

基于细菌觅食优化算法的支持向量机 在土壤墒情预测中的应用

丁辉¹, 仲跃², 张俊³, 钱建中², 谢能刚¹

(1. 安徽工业大学 机械学院, 安徽 马鞍山 243032;

2. 盐城市水利局, 江苏 盐城 224005; 3. 盐城市盐都区水务局, 江苏 盐城 224005)

摘要: [目的] 对基于细菌觅食优化算法的支持向量机在土壤墒情预测中的应用进行探讨, 为现代农业研究中土壤墒情预测及农业生产提供支持。[方法] 基于支持向量回归机方法建立土壤墒情预测模型, 利用细菌觅食优化算法优化支持向量机预测模型的相关参数。根据从种植区采集的田间数据对模型进行建模和测试。[结果] 与仅利用支持向量回归机和利用粒子群优化的支持向量回归机分别建立的模型进行对比, 发现本研究所提算法建立的预测模型的预测效果更佳。[结论] 该模型预测效果较好, 所建模型已应用于实际项目, 预测精度基本满足要求, 且运行稳定。进而证明了该研究所提算法的有效性和可行性。

关键词: 细菌觅食优化算法; 支持向量回归机; 土壤墒情预测

文献标识码: B

文章编号: 1000-288X(2016)06-0131-05

中图分类号: S152.7

文献参数: 丁辉, 仲跃, 张俊, 等. 基于细菌觅食优化算法的支持向量机在土壤墒情预测中的应用[J]. 水土保持通报, 2016, 36(6): 131-135. DOI: 10.13961/j.cnki.stbctb.2016.06.022

Application of Support Vector Regression Machines in Soil Moisture Prediction Based on Bacteria Foraging Optimization Algorithm

DING Hui¹, ZHONG Yue², ZHANG Jun³, QIAN Jianzhong², XIE Nenggang¹

(1. School of Mechanical Engineering, Anhui University of Technology,

Ma'anshan, Anhui 243032, China; 2. Yancheng Water Conservancy Bureau, Yancheng,

Jiangsu 224005, China; 3. Yandou Water Conservancy Bureau, Yancheng, Jiangsu 224005, China)

Abstract: [Objective] The application of support vector regression machines in soil moisture prediction based on bacteria foraging optimization algorithm (BFOA) was discussed to provide supports for the prediction of soil moisture of modern agriculture and agricultural production. [Methods] The soil moisture prediction model based on support vector regression machines (SVR) was established. And the related parameters of SVR were optimized by using bacteria foraging optimization algorithm (BFOA). Then the model was set up and tested according to the collected data of growing region. [Results] The proposed algorithm was compared with the established model using SVR and SVR based on particle swarm optimization, respectively. The results showed that the prediction model established by the proposed algorithm performed better. [Conclusion] The model had been applied to the actual project. The prediction accuracy of the model was testified well and the operation was stable. The validity and feasibility of the proposed algorithm had been proved.

Keywords: bacteria foraging optimization algorithm (BFOA); support vector regression machines (SVR); soil moisture prediction model

土壤墒情信息主要指土壤的水分状况, 是农业种植中至关重要的指标之一, 对农作物生长发育的影响尤为突出。因此, 及时了解田间的土壤墒情信息, 让

作物生长在适宜的环境中, 可有效地保证农作物的产量。随着农业现代化步伐的加快, 信息技术应用于现代农业已成为必然趋势, 及时、准确地获取田间的土

收稿日期: 2016-03-28

修回日期: 2016-06-01

资助项目: 江苏省水利科技项目“沿海地区规模化节水灌溉集成技术推广示范基地”(2014078)

第一作者: 丁辉(1988—), 男(汉族), 安徽省阜阳市人, 硕士研究生, 研究方向为多传感器数据融合, 物联网技术, 神经网络等。E-mail: ding-huihut@163.com.

