Computers and Electronics in Agriculture 148 (2018) 299–308

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

A zone-based approach for processing and interpreting variability in multi-temporal yield data sets

Corentin Leroux[a,b,*], Hazaël Jones[b], James Taylor[c], Anthony Clenet[a], Bruno Tisseyre[b]

[a] SMAG, Montpellier, France
[b] ITAP, Montpellier SupAgro, Irstea, Univ Montpellier, Montpellier, France
[c] School of Natural and Environmental Sciences, Newcastle University, Cockle Park Farm, Ulgham, Morpeth NE65 9EG, UK

ARTICLE INFO

Keywords:
Co-occurrence matrices
Historical yield data
Temporal stability
Segmentation
Within-field yield zones

ABSTRACT

The availability of combine yield monitors since the early 1990′s means that long time-series (10+ years) of yield data are now available in many arable production systems. Despite this, yield data and maps are still under-exploited and under-valued by professionals in the agricultural sector. These historical data need to be better considered and analyzed because they are the only audited means by which growers and practitioners can assess the spatio-temporal yield response within a field. When done, time-series of yield maps are mostly processed by classification-based algorithms to generate spatial and temporal yield stability maps or to provide yield or management classes. This work details an alternate segmentation-based methodology to first generate and then characterize contiguous within-field yield zones from historical yield data. It operates on the yield data rather than interpolated yield maps. A seeded region growing algorithm is proposed that enables both the specification of seeds and zone segmentation in a multivariate (multi-temporal yield) attribute space. Novel metrics to assess the yield zoning are proposed that are derived from textural image analysis. The zoning algorithm and metrics were applied to two fields with long time-series (6+ years) of yield data in combinable crops. The two case studies showed that the proposed zone-based approach was effective in delimitating relevant within-field yield zones. The generated zones had differing temporal yield responses between neighbouring zones that were of agronomic significant and interest to the production systems. As this is a first attempt to apply a segmentation algorithm to yield data, areas for future development applications are also proposed.

1. Introduction

Yield monitors mounted on combine harvesters have been available since the early 1990′s. However, yield data still have difficulties in being a decisive component of the decision-making process in precision agriculture studies. In terms of the utility of yield data, multiple issues have been reported by the scientific community. First of all, it is acknowledged that the yield temporal variability is often stronger than the yield spatial variability, which can hinder analyses over short and long-time periods (Blackmore et al., 2003; Bramley and Hamilton, 2004; Eghball and Power, 1995; Lamb et al., 1997). This temporal variability is essentially due to non-stable factors, such as climate patterns or the type of crops being grown each year (Basso et al., 2012). Multiple authors have stated that the number of years of yield data available to conduct yield temporal analyses was critical (Bakhsh et al., 2000; Kitchen et al., 2005) and some have even tried to propose a minimum number of years necessary to obtain reliable results (Ping and Dobermann, 2005). Secondly, it is clear that the spatial yield pattern results from an interaction of management, climate and soil conditions within a cropping season, which means that it is not possible to derive variable-rate application maps directly for a year n by solely relying on yield data in year $n - 1$. On top of that, yield data often come with a large number of defective observations resulting from the pass of the combine harvester inside the fields. Some of these errors are widely reported in the literature, e.g. flow delay, filling and emptying times, abrupt speed changes or unknown cutting width when entering the crops (Arslan and Colvin, 2002; Sudduth and Drummond, 2007). These errors, if not accounted for, can influence agronomical decisions over the fields (Griffin et al., 2008).

However, from a precision agriculture standpoint, these high-resolution yield data are a very valuable source of information that would be aberrant not to consider (Florin et al., 2009). Yield spatial patterns are a valuable piece of information to better characterize the sources of spatial variability across the fields. Growers are interested to know about the mean yield spatial and temporal patterns over their fields so they can make informed and reliable management decisions. It has been

* Corresponding author at: SMAG, Montpellier, France.
E-mail address: cleroux@smag-group.com (C. Leroux).

https://doi.org/10.1016/j.compag.2018.03.029
Received 26 October 2017; Received in revised form 29 November 2017; Accepted 22 March 2018
0168-1699/ © 2018 Elsevier B.V. All rights reserved.

shown that, despite a strong temporal variability, it was often possible to detect consistent yield spatial patterns across years (Kitchen et al., 2005; Taylor et al., 2007). Be aware that some patterns were found consistent even under different crops and varying climate conditions. Furthermore, yield spatial patterns can deliver relevant information with respect to soil characteristics within the field or can help depict the influence of other external factors, such as management practices and weather conditions (Diker et al., 2004). For instance, Taylor et al. (2007) showed that, in specific portions of their field study, crop rotation management in previous years originated variations in yield spatial patterns. Other authors have found that high-yielding areas in dry years could, at the same time, be low-yielding areas in wet years which could give critical information with respect to within-field soil characteristics (Colvin et al., 1997; Sudduth et al., 1997; Taylor et al., 2007). Another strong advantage of these yield datasets is their accessibility. Indeed, in most cases, harvest has to be made which means that these data can be collected yearly once growers have invested in yield monitors.

