

王丽爱,谭昌伟,马 昌,等. 农情信息遥感监测预报模型构建算法研究进展[J]. 江苏农业科学,2013,41(11):1-5.

农情信息遥感监测预报模型构建算法研究进展

王丽爱,谭昌伟,马 昌,童 璐,杨 昕,崔怀洋,郭文善

(扬州大学江苏省作物遗传生理重点实验室/农业部长江中下游作物生理生态与栽培重点开放实验室,江苏扬州 225009)

摘要:综合介绍了线性逐步回归算法、偏最小二乘回归算法、神经网络算法、支持向量机回归算法在农作物信息遥感预测建模中的应用,阐述了这些算法的基本思想,给出了每种算法的求解步骤,归纳了这些算法的优点,分析了每种算法的不足之处,并指明了农情信息遥感预测建模的研究发展方向。

关键词:遥感;建模;逐步多元回归;偏最小二乘;神经网络;支持向量机

中图分类号: S127 **文献标志码:** A **文章编号:** 1002-1302(2013)11-0001-05

农业是国民经济的基础,农业高技术引领农业的未来。遥感作为现代信息技术的前沿,能够快速准确地收集农业资源和生产信息,遥感技术在农业领域得到了越来越多的应用,如利用遥感信息进行农作物估产、长势、品质、病虫害等遥感监测;在计算机系统的支持下,通过建模将获取的农业遥感信息与农业目标参量联系,定量地反演或推算出农学参量(即定量遥感)。定量遥感模型概括起来分为3类:物理模型、统计模型和半经验模型^[1],其中统计模型是对一系列观测数据作经验性的统计描述或相关分析,建立遥感参数与农情观测数据之间的回归模型。

近年来,文献资料运用逐步多元回归算法、偏最小二乘回归算法、神经网络回归算法、支持向量机回归算法,在计算机系统支持下,建立多种农情信息遥感预测统计模型。计算机系统操作的目的是对数据进行加工处理,以得到期望的结果,其中算法是灵魂,算法是为了解决一个问题而采取的方法和步骤。因此,本文介绍这些算法应用于遥感建模预测多种农作物的多种农学参量情况的同时,重点阐述了这些算法的基本思想,给出了每种算法的求解步骤,总结了这些算法的优点。每种算法都有不足之处,针对每种算法的缺陷加以改进或多种算法结合使用,可以有针对性地避免原算法的缺陷,达到取长补短的优势互补效果。因此,传统算法的改进算法、多种算法的组合算法成为了农情信息遥感预测建模算法的研究趋势,最后指明了农情信息遥感预测建模的研究发展方向。

1 逐步多元回归算法(stepwise multiple regression, SMR)

在解释变量个数较多的情况下建立多元线性回归预测模型,需要选择解释变量,以使在回归方程中包含所有对被解释变量影响显著的变量而不包含影响不显著的变量。

逐步回归正是为解决这一问题而提出的一种算法。

Jin等运用ASD光谱仪测得小麦不同生育时期的冠层光谱数据,运用SMR算法建立了基于光谱参数估计小麦叶片叶绿素含量的回归模型^[2]。杨晓华等测定了水稻不同生育期的冠层高光谱反射率,分别以高光谱反射率的4种不同变换形式(光谱反射率、反射率一阶导数、反射率二阶导数和反射率对数变换)为因变量,以叶面积指数(LAI)和叶绿素密度(GLCD)为自变量进行分析,评价了SMR对LAI和GLCD的估算能力^[3]。周冬琴等用SMR算法建立了水稻成熟籽粒蛋白质含量监测模型,模型预测值与实测值之间符合度较高,对水稻成熟籽粒蛋白质含量具有较好的预测性^[4]。宋开山等利用不同生长季大田玉米的冠层高光谱反射率,采用SMR算法,建立了玉米地上鲜生物量高光谱遥感估算模型,模型预测效果较理想^[5];利用不同生长季的大田玉米、大豆的冠层高光谱反射率,采用SMR算法,建立了玉米、大豆LAI高光谱遥感估算模型,结果表明以多波段逐步回归构造的模型预测效果很好^[6]。牛铮等利用小麦及11种植物的叶片样本光谱测量数据,采用SMR算法,分析了鲜叶片中7种化学组分含量与其光谱特性的统计关系,并分别建立了叶片反射率及其变换形式与鲜叶片中化学组分含量的SMR模型^[7]。