壤墒情信息显得尤为重要。实时地获取墒情信息对于指导水资源的合理配置、提高作物的产量均具有重要意义^[1-2]。目前,较为常用的土壤墒情预测模型与方法包括:经验法、水量平衡模型、时间序列模型、土壤水分动力学模型、遥感监测方法和神经网络模型等^[3]。因神经网络模型具有较好的非线性拟合能力和学习功能,在众多方法与模型中脱颖而出。其中,支持向量机模型近年来被广泛的关注,该模型已被应用于时间序列分析^[4]、模式识别^[5]、图像处理^[6]、城市用水量预测^[7]等。支持向量机模型在土壤墒情预测中的应用已成为众多学者近年来研究的焦点。薛晓萍等^[8]利用支持向量机模型建立起不同土层深度的土壤水分预测模型,该模型在 0—10 cm 土层的土壤水分预测中效果较好,土层越深则预测效果越差,并对历史数据进行分析,发现各层与 0—10 cm 层具有较高的相关关系,由此建立了较精确的预测模型;Kashif 等^[9]利用支持向量机相关原理,将影响土壤墒情的各种参数作为输入,例如空气湿度、空气温度、土壤温度以及土壤水分等,对模型进行训练,并作出相应的预测,对试验区的分析表明,模型的预测效果较佳;Zhao 等^[10]利用粒子群优化最小二乘支持向量机预测土壤含水量,通过分析确定 0—10 cm 土壤含水量、饱和含水量和有机物含量为输入量,土壤含水量作为输出量,并与广义回归神经网络、小波神经网络和 BP 神经网络预测效果进行对比分析,发现该方法效果更佳。以上方法在土壤墒情预测中均表现出较好的预测效果,然而以上方法,在输入参数的选择方面,主要依靠科研人员的经验,人为因素影响较大。尽管有些方法进行了相应分析,但需改变原始数据,本文为充分利用原始数据,在输入参数的选择上,利用灰色系统理论中的灰色关联分析计算各参数对土壤含水量的贡献率,将贡献率大的参数作为输入参数,并利用细菌觅食优化算法优化支持向量机模型的相关参数,以提高支持向量机模型的性能。

1 数据来源

在安徽工业大学校园内建立了一个小型的草莓种植试验区,并搭建土壤墒情监测系统,通过埋设在田间的土壤水分、土壤温度、空气温湿度和光照度传感器,实时监测种植区的土壤水分、土壤温度、空气温度、空气湿度和光照度等作物的生长环境因素。为有效监测种植区的墒情信息,种植区中,每个检测量同一时间利用 3 组传感器进行数据采集,每隔 30 min 采集 1 组数据。采集从 11 月 09 日至 12 月 1 日期间的空气温湿度、光照度、土壤温度信息,其采集值分别

