

基于多源数据和模糊 k -均值方法的农田土壤管理分区研究*

郭 燕¹ 田延峰¹ 吴宏海² 史 舟^{1†}
(1 浙江大学环境与资源学院, 杭州 310058)
(2 浙江大学地球科学系, 杭州 310027)

摘 要 以近地传感器(EM38)快速获取的土壤表观电导率数据(E_{Ca})、高分遥感影像提取的归一化植被指数(NDVI)和雷达影像提取的后向散射系数为数据源,以滨海滩涂区为研究对象,利用聚类法进行农田土壤管理分区研究。在对不同的数据源进行空间变异分析的基础上,结合滨海滩涂区土壤盐分的空间变异规律,利用非监督分法—模糊 k -均值聚类方法进行特征相似性归类,划分不同的管理分区。结果表明,最佳的分区数目为 3 个。根据每一分区的特点,可以采取相应的管理措施进行土壤改良和农田精确管理。分区结果不但可以指导采样,而且可用于实施变量投入和精确施肥推荐,为样区大面积农田土壤管理提供科学的决策依据。

关键词 模糊 k -均值聚类;滨海盐土;管理分区;高分遥感;土壤表观电导率(E_{Ca})

中图分类号 S159 **文献标识码** A

数字化采集和精准化作业是当前包括土壤和作物精确管理的重要研究方向。特别是随着遥感技术和近地传感技术的发展,高分辨率遥感影像和农田快速实时获取的信息在土壤学科中的应用越来越广泛。2008 年国际土壤科学联合会还专门成立了土壤近地传感器技术工作组,大力推进先进传感获取技术与土壤学科的交叉应用,重点解决农田土壤信息的多源获取与处理的两大瓶颈问题^[1-2]。当前,农田土壤信息快速采集手段还不多,地面手段主要集中在大地电导率仪(EM)、探地雷达、时域反射仪(TDR)等。如 Rodríguez-Pérez 等^[3]利用土壤电导率数据、TDR 测量的水分数据和室内分析的理化数据,结合 GIS 框架进行了高黏土含量特征的土壤制图研究。卫星遥感方面,由于受到农田作物和残茬覆盖的影响,单一应用遥感技术进行农田土壤信息获取还存在许多的不确定性,因此充分发挥星地多源传感器的优势是当前发展的新方向。

利用多源数据进行科学的农田管理分区是精准化作业的重要基础。管理分区的合理划分是进行正确管理土壤空间变异的重要组成部分。目前虽然有

多种不同的方法可以进行管理分区定义^[4-5],但最常用的方法还是聚类算法,如人工神经网络法^[6]、 k -均值法^[7]、模糊 c -均值法^[8]等。在农田管理分区中,所采用的土壤特征数据一般均是连续的,数量较多,在进行相似特征归类的过程中采用非监督模糊聚类方法进行分区则更能反映客观实际情况。近年来,非监督分类方法之一——模糊 k -均值聚类法由于其采用软分类思想、最佳分类数目可计算、聚类信息丰富等优势,从而在空间分析和图像处理等领域被广泛使用^[9-10]。如 Triantafyllis 等^[11]利用 EM34 获取的土壤体积电导率通过对纳莫伊河谷盐渍化土壤地下水特征的空间变异性的研究,采用模糊 k -均值聚类法对地上景观植被进行了分区研究。

本研究针对研究区滨海围垦区土壤盐分、水分等关键影响因子,利用近地传感器和主被动高分遥感等多种手段进行快速大面积信息获取和解译方法的研究,并根据获取信息的空间变异特性采用模糊 k -均值聚类法进行区域内农田土壤管理分区的划分,可以因地制宜进行土壤改良,提高农田土壤精确管理的科学性。

* 国家自然科学基金项目(40571066,40871100)、国家科技支撑课题(2011BAD21B04)资助

† 通讯作者, E-mail: shizhou@zju.edu.cn

作者简介:郭 燕(1983—),女,博士研究生,主要从事土壤遥感与信息技术方面的研究。E-mail: guoyan8372@163.com

收稿日期:2012-06-27;收到修改稿日期:2012-09-25

1 材料与方 法

1.1 研究区域

研究区位于浙江省上虞市西北杭州湾南岸的海涂实验农场。该地属于亚热带中部区域,自然植被为常绿阔叶林,年平均温度 16.5℃,年均降水量

约为 1 300 mm。围垦土壤主要来自于河海相沉积物,土壤质地以轻壤土或者砂壤土为主。土壤类型属于海积潮湿正常盐成土亚类,剖面发育不明显,同一剖面上下层质地较为均匀。近年来该区已被连续垦种,但是由于围垦措施和种植利用的不同,土壤水盐交互作用强烈,田块内土壤盐分含量具有差异性。本研究区围垦于 1996 年,面积为 4.5 hm²(图 1)。