SMR算法的主要思想是双向筛选进入模型的解释变量:在所考虑的全部解释变量中按照对被解释变量的贡献大小逐个引入回归方程,已被引入回归方程的变量在引入新变量后也可能失去重要性,而需要从回归方程中剔除出去。引入一个变量或者从回归方程中剔除一个变量都要进行 F 检验,以保证在引入新变量前回归方程中只含有对被解释变量影响显著的变量,而不显著的变量已被剔除^[8-9]。

SMR算法步骤如下:(1)求被解释变量 y 与每一个自变量 x 的一元线性回归方程。(2)通过 F 检验,选择 F 最高者作为第1个进入模型的自变量。(3)对模型外的变量分别进行偏 F 检验,在若干通过偏 F 检验的变量中,选择 F 最大者进入模型。(4)对模型中的自变量分别进行偏 F 检验,如果所有自变量都通过偏 F 检验,转入步骤(3);如果有若干自变量没有通过偏 F 检验,选择 F 最小的自变量,将它从模型中删除,再重新建立剩余自变量与 y 的拟合模型。重复步骤(1)和(2)。(5)重复步骤(3)和(4),直至无法剔除已引入的变量、也无法再引入新的自变量,则算法终止^[10]。

收稿日期:2013-04-23

基金项目:国家自然科学基金(编号:41271415、40801122、61003180);江苏省高校自然科学基金(编号:12KJB520018);江苏省自然科学基金(编号:BK201031)。

作者简介:王丽爱(1975—),女,山西祁县人,博士研究生,研究方向为农业遥感与信息技术应用。E-mail:wla001@163.com。

通信作者:郭文善,教授,博士生导师,主要从事作物栽培生理与信息农业研究。E-mail:guowa@yzu.edu.cn。

SMR 算法是一种建立多元回归模型的有效方法,应用极广。利用 SMR 算法建立模型的优点是:模型中包含了尽可能多的自变量,特别是包含那些对被解释变量 y 有显著作用的自变量;为了使模型更加有效和方便,最终模型中不包含对 y 作用不显著的自变量^[11]。其缺点是:在逐步回归中每剔除或引入一个变量都需要计算 F 统计量,这需要一定的工作量,同时容易造成筛选过度,影响建模效果^[12];逐步回归中所用的 F 检验对众多初学者和应用者来说难以理解和把握^[9];逐步线性回归基于线性筛选变量结果存疑,特别在变量间存在多重相关时更是如此^[13]。

2 偏最小二乘回归算法 (partial least - squares regression, PLS)

PLS 是由 Wold 等在 1983 年首次提出的一种新型的多元统计分析方法,集多元线性回归分析、典型相关分析和主成分分析的功能和优点于一体。它主要研究的是多因变量对多自变量的回归建模,特别当各变量内部高度线性相关时,用偏最小二乘回归算法建模更有效^[13]。

目前 PLS 算法被认为是利用近红外光谱建立预测模型的最可靠方法之一^[14]。Ecarnot 等运用 PLS 算法构建了运用近红外光谱预测硬质小麦叶片含氮量及比叶重的回归模型^[15]。Li 等运用 PLS 算法,通过不同时期的小麦生物量、叶绿素密度、叶片等价水厚度和叶片氮密度等农学参数估计成熟期籽粒蛋白质含量^[16]。这些农学参数可以通过多时相遥感数据获取,进而基于 PLS 算法建立小麦的多时相多光谱 GPC 遥感预测间接模型。邵田田等运用 PLS 算法构建了玉米高光谱数据与玉米 FPAR 的定量估算模型,FPAR 模型反演精度优于植被指数法构建的模型^[17]。张浩等通过水稻盆栽试验,测定水稻叶片氮素、籽粒蛋白质含量和冠层光谱,分别采用人工神经网络和 PLS 算法建立叶片氮素、籽粒蛋白质含量高光谱估算模型,结果表明,这 2 种模型预测效果良好,且 PLS 回归模型的预测效果最好,可以实现水稻氮素营养和籽粒品质的高光谱估测^[18]。付元元等采用 PLS 算法建立了基于原始光谱反射率与 LAI 相关性和基于光谱曲线特征的 2 种波段选择方式来预测大田冬小麦叶面积指数的估算模型^[19]。谭昌伟等运用 PLS 构建了遥感预测冬小麦籽粒蛋白质含量模型,为提高遥感预测小麦品质的精度提供了一种有效途径^[20]。王圆圆等测定了冬小麦 3 个发育期(孕穗、开花、乳熟)的冠层光谱和叶片含水量(LWC),针对每期数据,结合 PLS 和迭代特征去除,建立了基于诊断波段的 LWC 回归模型,结果表明叶片水分的光谱响应及反演精度受小麦生长状态的影响^[21]。王圆圆等通过人工田间诱发不同等级条锈病,在不同生育期(挑旗期、抽穗期、灌浆期和成熟期)测定冬小麦感染条锈病严重程度和冠层光谱,采用 PLS 算法建立了冠层光谱和条锈病严重度的回归模型,结果显示 PLS 回归建模不仅精度高,而且需要的人为干预少,自动化水平高^[22]。王纪华等运用 PLS 测定了田间小麦冠层可见光和近红外光谱不同波长处的冠层光谱反射率,将其与小麦不同层次的叶绿素、叶片全氮含量之间进行建模,通过建模和验证试验表明运用 PLS 算法能够较好地通过作物冠层光谱反演叶片全氮的垂直分布^[23]。

