

三江平原土壤质量变化评价的模糊神经网络模型构建及应用^{*}

杨先野, 付 强, 王宝华, 姜秋香
(东北农业大学 水利与建筑学院, 哈尔滨 150030)

摘 要: 针对土壤质量变化评价指标值的选取具有模糊性的特点, 将 BP 网络和模糊数学相结合建立了土壤质量评价的模糊 BP 网络模型。选取三江平原具有典型性土壤质量变化趋势的试验点——富锦市永富乡土壤进行取样分析, 选择与土壤质量相关的指标因子和相应的等级值建立评价模型。结果表明, 该模型具有很强的学习能力, 对参数的改变也有较强的鲁棒性, 而且各个评价指标的权重都是由网络通过学习自动生成的, 避免了主观赋权的人为干扰, 因此评价结果更具有客观性和准确性, 同时也为该领域研究提供了新的思路。

关键词: 土壤质量; 评价; BP 网络; 模糊数学

中图分类号: S159

文献标识码: A

文章编号: 1005-3409(2008)03-0054-04

Construction and Application of Fuzzy Neural Networks Model in Evaluation on the Soil Quality Changes of Sanjiang Plain

YANG Xiann ye, FU Qiang, WANG Baohua, JIANG Qiuxiang

(College of Water Conservancy & Architecture, Northeast Agricultural University, Harbin 150030, China)

Abstract: Choosing evaluation targets of soil quality changing has fussiness, so fussy BP network model is created through the combination of BP network and fussy mathematics. Experimental area is in Yongfu town, Fujin city, in Sanjiang plain where the trend of soil quality changing is typical. Evaluation model is created through choosing the factor of soil quality changing and corresponding class value. The results show learning capability of the model is well, and the model manifests well robustness with parameter changing. The weight of evaluation target generates automatically through network itself studying, avoids the human interferences of weighting subjectively. So the evaluation result is objective and accurate. The model offered new idea for the research in this domain.

Key words: soil quality; evaluation; BP network; fuzzy mathematics

1 引 言

土壤质量是土壤特性的综合反映, 也是揭示土壤条件动态变化的最敏感的指标。近年来, 国内、国外学者将人工智能方法逐渐应用在土壤质量的评价中, 并取得了较好效果。赵小勇等针对传统评价方法存在指标权重确定没有统一的理论和计算方法的不足, 提出用基于实码加速遗传算法的投影寻踪等级评价模型^[1]; 刘梦云等综合了多变量指标克立格法、土壤质量综合评分法和土壤相对质量评价法各种方法的优点, 制定了有区域特色的层次分析法来评价土壤质量^[2]; 胡焱弟等针对土壤环境质量评价缺乏精确客观的评价数值作为目标集的难点, 提出了径向基函数神经网络^[3], 还有一些其他方法^[4-7]在土壤质量评价上也取得了良好的效果。

三江平原位于黑龙江省的东北部, 是我国重要的商品粮基地, 土地面积达 10.89 万 km²。由于 20 世纪 80 年代以前的

土地开发缺乏环境与水土保持意识和盲目追求经济效益, 导致土壤退化, 直接威胁到该区土壤的可持续利用和农业的持续发展。选取三江平原具有典型性土壤质量变化趋势的试验点——富锦市永富乡耕作土壤进行剖面取样分析, 选择与土壤质量相关的指标因子(有机质、全氮、全磷、全钾、速效氮、速效磷、速效钾、CEC、pH 值)和相应的等级值建立了土壤质量评价的模糊神经网络(Fuzzy Neural Network, 简称 FNN)模型, 评价结果表明, 该模型取得了很好的评价效果, 为土壤质量评价的研究提供了新的研究方法, 同时也为本市乃至整个三江平原地区今后土地开发提供一定的科学决策依据。

2 模糊神经网络的构建

2.1 输入数据的标准化

由于土壤质量评价的各指标因子之间大小相差很大, 因此应将指标因子采用公式(1)进行标准化。

^{*} 收稿日期: 2007-08-25

基金项目: 国家自然科学基金项目(30400275)

