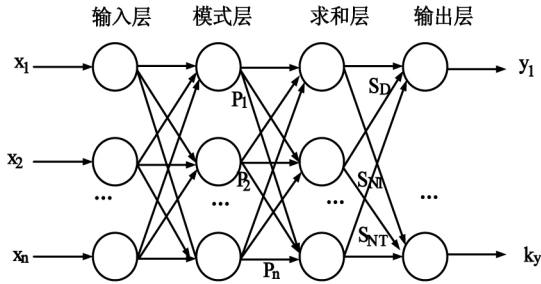


图1 GRNN网络结构图

Fig.1 GRNN net structure

GRNN 由输入层、模式层、求和层和输出层四层网络构成。对应网络输入为 $X = [x_1, x_2, \dots, x_n]^T$ 输出为 $Y = [y_1, y_2, \dots, y_k]^T$ 。

输入层。学习样本中输入向量的维数和输入层神经元的数目是相同的,各神经元是简单的分布单元,将输入变量直接传递给模式层。

模式层。学习样本的数目和该层神经元数目是相同的,每个神经元对应不同的学习样本,该层神经元传递函数为

$$P_i = \exp \left[- \frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right] \quad i = 1, 2, \dots, n \quad (1)$$

式中 X - 网络输入变量; X_i - 第 i 个神经元对应的学习样本; σ - 光滑因子。

求和层。对求和层进行求和时应用两种类型的神经元。

第一类计算公式为

$$\sum_{i=1}^n \exp \left[- \frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right] \quad (2)$$

式中所有的模式层中的神经元的输出算术求和,模式层与各神经元的连接权值为 1,传递函数

为 $S_D = \sum_{i=1}^n P_i$ 。

第二类计算公式为

$$\sum_{i=1}^n Y_i \exp \left[- \frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right] \quad (3)$$

式中对所有模式层中的神经元加权求和,模式层中第 i 神经元和求和层中第 j 个分子求和神经元之间的连接权值是第 i 个输出样本 Y_i 中的第 j 个

元素,传递函数为 $S_{Nj} = \sum_{i=1}^n y_{ij} P_i \quad j = 1, 2, \dots, k$ 。

输出层。输出层中神经元数目等于学习样本中输出向量的维数 k ,各神经元把求和层的输出相除,神经元 j 输出对应估计结果 $\hat{y}(x)$ 的第 j 个元素,即 $\hat{y}_i = \frac{S_{Nj}}{S_D} \quad j = 1, 2, \dots, k$ 。

1.2 NPSO 算法

粒子群算法最初是由美国社会心理学家 Kennedy 和电气工程师 Eberhart 在 1995 年共同提出的^[8]。在粒子群算法中,我们将每个个体看做是一个在 D 维搜索空间中不存在质量和体积的粒子,并且拥有飞行速度^[9-10]。该算法中粒子的速度和位置进化方程分别为

$$v_j(t+1) = \omega v_j(t) + c_1 r_1 (p_j(t) - x_j(t)) + c_2 r_2 (p_g(t) - x_j(t)) \quad (4)$$

$$x_j(t+1) = v_j(t+1) + x_j(t) \quad (5)$$

式(4)和式(5)中 $v_j(t)$ - 粒子 j 在第 t 代的速度; ω - 惯性权重; c_1 - 认知系数; r_1, r_2 - 服从均匀分布的随机数; $p_j(t)$ - 粒子 j 的个体历史最优位置; $x_j(t)$ - 粒子 j 在第 t 代的位置; c_2 - 社会系数; $p_g(t)$ - 群体历史最优位置; $x_j(t+1)$ - 粒子 j 在第 $t+1$ 代的位置。

2002 年, Brits 在粒子群优化算法中引入了小生境技术,提出了 NPSO 算法。为保持粒子群的多样性,如果某个粒子在运算连续多次迭代中对应的适应值变化量很小,则以此粒子为中心,以此粒子与其最近的粒子的距离半径构造一个圆形小生境^[11]。将小生境粒子群的半径定义为

$$R_{s_j} = \max \{ \|x_{s_j g} - x_{s_j i}\| \} \quad (6)$$

式中 $x_{s_j g}$ - 子粒子群 S_j 中的最优粒子; $x_{s_j i}$ - 子粒子群 S_j 中任一非最优粒子。

算法有两个核心操作: (1) 如果粒子 x_i 进入子粒子群 S_j 的范围内,即 $\|x_i - x_{s_j i}\| \leq R_s$,则此小生境粒子群将会把该粒子吸收; (2) 如果两个子粒子群 S_j, S_k 的范围相交,即 $\|x_i - x_{s_j i}\| \leq |R_{s_j} - R_{s_k}|$,则两个子粒子群将被合并成一个。