PLS 算法思想是:假定有 q 个因变量 $\{y_1, y_2, \dots, y_q\}$ 和 p 个自变量 $\{x_1, x_2, \dots, x_p\}$, 构成自变量和因变量数据表 x 和 y , 样本容量为 n 。运用 PLS 分别在 x 和 y 中提取成份 t_1 和 u_1 (t_1 是 x_1, x_2, \dots, x_p 的线性组合; u_1 是 y_1, y_2, \dots, y_q 的线性组合)。在提取成分时, t_1 和 u_1 应尽可能大地携带各自数据表中的变异信息,且 t_1 和 u_1 相关程度达到最大。在第 1 个成分提取后,分别实施 x 对 t_1 的回归和 y 对 u_1 的回归,如果回归方程满足预设精度,则算法停止;否则,利用 x 被 t_1 解释后的残余信息以及 y 被 u_1 解释后的残余信息进行第 2 轮的成分提取,如此反复,直到精度满足要求为止。若最终对 x 提取了 m 个成分 (t_1, t_2, \dots, t_m), 最后实施每个因变量 y_k ($k=1, 2, \dots, q$) 对 t_1, t_2, \dots, t_m 的回归,然后再转化为 y_k 关于原变量 x_1, x_2, \dots, x_p 的回归方程,则 PLS 回归建模结束^[10,22]。

PLS 回归模型计算步骤如下:(1)将数据标准化,记标准化后的因变量为 F_0 , 自变量为 E_0 。(2)记 t_1, u_1 分别是 E_0, F_0 的第 1 个主成分, $t_1 = E_0 \times w_1, u_1 = F_0 \times c_1$ 。其中, w_1 和 c_1 是 E_0 和 F_0 的第一主轴, $\|w_1\| = 1, \|c_1\| = 1$ 。 w_1 和 c_1 分别是矩阵 $E_0'F_0F_0'E_0$ 和矩阵 $F_0'E_0E_0'F_0$ 最大特征值的单位特征向量。 $E_0 = t_1p_1' + E_1$ 和 $F_0 = t_1r_1' + F_1$, 式中 $p_1 = E_0't_1/\|t_1\|^2, r_1 = F_0't_1/\|t_1\|^2$ 。(3)用残差矩阵 E_1, F_1 取代 E_0, F_0 , 求第二主轴和第二成分。如此计算下去,设提取 A 个成分 (t_0, t_1, \dots, t_A), 则有 $E_0 = t_1p_1' + \dots + t_Ap_A', F_0 = t_1r_1' + \dots + t_Ar_A' + F_A$ 。(4)在第(3)步中,第 k ($k \geq 1$) 次运算完成之后,判断回归模型是否有效,若有效,则运算中止,否则,进行第 $k+1$ 次运算^[10,24]。

PLS 是近几年发展起来的建模方法,其突出的优点包括:能够在自变量存在多重相关性的条件下进行回归建模;允许在样本点个数少于变量个数的条件下进行回归建模;PLS 回归模型中包含原有的所有自变量;更易辨识系统信息与噪音;PLS 模型中,每一个自变量的系数更容易解释。其缺点包括:PLS 算法本质上是对数据平均趋势的估算,无法避免“依据错误的模型得到错误的模型”的根本问题^[25];计算速度相对较慢,需要多次迭代,计算过程复杂^[19,26];不能减少自变量个数以减轻测试强度^[27]。

3 人工神经网络 (artificial neural network, ANN)