作者简介: 杨先野(1982-), 男, 黑龙江富锦人, 硕士, 主要从事系统工程建模、水土资源优化利用与系统分析研究。E-mail: xianyeyang@163.com

通信作者: 付强(1973-), 男, 辽宁锦州人, 教授, 博士生导师, 主要从事农业水土资源系统分析、节水灌溉及农业系统工程建模与优化技术研究。E-mail: fuqiang100@371.net

$$y = \left(\frac{0.9 - 0.1}{x_{\max} - x_{\min}} \right) \times x + \left[0.9 - \left(\frac{0.9 - 0.1}{x_{\max} - x_{\min}} \right) \times x_{\max} \right] \quad (1)$$

式中: x_{\max} , x_{\min} —— 指标因子的最大值和最小值; y —— 标准化后的指标因子。

2.2 指标因子 y 的模糊化

设由 y 构成的征兆域向量 $K = [K_1, K_2, \dots, K_n]$, 考虑征兆域向量中分量 $K_q (q = 1, 2, \dots, n)$ 的具体变化趋势和其极限值, 应用模糊理论^[8], 确定隶属函数为梯形分布函数, 描述 K_q 的语义值的模糊集为 {低(L), 高(H)}。

梯形分布函数的解析式如下:

偏低型, 即 $\mu_L(K_q)$:

$$\mu_{L1}(x) = \begin{cases} 1 & x < a \\ \frac{b-x}{b-a} & a \leq x \leq b \\ 0 & x > b \end{cases} \quad (2)$$

偏高型, 即 $\mu_H(K_q)$:

$$\mu_{L2}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases} \quad (3)$$

a, b 分别为上述分布函数中具有特殊几何意义的点, 取值可根据其在隶属函数中的特殊几何含义针对征兆域中的不同分量的具体范围分别予以确定, 但应确保测量的任一数据分量 K_q 的 $\mu_L(K_q)$ 和 $\mu_H(K_q)$ 中至少有 1 个必大于 0.5 (其中 $\mu_L(K_q)$ 和 $\mu_H(K_q)$ 分别表示该分量隶属于“低”和“高”的隶属度)。隶属函数是通过实践来检验, 利用信息反馈不断调整, 使隶属函数逐步达到符合客观规律。

2.3 神经网络训练

知识的获取属多层前馈神经网络的训练精度高, 且泛化结果令人满意, 而且三层前馈神经网络可以以任意期望精度逼近任意函数, 如三层 BP 网络 (Back Propagation, 简称 BP)。以指标因子 y 模糊化后的 K_q 作为神经网络的输入, 相应的等级值作为网络的输出, 建立 BP 网络。

设 BP 网络的输入层、隐含层和输出层的节点数分别为 n, p, q , w_{ij} 是输入层第 i 单元与隐含层第 j 单元间的联接权重, θ_j 是隐含层第 j 单元的阈值, w_{jt} 是隐含层第 j 单元与输出层第 t 单元之间的联接权重, θ_t 是输出层第 t 单元的阈值, $f(\cdot)$ 为隐含层到输出层的激活函数, 取单极性的 Sigmoid

函数; BP 网络在运行时分为 2 个阶段:

(1) 第一阶段——样本 K_q 的正向运算 (用 S 函数计算):

隐含层到输出层的输出为

$$y_j = f(I_j) = f\left(\sum_{i=1}^n W_{ij} x_i - \theta_j\right) \quad j = 1, 2, \dots, p \quad (4)$$

输入层到隐含层的输出为

$$y_t = f(I_t) = f\left(\sum_{j=1}^p W_{jt} y_j - \theta_t\right) \quad t = 1, 2, \dots, q \quad (5)$$

其中

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

(2) 第二阶段——误差反向传播:

由于期望输出与实际输出不一致, 因而产生误差, 通常用方差来表示这一误差:

$$e_t = \frac{1}{2} \sum_{i=1}^q (d_i - y_i)^2 \quad (7)$$

所以用下面公式来修正输出层及隐含层各单元的权重及阈值:

隐含层各单元的调整误差 δ_j :

$$\delta_j = - \frac{\partial e_t}{\partial I_j} = - \frac{\partial e_t}{\partial y_j} y_j (1 - y_j) = y_j (1 - y_j) \cdot \sum_{t=1}^q \delta_t W_{jt} \quad (8)$$

