

基于遗传算法最小二乘支持向量机的耕地变化预测

张豪¹, 罗亦泳², 张立亭², 陈竹安²

(1. 浙江工业大学建筑工程学院, 杭州 310032; 2. 东华理工大学长江学院, 南昌 330013)

摘要: 针对耕地变化内部规律与模拟方法进行研究, 提出最小二乘支持向量机耕地变化预测方法, 有效构建耕地变化与耕地变化影响因子之间复杂的非线性关系模型。利用遗传算法全局寻优功能优化最小二乘支持向量机内部参数, 提高最小二乘支持向量机耕地变化预测模型精度。利用该模型对江苏无锡市 1987—2000 年期间耕地变化进行预测, 并与多元回归、GM(1,1)、BP 网络、支持向量机 (SVM) 耕地预测模型和实际调查耕地变化数据进行比较分析。预测精度评价结果证实, 该方法耕地预测精度远高于多元回归、GM(1,1)、BP 网络模型, 略高于 SVM 模型, 但算法复杂度和计算效率远优于 SVM 预测模型, 是一种有效的耕地变化预测方法。

关键词: 最小二乘支持向量机, 遗传算法, 耕地预测, 影响因子, 精度分析

doi: 10.3969/j.issn.1002-6819.2009.07.041

中图分类号: O241.5, F301.24

文献标识码: A

文章编号: 1002-6819(2009)-7-0226-06

张豪, 罗亦泳, 张立亭, 等. 基于遗传算法最小二乘支持向量机的耕地变化预测[J]. 农业工程学报, 2009, 25(7): 226—231.

Zhang Hao, Luo Yiyong, Zhang Liting, et al. Cultivated land change forecast based on genetic algorithm and least squares support vector machines[J]. Transactions of the CSAE, 2009,25(7): 226—231.(in Chinese with English abstract)

0 引言

耕地是人类赖以生存的基本条件, 是社会经济发展的重要物质基础, 是最为重要的农业自然资源。随着人口规模扩大、社会经济的快速发展, 特别是城市化、工业化进程的加速推进, 中国建设用地规模迅速增加, 耕地总量不断减少, 人地矛盾将日益突出。与此同时, 作为一个农业人口占 78% 的农业大国来讲, 中国国民生产总值的 72% 直接或间接地来源于耕地, 农民年平均收入的 60% 以上都是由耕地来提供。然而, 中国的耕地资源又相当有限, 国土资源部发布的 2006 年度全国土地利用变更调查结果报告显示, 截至 2006 年 10 月 31 日, 全国耕地面积为 1.218 亿 hm^2 , 比上年度末的 1.221 亿 hm^2 净减少 30.68 万 hm^2 。这意味着, 中国目前人均耕地仅 0.093 hm^2 , 不到世界人均水平的 40%, 人地矛盾相当尖锐^[1]。耕地资源的短缺问题将成为制约中国未来发展特别是农业发展的重要问题, 将对中国的粮食生产和粮食安全构成重大隐患。耕地的数量和质量反映了一个国家的基本国情, 大体上决定了国家的人口承载量和可持续发展能力。因此, 对耕地数量变化进行研究对保证粮食安全和维持区域生态平衡具有十分重要的意义, 为实现可持续发展提供重要保障, 耕地变化预测将成为区域土地利用变化研究的核心问题。