ANN 是 20 世纪 80 年代中后期世界范围内迅速发展起来的一个前沿领域,因其良好的预测性和实用性被广泛应用于各个领域。ANN 模拟人脑学习的过程,通过重复的输入和输出训练来增强和修改输入和输出数据之间的联系。

目前,在 ANN 的实际应用中,绝大多数模型都是采用 BP (back propagation) 网络及其变化形式^[28]。Yang 等用 ASD 光谱仪测得玉米高光谱反射率,运用 BPNN 算法建立了基于光谱数据的中国东北地区玉米 LAI 估算模型^[29]。姚霞等运用 BPNN 算法建立了小麦叶片氮含量的鲜叶和粉末状干叶近红外光谱估测模型,该模型能准确地估测小麦叶片氮含量^[30]。张浩等分析了试验点晚稻穗期叶片氮素与 CBERS-1 影像冠层光谱信息之间的关系,运用 BPNN 算法反演水稻穗期叶片氮素含量,发现 BPNN 模型具有很强的非线性拟合能力^[31]。张浩等在浙江省海宁市试验田同步采集了乳熟期水稻籽粒光谱反射率和蛋白质数据,利用 BPNN 算法,建立了水稻籽粒蛋

白质含量的高光谱估测模型,预测精度良好^[32]。王大成等测得小麦冠层反射光谱数据及相应的小麦生物量,基于 BPNN 算法,建立了利用高光谱遥感反演小麦生物量的模型,结果表明 BPNN 对作物小麦的生物量高光谱遥感诊断是一种实时高效的方法^[33]。宋开山等实测了不同生长季大田玉米的冠层高光谱数据与相应的作物地上鲜生物量,采用 SMR 以及 BPNN 算法,建立了玉米地上鲜生物量高光谱遥感估算模型,结果表明通过 BPNN 算法建立的遥感估算模型结果最为理想,能够准确预测玉米的地上鲜生物量^[5]。

BPNN 是一种具有 3 层或 3 层以上神经元的神经网络,包括输入层、中间层和输出层,上下层之间实现全连接,同一层的神经元之间无连接。

BPNN 算法的基本思想是学习过程由信号的正向传播与误差的反向传播 2 个过程组成。正向传播时,从输入层传入样本,经各隐层逐层处理后,传至输出层。若输出层的实际输出与期望输出不符,则转入误差反向传播阶段,将输出误差以某种形式通过隐层向输入层逐层反传,并将误差分摊给各层的所有单元,进行权值修正。这种信号正向传播与误差反向传播的各种权值调整过程周而复始地进行,权值不断调整的过程,也就是网络的学习训练过程,该过程一直进行到网络输出的误差减少到可接受的程度,或进行到预先设定的学习次数为止^[34]。

BPNN 算法步骤如下:(1)确定神经网络的结构,用小随机电数初始化权值;(2)从训练样本集中随机取 1 个样本;(3)前向过程求网络输出;(4)反向过程调整网络权值;(5)转入步骤(2),直到所有样本都完成步骤(2)至(4),一次训练周期结束;(6)计算误差:若达到精度要求,则算法训练成功,存储训练完的权值,算法结束。若未达到精度要求,则判断是否超过最大学习次数,即如果未超过,则转入步骤(2),开始下一个训练周期;如果超过,则算法训练失败,不能在设定学习次数内收敛,算法结束。

ANN 采用并行分布处理方法,可以快速地处理大量运算,具有很好的非线性逼近能力;它是自适应学习,可以学习不知道或不确定的系统;所有定量或定性的信息都等势分布贮存于网络内的各神经元,所以有很强的鲁棒性和容错性。这些特性使 ANN 具有独特的信息处理和计算能力,在机制尚不清楚的高维非线性系统上体现出强大优势^[35]。

ANN 还存在一些缺陷:网络结构需要事先指定或应用启发算法在训练过程中修正,这些启发算法难以保证网络结构的最优化;网络权系数的调整方法存在局限性;神经网络易陷入局部最优,有些甚至无法得到最优解;过分依赖学习样本;目前尚无一种理论能定量分析神经网络训练过程的收敛速度及其决定条件,并对其加以控制;神经网络的优化目标是基于经验风险最小化,不能保证网络的泛化能力;神经网络本身是一个黑盒子结构,不能有效解释输出结果与输入之间的关系^[36]。

4 支持向量机算法 (support vector machine, SVM)