输出层各单元的调整误差 δ_t :

$$\delta_t = - \frac{\partial e_t}{\partial I_t} = (d_t - y_t) f'(I_t) = (d_t - y_t) y_t (1 - y_t) \quad (9)$$

隐含层修正的权重 W 和阈值 θ :

$$\begin{cases} W_{ij}(k+1) = W_{ij}(k) + \lambda \delta_j x_i \\ \theta_j(k+1) = \theta_j(k) + \lambda \delta_j \end{cases} \quad (10)$$

输出层修正的权重 W 和阈值 θ :

$$\begin{cases} W_{jt}(k+1) = W_{jt}(k) + \lambda \delta_t y_j \\ \theta_t(k+1) = \theta_t(k) + \lambda \delta_t \end{cases} \quad (11)$$

对 BP 网络进行训练, 直至网络全局误差函数小于设定值, 即网络收敛为止。

2.4 FNN 结构

FNN 是神经网络与模糊数学相结合的产物, 具有学习、联想、自适应性, 又能进行模糊推理^[11-12]。本文建立了一个 FNN, 如图 1 所示。

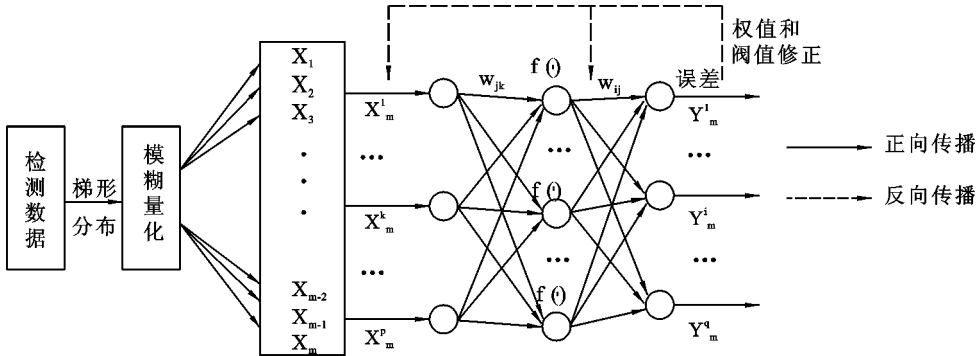


图 1 模糊神经网络结构原理

该结构采用的是模糊 BP 网络: 在模糊量化中, 它的每一个节点代表一个输入变量(K_q), 通过梯形隶属函数将输入变量模糊化; 在 BP 网络中有 3 层, 其中第 1 层为量化输入层, 即为模糊化的输出; 第 2 层为 BP 网络的隐含层, 用于实现输入变量模糊值到输出变量模糊值的映射; 第 3 层为量化输出层, 其输出是该网络最终的输出。每个输出神经元代表着一个土壤等级。

3 应用实例

主要借用文献[1] 的资料, 选择与土壤质量相关的物理、化学指标(生物学指标难以测定), 指标因子包括: 有机质、全氮、全磷、全钾、速效氮、速效磷、速效钾、CEC、pH 值。各个评价指标的分级标准见表 1^[1]。富锦市永富乡表层土壤的

原始实测数据见表 2^[1]。

表 1 三江平原土壤(表层土) 质量评价标准

评价指标	I	II	III	IV	V
有机质/($\text{g} \cdot \text{kg}^{-1}$)	> 80	60~ 80	40~ 60	20~ 40	< 20
全氮/($\text{g} \cdot \text{kg}^{-1}$)	> 5	3. 5~ 5	2~ 3. 5	0. 5~ 2	< 0. 5
全磷/($\text{g} \cdot \text{kg}^{-1}$)	> 2	1. 5~ 2	1~ 1. 5	0. 5~ 1	< 0. 5
全钾/($\text{g} \cdot \text{kg}^{-1}$)	> 25	17. 5~ 25	10~ 17. 5	2. 5~ 10	< 2. 5
速效氮/($\text{mg} \cdot \text{kg}^{-1}$)	> 350	275~ 350	200~ 275	125~ 200	< 125
速效磷/($\text{mg} \cdot \text{kg}^{-1}$)	> 100	70~ 100	40~ 70	10~ 40	< 10
速效钾/($\text{mg} \cdot \text{kg}^{-1}$)	> 350	270~ 350	190~ 270	110~ 190	< 110
CEC/($\text{cmol} \cdot \text{kg}^{-1}$)	> 200	150~ 200	100~ 150	50~ 100	< 50
pH	6~ 6. 5	5. 5~ 6, 6. 5~ 7	5~ 5. 5	7~ 7. 5	< 5, > 7