目前耕地变化预测方法主要有线性回归模型、指数

平滑法、马尔柯夫链和灰色理论等。这些方法不能充分考虑到自然因素和社会因素对耕地总量影响的后效性, 不能很好地分析和拟合高度非线性多因素的耕地变化系统, 具有解释力不足的缺陷^[2]。胡喜生提出基于 BP 神经网络的福建省耕地预测模型^[3], 较好地拟合耕地变化与自然社会经济因素之间的非线性关系。但是神经网络模型基于经验风险最小原则, 容易陷入局部最优。本文针对现有耕地预测模型不足进行分析研究, 提出基于最小二乘支持向量机 (least squares support vector machines, 简称 LS-SVM) 的耕地变化预测模型, 该模型基于结构风险最小原则, 保证 LS-SVM 模型具有较好的拟合精度和推广性能, 很好地克服了传统方法的缺点, 是一种很好地处理多因素非线性耕地变化系统的新方法。利用遗传算法 (GA) 优化 LS-SVM 参数, 大大提高 LS-SVM 模型预测精度, 很好地解决了 LS-SVM 参数确定难这一难点。因此对遗传最小二乘支持向量机耕地预测模型研究具有重要意义, 可以为当地经济良性快速发展和合理制定土地利用规划提供可靠理论支持。

2 最小二乘支持向量机基本原理

支持向量机 (support vector machines, 简称 SVM) 由 V.Vapnik 等人于 1995 年在统计学习理论上提出的一种新的通用学习方法^[4]。SVM 模型可以有效地实现对基于小样本的高维非线性系统精确拟合, 并且采用结构风险最小原则, 具有很好的推广性。在脸检测、手写体数字识别、文本自动分类、非线性回归建模与预测等方面得到初步应用并取得较好的效果。由于 SVM 模型最终转化为求解一个二次凸规划问题, 所以当训练数据量大和维数高时, SVM 模型计算速度较慢。因此, J.A.K.Suykens

收稿日期: 2009-01-21 修订日期: 2009-06-11

基金项目: 国家自然科学基金资助项目 (40874010)

作者简介: 张豪 (1961—), 男, 江苏南京人, 高级工程师, 博士, 硕士生导师, 主要从事土木工程测量、GPS 理论及应用、土地信息技术研究。杭州 浙江工业大学建筑工程学院, 310032。Email: zhanghao@zjut.edu.cn

在标准 SVM 的目标函数中增加了误差平方和项，提出 LS-SVM 方法^[5]。它用等式约束条件代替不等式约束条件，求解过程变成解一组等式方程，避免了求解耗时的二次规划问题，解算速度加快，而且 LS-SVM 不再需要指定逼近精度，增强了标准 SVM 模型实用性。

已知一组训练集 $D=\{(x_1, y_1), \dots, (x_l, y_l)\}$, l 为样本数量, $x_i \in R^n$, $y_i \in R$, $i=1,2,\dots,l$, n 为 x_i 向量维数, R 为实数集。对于非线性问题可以通过非线性变换将输入向量映射到高维特征空间, 转化为类似的线性回归问题加以解决。设原训练集空间为 D , 通过映射 $z = \varphi(x)$ 变为高维空间 Z , 确定一个基于训练集 Z 的函数

$$f(x, \alpha) = w \cdot x + b \tag{1}$$

来逼近未知的回归函数, 其中 w 为权向量。把回归估计问题定义为对一个损失函数进行风险最小化的问题。利用结构风险最小化原则^[6-7] (简称 SRM) 进行风险最小化时, 最后的回归函数是在一定的约束条件式 (3) 下最小化泛函^[8-9] (2)

$$\min_{w,b,e} \frac{1}{2} \|w\|^2 + \gamma \frac{1}{2} \sum_{i=1}^l e_i^2 \tag{2}$$

$$y_i = w\varphi(x_i) + b + e_i \quad (i=1,2,\dots,l) \tag{3}$$

式中: γ ——正则化参数; e ——误差向量; b ——偏置量。

对最小二乘支持向量机问题, 定义 Lagrange 函数为

$$L(w,b,e,\alpha) = \frac{1}{2} \|w\|^2 + \gamma \frac{1}{2} \sum_{i=1}^l e_i^2 - \sum_{i=1}^l \alpha_i \{w\varphi(x_i) + b + e_i - y_i\} \tag{4}$$