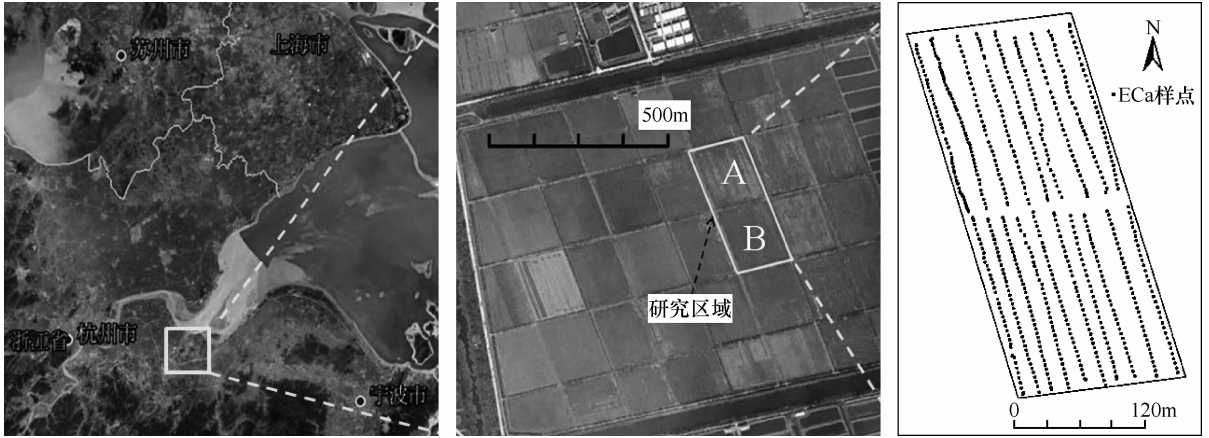


图 1 研究区域及 ECa 采样点

Fig. 1 Study region and locations for ECa sampling

1.2 地面数据采集与处理

地面主要利用大地电导率仪 EM38 (Canada, Genicos Co, Ltd) 进行农田土壤表观电导率 (ECa) 的采集,采用水平模式 (EM38 水平模式下能够快速获取 0~1.5 m 土体的电导率数据) 系统于 2010 年 4 月 26 日获得 711 个电导率数据,其中 A 田块 338 个, B 田块 373 个 (图 1)。EM38 工作的同时连接 GPS 进行采样点位置的确定。数据采用 Arcgis9.3 和 Minitab15 软件进行分析处理。

1.3 卫星数据收集与处理

多光谱高分遥感数据采集有两个时相,分别为春季 (2010 年 4 月 17 日) 的 WorldView-II 和秋季 (2010 年 8 月 21 日) 的 GeoEye-1 多光谱遥感影像,分辨率均为 2 m。数据通过 ENVI4.7 软件进行经过几何校正、归一化植被指数 (NDVI) 的计算,其中 NDVI 的计算通过式 (1) 完成。

$$NDVI = (NIR - R) / (NIR + R) \quad (1)$$

式中, NIR 表示经红外波段的反射率, R 表示红光波段的反射率。

雷达影像获取于 2010 年 11 月 21 日过境的 ALOS 卫星,它能同时提供全色、多光谱以及 SAR 数据。本研究中选取 PALSAR 的 HH 和 HV 两种极化方式的数据,分辨率为 12.5 m,利用前述的 World-

View-II 全色影像为参考进行几何校正,然后利用式 (2) 计算后向散射系数 (σ^0)。

$$\sigma^0 = 10 \times \log_{10} (DN)^2 + \varepsilon \quad (2)$$

式中, ε 为雷达定标系数, DN 为相元的灰度值,对于 ALOS/PALSAR Level 1.5 数据, $\varepsilon = -83.0$ [12]。

1.4 模糊 k-均值聚类法

模糊 k-均值聚类是一种常用的非监督聚类方法,它的基本思想是将一个数据集分成 k 个类别,寻找目标函数的迭代最小化,在气候、地质、土壤景观等研究中已得到了广泛的应用 [13]。最常用的目标函数为 [14]:

$$J_{\phi}(M, C) = \sum_{i=1}^n \sum_{j=1}^k \mu_{ij}^{\phi} d_{ij}^2 \quad (3)$$