表 2 原始实测数据

采样地点	开垦 年限/a	有机质	全氮	全磷	全钾	速效氮	速效磷	速效钾	CEC/	pH
		$\text{g} \cdot \text{kg}^{-1}$				$\text{mg} \cdot \text{kg}^{-1}$			$(\text{cmol} \cdot \text{kg}^{-1})$	
富锦市 永富乡	荒地	96. 54	6. 23	1. 30	5. 41	840. 4	13. 2	217. 6	92. 7	6. 4
	5	49. 62	3. 35	0. 89	6. 90	360. 3	10. 50	281. 6	85. 7	6. 1
	10	34. 00	2. 61	1. 15	8. 05	289. 1	7. 28	125. 3	98. 4	5. 9
	25	22. 15	1. 46	0. 89	6. 63	317. 8	8. 91	65. 7	87. 9	6. 2

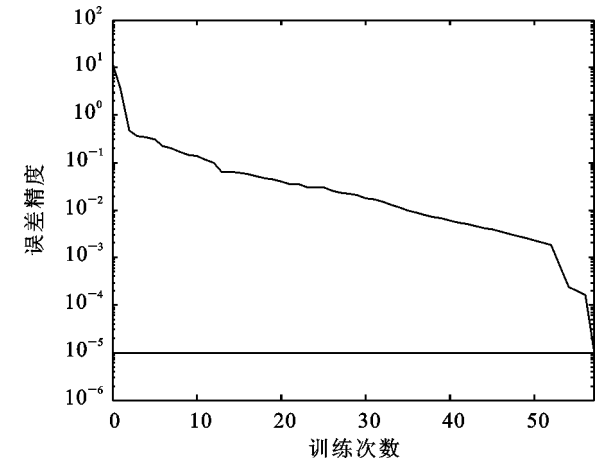


图 2 网络训练误差变化曲线

根据表 1 的土壤质量评价标准, 在各等级取值范围内均匀随机产生各 5 个样本, 经由标准化和模糊化计算出隶属度值作为 BP 网络的输入, 将其对应的等级值作为网络输出。经过预备试验, 确定网络的输入层结点为 18, 输出层结点个数为 1, 当隐含层结点个数为 18, 训练次数为 57 次时, 网络收敛且误差精度达到 1×10^{-5} , 训练过程如图 2, 网络输出的值见表 3。

由表 3 可知, FNN 的输出值与经验值非常接近, 所以将表 2 的实测数据也经过标准化和模糊化处理后输入到训练好的网络中, 最后网络输出的值还有文献[1] 的评价结果见表 4。

由表 4 可知, 本文的评价结果与文献[1] 的评价结果一致。从各开垦年限土壤质量等级 FNN 的输出值可知, 富锦市永富乡表层土壤质量呈下降趋势: 富锦市永富乡富锦市永