式中 α_i ——Lagrange 乘子。

分别求 $L(w,b,e,\alpha)$ 对 w, b, e, α 的偏微分, 可以得到式 (4) 的最优条件

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \\ \frac{\partial L}{\partial b} = 0 \\ \frac{\partial L}{\partial e_i} = 0 \\ \frac{\partial L}{\partial \alpha_i} = 0 \end{cases} \tag{5}$$

由式 (5) 求解可得

$$\begin{cases} w = \sum_{i=1}^l \alpha_i \varphi(x_i) \\ \sum_{i=1}^l \alpha_i = 0 \\ \alpha_i = \gamma e_i \\ w\varphi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{6}$$

消去上式中的 w 和 e 可得到

$$\begin{bmatrix} 0 & I^T \\ I & ZZ^T + \gamma^{-1}E \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{7}$$

其中 $y = [y_1, y_2, \dots, y_l]^T$, $I = [1, \dots, 1]^T$, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$, E 为 $l \times l$ 维的单位矩阵, $Z = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_l)]^T$ 。

解方程组 (7) 可以求得最小二乘支持向量机系数 a 和 b , 进而得到最小二乘支持向量机 (LS-SVM), 最小二乘支持向量机回归函数为

$$f(x) = \sum_{i=1}^l \alpha_i (\varphi(x)\varphi(x_i)) + b \tag{8}$$

在高维线性空间 Z 中直接求解内积要求知道 $z = \varphi(x)$ 的具体形式, 并且容易造成“维数灾难”问题。因此引入核函数 $K(x, x_i) = \varphi(x)\varphi(x_i)$, 巧妙地解决了以上两难题。常用的核函数有: 多项式核函数、径向基核函数 (RBF) 和 Sigmoid 核函数^[10]。通过试验证实 RBF 核函数效果较好。

3 基于遗传算法的 LS-SVM 耕地变化预测模型建立

由最小二乘支持向量机模型基本原理可知, γ 正则化参数和核函数参数确定是影响 LS-SVM 模型预测精度的主要因素, 是目前支持向量机模型研究难点与热点。本文具体分析了遗传算法与 LS-SVM 模型融合的可行性与具体方法。

3.1 遗传算法优化 LS-SVM 参数分析

遗传算法 (简称 GA) 的基本原理来源于 Darin 的进化论和 Mendel 的基因遗传学原理^[11-12]。遗传算法将问题的求解表示成“染色体” (用编码表示字符串)。该算法从一群“染色体”串出发, 根据适者生存的原则, 从中选择出适应度高的“染色体”进行复制, 通过交叉、变异两种基因操作产生出新一代的更适应环境的“染色体”种群。随着遗传算法的一代一代地进行, 那些适应度高的模式将在后代中呈指数级增长, 最终得到适应度最高的染色体, 即优化问题的最优解。

本文充分利用遗传算法的强大全局寻优功能、隐含并行性、算法的高度稳定性, 对 SVM 模型的 γ 和核函数参数 σ (本文采用 RBF 核函数) 在参数空间中自动确定最优或者近似最优参数, 这样大大提高 LS-SVM 模型精度和使用范围, 具体结合方式如图 1 所示。