1.3 NPSO - GRNN 需水量预测具体步骤

步骤 1 种群初始化及参数初始化设置,粒子规模 N ,认知系数和社会系数 c_1, c_2 ,迭代终止条件。

步骤 2 鉴于影响需水量预测的各个因素的量纲和数量级不同,要先对原始数据进行归一化处

理再进行训练。本文对原始数据应用比例压缩法进行处理,公式如下:

$$T = T_{\min} + \frac{T_{\max} - T_{\min}}{X_{\max} - X_{\min}}(X - X_{\min}) \quad (7)$$

式中 X - 原始数据; X_{\max} - 原始数据的最大值; X_{\min} - 原始数据的最小值; T - 变换后的数据,也称之为目标数据; T_{\max} , T_{\min} - 目标数据的最大值和最小值。

步骤 3 确定适应度函数: 把训练样本的输出值与实际值的均方差 F 看做适应度函数寻找最优 $\sigma^{[12]}$ 。适应度函数为

$$F = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (8)$$

式中 y_i - 实际值 \hat{y}_i - 计算值。训练样本的 F 越小,越有利于算法迭代停止。

步骤 4 将学习样本和粒子带入 GRNN。

步骤 5 对每个粒子的适应值进行计算,对最优的适应值和个体进行保留,检验是否满足优化条件,如果满足误差精度,则结束。否则,进入下一个粒子的小生境群体进行优化,全局极值对应

的粒子种群当前的最优解为 GRNN 的光滑因子。

步骤 6 若最优值没有找到,则对每个粒子的小生境群体保留的最优个体组成新的群体空间,重新确定小生境种群个体,重复步骤(4)。

步骤 7 通过小生境粒子群算法优化后,算法终止时,全局极值点的位置对应的值即是需水量预测 GRNN 模型中的光滑因子的取值,将其代入 GRNN 模型中进行训练学习,就可以用来进行预测模型的求解。

2 模型应用实例分析

2.1 数据来源及分析

通过筛选《北京统计年鉴》中提供的相关用水资料,1988—2012 年北京市城市生活用水量相关数据如表 1 所示。

本文依据表 1 中 1988—2012 年影响城市生活需水量的城镇人口、第三产业产值、人均居住面积、人均可支配收入作为 NPSO - GRNN 的输入,输出则为城市生活需水量,网络训练样本为 1988

表 1 北京市 1988—2012 年城市生活用水量相关数据表

Tab. 1 The related data tables of urban domestic water consumption in Beijing from 1988 to 2012

年份	城市生活用水量/ 10^8 m^3	城镇人口/万人	第三产业产值/亿元	人均居住面积/ m^2	人均可支配收入/元
1988	5.09	650.0	151.88	7.17	1 437
1989	5.14	664.0	165.22	7.45	1 597.1
1990	5.86	798.0	194.55	7.72	1 787.1
1991	5.66	808.0	261.84	8.01	2 040.4
1992	5.84	819.0	314.52	8.31	2 363.7
1993	6.26	831.0	412.90	8.51	3 296
1994	6.73	846.0	560.20	8.73	4 731.2
1995	7.05	946.2	788.40	8.87	5 868.4
1996	6.99	957.9	999.50	9.17	6 885.5
1997	7.56	948.3	1 218.00	9.49	7 813.1
1998	7.87	957.7	1 458.70	9.87	8 472
1999	8.27	971.7	1 693.20	10.43	9 182.8
2000	8.49	1 057.4	2 049.1	11.5	10 349.7
2001	12	1 081.2	2 484.8	15.4	11 577.8
2002	10.8	1 118.0	2 982.6	19.22	12 463.9
2003	13	1 151.3	3 435.9	19.71	13 882.6
2004	12.8	1 187.2	4 092.2	21.49	15 637.8
2005	13.4	1 286.1	4 854.3	22.03	17 653
2006	13.7	1 350.2	5 837.6	23.65	19 978
2007	13.9	1 416.2	7 236.1	24.77	21 989
2008	14.7	1 503.6	8 375.8	26.9	24 725
2009	14.7	1 581.1	9 179.2	27.69	26 738
2010	14.8	1 686.4	10 600.8	28.94	29 073
2011	15.6	1 740.7	12 363.1	29.38	32 903
2012	16	1 783.7	13 669.9	29.26	36 469