式中, n 为属性数据的个数; k 为类别个数; μ_{ij} ($1 \leq i \leq n, 1 \leq j \leq k$) 表示数据矩阵 X 中第 i 个样本 x_i 属于聚类中心矩阵 C 中第 j 个聚类中心 c_j 的隶属度; M 为隶属度矩阵; $(d_{ij})^2$ 等于 x_i 与 c_j 在特征向量上距离的平方; ϕ 为模糊加权指数 ($\phi \geq 1$), 它控制了不同类别间共用数据的数目,当 $\phi = 1$ 时,就会发生硬聚类,即没有数据共用现象,此时为一般的 k-均值聚类法,即用最小二乘法找出一个恰当的 k 组分类矩阵 M 和一个由恰当的聚类中心向量 c_j 组成的矩阵 C,使得目标函数 $J_{\phi}(M, C)$ 达到最小;当 $\phi > 1$ 时用加

权最小二乘法找出一个恰当的 M 和 C 使得目标函数达到最小, ϕ 越大, 共用数据数目增加, 最终的分类变得较不明显。

2 结 果

2.1 地面获取数据分析

在滨海滩涂围垦区, 土壤盐分是影响作物生长的关键因素, 利用 EM38 设备获取的土壤表观电导率可以间接地反映土壤盐分的空间变异特征。首先对两个田块内 ECa 进行统计分析, 其中 A 田块采集到数据为 338 个, B 田块为 373 个。A 和 B 块田 ECa 的变幅分别为 45.5 ~ 226.6 mS m^{-1} 和 18.3 ~ 223.5 mS m^{-1} , 平均值分别为 147.0 mS m^{-1} 和 146.7 mS m^{-1} , t 检验结果表明, 田块之间在 0.05 水平上差异不显著。

由 ECa 的统计分析知, 田块之间的差异主要是微地形因素和耕作方式不同造成的, 因此可将 A 和 B 田块合并进行统一分析。对 ECa 值的空间变异特性进行分析, 半方差函数模型以指数模型拟合最佳, 块金值 C_0 为 340, 基台值 ($C + C_0$) 为 3 508, 决定系数为 0.901, 块金值与基台值的比值为 0.10。块金值和基台值的比值表示空间相关性程度, 如果比值小于 0.25, 说明变量具有强烈的空间相关性; 在 0.25 ~ 0.75 之间, 变量具有中等程度的空间自相关; 大于 0.75 时, 变量空间相关性很弱; 如果比值接近 1, 变量在整个尺度上具有恒定的变异^[15-16], 土壤表观电导率具有强烈的空间相关性。本研究中块金值和基台值的比值为 0.10, 土壤表观电导率在空间上具有强烈的相关性。

2.2 卫星遥感数据分析

2.2.1 NDVI 数据分析 研究区围垦于 1996 年, 主要土壤类型属于海积潮湿正常盐成土亚类, 含砂量高, 土壤盐分含量也相对较高, 主要利用类型相对简单, 农田主要是以水稻和旱作轮作为主。尤其是在水稻种植和旱作交替种植区域, 水分、盐分胁迫比较明显, 冬季寒冷, 雨水少, 盐分会随着毛管水上升至地表, 种植作物主要是蔬菜(榨菜, 芥菜, 主要用来腌制咸菜), 盐分胁迫明显; 夏季高温湿热, 8 月份降雨减少, 耕层会产生一个强烈的返盐过程, 此时种植作物水稻处于抽穗期, 田间灌水会抑制盐分向上迁移运动, 这种轮作的种植方式能够有效地抑制盐分对作物的伤害^[17]。对获取于春季的 WorldView-II 高分遥感影像和获取于秋季的 Geo-

Eye-1 高分遥感影像进行统计分析, 春季芥菜的 NDVI 植被指数变幅为 0.32 ~ 0.60, 平均值为 0.52, 秋季水稻的 NDVI 变幅为 0.36 ~ 0.70, 平均值为 0.64。频数分析表明两个时期的 NDVI 分布特征均表现为右偏斜的尖峭峰(图 2), 秋季的较春季的更右偏, 且中间值所占的比例较大。这种差异性使得春季和秋季时相的遥感影像均作为分区管理变量输入是非常必要的。