The delineation of management zones or management units has long been a subject of interest in precision agriculture because it provides growers with a simple representation of their field. Such zones are defined as spatially contiguous areas over which specific management decisions are to be considered. More than often, management zones are found fragmented in space. This originates from a confusion between the concepts of management classes and management zones (Pedroso et al., 2010). In fact, management classes gather all the management zones over which a specific management decision is to be considered. Authors mostly use classification-based techniques, mostly *k*-means clustering and its fuzzy variant, the fuzzy c-means algorithm (Li et al., 2007; Moral et al., 2010) to delineate these management units. Some others have also proposed some post-processing methods to overcome the fragmentation issue (Ping and Dobermann, 2003). However, as non-spatial algorithms, classification-based methods do not seem to be the most relevant approaches to delineate spatially contiguous areas. One solution could be to make use of object-oriented methodologies from the image processing domain, which aim at delineating objects inside an image (Leroux et al., 2017; Pedroso et al., 2010; Roudier et al., 2008).

Despite the availability of yield data, spatio-temporal yield pattern analysis is not widely done, and when done, is typically applied in an *ad-hoc* or qualitative manner. The industry is missing effective and easily implemented approaches for spatio-temporal yield pattern analysis. The major contribution of this work is to propose a new methodology to analyze historical yield data so that growers and agronomic advisors can better understand the spatio-temporal yield variability in their fields. It must be clear that that the objective of this study is only to look information contained within yield data. It is not, as is typically done with management units, an approach to integrate and simplify crop and environmental variables. In the first instance, the method utilizes a novel multi-dimensional segmentation algorithm that can be applied directly to yield data to define within-field yield zones. The method is therefore not dependent on map production or co-location of information from disparate years. To assess the magnitude and the temporal stability of the yield response within the yield zones, novel metrics adapted from co-occurrence matrix and image textural analyses are then introduced. The algorithm and metrics are derived and then applied to two case studies from arable production systems in France and the UK. The applicability of this novel approach is then discussed including the ability to deliver the processing within a simplified framework that is applicable to non-scientific end-users. Finally, the questions and concerns requiring further work are discussed in the last section.

2. Material and methods

2.1. Study sites

The study was conducted on a 20-ha field in England near Amble, Northumberland (WGS84 datum: E: −1.62, N: 55.37) and on a 31-ha field in the north of France near Evreux (WGS84 datum: E: 0.78, N: 48.95). Both fields are cropped in a wheat (*Triticum aestivum*) and canola (*Brassica napus*) rotation and exhibit a relatively strong yield spatial structure. For the English field, wheat yield data were acquired for six years between 2004 and 2015 with a Case combine harvester operating a 10-m cutting front. For the French field, eight years of yield mapping were available spanning the 2003–2011 period. Over the years, the field was mostly harvested with a Claas combine using a 6-m front.

2.2. Pre-processing multi-year yield data

Yield data were first cleaned to remove technical errors commonly reported in the literature, e.g. speed changes, unknown cutting width when entering the crop, filling and emptying times and abnormal yield values among others (Arslan and Colvin, 2002; Sudduth and Drummond, 2007). To compare yield data from multiple years with possible significant temporal variations, yield observations were standardized for each year *m* with a mean of zero and a variance of one (Eq. (1)):

$$\widetilde{Y}_m(i) = \frac{Y_m(i)-\overline{Y_m}}{\sigma_m} \tag{1}$$

where $\widetilde{Y}_m(i)$ is the ith scaled and centered yield observation in year *m*, $Y_m(i)$ is the ith yield observation in year *m*, $\overline{Y_m}$ is the mean yield in year *m* and σ_m is the yield standard deviation in year *m*.

Following a methodology proposed by Blackmore et al. (2003) and Marques da Silva (2006), a grid composed of 20 × 20 m pixels, and whose orientation followed that of the harvested rows, was superimposed on the yield data. For each pixel of the grid, yield values within the pixel were first averaged by year so as to obtain one yield value for each pixel and each year. The objective was to make sure that each year had the same influence in each pixel even if the number of observations falling into each pixel was different from year to year. Empty pixels in specific years due to missing yield observations were given the mean yield value over the years in the same pixel.

2.3. Delineating within-field yield zones

2.3.1. General description of the algorithm

The objective is to delineate within-field yield zones using a time series of yield data. Within-field yield zones were derived from a seeded region growing algorithm (Adams and Bischof, 1994; Mehnert and Jackway, 1997). This procedure, arising from the image processing domain, starts by selecting a set *S* of *k* observations [S_1, S_2, …, S_k], referred to as the seeds, from which zones are grown. Once the seeds have been chosen, the remaining observations inside the field, i.e. the non-seeds, are recursively associated to an existing seed, given similarity measures between observations. As a consequence, this process expands and grows the zones from the selected seeds. The growing algorithm stops when all the observations have been associated to a zone. Such a procedure has already been applied in the precision agriculture domain but solely with regard to one single agronomic variable (Leroux et al., 2017; Pedroso et al., 2010; Zane et al., 2013). Here, the objective is to extend the procedure to a multi-dimensional case for which there is a need to account for several yield data at the same time. Note that the proposed methodology presents some similarities with that of Leroux et al. (2017).

Download English Version:

https://daneshyari.com/en/article/6539582

Download Persian Version:

https://daneshyari.com/article/6539582

Daneshyari.com