如图 1 所示。利用采集的 1 000 余组数据对模型进行建模与预测分析。

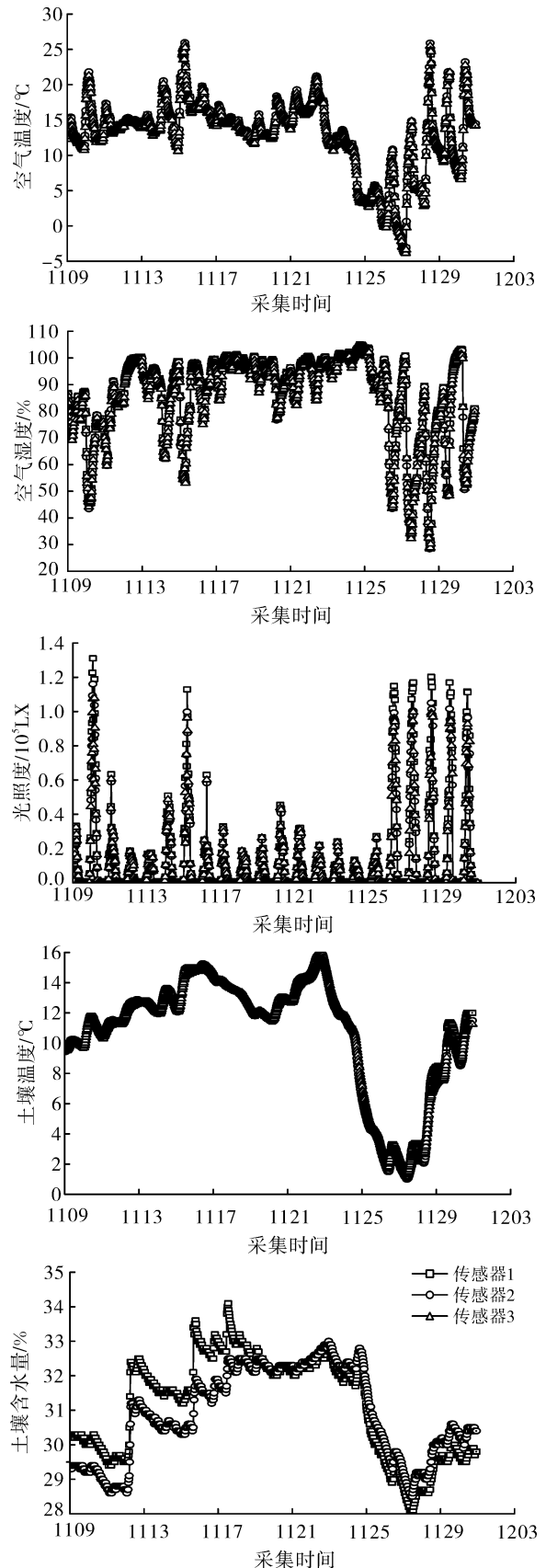


图 1 试验区土壤各指标采集值

2 灰色关联分析

灰色关联分析是通过灰色关联度来分析和确定系统诸因素间的影响程度或各因素对系统主行为的贡献测度的一种方法。灰色关联分析属于灰色系统理论的一个分支,它的基本思想是依据时间序列曲线之间的几何形状的相似程度判断各序列相互之间的联系程度^[11]。序列折线的几何形状越接近,其相对应的序列的关联度也就越大,反之则说明其关联度越小。

由文献[11]可知, k 时刻 x_i 序列对 x_0 序列的关联系数为:

$$\xi(x_0(k), x_i(k)) = \frac{\min_{i \in N} \min_{k \in n} |x_0(k) - x_i(k)| + \rho \max_{i \in N} \max_{k \in n} |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_{i \in N} \max_{k \in n} |x_0(k) - x_i(k)|}$$

式中: ρ ——分辨系数,在实际应用中分辨系数 ρ 一般取为0.5。

通过上式计算出关联系数后,便可求得 x_i 序列对 x_0 序列的关联度为:

$$r(x_0, x_i) = \frac{1}{n} \sum_{k=1}^n \xi[x_0(k), x_i(k)]$$

3 细菌觅食优化算法

细菌觅食优化算法是Passion受到大肠杆菌觅食行为的启发于2002年提出的。该算法包含了模拟大肠杆菌真实觅食行为的3个主要操作:趋向性操作、复制性操作和迁徙性操作。在细菌系统中,每个细菌都可以看作是优化问题的一个寻优解,大肠杆菌寻找最佳路线觅食的过程则可以认为是其中的一个寻优解寻找最优解的过程。该算法的算法流程见参考文献[12-13]。

4 回归型支持向量机

建立在统计学习理论基础之上的支持向量机由Vapnik等首次提出,该模型最初用于分类问题。随

着实际应用的需要,Vapnik等将 ϵ 不敏感损失函数引入支持向量机,建立了支持向量回归机(SVR)。支持向量回归机的基本原理是寻找到所有训练的样本点距离某个分类面的误差最小的最优分类面。支持向量回归机的提出使得支持向量机理论进一步完善,应用范围也进一步拓展^[14-15]。