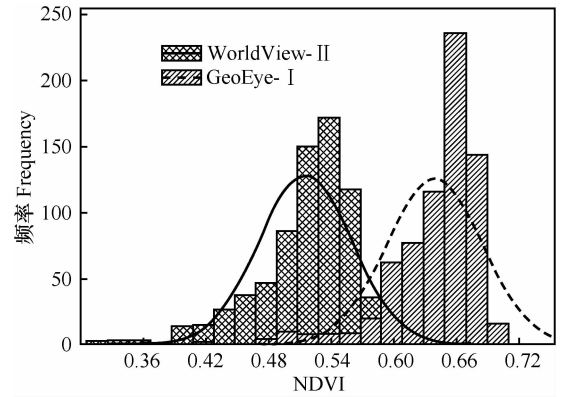


图 2 NDVI 频数分布图

Fig. 2 Frequency distribution of NDVI

2.2.2 雷达后向散射系数分析 雷达遥感具有全天时、全天候、及对某些地物的穿透能力, 这在阴雨天较多的浙江上虞地区的优势更加明显。另一方面, 在滩涂区, 土壤水分和盐分之间存在着密切关系^[18], 水分在农田利用中的影响也必须作为一个重要的影响因子进入到精细化农田作业管理中。利用 ALOS/PALSAR L 波段的雷达波对土壤水分的反应比较敏感的这种特性, 金希^[19]采用经验模型分别用 HH 和 HV 两种极化方式的数据对裸土土壤水分含量进行了反演研究。也有人基于雷达不同极化方式数据的特征, 采用比值等进行了土壤含水量的反演研究^[20]。本研究采用 HH 极化和 HV 极化数据之差来研究滩涂盐碱地地表水分的散射情况。统计分析表明, (HH-HV) 差值的变幅范围为 4.88 ~ 17.85 dB, 平均值为 10.60 dB, 标准差为 2.04 dB, 偏度系数为 0.442, 峰度系数为 0.155, 表现为正偏斜尖峭峰分布特征。进一步采用半方差函数进行模拟分析, 最优的拟合效果函数是高斯模型, 块金值 C_0 为 1.78, 基台值 ($C + C_0$) 为 20.05, 决定系数为 0.936, $C_0 / (C_0 + C)$ 为 0.09。这表明两种极化方式后向散射系数的差值具有强烈的空间相关性, 为进一步研究滩涂土壤空间变异性奠定了基础。

2.3 管理分区划分及精度检验

聚类分析能有效地将具有相似的土壤肥力特性、地形特征或者作物产量的区域进行分类而被较多地用于精确农业管理分区中。利用不同的数据源,譬如土壤属性(肥力、电导率、有机质、水分、质地等)、遥感影像、产量数据等进行差异化种植和管理,可以提高生产效率,增大效益。本研究中综合利用近地传感器 EM38 非侵入式快速获取的土壤 ECa 数据,高分遥感影像提取的 NDVI 指数和 ALOS/PALSAR 雷达影像的后向散射系数,借助模糊 k -均值聚类方法进行管理分区研究。

将数据归一化处理,作为输入变量用于模糊 k -均值聚类分析中,此过程采用 FuzME^[21] 软件实现。取最大迭代次数为 300,收敛阈值为 0.001,模糊加权指数为 1.5。为了找出最佳分区数目,分别产生 2、3、4、5、6、7、8、9、10 个类别,计算并比较每一类的模糊性能指数(FPI)、改进分类熵(MPE)和聚类独立性指数(S)的值。

在进行模糊聚类分析时,一个关键的问题是划分多少类是最合适的。本文借助于 FPI、MPE 和 S 来确定合适的聚类数。FPI^[9] 的值在 0~1 之间变动。若该值接近 0,表示聚类时共用数据较少,类的划分明显;若该值接近 1,则表示具有较多的共用数据,类的划分不明显。MPE^[9] 表示模拟数据矩阵的模糊分区的分解量,是度量各类间模糊程度的一个综合指标。S 表示类与类之间相对分明性。三者均是越小越好。FPI 越小表示聚类中的共用像元数越小,MPE 越小则表示分解量越大,S 越小表示分类精度越高。若 FPI 和 MPE 的值在某个类别时均最小,表示该类是研究区的最佳分类数。图 3 是将研究区划分为 2、3、4、5、6、7、8、9、10 时得到的 FPI、MPE 和 S 值。可以看出,当分类数目为 3 时,FPI、MPE 和 S 值均为最小,因此对本研究区,最佳的分区数目为 3。

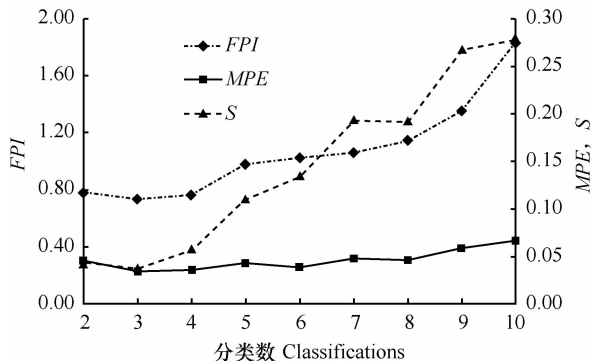


图 3 不同类别对应的 FPI、MPE 和 S 值

Fig. 3 Calculated FPI, MPE, and S corresponding to various classes of soils

另一个度量指标是混淆指数(CI)^[22]表示属性空间在类之间的重叠程度,是检验单个样本分类好坏的一个重要指标。CI 是将模糊隶属度图转换至易于理解的破碎区间,CI 越接近 0,观测数据属于占优势类的可能性越大;而如果 CI 接近 1 时,这个特殊观测值的分类产生了混淆,占优势地位和次要地位的两个类之间的差异可以忽略,因此 CI 越小越好。