支持向量机是由 Vapnik 等提出的一种建立在统计学习理论基础上的基于结构风险最小化的学习方法^[37],最初用于语音识别等模式识别即支持向量分类 (support vector classification, SVC),现已扩展到回归预测即支持向量回归 (support

vector regression, SVR)^[38]。目前在理论研究和实际应用两方面支持向量机正处于飞速发展的阶段。

吕杰等利用地物光谱仪测量研究区水稻冠层反射光谱,同步测量水稻叶片镉含量,应用 SVR 算法建立了水稻镉含量高光谱预测模型^[39]。梁亮等以最小二乘支持向量回归 (least squares support vector regression, LS-SVR) 算法分别建立了小麦冠层叶绿素含量、小麦冠层含氮率反演模型^[40-41];提出了微分归一化氮指数 (FD-NDNI) 作为估测小麦冠层氮含量的新高光谱指数,并采用 LS-SVR 算法对已建立的指数模型进行改进^[42]。黎锐等利用北京郊区冬小麦主要生育期多时相 LandsatTM 影像生成的归一化植被指数,运用 SVR 构建了小麦遥感估产模型,并与常规的多元回归方法进行对比,结果表明利用 SVR 方法构建的遥感估算模型有效地提高了估算精度^[43]。杨晓华等在浙江大学水稻试验区利用高光谱反射率的 4 种变换 (反射率、反射率一阶导数、反射率二阶导数和反射率对数变换) 分析了 SVR 对水稻叶面积指数和叶绿素浓度的预测能力,结果表明 SVR 对水稻生物物理参数的估算能力不但优于 SMR 模型,而且对分析不同形式的光谱反射率数据具有独到的功效^[2]。杨晓华等选取了与水稻叶面积指数相关性较好的 3 种植被指数,通过具有不同核函数的 SVR 模型对 LAI 进行估算,结果表明 SVR 方法用于水稻 LAI 高光谱估算具有良好的学习能力和鲁棒性^[44]。

SVR 算法的基本思想起源于使用 SVC 解决线性可分情况下 2 类样本的分类问题, SVC 核心思想是找到一个最优分类超平面,使 2 类样本的分类间隔最大化。SVR 与 SVC 相似,但 SVR 所求超平面是使所有样本点到超平面的距离为最小。对线性回归而言,问题变为寻求一个最优超平面,使得在给定精度 ε 条件下,所有样本点到最优超平面的距离都不大于 ε ; 对非线性回归而言,可通过核函数变换将样本映射到一个高维特征空间中用线性回归解决^[45]。

SVR 算法步骤如下:(1)给定训练集 $T = \{(x_1, y_1), \dots, (x_l, y_l)\} \subset (R^n \times y)^l$, 其中: $x_i \in R, y_i \in y = R, i = 1, \dots, l$;(2)选取适当的核函数 $K(x, x')$ 及精度 $\varepsilon > 0$ 和惩罚参数 $C > 0$;(3)求下列优化问题的解:

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*), \quad (1)$$

$$s. t. \begin{cases} [w \cdot \Phi(x_i) + b] - y_i \leq \varepsilon + \xi_i, i = 1, \dots, l & (2) \\ y_i - [w \cdot \Phi(x_i) + b] \leq \varepsilon + \xi_i^*, i = 1, \dots, l & (3) \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, l & (4) \end{cases}$$

求解步骤(3)的通常做法是解其对偶问题:通过使用拉格朗日乘子 $\alpha_i, i = 1, \dots, l$, 将线性约束条件(2)至(4)引入(1), 计算库恩-塔克条件, 得到最终解 $y = g(x) = \sum_{i=1}^l (\bar{\alpha}_i^* - \bar{\alpha}_i) K(x_i, x) + \bar{b}_0$ 。

SVR 算法具有以下特点:是一种有坚实理论基础的新颖的小样本学习方法,结构简单,便于应用;最终预测模型只由少数的支持向量确定,计算的复杂性取决于支持向量的数量,而不是样本空间的维数,这在某种意义上避免了“维数灾难”,适于处理高维问题,而且具有较好的鲁棒性;SVR 通过选择不同的核函数和参数可以获得不同的特性,灵活性好。

SVR 存在缺点:随着样本数量的增大,所需的计算时间和空间存储资源都会成几何级数增加;建模的核心问题是核

函数和参数的确定^[46],单一的 SVR 由于存在核函数和惩罚因子等参数选择的限制,其应用受到一定的限制。目前研究通过引入遗传算法、小波变换、模拟退火算法、混沌等全局寻优功能优化支持向量机内部参数,提高 SVR 预测模型精度;此外,还采用如主成分、模糊数学 EMD 和小波分析方法及混沌理论对 SVR 模型的输入进行一定的预处理,即采用主成分分析和模糊数学分析提取预测模型的输入因子,降低输入的维数,或采用粗集理论、小波和混沌理论重构相空间序列作为模拟的输入^[47]。