5 基于细菌觅食优化算法的支持向量回归机预测模型

选用的支持向量机核函数为径向基核函数,进而需要优化的参数为支持向量机惩罚因子 C 和径向基函数宽度 σ (算法流程图略)。

6 预测模型的建立与结果分析

6.1 模型的输入输出参数选择

对于预测模型输入量的确定,在学术界至今还未有统一的规范,一般处理方法将普遍认为对土壤含水量有影响的参数作为输入量,很明显这样选择支持向量机预测模型的输入量具有较大的主观随意性,很难让人信服。本研究除采集作为输出量的土壤含水量外,还采集了影响土壤含水量变化的土壤温度、空气湿度、空气温度和光照度这四组数据。本文的研究是利用建立的模型预测3d后的土壤含水量,因有些因素难以考虑,将这些因素隐含于历史含水量中,故将其作为一个影响因素。至此,影响土壤水分含量的参数有土壤温度、空气湿度、空气温度、光照度和历史含水量。为避免个人的主观随意性,本研究利用灰色关联分析,计算各参数对土壤含水量的贡献率,将各参数中对土壤含水量贡献较大的参数作为输入参数。经计算空气温度、空气湿度、光照度、土壤温度和历史土壤含水量对土壤含水量的灰色关联度分析结果详见表1。由表1可以观察到,各参数对土壤含水量的关联度均很高,说明各参数对土壤含水量的贡献均很大。故将空气温度、空气湿度、光照度、土壤温度和历史土壤含水量这5个参数作为预测模型的输入量。

表1 各参数对土壤含水量的灰色关联分析结果

| 项目 | 空气温度 x_1 | 空气湿度 x_2 | 光照度 x_3 | 土壤温度 x_4 | 历史土壤含水量 x_5 |
|-------|------------|------------|-----------|------------|---------------|
| 灰色关联度 | 0.995 2 | 0.997 6 | 0.977 0 | 0.995 7 | 0.999 3 |

6.2 模型的参数寻优

预测模型的输入、输出参数确定后,便可利用采集数据建立预测模型。然而,在利用支持向量机进行预测建模时,相关参数的合理选择对于模型的预测性

能至关重要。故而需要找到合适的支持向量机惩罚因子 C 和径向基函数的宽度参数 σ 。在利用细菌觅食优化算法寻优之前,首先要设置算法中各参数的初始值。本研究中设置细菌觅食优化算法的主要参数

为:细菌个数 S 为 20,趋向性操作循环次数 N_c 为 20,趋向性操作单向运动的最大步数 N_s 为 4,复制性操作循环次数 N_{re} 为 4,迁徙性操作循环次数 N_{nd} 为 2,迁徙概率 P_{nd} 为 0.25。参数设定后,利用采集数据对其进行寻优,得到寻优后的结果为: $C=1.2160, \sigma=37.2234$ 。

6.3 预测模型的结果分析

利用细菌觅食优化算法寻优得到惩罚因子 C 和径向基函数的宽度参数 σ ,建立支持向量机预测模型。在进行支持向量机预测建模中,本文将采集的土壤墒情数据分为两组,分别称其为训练集和测试集。一般情况下,利用分析数据的 80% 左右用于训练,20% 用于测试模型的性能。为消除时间序列分析中长期趋势等的影响,对采集数据作随机乱序处理。因此,首先利用训练集数据对支持向量机模型进行学习,然后利用测试集数据对模型进行预测并检验预测的效果。为证明本文算法的可行性,分别分析了仅利用支持向量机建立预测模型(SVR)、利用粒子群优化支持向量机建立预测模型(PSO-SVR)以及本文所提的利用细菌觅食优化算法优化支持向量机建立预测模型(BFOA-SVR)的效果,并将这些模型的预测性能进行对比。

经计算,分别利用 BFOA-SVM, PSO-SVM 以及 SVR 建立预测模型后的模型性能参数结果详见表 2。分析表 2 中各模型的性能参数,不难看出仅用 SVR 进行建模,虽然样本训练的效果很好,但是预测的效果却很差,这说明模型的泛化能力较差,不适宜用于预测。利用粒子群优化算法对支持向量机模型参数进行寻优,并建立起预测模型,该模型整体性能很不理想,极有可能陷入了局部最优解,进而导致训练和预测的性能较差。然而,利用 BFOA-SVR 建立预测模型可有效避免局部最优解,进而模型的训练和预测性能效果较好。这便证明本文所提算法的有效性和可行性。