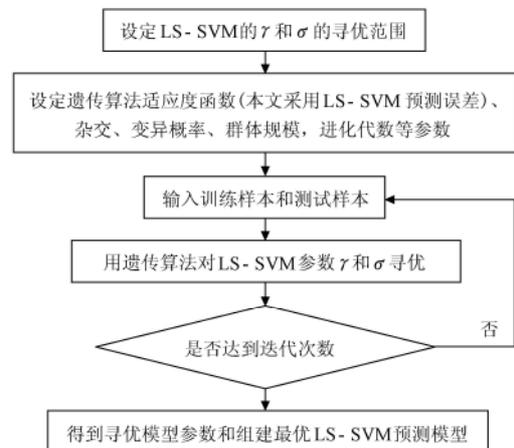


图 1 GA 优化 LS-SVM 流程图
Fig.1 Flow chart of GA-LS-SVM

3.2 耕地变化影响因子的选择和确定

耕地变动的驱动因素较多, 主要由自然驱动力和人

文驱动力构成,存在时空差异^[13]。自然驱动力包括环境变化、气候、自然灾害和地形、坡度等驱动因子;人文驱动力包括人口增长、经济发展、技术进步、社会行为、土地利用者主体行为等驱动因子和对土地利用方式起到重要影响、引导和规范作用的方针、政策,如生态退耕政策、严格土地管理政策、产业结构调整政策等^[14]。从纵多影响因素中选取主要因素是建立分析耕地变化和合理预测的基础,主要影响选取方法主要有相关分析法、灰色关联度分析法、主成分分析法^[15]。当前有关耕地变化驱动力研究较多,因此不做详细分析。

3.3 遗传 LS-SVM 耕地变化预测模型建立

LS-SVM 耕地变化预测模型本质就是充分利用 LS-SVM 具有强大的处理非线性多因素系统能力,建立耕地变化与各影响因素之间复杂的非线性关系,挖掘内部规律,从而可以实现利用未来影响因素的变化对相应耕地变化做精确的预测。利用 GA 优化 LS-SVM 参数,可以解决 LS-SVM 参数难以确定从而大大影响预测模型精度这一难点。具体耕地预测模型建立步骤如下:

- 1) 建模数据准备。将连续若干年的耕地变化和相应影响因子数据分成训练集和测试集。分别用于建立遗传 LS-SVM 模型和验证模型预测精度。
- 2) 利用 MATLAB 语言编程,确定 GA 遗传算子及相应参数。
- 3) 利用 K-折交叉验证法对每个染色体对应参数的 LS-SVM 预测预测进行估计,并作为染色体适应度计算函数,用于计算每个染色体适应度。
- 4) 进行遗传操作,并判断是否达到终止进化条件。
- 5) 得到最优 LS-SVM 耕地预测模型,利用测试集进行模型精度评价。

4 耕地变化预测模型预测实例及精度验证

本文针对江苏省无锡市耕地变化实际情况,利用相关分析法、灰色关联度分析法、主成分分析法最终确定地区生产总值 X_1 (万元)、总人口 X_2 (万人)、固定资产投资额 X_3 (万元)、第三产业产值 X_4 (万元)、城市化水

平 X_5 (%) 及耕地保护政策 X_6 为无锡市耕地变化的六大影响因子^[16],收集了 1987—2000 年的耕地变化影响因子与对应耕地面积数据,具体数据见表 1。其中人口数量、固定资产投资和第三产业产值数据由《无锡市统计年鉴》^[17]中直接获取;地区生产总值用市区居民消费价格指数修正;城市化水平通过非农人口占总人口的比重计算求得;便于建模,用数字表示不同时期耕地保护政策强度,最严厉耕地资源保护状态为“10”。1987—1994 年开始实施《土地管理法》政策变量定义为“3”,1995—1996 年开始实施基本农田保护制度定义为“5”,1997 年以后开始实施土地用途管制制度定义为“8”;1996 年以前耕地面积为耕地统计数的修正数^[18],1996 年以后为无锡市土地详查和变更调查数据。1987—1995 年作为训练数据,分别建立多元回归、GM(1,1)、BP 网络、SVM 与 LS-SVM 耕地预测模型。把 1996—2000 年作为测试,利用各耕地预测模型预测 1996—2000 年耕地面积,并与 1996—2000 年期间实际调查耕地面积进行比较,对各耕地预测模型进行精度评价与验证,具体各预测模型预测结果与精度分析见图 2、表 2 和表 3。