5 总结与展望

遥感作为一门新兴的学科,在世界范围内获得了飞速发展。遥感技术在农业方面的应用日趋深入。遥感信息模型是遥感应用深入发展的关键,应用遥感信息模型,可计算和反演对实际应用非常有价值的农业参数。算法在建立模型过程中发挥了重要作用,本文分别介绍了 SMR、PLS、ANN、SVR 算法,综述了这些算法近年来在建立作物信息遥感预测模型中的应用。

这些算法有助于建模精度的提高,但每种算法仍然存在着诸多不足,近年来提出了多种完善和改进算法,如基于 t 检验的逐步回归算法^[9]、基于惩罚的 PLS 回归^[48]、核 PLS 算法^[49-50]、BiPLS、FiPLS、MwPLS 算法^[51-52]、多核 SVM^[53]、GSA-SVM^[54]、加权 SVM^[55]、PSVM^[56-57] 等。这些改进算法或多种算法混合使用是农情信息遥感预测建模算法可供选择的。大多数的农学参数都会受到外界环境的影响,今后的建模工作应该综合考虑不同年份的土壤、气候等环境因子,并使这些因子参与到建模过程中,有助于提高模型的稳定性和广泛适用性。目前运用这些算法建模时,分析遥感参数与实际农学参数之间的关系,多数基于相关性对数据进行分析。在今后的工作中,应该将遥感知识和农学知识有效结合,充分利用多源多时相的遥感数据所提供的多种信息,从不同角度、不同层面采用更多更好的方法分析数据。为了方便地通过建立的模型进行农情信息遥感预测,可以软件工程的思想进行分析,以面向对象的编程工具进行开发,编制一系列容易操作的、基于图形界面的遥感预测软件,从而进一步实现作物遥感预测操作的普及化、高效化、便捷化。目前针对某个生育期建立的农学参数反演模型,在整个此生育期内不能反映此农学参数的动态变化规律。将回归模型和马尔科夫模型等结合,采用多元数据融合、多层次信息提取等技术方法,对农情变化过程进行描述、分析以及预测,建立农情信息遥感预测动态模型,这也是预测建模的研究趋势。

总之,实际工作应该学科交叉,将计算机技术、遥感技术、农业技术等有效结合,建立适宜的农作物遥感预测模型,最终以所建模型作为工具,帮助农民提高种植效益,增加收入。

参考文献:

