Computers and Electronics in Agriculture 102 (2014) 51–63

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Modeling the spatial distribution of crop sequences at a large regional scale using land-cover survey data: A case from France

Ying Xiao [a,*], Catherine Mignolet [a], Jean-François Mari [b], Marc Benoît [a]

[a] INRA SAD UPR 055 ASTER, 662 Avenue Louis Buffet, 88500 Mirecourt, France
[b] Université de Lorraine, LORIA, UMR 7503, 54506 Vandoeuvre-lès-Nancy, France

ARTICLE INFO

Article history:
Received 1 July 2013
Received in revised form 8 January 2014
Accepted 14 January 2014

Keywords:
Crop sequences
Cropping patterns
Cropping systems
Hidden Markov models
Agricultural land-use
Teruti survey

ABSTRACT

Assessing the environmental impacts of agricultural production systems requires spatially explicit information regarding cropping systems. Projecting changes in agricultural land use that are caused by changes in land management practices for analyzing the performance of land activity-related policies, such as agricultural policies, also requires this type of data for model inputs. Crop sequences, which are vital and widely adopted agricultural practices, are difficult to directly detect at a regional scale. This study presents innovative stochastic data mining that was aimed at describing the spatial distribution of crop sequences at a large regional scale. The data mining is performed by hidden Markov models and an unsupervised clustering analysis that processes sequentially observed (from 1992 to 2003) land-cover survey data on the French mainland named Teruti. The 2549 3-year crop sequences were first identified as major crop sequences across the entire territory, which included 406 (merged) agricultural districts, using hidden Markov models. The 406 (merged) agricultural districts were then grouped into 21 clusters according to the similarity of the probabilities of occurrences of major 3-year crop sequences using hierarchical clustering analysis. Four cropping systems were further identified: vineyard-based cropping systems, maize monoculture and maize/wheat-based cropping systems, temporary pasture and maize-based cropping systems and wheat and barley-based cropping systems. The modeling approach that is presented in this study provides a tool to extract large-scale cropping patterns from increasingly available time series data on land-cover and land-use. With this tool, users can (a) identify the homogeneous zones in terms of fixed-length crop sequences across a large territory, (b) understand the characteristics of cropping systems within a region in terms of typical crop sequences, and (c) identify the major crop sequences of a region according to the probabilities of occurrences.

1. Introduction

Currently, 43% of the area of Europe (Eurostat, 2010) and 36% of the total area of the world (FAOSTAT, 2011) are dominated by agricultural land use, including both cropland and grassland. The current challenge for agronomists, farmers and their allied partners is to satisfy humanity's need for food and fiber, as well as the accelerating demand for biomass, in an ecologically sustainable way through socially accepted production systems (Miller, 2008).

Over the past decade, in the land change scientific community, the scientific interest regarding the investigation of land-cover modifications that are caused by changes in land management practices has increasingly been noticed by researchers. As noted by Lambin et al. (2000), changes in agricultural land-use management, e.g., changes in input levels, and their effect on profitability or the periodicity of complex land-use trajectories such as fallow cycles and rotation systems frequently drive land-cover modification. Incorporating the representation of agricultural land management practices and their changes into land system models will improve our understanding of the endogenous driving forces of land-cover modification. Several land system models have integrated the module for simulating farmer management practices and decision-making processes (Rounsevell et al., 2003). Agent-based models were specially developed and applied to represent human behavioral and decisional processes in the land system (Matthews et al., 2007). As one of the most significant forms of land-cover modification, agricultural land intensification has recently been studied using different land-use intensity indicators, such as livestock density and nitrogen input to UAA (utilized agricultural area), in relation to land management practices (Herzog et al., 2006). For instance, Temme and Verburg (2011) mapped and modeled agricultural land-use intensity in terms of nitrogen input at the European Union scale. A multi-scale modeling approach for exploring the spatial–temporal dynamics of European livestock distribution was proposed by Neumann et al. (2011).

* Corresponding author. Tel.: +33 (0)329385500; fax: +33 (0)329385519.
E-mail address: yingxiao0201@gmail.com (Y. Xiao).

0168-1699/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compag.2014.01.010

However, crop rotations, which are a vital agricultural land management practice, are rarely integrated into a land-use modeling framework at regional to global scales (Schönhart et al., 2011).

Crop rotations are defined as the practice of growing a sequence of crops on the same land (Wibberley, 1996). The term 'rotation' implies a cycle and is characterized by the identified starter crops and by the cycle period (e.g., biannual, triennial, and quadrennial) (Leteinturier et al., 2006). Because of the multiple benefits of crop rotations, such as increasing crop yields, decreasing the incidence of plant diseases and weeds, maintaining soil fertility, improving the soil structure, and preserving biodiversity, crop rotations are an old widespread practice. In the context of the establishment of new economic, agronomic and governmental policies, farmers will be paid for re-establishing and increasing ecosystem services on agricultural land (Miller, 2008). The positive effect of crop rotations has once more come to the attention of researchers (Merrill et al., 2012; Le Féon et al., 2013).