表 3 土壤质量样本等级的经验值和 FNN 模型的输出值的对比结果

土样 序号	土壤质量指标								土壤质量等级		
	有机质/ ($\text{g} \cdot \text{kg}^{-1}$)	全氮/ ($\text{g} \cdot \text{kg}^{-1}$)	全磷/ ($\text{g} \cdot \text{kg}^{-1}$)	全钾/ ($\text{g} \cdot \text{kg}^{-1}$)	速效氮/ ($\text{mg} \cdot \text{kg}^{-1}$)	速效磷/ ($\text{mg} \cdot \text{kg}^{-1}$)	速效钾/ ($\text{mg} \cdot \text{kg}^{-1}$)	CEC/ ($\text{cmol} \cdot \text{kg}^{-1}$)	pH	等级值	输出值
1	156. 01	8. 81	3. 23	35. 14	387. 63	120. 28	355. 35	283. 73	6. 42	1	1. 0025
2	114. 31	6. 51	3. 72	45. 54	538. 32	183. 85	505. 71	304. 52	6. 13	1	1. 0015
3	103. 9	9. 94	2. 87	41. 01	649. 71	105. 03	618. 78	299. 66	6. 36	1	0. 9950
4	99. 70	8. 04	2. 09	42. 08	360. 24	136. 76	509. 02	339. 84	6. 23	1	1. 0012
5	144. 39	5. 39	2. 34	34. 13	648. 09	187. 29	654. 63	378. 06	6. 08	1	1. 0019
6	64. 62	4. 18	1. 90	24. 52	298. 23	75. 96	329. 74	192. 31	6. 84	2	2. 0009
7	66. 09	4. 31	1. 93	22. 34	297. 30	87. 04	325. 57	194. 01	5. 94	2	1. 9999
8	73. 23	4. 37	1. 61	19. 07	315. 43	82. 46	347. 67	160. 70	6. 87	2	2. 0003
9	76. 78	4. 44	1. 51	18. 20	271. 31	88. 94	305. 35	186. 38	5. 86	2	2. 0006
10	78. 17	4. 46	2. 01	20. 45	339. 59	77. 14	270. 79	186. 75	6. 71	2	1. 9994
11	52. 17	2. 03	1. 46	16. 88	260. 99	58. 11	255. 61	126. 26	5. 19	3	2. 9961
12	43. 79	2. 23	1. 30	16. 13	240. 06	51. 11	239. 70	108. 65	5. 07	3	2. 9980
13	45. 69	2. 64	1. 29	12. 85	259. 57	49. 15	269. 21	132. 17	5. 13	3	2. 9943

续表 3 土壤质量样本等级的经验值和 FNN 模型的输出值的对比结果

土样 序号	土壤质量指标								土壤质量等级		
	有机质/ (g·kg ⁻¹)	全氮/ (g·kg ⁻¹)	全磷/ (g·kg ⁻¹)	全钾/ (g·kg ⁻¹)	速效氮/ (mg·kg ⁻¹)	速效磷/ (mg·kg ⁻¹)	速效钾/ (mg·kg ⁻¹)	CEC/ (cmol·kg ⁻¹)	pH	等级值	输出值
14	52.58	2.56	1.16	10.27	214.26	61.53	218.26	123.92	5.34	3	2.9996
15	44.64	2.28	1.22	14.44	270.07	59.38	200.96	134.37	5.43	3	2.9969
16	29.72	1.73	0.87	5.58	125.74	18.17	184.55	60.13	7.41	4	3.9970
17	23.87	1.55	0.75	7.45	179.53	31.08	173.58	98.99	7.01	4	3.9942
18	29.38	1.27	0.88	8.38	129.44	36.23	153.11	66.0	7.22	4	3.9994
19	22.68	1.36	0.51	7.09	169.02	30.78	122.29	77.74	7.13	4	3.9948
20	24.79	1.77	0.67	3.40	144.83	39.01	175.50	67.31	7.25	4	3.9959
21	17.83	0.22	0.09	2.23	17.36	1.99	51.26	33.61	8.01	5	5.0063
22	13.64	0.19	0.45	0.85	38.66	5.47	105.25	13.57	4.48	5	4.9973
23	1.30	0.17	0.26	1.70	75.36	0.15	48.25	48.00	4.67	5	4.9971
24	4.14	0.23	0.19	1.52	7.20	0.84	74.32	6.05	5.34	5	4.9982
25	1.01	0.09	0.16	0.10	20.04	6.65	47.32	8.30	9.18	5	4.9975

表 4 土壤质量的评价等级

开垦年限/a	FNN 输出值	土壤等级	文献[1]的评价结果
荒地	1.9976	II	II
5	2.9975	III	III
10	3.9968	IV	IV
25	3.9972	IV	IV