表 2 各模型的性能参数

| 模型类型 | 训练集 | | | 测试集 | | |
|----------|---------|---------|---------|---------|---------|---------|
| | MSE | MARE | R^2 | MSE | MARE | R^2 |
| SVR | 0.000 2 | 0.000 9 | 0.999 6 | 0.138 9 | 0.013 4 | 0.488 3 |
| PSO-SVR | 0.069 9 | 0.029 0 | 0.161 7 | 0.066 6 | 0.031 2 | 0.151 5 |
| BFOA-SVR | 0.015 3 | 0.003 9 | 0.942 8 | 0.016 7 | 0.006 3 | 0.933 0 |

注: MSE 为均方根误差, MARE 为平均相对误差, R^2 为相关系数。

现已证明了本文算法的有效性,便可根据本文所提算法流程建立土壤墒情预测模型。利用本文算法所建模型的分析结果如图 2—7 所示。

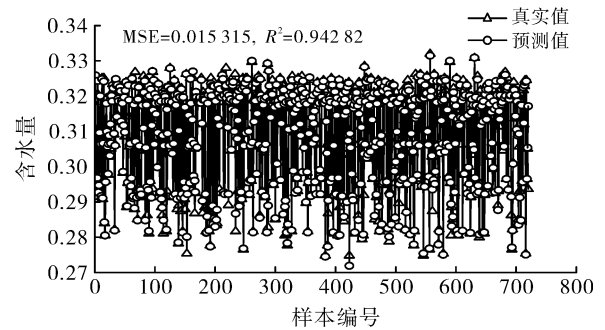


图 2 训练集预测结果对比

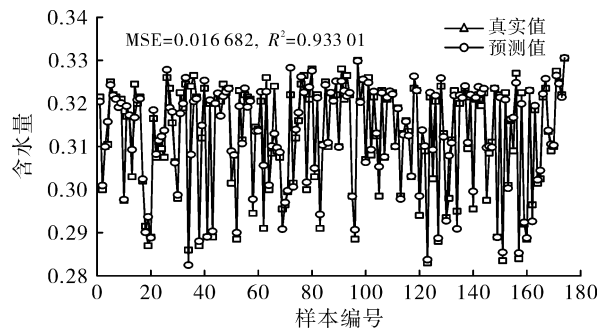


图 3 测试集预测结果对比

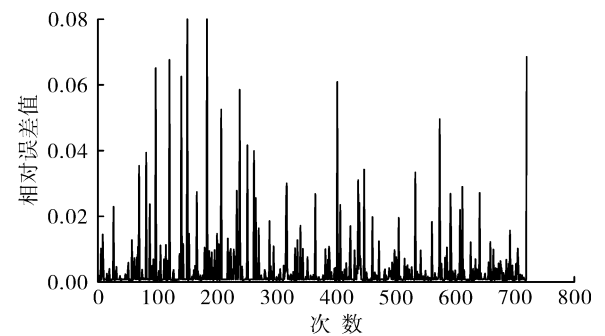


图 4 训练集相对误差

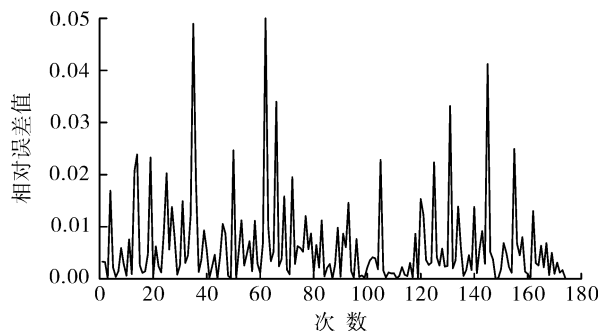


图 5 测试集相对误差

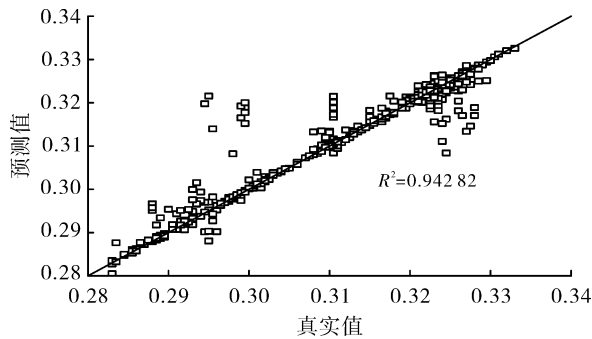


图6 训练集预测结果检验

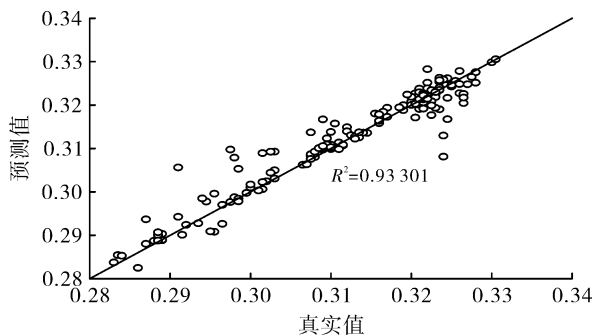


图7 测试集预测结果检验

7 结论

在实际的农业生产中,建立有效的土壤墒情预报系统对保证作物的产量、避免资源的不必要浪费至关重要。本文利用新型的智能优化算法——细菌觅食优化算法对支持向量回归机模型的相关参数进行优化,以提高模型的预测性能。在模型的输入参数选择上,本研究利用灰色关联分析计算各参数对输出参数的贡献率,进而避免了在输入参数选择上的人为因素影响。利用试验区采集数据建立本文所提模型,经计算分析后,证明所提模型预测效果较好。所建模型已应用于实际项目,预测精度基本满足要求,且运行稳定,进而证明了本研究所提算法的有效性和可行性。

[参考文献]

- [1] 许秀英,衣淑娟,黄操军.土壤含水量预报现状综述[J].农机化研究,2013,35(7):11-15.