- [1] 王纪华,赵春江,黄文江,等. 农业定量遥感基础与应用[M]. 北京:科学出版社,2008.
- [2] Jin X L, Wang K R, Xiao C H, et al. Comparison of two methods for estimation of leaf total chlorophyll content using remote sensing in wheat[J]. Field Crops Research, 2012, 135: 24-29.
- [3] 杨晓华,吴耀平,黄敬峰,等. 基于支持向量机的水稻生物物理参数遥感估算[J]. 中国科学(C辑):生命科学, 2009(11): 1080-1091.
- [4] 周冬琴,朱艳,姚霞,等. 基于水稻冠层光谱特征构建粳型水稻籽粒蛋白质含量预测模型[J]. 作物学报, 2007, 33(8): 1219-1225.
- [5] 宋开山,张柏,于磊,等. 玉米地上鲜生物量的高光谱遥感估算模型研究[J]. 农业系统科学与综合研究, 2005, 21(1): 65-67.
- [6] 宋开山,张柏,王宗明,等. 玉米和大豆 LAI 高光谱遥感估算模型研究[J]. 中国农学通报, 2005, 21(1): 318-322.
- [7] 牛铮,陈永华,隋洪智,等. 叶片化学组分成像光谱遥感探测机理分析[J]. 遥感学报, 2000, 4(2): 125-130.
- [8] 周纪芎. 实用回归分析方法[M]. 上海:上海科学技术出版社, 1990.
- [9] 刘明,王仁曾. 基于 t 检验的逐步回归的改进[J]. 统计与决策, 2012(6): 16-19.
- [10] 边重. 逐步回归方法介绍[J]. 数理统计与管理, 1983(3): 39-48.
- [11] 张小海,金家善,耿俊豹,等. 基于粗糙集的偏最小二乘回归方法[J]. 上海交通大学学报, 2010, 44(12): 1678-1686.
- [12] 王惠文. 偏最小二乘回归方法及其应用[M]. 北京:国防工业出版社, 1999.
- [13] 王惠文,吴载斌,孟洁. 偏最小二乘回归的线性与非线性方法[M]. 北京:国防工业出版社, 2006.
- [14] Serbin S P, Dillaway D N, Kruger E L, et al. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature[J]. Journal of Experimental Botany, 2012, 63(1): 489-502.
- [15] Ecartot M, Compan F, Roumet P. Assessing leaf nitrogen content and leaf mass per unit area of wheat in the field throughout plant cycle with a portable spectrometer[J]. Field Crops Research, 2013, 140: 44-50.
- [16] Li C J, Wang J H, Wang Q, et al. Estimating wheat grain protein content using multi-temporal remote sensing data based on partial least squares regression[J]. Journal of Integrative Agriculture, 2012, 11(9): 1445-1452.
- [17] 邵田田,宋开山,杜嘉,等. 基于偏最小二乘法的玉米 FPAR 高光谱反演模型研究[J]. 地理与地理信息科学, 2012, 28(3): 27-31.
- [18] 张浩,胡昊,陈义,等. 水稻叶片氮素及籽粒蛋白质含量的高光谱估测模型[J]. 核农学报, 2012, 26(1): 135-140.
- [19] 付元元,杨贵军,冯海宽,等. 基于高光谱维数约简与植被指数估算冬小麦叶面积指数的比较[J]. 农业工程学报, 2012, 28(23): 107-113.
- [20] 谭昌伟,王纪华,黄文江,等. 基于 TM 和 PLS 的冬小麦籽粒蛋白质含量预测[J]. 农业工程学报, 2011, 27(3): 388-392.
- [21] 王圆圆,李贵才,张立军,等. 利用偏最小二乘回归从冬小麦冠层光谱提取叶片含水量[J]. 光谱学与光谱分析, 2010, 30(4): 1070-1074.
- [22] 王圆圆,陈云浩,李京,等. 利用偏最小二乘回归反演冬小麦条锈病严重度[J]. 国土资源遥感, 2007(1): 57-60.
- [23] 王纪华,黄文江,劳彩莲,等. 运用 PLS 算法由小麦冠层反射光谱反演氮素垂直分布[J]. 光谱学与光谱分析, 2007, 27(7): 1319-1322.