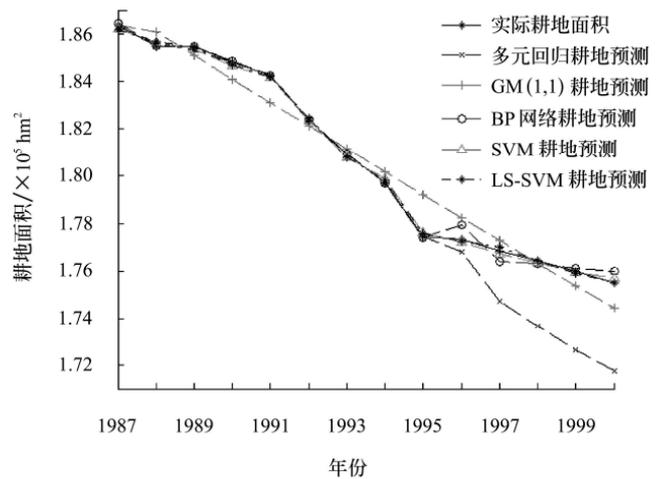


图 2 各耕地预测模型预测曲线
Fig.2 Forecast curve for each cultivated land forecast model

表 1 遗传 LS-SVM 建模数据

Table 1 Data of modeling based on GA-LS-SVM

时间/年	X_1 /万元	X_2 /万人	X_3 /万元	X_4 /万元	X_5 /%	X_6	Y/hm^2
1987	3 989 831	400.86	289 853	193 845	27.4	3	186 354.4
1988	4 083 435	406.80	424 617	261 416	28.7	3	185 471.9
1989	3 815 158	412.42	341 137	294 685	29.9	3	185 471.9
1990	4 038 387	417.71	353 027	352 446	34.5	3	184 837.2
1991	4 252 255	421.22	565 941	412 356	35.1	3	184 242.3
1992	6 412 247	423.90	1 305 696	828 107	35.7	3	182 348.3
1993	7 738 068	426.53	1 829 668	1 314 557	39.2	3	180 969.9
1994	8 522 866	428.32	1 899 938	1 983 538	39.7	3	179 710.6
1995	9 232 954	429.19	2 495 883	2 669 906	42.1	5	177 459.6
1996	9 440 016	430.82	2 506 959	3 110 603	42.7	5	177 335.4
1997	10 103 406	432.29	2 410 100	3 507 301	40.2	8	176 836.5
1998	10 972 830	432.21	2 772 200	3 963 100	39.6	8	176 435.7
1999	11 881 755	433.40	3 108 200	4 351 400	40.7	8	175 969.1
2000	12 530 706	434.61	3 501 000	4 689 200	42.2	8	175 513.2

表 2 各模型预测结果
Table 2 Forecast results for each model

时间/年	实际耕地面积/hm ²	多元回归预测/hm ²	GM (1, N) 预测/hm ²	BP 预测/hm ²	SVM 预测/hm ²	LS-SVM 预测/hm ²
1987	186 354.4	186 290.7	186 354.4	186 450.4	186 185.9	186 157.9
1988	185 471.9	185 643.4	186 090.4	185 498.5	185 640.4	185 669.2
1989	185 471.9	185 329.5	185 089.1	185 475.2	185 368.1	185 429.9
1990	184 837.2	184 836.6	184 093.2	184 846.3	184 668.7	184 714.7
1991	184 242.3	184 285.7	183 102.7	184 248.4	184 154.9	184 217.1
1992	182 348.3	182 381.0	182 117.4	182 372.1	182 393.4	182 463.2
1993	180 969.9	180 901.5	181 137.5	180 949.3	180 801.4	180 798.6
1994	179 710.6	179 738.2	180 162.9	179 699.5	179 879.1	179 821.6
1995	177 459.6	177 459.6	179 193.5	177 418.8	177 628.1	177 594.0
1996	177 335.4	176 797.1	178 229.3	177 943.0	177 209.2	177 238.8
1997	176 836.5	174 710.1	177 270.3	176 399.2	176 736.9	177 014.0
1998	176 435.7	173 678.9	176 316.4	176 319.2	176 340.4	176 454.2
1999	175 969.1	172 648.8	175 367.7	176 105.9	175 962.6	175 903.4
2000	175 513.2	171 734.0	174 424.1	175 976.0	175 690.5	175 488.8