In the research community, which assesses the environmental impacts of agricultural systems, modeling frameworks are increasingly incorporating crop rotations instead of single crops for representing cropping patterns. These modeling approaches are related to nitrate leaching in intensive agriculture (Beaudoin et al., 2005), the impacts of agricultural management on the reduction of nitrogen content (Rode et al., 2009), the impact of farming on water resources (Graveline et al., 2012), etc. The manner of cropping system representation in terms of crop rotations in these studies was often simplified by expert knowledge, which was based on their own specific field observations or on interviews with farmers. A limited number of representative crop rotations were used for describing the cropping patterns in a spatial unit. For allocating these crop rotations within their study area, a crop rotation was usually stochastically assigned to a field, similar to the study by Rode et al. (2009). This simplified approach of representing cropping patterns is due to a lack of information concerning the allocation of crop rotations (Rode et al., 2009). Furthermore, a 'crop generator' was proposed for producing spatial and temporal crop distributions under certain conditions, such as soil types, agronomic rules or expert knowledge, and possibly calibrated with observed data (Dogliotti et al., 2003; Schönhart et al., 2011). A crop generator was included as an additional module in several hydrological models (Wechsung et al., 2000; Klöcking et al., 2003). The limitation of agronomic rule-based crop generators is that these generators create theoretical crop rotations according to the agronomic suitability; however, actual crop rotation practices at the field level are primarily influenced by economic conditions, and biophysical conditions play only a secondary role (Klöcking et al., 2003). Meanwhile, a study regarding uncertainty in the simulation of nitrate leaching at a large regional scale notes that the lack of information about agricultural land use management presents the greatest uncertainty and underlines its importance (Schmidt et al., 2008). All these reviewed modeling approaches represented cropping patterns from the field to regional mesoscale. For representing cropping patterns at a large regional scale or at a global scale, no modeling work has been proposed in the literature. In contrast to the existence of various models at a field scale for designing sustainable cropping systems, the lack of cropping system models at a regional or at a global scale results from the unavailability of spatially and temporally explicit information regarding crop rotations and their associated crop management system (Therond et al., 2011).

The aim of our study is to present an innovative stochastic data mining methodology for describing the spatial distribution of crop sequences at a large regional scale. The data mining is performed using hidden Markov models and an unsupervised clustering analysis that processes sequentially observed (from 1992 to 2003) land-cover data on the French mainland.

Our study can be considered an empirical analysis of historical cropping patterns at a large regional scale, which will contribute to the creation of scenarios of agricultural land-use changes that are caused by changes in land management practices to analyze the performance of land activity-related policies and land planning. This study also provides a tool to extract large-scale spatially explicit cropping pattern data from increasingly available time series data regarding land-cover and land-use, which will improve the accuracy of the assessment of environmental impacts of agricultural systems. In this study, we define 'crop sequences' as the order of appearance of crops during a fixed period. Crop sequences are strictly synonymous with crop successions. Crop sequences are the partial or total development of a cycle of rotation or even the basis of several cycles (Leteinturier et al., 2006). As noted by the field survey-based study, farmers have grown different crops over the years in their farm fields without necessarily designing strict rotations (Joannon et al., 2008). For the study of cropping patterns at a national scale, we have limited our investigation to major crop sequence-related cropping patterns.

We present our modeling approach as follows. First, we describe our study area and the available land-cover data sources. Next, we briefly introduce the temporal data-mining tool. We then apply our modeling approach, which uses this historical national land-cover survey data for clustering the French agricultural districts in terms of the similarity of occurrences of crop sequences. Finally, we further characterize the clusters of agricultural districts using both the typical regional crop sequences and the major crop sequences of a region.

2. Materials and methods

2.1. Study area

Our study area is the French mainland (the island of Corsica is not included) in Western Europe, which covers 552 thousand square kilometers. The percentage of the total land area in mainland France that was considered agricultural was 55.4% in 1992 and 54.2% in 2003 (FAOSTAT, 2011). The area of the main agricultural land use at the beginning and end of our study period is described in Table 1. Because of the variation in environmental and socio-economic conditions across the entire territory, the French agricultural production systems reveal their regularity through the spatial distribution. Fig. 1 describes the spatial distribution of farm typology, which is based on the community typology of agricultural holdings in France in 2000, which was performed by the French Ministry of Agriculture. This EU farm typology is based on economic criteria, such as economic size and type of farming. This typology provides a glimpse of the spatial distribution of farming systems across the French territory. The main cropping zones for cereal and oilseed production are in central, northern and southwestern France. The livestock zone is situated mainly in northwestern and in the Massif Central of France. The mixed cropping and livestock zone is located mainly in southwestern France.

2.2. Data source

The sequential land-cover data that were used in this study were derived from Teruti databases. Teruti is a two-level sampling survey of land-cover, which was conducted by the French Ministry of Agriculture (Ledoux and Thomas, 1992). Fig. 2 illustrates the sampling method that was performed in this survey. At the first sampling level, the entire territory was segmented into 4700 grids, with an area of 12 × 12 km per grid (Fig. 2a). In most regions, four aerial photos among eight at the positions numbered in 1, 2, 3, 4