- [2] 陈天华,唐海涛.基于ARM和GPRS的远程土壤墒情监测预报系统[J].农业工程学报,2012,28(3):162-166.
- [3] 李茂涵,方丽,贺京,等.基于前期降水量和蒸发量的土壤湿度预测研究[J].中国农学通报,2012,28(14):252-257.
- [4] 陈果.基于遗传算法的支持向量机时间序列预测模型优化[J].仪器仪表学报,2006,27(9):1080-1084.
- [5] 沈明华,肖立,王飞行.支持向量机在模式识别中的应用[J].电讯技术,2006,46(4):9-12.
- [6] 吕峰,高春林.支持向量机在皮肤症状图像识别中的应用研究[J].计算机仿真,2010,27(11):267-269.
- [7] 苗志刚,付强.基于灰色支持向量机的城市用水量预测研究[J].计算机仿真,2012,29(8):196-199.
- [8] 薛晓萍,王新,张丽娟,等.基于支持向量机方法建立土壤湿度预测模型的探索[J].土壤通报,2007,38(3):427-433.
- [9] Gill M K, Asefa T, Kembrowski M W, et al. Soil moisture prediction using support vector machines 1[J]. Jawa Journal of the American Water Resources Association, 2006, 42(4):1033-1046.
- [10] Zhao Lixi, Shui Pengbo, Jiang Fang, et al. Using monitoring data of surface soil to predict whole crop-root zone soil water content with PSO-LSSVM, GRNN and WNN[J]. Earth Science Inform, 2014(7):59-68.
- [11] 刘思峰,蔡华,杨英杰,等.灰色关联分析模型研究进展[J].系统工程理论与实践,2013,33(8):2041-2046.
- [12] 李娜,雷秀娟.细菌觅食优化算法的研究进展[J].计算机技术与发展,2014,24(8):39-44.
- [13] 周雅兰.细菌觅食优化算法的研究与应用[J].计算机工程与应用,2010,46(20):16-21.
- [14] Vapnik V N. Statistical Learning Theory [M]. New York: John Wiley, 1988.
- [15] Vapnik V N. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1985.