表 3 各模型预测结果精度分析
Table 3 Precision analysis of forecast results for each model

精 度	多元回归预测	GM (1, N) 预测	BP 预测	SVM 预测	LS-SVM 预测
模型内符合均方根误差/hm ²	83.20	780.14	37.82	145.6	136.56
最大相对预测误差/%	2.15	0.98	0.34	0.11	0.10
平均相对预测误差/%	1.42	0.36	0.20	0.06	0.04
平均绝对预测误差/hm ²	2 504.20	627.5	352.19	100.97	76.54
均方根预测误差/hm ²	2746.6	714.1	401.7	115.97	96.01

由图 2、表 2 和表 3 中各耕地预测模型在 1987~1995 年预测结果与训练数据的符合程度可知,除 GM (1, 1) 模型符合精度差外,其余各预测模型符合程度均较高, BP 模型内符合程度最高,具体各模型内符合精度见表 3 模型内符合均方根误差。由各耕地预测模型在 1996—2000 年各预测结果可知,多元回归模型预测结果最差,其最大相对预测误差、平均相对预测误差、平均绝对预测误差和均方根预测误差等精度指标均较大,该模型不能很好地表达非线性的耕地变化系统;GM (1, 1) 模型没有综合考虑耕地变化影响因子,仅从耕地本身的变化趋势来进行预测,预测精度较低^[19];BP 预测模型具有很好的内符合精度,预测精度指标均高于前两种预测模型,但结构设计和参数确定困难,并且十分不稳定,很容易出现过拟合现象;LS-SVM 耕地预测模型内符合与各项预测精度指标略高于 SVM 耕地预测模型,LS-SVM 与 SVM 耕地预测模型远远高于其他预测模型,能很好地拟合耕地变化与各影响因子间的高度非线性关系,是耕地变化预测的新方法。LS-SVM 在目标函数中增加误差平方和项,用等式约束代替不等式约束,求解过程变成了解一组等式方程,避免了求解耗时的二次规划问题,解算速度明显加快,并不需要指定逼近精度,简化模型参数确定难度和参数优化复杂度,可以提高模型适应性和精度。本文依据 14 a 耕地数据利用 GA 优化 SVM 与 LS-SVM 参数,调用 SVM 258 次共耗时 44.39 s,调用 LS-SVM 模型 248 次共耗时 10.14 s。为进一步分析 LS-SVM 模型比 SVM 模型时间优越性,依据 50 期变形数据调用 SVM 272 次共

耗时 190.16 s,调用 LS-SVM 模型 236 次共耗时 11.97 s。因此当处理大型耕地数据时,LS-SVM 模型计算速度远远优于 SVM 模型,是一种快速、高精度的耕地变化预测新模型。

5 结 论

本文在对耕地变化规律和现有预测模型进行研究分析的基础上,提出遗传算法最小二乘支持向量机的耕地预测模型,通过理论研究和实例分析得到以下结论:

1) 耕地变化受多种因素影响,并且他们之间是高度非线性关系,传统单因素和线性的耕地预测模型不能满足实际耕地分析与预测要求。神经网络模型具有处理多因素非线性问题能力,但是该模型基于经验风险最小原则,不能保证好的推广性,如本文 BP 预测模型内符合精度高预测精度很差,并且该算法没有严格完善的理论支持。

2) 最小二乘支持向量机在目标函数中增加误差平方和项,用等式约束条件代替不等式约束条件,求解过程变成解一组等式方程,算法计算效率比标准支持向量机大大提高,适合处理大数据量的耕地变化数据。本文把 GA 遗传算法模型的强大寻优功能和最小二乘支持向量机融合,实现最小二乘支持向量机模型最优参数自动确定,很大程度解决最小二乘支持向量机参数确定难这一难点,提高了模型预测精度。