利用所得的分类结果将研究区划分为 3 个管理分区(图 4):分区 1,ECa 变幅范围 18.3~88.1 mS m⁻¹;分区 2,土壤盐分含量相对较高,ECa 变幅范围 89.2~153.5 mS m⁻¹;分区 3,土壤盐分含量较高,ECa 变幅范围 154.2~226.6 mS m⁻¹。由图 4 可知,各管理分区基本上是连续的,可以根据每一个分区的特点进行于精确农业作业管理,但分区结果中出现了边界效应和零星的小单元,这种现象主要由于田块周围的田埂和排水沟渠造成的,稻田的灌排作用使得靠近沟渠的地方盐分含量和周围相差较大。混淆指数则是评价单个管理单元分类效果好坏的一个重要指标,基于此进行管理单元的量化(图 5),可以看出,在分区的边界和零星的小单元的区域混淆指数值也较大,即在分区时出现了类之间的重叠,且重叠程度较高。对于这些区域混淆指数值较高的单元可按照重叠的类的管理方式进行管理。那么这种重叠方式管理的误差是多大呢?利用判别分析对分区的错分误差进行评估(表 1),有 705 个数值数分类是正确的,总体正确率为 0.992。

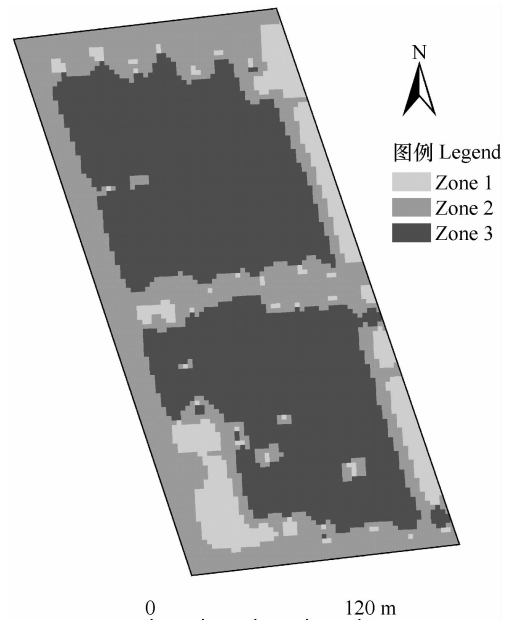


图 4 分区结果图

Fig. 4 Soil management zoning map

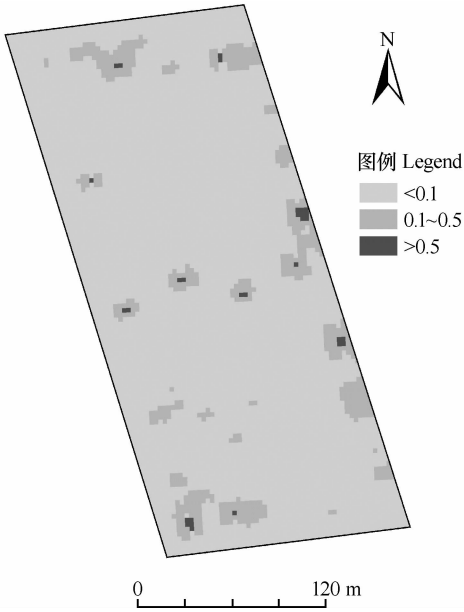


图 5 混淆指数图

Fig. 5 Confusion index map

表 1 模糊 k-均值聚类判别分析结果

Table 1 Discriminant analysis of Fuzzy-k means classification

实际组 Actualgroup	置入组 Placed group			合计数 Total number	正确数 Correct number	比率 Ratios
	分区 1 Zone 1	分区 2 Zone 2	分区 3 Zone 3			
分区 1 Zone 1	145	1	0	146	145	0.993
分区 2 Zone 2	4	137	0	141	137	0.972
分区 3 Zone 3	0	1	423	424	423	0.998

3 讨论

模糊聚类是常用的一种非监督分类方法,广泛应用于土壤、地形地貌、遥感数据等的分类中,能有效地将具有相似特征的属性数据归类,计算简单且运行速度较快,在描述连续性属性的中介性上具有优势,能客观反映现实世界。