富乡开荒前表层土壤质量属于 II 级; 开荒 5 a 表层土壤质量 III 级; 开荒 10 a 表层土壤质量属于 IV 级; 开荒 25 a 后表层土壤质量属于 IV 级。本文选择的是三江平原典型性土壤质量变化趋势的试验点, 所以可以得知三江平原大面积开荒后, 随着开垦年限的增长, 土壤质量一般都呈现下降趋势, 土壤质量发生了较大变化。由于土壤质量的变化直接影响着粮食的增产, 而三江平原又是我国重要的商品粮出产地, 所以三江平原土壤的下降应当引起有关部门的高度重视。

4 结 论

(1) 以土壤等级划分实例为知识源, 应用三层前馈 BP 神经网络从实例中抽取知识的方法, 使生成的知识具有良好的数据支持, 由于其具有很强的学习能力, 各个评价指标的权重都是由网络通过学习自动生成的, 避免了主观赋权的人为干扰, 因此更具客观性和准确性。

(2) 将模糊数学方法引入神经网络中, 对土壤质量指标值采用隶属函数进行量化, 建立了 FNN 模型, 同时 FNN 模型还考虑到了土壤质量评价标准中各等级取值所具有的模糊性。评价结果表明, FNN 有较好的效果。

(3) 借助文献[1]数据建立的 FNN 模型结果表明, FNN 模型结果与文献[1]的基于 RAGA 的 PPC 模型的评价结果吻合, 都反映富锦市永富乡土壤质量下降的趋势。同时 FNN 对参数的改变有较强的鲁棒性, 从而为该方面研究提供了一条新的方法与思路。

参考文献:

[1] 赵小勇, 付强, 刑贞相. 投影寻踪等级评价模型在土壤质量变化综合评价中的应用[J]. 土壤学报, 2007, 44(1): 164-168.

[2] 刘梦云, 安韶山, 常庆瑞, 等. 宁南山区不同土地利用方

式土壤质量评价方法研究[J]. 水土保持研究, 2005, 12(3): 41-43.

[3] 胡焱弟, 赵玉杰, 白志鹏, 等. 土壤环境质量评价的径向基函数神经网络的模型设计与应用[J]. 农业环境科学学报, 2006, 25(增刊): 5-12.

[4] 杨海东, 胡月明, 邓飞其. 基于 DNA 编码的人工免疫模型在土壤质量评价中的应用[J]. 农业工程学报, 2005, 21(6): 40-44.

[5] Hussaim L, Olson K R, Wander M M, et al. Adaptation of soil quality indices and appliciation to the tillage systems in southern llinoid[J]. Soil and Tillage Research, 1999, 50: 237-249.

[6] Kwansoo K, Bmdford L, Barham L. Coxhead Measur indg soil quality dynamic. A role for economists, 2001, 25: 13-26.

[7] Berit M, Christel C, Lisa B. Agricultural land use in life cycle assessment (LCA): case studies of three vegetable oil crops[J]. Journal of Cleaner Production, 2000, 8: 283-292.

[8] 谢季坚, 刘承平. 模糊数学方法应用[M]. 武汉: 华中科技大学出版社, 2000.

[9] 邢美凤, 冯斌. 基于模糊神经网络的核桃病虫害诊断专家系统的设计[J]. 雁北师范学院学报, 2005, 21(2): 28-30.

[10] 何勇, 鲍一丹. 基于神经网络的发动机故障多媒体诊断专家系统[J]. 内燃机学报, 2001, 19(3): 245-248.

[11] 付强. 数据处理方法及其农业应用[M]. 北京: 科学出版社, 2006.

[12] 杨昆, 欧阳光耀. 燃油系统模糊神经网络故障诊断专家系统研究[J]. 内燃机学报, 2006, 24(2): 25-27.

[13] 刘树文, 王庆伟, 何东健, 等. 基于模糊神经网络的葡萄病害诊断系统研究[J]. 农业工程学报, 2006, 22(9): 144-147.

[14] 卢铁光, 杨广林, 王立坤. 基于相对土壤质量指数法的土壤质量变化评价与分析[J]. 东北农业大学学报, 2003, 34(1): 56-59.