3) 基于 GA 优化的最小二乘支持向量机耕地变化预测模型各项精度评价指标远远优于多元回归模型、灰色

模型和 BP 网络模型等预测模型, 略高于支持向量机模型, 但是算法复杂度和计算效率远优于支持向量机模型。且基于 GA 优化的最小二乘支持向量机耕地变化预测模型具有预测精度高、实用和有效性等特点。

[参 考 文 献]

- [1] 林英. 我国耕地面积人均不足 1.4 亩[N]. 光明日报, 2007-04-13.
- [2] 苗作华, 刘耀林, 王海军. 耕地需求量预测的加权模糊-马尔可夫链模型[J]. 武汉大学学报(信息科学版), 2005, 30(4): 305-308.
Miao Zuohua, Liu Yaolin, Wang Haijun. Fuzzy-Markov chain model with weights for prediction of Cultivated Land Demand[J]. Geomatics and Information Science of Wuhan University, 2005, 30(4): 305-308. (in Chinese with English abstract)
- [3] 胡喜生, 洪伟, 吴承祯. 基于 BP 神经网络的福建省耕地预测模型[J]. 福建农林大学学报(自然科学版), 2008, 37(4): 425-427.
Hu Xisheng, Hong Wei, Wu Chengzhen. Prediction model of cultivated land in Fujian province based on BP neural network[J]. Journal of Fu Jian Agriculture and Forestry University (Natural Science Edition), 2008, 37(4): 425-427. (in Chinese with English abstract)
- [4] Vapnik V. Estimation of Dependencies Based on Empirical Data Berlin: Springer-Verlag, 1982.
- [5] 罗伟, 习华勇. 基于最小二乘支持向量机的降雨量预测[J]. 人民长江, 2008, 38(19): 29-31.
Luo Wei, Xi Huayong. Prediction of precipitation based on LS-SVM[J]. Yangtze Rive, 2008, 38(19): 29-31. (in Chinese with English abstract)
- [6] Vapnik V. The Nature of Statistical Learning Theory[M]. New York: Springer, 1995.
- [7] Vapnik V. 张学工译. 统计学习理论的本质[M]. 北京: 清华大学出版社, 2000.
Vapnik V. Zhang Xuegong translation. The Nature of Statistical Learning Theory[M]. BeiJing: Tsinghua University Press, 2000. (in Chinese)
- [8] Suykens J A K, Vandewall J. Least squares support vector machine classifiers[J]. Neural Processing Letters (S1370-4621), 1999, 9(3): 293-300.
- [9] Pelckmans K, Suykens J A K, De Moor B. Building sparse representations and structure determination on least squares support vector machine substrates[J]. Neurocomputing (S0925-2312), 2005, 64(S): 137-159.
- [10] 李波, 徐宝松, 武金坤, 等. 基于最小二乘支持向量机的大坝力学参数反演[J]. 岩土工程学报, 2008, 30(11): 1722-1725.
Li Bo, Xu Baosong, Wu Jinkun, et al. Back analysis of dam mechanical parameters based on least squares support vector machine[J]. Chinese Journal of Geotechnical Engineering, 2008, 30(11): 1722-1725. (in Chinese with English abstract)
- [11] 周明, 孙树栋. 遗传算法原理及应用[M]. 北京: 国防工业出版社, 1999.
- Zhou Ming, Sun Shudong. Theories and Application of Genetic Algorithms[M]. Beijing: National Defense Industry Press, 1999. (in Chinese)
- [12] 雷英杰, 等. MATLAB 遗传算法工具箱及应用[M]. 西安: 西安电子科技大学出版社, 2005.
Huo Yingjie, et al. Matlab GA Toolbox and Application[M]. Xi'an: Xi Dian University Prss, 2005. (in Chinese)
- [13] 李伟, 郝晋珉, 冯婷婷, 等. 基于计量经济模型的中国耕地数量变化政策与资产因素分析[J]. 农业工程学报, 2008, 24(6): 115-118.
Li Wei, Hao Jinmin, Feng Tingting, et al. Land asset and policy factors for acreage changes of cultivated land in China based on econometric model[J]. Jinmin Transactions of the Chinese Society of Agricultural Engineering, 2008, 24(6): 115-118. (in Chinese with English abstract)
- [14] 张勇, 於忠祥. 合肥市耕地资源变化驱动力分析及其保护研究[J]. 国土资源科技管理, 2008, 25(3): 74-79.
Zhang Yong, Yu Zhongxiang. A study of driving force of arable land resources change and protection in Hefei city[J]. Scientific and Technological Management of Land and Resources, 2008, 25(3): 74-79. (in Chinese with English abstract)
- [15] 林建平, 赵小敏, 邓爱珍, 等. 城镇建设用地规模影响因素分析及预测[J]. 国土资源科技管理, 2008, 25(2): 102-106.
Lin Jianping, Zhao Xiaomin, Deng Aizhen, et al. An analysis and prediction of factors affecting scale of urban construction land[J]. Scientific and Technological Management of Land and Resources, 2008, 25(2): 102-106. (in Chinese with English abstract)
- [16] 朱健宁, 吴群. 耕地资源数量变化及其驱动力分析——以无锡市为例[J]. 土壤, 2008, 40(2): 193-199.
Zhu Jianning, Wu Qun. Variation of acreage of cultivated land and its driving forces--A case study of Wuxi City[J]. Soils, 2008, 40(2): 193-199. (in Chinese with English abstract)
- [17] 无锡市统计局. 无锡市统计年鉴(2005)[M]. 北京: 中国统计出版社, 2005.
Wu Xi Statistics Bureau. Statistical Yearbook System of Chuzhou city(2005)[M]. Beijing: China Statistics Press, 2005. (in Chinese)
- [18] 蔡运龙, 汪涌, 李玉平. 中国耕地供需变化规律研究[J]. 中国土地科学, 2009, 23(3): 13-17.
Cai Yunlong, Wang Yong, Li Yuping. Study on changing relationship of demand and supply of cultivated land in China[J]. China Land Science, 2009, 23(3): 13-17. (in Chinese with English abstract)
- [19] 王良健, 师迎春, 林目轩. BP 神经网络结合小波处理在城市扩张预测中的应用——以长沙市区为例[J]. 中国土地科学, 2008, 22(1): 39-47.
- [20] Wang Liangjian, Shi Yingchun, Lin Muxuan. Applications of BP neural network theory on urban expansion forecasts with wavelet treatment: a case of Chang Sha urban area[J]. China Land Science, 2008, 22(1): 39-47. (in Chinese with English abstract)