在模糊 k-均值聚类算法中,影响分类效果的两个关键参数是聚类数目 k 和模糊加权指数 ϕ ,其中分类数目 k 可以通过设定一个范围进行比较选择,此时模糊指数 ϕ 就显得尤为重要。当 ϕ 趋近于 1 时,模糊 k-均值聚类算法逐渐退化为 k-均值聚类算法;当 $\phi > 1$ 时,趋近于无穷大时,模糊 k-均值聚类算法的聚类结果越来越模糊。如何确定 ϕ 的值? Bezdek 等^[23]从算法的收敛性角度着手得出 ϕ 的取值要大于 $n/(n-2)$,其中 n 为样本数目;陈彦和吕新^[24]则以 $c-\phi$ 多次组合最优取值法寻找适宜的模糊控制参数,确定了模糊指数为 1.5 时的结果最优;Pal 等^[25]从聚类有效性角度得出 ϕ 的取值范围为 1.5~2.5;Bezdek^[23]给出经验数值为 1.1~5.0。本文对模糊指数 ϕ 为 1.3、1.4、1.5、2.0、2.5、3.0、3.5、4.0、4.5、5.0 分别进行了拟合(图 6),结果表明, $\phi = 1.5$ 时, FPI 、 MPE 和 S 的值均为最小,因此本研究选取 $\phi = 1.5$ 进行聚类分析、划分管理分区。

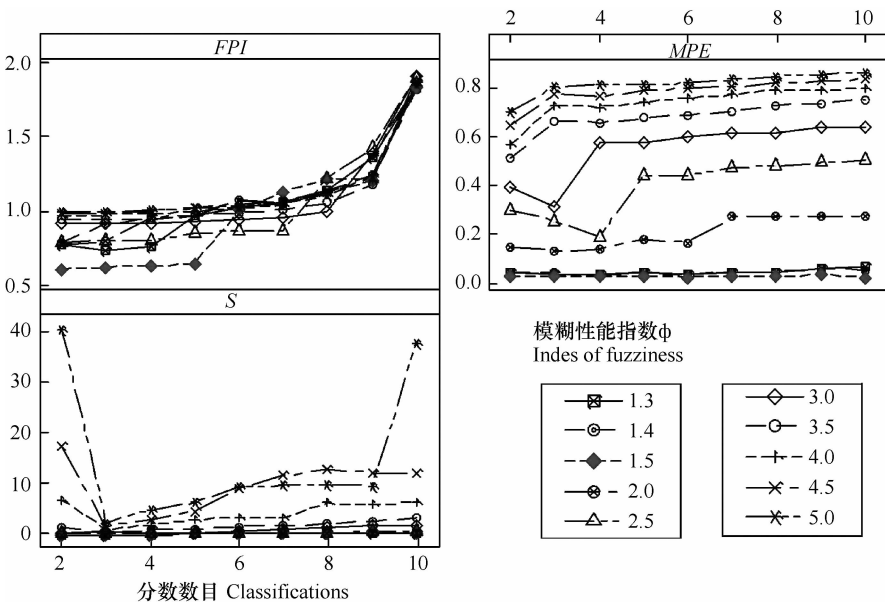


图 6 不同模糊指数下 FPI 、 MPE 和 S 的值

Fig. 6 FPI , MPE , and S values as affected by fuzziness index

对得到的 3 个分区,可以采取针对性的土壤管理措施。分区 1 土壤盐分含量较低(ECa 变幅范围 $18.3 \sim 88.1 \text{ mS m}^{-1}$, 利用 ESAP^[26] 转换成盐度后为 $0.54 \sim 2.74 \text{ dS m}^{-1}$), 根据 FAO^[27] 和 Tanji^[28] 对耐盐性指标的研究, 水稻(耐盐阈值 3.0 dS m^{-1})、小麦(耐盐阈值 4.0 dS m^{-1}) 及部分蔬菜能正常生长, 可以采取合理灌溉和施肥为主的综合农业措施; 分区 2 土壤盐分含量相对较高(ECa 变幅范围 $89.2 \sim 153.5 \text{ mS m}^{-1}$, 转换成盐度后为 $2.78 \sim 4.81 \text{ dS m}^{-1}$), 但是大豆(耐盐阈值 5.0 dS m^{-1})、豇豆(耐盐阈值 5.0 dS m^{-1}) 等作物也能正常生长, 可以采取以生物改良为主, 配合灌溉洗盐、地膜覆盖、秸秆还田和增施有机肥等措施; 分区 3 土壤盐分含量较高(ECa 变幅范围 $154.2 \sim 226.6 \text{ mS m}^{-1}$, 转换成盐度后为 $4.83 \sim 7.12 \text{ dS m}^{-1}$), 耐盐作物棉花(耐盐阈值 7.5 dS m^{-1})、大麦(耐盐阈值 8.0 dS m^{-1}) 能正常生长, 可以采取灌排网建设、农田平整等工程措施为主, 结合耐盐作物种植、施用土壤改良剂等措施。考虑到本研究区的实际情况和农田作业的方便, 一方面可以平整土地, 减少微地形对盐分迁移的影响; 另一方面可以通过水旱轮作和灌排水工程加快土壤脱盐过程。此方法同样可以用在大面积的分区管理研究中, 此时则可以在不同的分区中采用不同的作物种植或者是不同的农业措施来管理。