- [24] 丁涛, 杨慧中. 偏最小二乘回归模型的泛化能力改进研究[J]. 控制工程, 2008, 15(2): 150-153.
- [25] 张小海, 金家善, 耿俊豹. 用 DEA 优化偏最小二乘回归建模及应用[J]. 浙江大学学报: 工学版, 2011, 45(9): 1688-1692.
- [26] Cho M A, Skidmore A, Corsi F, et al. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression[J]. International of Journal of Applied Earth Observed and Geoinformation, 2007, 9(4): 414-424.
- [27] 袁哲明, 谭显胜. 基于支持向量机非线性筛选水稻苗期抗旱性指标[J]. 作物学报, 2010, 36(7): 1176-1182.
- [28] 邹华东, 陈树人, 陈刚, 等. 基于人工神经网络的稻田杂草稻光谱识别[J]. 农机化研究, 2013, 35(1): 156-158, 163.
- [29] Yang F, Sun J L, Fang H L, et al. Comparison of different methods for corn LAI estimation over northeastern China[J]. International Journal of Applied Earth Observation and Geoinformation, 2012, 18: 462-471.
- [30] 姚霞, 汤守鹏, 曹卫星, 等. 应用近红外光谱估测小麦叶片氮含量[J]. 植物生态学报, 2011, 35(8): 844-852.
- [31] 张浩, 姚旭国, 张小斌, 等. 区域水稻穗期叶片氮素的遥感估测初探[J]. 核农学报, 2009, 23(3): 364-368.
- [32] 张浩, 姚旭国, 张小斌, 等. 水稻籽粒蛋白质含量的高光谱估测研究[J]. 中国粮油学报, 2009, 24(11): 1-5.
- [33] 王大成, 王纪华, 靳宁, 等. 用神经网络和高光谱植被指数估算小麦生物量[J]. 农业工程学报, 2008, 24(S2): 196-201.
- [34] 栾福明, 张小雷, 熊黑钢, 等. 基于不同模型的土壤有机质含量高光谱反演比较分析[J]. 光谱学与光谱分析, 2013, 33(1): 196-200.
- [35] Rogers S K, Kabrisky M. An introduction to biological and artificial neural networks for pattern recognition[M]. Bellingham; Society of Photo Optical, 1991.
- [36] 安琼, 杨邦杰, 焦险峰. 作物遥感识别中的数据挖掘技术[J]. 农业工程学报, 2007, 23(8): 181-186.
- [37] Bharghavan V, Demers A, Shenker S, et al. MACAW: a medium access protocol for wireless LANs[C]//New York; SIGCOMM'94 Proceedings of the Conference on Communications Architectures, Protocols and Applications, 1994: 212-225.
- [38] Liang Y C, Sun Y F. An improved method of support vector machine and its applications to financial time series forecasting[J]. Prog Nat Sci, 2003, 13(9): 696-700.
- [39] 吕杰, 刘湘南. 利用支持向量机构建水稻镉含量高光谱预测模型[J]. 应用科学学报, 2012, 30(1): 105-110.
- [40] 梁亮, 杨敏华, 张连蓬, 等. 基于 SVR 算法的小麦冠层叶绿素含量高光谱反演[J]. 农业工程学报, 2012, 28(20): 162-171.
- [41] 梁亮, 杨敏华, 臧卓. 基于小波去噪与 SVR 的小麦冠层含氮率高光谱测定[J]. 农业工程学报, 2010, 26(12): 248-253.
- [42] 梁亮, 杨敏华, 邓凯东, 等. 一种估测小麦冠层氮含量的新高光谱指数[J]. 生态学报, 2011, 31(21): 6594-6605.
- [43] 黎锐, 李存军, 徐新刚, 等. 基于支持向量回归(SVR)和多时相遥感数据的冬小麦估产[J]. 农业工程学报, 2009, 25(7): 114-117.
- [44] 杨晓华, 黄敬峰, 王秀珍, 等. 基于支持向量机的水稻叶面积指数高光谱估算模型研究[J]. 光谱学与光谱分析, 2008, 28(8): 1837-1841.
- [45] Vapnik V N. The nature of statistical learning theory[M]. New York; Springer-Verlag Press, 1995.
- [46] 乌玲瑰, 徐旻, 蔡晓晓, 等. 基于机器学习的水稻发育期预测模型构建[J]. 扬州大学学报: 农业与生命科学版, 2012, 33(3): 44-50.
- [47] 卢敏, 蒋浩, 杨岱根. 支持向量机在农业水土领域的应用与研究进展[J]. 安徽农业科学, 2012, 40(29): 14163-14164, 14190.
- [48] Yin H, Wang B B. Penalized partial least squares[J]. Journal of Mathematics, 2013, 33(1): 63-74.
- [49] Rosipal R, Trejo L J, Matthews B. Kernel PLS_SVC for linear and nonlinear classification[C]//Washington DC; 20th International Conference on Machine Learning, 2003.
- [50] Zhang Y W, Teng Y D, Zhang Y. Complex process quality prediction using modified kernel partial least squares[J]. Chemical Engineering Science, 2010, 65(6): 2153-2158.
- [51] Zou X B, Zhao J W, Li Y X. Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models[J]. Vibrational Spectroscopy, 2007, 44(2): 220-227.
- [52] Xiaobo Z, Jiewen Z, Povey M J, et al. Variables selection methods in near-infrared spectroscopy[J]. Analytica Chimica Acta, 2010, 667(1/2): 14-32.
- [53] 万源, 童恒庆, 朱映映. 基于遗传算法的多核支持向量机的参数优化[J]. 武汉大学学报: 理学版, 2012, 58(3): 255-259.
- [54] Sarafrazi S, Nezamabadi-pour H. Facing the classification of binary problems with a GSA-SVM hybrid system[J]. Mathematical and Computer Modelling, 2013, 57(1/2): 270-278.
- [55] Wang D F, Yeung D S, Tsang E C. Weighted mahalanobis distance kernels for support vector machines[J]. IEEE Trans on Neural Networks, 2007, 18(5): 1453-1462.
- [56] Mangasarian O L, Wild E W. Multisurface proximal support vector machine classification via generalized eigen values[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2006, 28(1): 69-74.
- [57] 陶晓燕, 姬红兵, 马志强. 基于样本分布不平衡的近似支持向量机[J]. 计算机科学, 2007, 34(5): 174-176.