Cultivated land change forecast based on genetic algorithm and least squares support vector machines

Zhang Hao¹, Luo Yiyong², Zhang Liting², Chen Zhuan²

(1. College of Civil Engineering and Architecture, Zhejiang University of Technology, Hangzhou 310032, China;

2. East China University of Technology Yangtze College, Nanchang 330013, China)

Abstract: A prediction method of cultivated land change based on least squares support vector machines (LS-SVM) was developed by studying the inherent tendency toward land change and simulating the trajectories of changes in land use. A nonlinear dynamic model of cultivated land change and influence factors was introduced. The prediction accuracy was improved by using the genetic algorithm to automatically determine the optimal parameters of least squares support vector machines. The proposed model has been thoroughly tested on predicting the cultivated land change during the period of 1987—2000 in Wuxi, Jiangsu. The results were compared and analyzed with those obtained from multiple regression, GM(1,1), BP algorithm, support vector machines(SVM) and the survey data on cultivated land change. The evaluation of prediction precision showed that the method based on LS-SVM was far more accurate than multiple regression, GM(1,1) and BP network model. Compared with the support vector machines model, the method was even slightly better and possesses much less algorithm complexity and higher computational efficiency. The overall performance suggests that the method is effective in predicting land change.

Key words: least squares support vector machines, genetic algorithm, cultivated lands change forecast, influence factors, accuracy analysis