4 结 论

管理分区是实施精准农业变量投入的一个经济有效的手段。模糊 k -均值聚类作为模糊分类和模式识别方法, 具有聚类信息丰富等特点, 从而在数据挖掘、图像处理领域和聚类分析中被广泛使用。本文利用反映滩涂区土壤质量的电导率因子、反映地上植被特征的 NDVI 和反映土壤水分特征的后向散射系数作为输入变量, 采用模糊 k -均值聚类算法来定义田间管理分区, 并利用模糊性能指数、改进分类熵和聚类独立性指数有效地确定了最佳聚类效果和分区数目, 结果表明, 最佳分区数目为 3。用混淆指数评价单个管理单元分区效果的好坏, 为对海涂区盐碱地进行田间管理分区和土壤改良提供了依据。通过划分管理分区来指导土壤采样, 能通过少量的土壤样品调查获取各管理分区的土壤养分变异特征, 同时也可以为分区进行精确农业管理提供决策依据, 具有重要的理论和应用价值。

参 考 文 献

- [1] Viscarra Rossel R A, McBratney A B, Minasny B. Proximal soil sensing. Springer Science + Business Media B V, 2010
- [2] 史舟, 郭燕, 金希, 等. 土壤近地传感器研究进展. 土壤学报, 2011, 48(6): 1274—1281. Shi Z, Guo Y, Jin X, et al. Advancement in study on proximal soil sensing (In Chinese). Acta Pedologica Sinica, 2011, 48(6): 1274—1281
- [3] Rodríguez-Pérez J R, Plant R E, Lambert J-J, et al. Using apparent soil electrical conductivity (ECa) to characterize vineyard soils of high clay content. Precision Agriculture, 2011, 12(6): 775—794
- [4] Brock A, Brouder S M, Blumhoff G, et al. Defining yield-based management zones for corn-soybean rotations. Agronomy Journal, 2005, 97: 1115—1128
- [5] Moral F J, Terrón J M, Rebollo F J. Site-specific management zones based on the Rash model and geostatistical techniques. Computers and Electronics in Agriculture, 2011, 75: 223—230
- [6] Fortin J G, Anctil F, Parent L-E, et al. Site-specific early season potato yield forecast by neural network in Eastern Canada. Precision Agriculture, 2011, 12(6): 905—923
- [7] Altdorff D, Dietrich P. Combination of electromagnetic induction and gamma spectrometry using k -means clustering: A study for evaluation of site partitioning. Journal of Plant Nutrition and Soil Science, 2012, 175: 345—354
- [8] 李艳, 史舟, 吴次芳, 等. 基于模糊聚类分析的田间精确管理分区研究. 中国农业科学, 2007, 40(1): 114—122. Li Y, Shi Z, Wu C F, et al. Definition of management zones based on fuzzy clustering analysis in coastal saline land (In Chinese). Scientia Agricultura Sinica, 2007, 40(1): 114—122
- [9] McBratney A B, De Grujter J J. A continuum approach to soil classification by modified fuzzy k -means with extragrades. Journal of Soil Science, 1992, 43(1): 159—175
- [10] 李翔, 潘瑜春, 马景宇, 等. 基于多种土壤养分的精准管理分区方法研究. 土壤学报, 2007, 44(1): 14—20. Li X, Pan Y C, Ma J Y, et al. Soil nutrients-based zoning for management of precision agriculture (In Chinese). Acta Pedologica Sinica, 2007, 44(1): 14—20
- [11] Triantafyllis J, Odeh I O A, Minasny B, et al. Elucidation of physiographic and hydrogeological features of the lower Namoi valley using fuzzy k -means classification of EM34 data. Environmental Modelling & Software, 2003, 18: 667—680
- [12] Shimada M, Isoguchi O, Tadono T, et al. PALSAR radiometric and geometric calibration. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(12): 3915—3932
- [13] Gorsevski P V, Gessler P E, Jankowski P. Integrating a fuzzy k -means classification and a Bayesian approach for spatial prediction of landslide hazard. Journal of Geographical Systems, 2003, 5: 223—251
- [14] Bezdek J C. Pattern recognition with fuzzy objective function algorithms. New York: Plenum Press, 1981: 256
- [15] Cambardella C A, Moorman T B, Novak J M, et al. Field-scale