表2 NPSO-GRNN 和 BP 神经网络训练结果及对比分析

Tab. 2 NPSO-GRNN and BP neural network training results and their comparative analysis

年份	实际值/ 10^8 m^3	NPSO-GRNN 模型		BP 神经网络模型	
		计算值/ 10^8 m^3	相对误差/%	计算值/ 10^8 m^3	相对误差/%
1988	5.09	5.11	-0.392 93	5.1	-0.196 46
1989	5.14	5.12	0.389 105	5.2	-1.167 32
1990	5.86	5.80	1.023 891	5.6	4.436 86
1991	5.66	5.81	-2.650 18	5.76	-1.766 78
1992	5.84	5.87	-0.513 7	5.94	-1.712 33
1993	6.26	6.13	2.076 677	6.32	-0.958 47
1994	6.73	6.64	1.337 296	6.52	3.120 357
1995	7.05	7.04	0.141 844	6.79	3.687 943
1996	6.99	7.14	-2.145 92	7.06	-1.001 43
1997	7.56	7.53	0.396 825	7.46	1.322 751
1998	7.87	7.85	0.254 13	7.68	2.414 231
1999	8.27	8.15	1.451 028	8.12	1.813 785
2000	8.49	8.49	0	8.86	-4.358 07
2001	12	12	0	11.48	4.333 333
2002	10.8	10.89	-0.833 33	10.45	3.240 741
2003	13	12.91	0.692 308	12.39	4.692 308
2004	12.8	12.80	0	12.29	3.984 375
2005	13.4	13.4	0	12.95	3.358 209
2006	13.7	13.7	0	13.49	1.532 847
2007	13.9	13.9	0	14.30	-2.8777
训练样本 MAPE		0.714 958		2.598 815	

表3 NPSO-GRNN 和 BP 神经网络预测结果及对比分析

Tab. 3 NPSO-GRNN and BP neural network prediction results and their comparative analysis

年份	实际值/ 10^8 m^3	NPSO-GRNN 模型		BP 神经网络模型	
		计算值/ 10^8 m^3	相对误差/%	计算值/ 10^8 m^3	相对误差/%
2008	14.7	14.58	0.816 327	14.78	-0.544 22
2009	14.7	14.65	0.340 136	14.97	-1.836 73
2010	14.8	14.78	0.135 135	15.06	-1.756 76
2011	15.6	15.56	0.256 41	16.30	-4.487 18
2012	16	15.96	0.25	16.58	-3.625
预测样本 MAPE		0.359 602		2.449 978	

—2007 年的 15 组数据,网络预测检验样本为 2008—2012 年的 5 组数据。

2.2 需水预测结果及分析

设置小生境粒子群中粒子规模 $N = 30$, 认知系数和社会系数 $c_1 = c_2 = 2$ 迭代终止条件为: 当训练误差达到 10^{-4} 或最大迭代次数 100。

运用训练好的 NPSO-GRNN 模型和传统的 BP 神经网络模型对北京市城市生活需水量进行预测,并以平均相对误差的绝对值 MAPE (Mean Absolute Percentage Error) 作为衡量预测精度的指标。MAPE 的公式为

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

式中 y_i - 实际值; \hat{y}_i - 计算值。

训练结果和预测结果见表 2 和表 3 相对误差

拟合图和训练预测拟合结果见图 2。

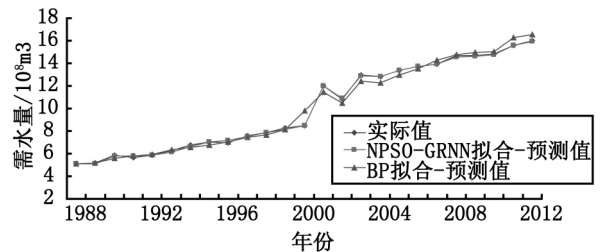


图2 需水量预测模型拟合-预测值

Fig.2 Fitting-prediction value of water demand forecasting model

通过分析表 2、表 3、图 2,说明: 无论是 NPSO-GRNN 模型还是 BP 神经网络模型,在训练和预测上两个模型的 MAPE 均控制在 5% 以内,达到了较高的精准度。表明: 建立的 NPSO-GRNN 城市生活需水预测模型是可行的,预测的精度高,算法稳定,与 BP 神经网络模型相(下转第 106 页)

综上可得,电站锅炉综合安全等级隶属于一般安全级别,与该电站锅炉系统实际情况进行对比,结果表明评价结果与企业实际相符。

3 结论

1) 综合评价结果的可靠性和准确性依赖于合理确定评价指标体系、指标的权重分配和综合评价的合成算子等。需要根据评价系统的特点来选取合适的评价模型和算法,使评价结果更加客观、科学和有针对性。

2) 各评价指标的权重值及单因素评价矩阵可根据各电站锅炉企业的具体情况而定,使模糊评价结果更符合企业的实际情况。

参考文献:

- [1]姜磊,杨俊保.电站锅炉故障诊断技术的发展[J].上海电力学院学报,2010,26(6):541-543.
 [2]彭小兰,范晓明.基于安全检查表法的锅炉厂质量保证体系的安全评价[J].安全与环境工程,2010,17(6):47-50.
 [3]程华瑞,栗继祖.运用TNT当量法及事故树法在分析

(上接第102页)比,还具有收敛速度快,不易陷入局部极小值和调整参数少等优势,在实际应用中有很大的前景。

3 结论

1) 构建了基于NPSO算法的NPSO-GRNN需水量预测模型,提高了网络的泛化能力和拟合性能。

2) NPSO-GRNN预测模型具有比传统BP神经网络模型预测精度高,不易陷入局部极小值,调整参数少等优点,在实际应用上有良好的前景,为城市生活需水量预测提供了新方法。

参考文献:

- [1]常淑玲,尤学一.天津市需水量预测研究[J].干旱区资源与环境,2008,02:14-19.
 [2]张雅君,刘全胜,冯萃敏.多元线性回归分析在北京城市生活需水量预测中的应用[J].给水排水,2003,04:26-29.
 [3]景亚平,张鑫,罗艳.基于灰色神经网络与马尔科夫链的城市需水量组合预测[J].西北农林科技大学学报:自然科学版,2011(7):229-234.
 [4]CUBERO R G. Neural networks for water demand time series forecasting [C]. Artificial Neural Networks Int .

锅炉爆炸的危害[J].太原理工大学学报,2013,44(3):361-365.

- [4]郭驰,杨扬,王超,等.模糊层次分析法在小额贷款信用评级中的应用[J].河北工程大学学报:自然科学版,2013,30(2):106-109.
 [5]王飞,胡静娴,黄晶.APH和模糊综合评判在绿色建筑中的评价研究[J].河北工程大学学报,2014,31(2):107-110.
 [6]魏胜桃,张树海,李文军,等.模糊综合评价在焦化企业安全评价中的应用[J].工业安全与环保,2010,36(7):58-59.
 [7]刘新宪,朱道立.选择与判断[M].上海:上海科学普及出版社,1990.
 [8]王雪颖,郭徽,房新亮.模糊综合评价法在煤与瓦斯突出危险性评估中的应用[J].工业安全与环保,2012,38(6):31-34.
 [9]杨纶标,高英仪.模糊数学原理及应用[M].华南理工大学出版社,2006.
 [10]鲍学英,王起才,宫文昌.基于模糊层次分析法的地铁工程施工安全评价[J].中国安全生产科学技术,2013,9(1):136-139.

(责任编辑 刘存英)

Workshop, IWANN, 1991, 453-460.

- [5]魏津瑜,苏思沁,施鹤南.基于小波分析的城市需水量预测[J].中南大学学报:自然科学版,2013(Z2):183-187.
 [6]黄崇珍,梁静国.基于GRNN的海上钻井平台建造质量预测研究[J].哈尔滨工程大学学报,2009(3):339-343.
 [7]陈伟根,奚红娟,苏小平,等.广义回归神经网络在变压器绕组热点温度预测中的应用[J].高电压技术,2012(1):16-21.
 [8]李玉军,汤晓君,刘君华.粒子群优化算法在混合气体红外光谱定量分析中的应用[J].光谱学与光谱分析,2009(5):1276-1280.
 [9]MATLAB中文论坛.MATLAB神经网络30个案例分析[M].北京:北京航空航天大学出版社,2010.
 [10]王晓敏,刘宏伟,李石妍.改进型混沌粒子群算法求解函数均值问题[J].河北工程大学学报:自然科学版,2011,28(3):100-104.
 [11]刘文颖,谢昶,文晶,等.基于小生境多目标粒子群算法的输电网检修计划优化[J].中国电机工程学报,2013(4):141-148.
 [12]李智勇,童调生.基于多种群进化小生境遗传算法的神经网络进化设计方法研究[J].控制与决策,2003(5):607-610.

(责任编辑 刘存英)