- variability of soil properties in Central Iowa soils. *Soil Science Society of America Journal*, 1994, 58: 1501—1511
- [16] Shi Z, Li Y, Wang R C, et al. Assessment of temporal and spatial variability of soil salinity in a coastal saline field. *Environmental Geology*, 2005, 48(2): 171—178
- [17] 董炳荣. 浙江新围海涂农业综合开发技术. 北京: 中国农业科技出版社, 1996. Dong B R. Comprehensive agricultural development technology in new reclaimed saline land in Zhejiang (In Chinese). Beijing: China Agricultural Science and Technology Press, 1996
- [18] 刘广明, 杨劲松. 土壤含盐量与土壤电导率及水分含量关系的试验研究. *土壤通报*, 2001, 32(6): 85—87. Liu G M, Yang J S. Study on the correlation of soil salt content with electric conductivity and soil water content (In Chinese). *Chinese Journal of Soil Science*, 2001, 32(6): 85—87
- [19] 金希. 高分辨率 SAR 影像裸土信息提取及土壤含水量反演初探. 杭州: 浙江大学, 2011. Jin X. Bare soil information extraction and soil moisture inversion based on high resolution SAR image (In Chinese). Hangzhou: Zhejiang University, 2011
- [20] 杜今阳. 多极化雷达反演植被覆盖地表土壤水分研究. 北京: 中国科学院研究生院, 2006. Du J Y. Soil moisture inversion with vegetation coverage by multi-polarization radar image (In Chinese). Beijing: Graduate University of Chinese Academy of Sciences, 2006
- [21] Minasny B, McBratney A B. FuzME version 3.0. Australian Centre for Precision Agriculture, The University of Sydney, Australia. 2002
- [22] Burrough P A, McDonnell R A. Principles of geographic information systems. Oxford: Oxford University Press, 1998: 333
- [23] Bezdek J C, Hathaway R J, Sabin M J, et al. Convergence theory for fuzzy c -means; Counterexamples and repairs. *IEEE Transactions on Systems, Man and Cybernetics*, 1987, 17(5): 873—877
- [24] 陈彦, 吕新. 基于模糊 c -均值聚类的绿洲农田精确管理分区研究. *生态学报*, 2008, 28(7): 3067—3074. Chen Y, Lü X. Definition of management zones based on fuzzy c -mean algorithm in oasis farmland (In Chinese). *Acta Ecologica Sinica*, 2008, 28(7): 3067—3074
- [25] Pal N R, Pal K, Keller J M, et al. A possibilistic fuzzy c -means clustering algorithm. *IEEE Transactions on Fuzzy Systems*, 2005, 13(4): 517—530
- [26] Lesch S M, Rhoades J D, Corwin D L. ESAP-95 version 2. 35R: User manual and tutorial guide. USDA-ARS, Brown G E Jr Riverside, CA: Salinity Laboratory, 2006
- [27] FAO. A framework for land evaluation. Rome: FAO and Agricultural Organization of the United Nations, 1976
- [28] Tanji K K. Agricultural salinity assessment and management. Reston: American Society of Civil Engineers, 1990

ZONING OF SOIL MANAGEMENT BASED ON MULTI-SOURCES DATA AND FUZZY- K MEANS

Guo Yan¹ Tian Yanfeng¹ Wu Honghai² Shi Zhou^{1*}

(1 College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China)

(2 Department of Earth Sciences, Zhejiang University, Hangzhou 310027, China)

Abstract Soil apparent electrical conductivity (ECa) acquired by proximal sensor EM38, normalized difference vegetation (NDVI) extracted from high resolution remote sensing images and backscattering coefficient obtained from radar images as the source data for farm field soil management zoning with the clustering method of the coastal saline for precision agriculture. So, based on spatial variability analysis of the data and the law for spatial variation of soil salinity in the area, zoning was done using the unsupervised classification method—fuzzy k -means cluster algorithm. Results show that the optimal number of zones was 3. In light of the characteristics of each zone, corresponding management practices should be adopted to ameliorate the soil and implement precision farmland management. The zoning is useful not only to guiding soil sampling, but also to recommendation of variable input and precision fertilization, and moreover, provides scientific basis for decision making in large-scale soil management.

Key words Fuzzy k -means classification; Coastal saline soil; Management zoning; High resolution remote sensing; Soil apparent electrical conductivity (ECa)

(责任编辑:陈德明)