Computational Statistics and Data Analysis 72 (2014) 30–44

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Reinforcement learning-based design of sampling policies under cost constraints in Markov random fields: Application to weed map reconstruction

Mathieu Bonneau[a,b], Sabrina Gaba[b], Nathalie Peyrard[a,*], Régis Sabbadin[a]

[a] INRA - UR875, Applied Mathematics and Computer Science Unit, 24 Chemin de Borde Rouge - Auzeville - CS 52627, 31326 Castanet Tolosan cedex, France

[b] INRA - UMR 1347, Agroécologie, 17 rue Sully BP 86510, F-21065 Dijon cedex, France

ARTICLE INFO

Article history:
Received 23 August 2012
Received in revised form 26 September 2013
Accepted 4 October 2013
Available online 19 October 2013

Keywords:
Sampling design
Markov decision process
Dynamic programming
Gibbs sampling
Least-squares linear regression
Weed mapping

ABSTRACT

Weeds are responsible for yield losses in arable fields, whereas the role of weeds in agro-ecosystem food webs and in providing ecological services has been well established. Innovative weed management policies have to be designed to handle this trade-off between production and regulation services. As a consequence, there has been a growing interest in the study of the spatial distribution of weeds in crops, as a prerequisite to management. Such studies are usually based on maps of weed species. The issues involved in building probabilistic models of spatial processes as well as plausible maps of the process on the basis of models and observed data are frequently encountered and important. As important is the question of designing optimal sampling policies that make it possible to build maps of high probability when the model is known. This optimization problem is more complex to solve than the pure reconstruction problem and cannot generally be solved exactly. A generic approach to spatial sampling for optimizing map construction, based on Markov Random Fields (MRF), is provided and applied to the problem of weed sampling for mapping. MRF offer a powerful representation for reasoning on large sets of random variables in interaction. In the field of spatial statistics, the design of sampling policies has been largely studied in the case of continuous variables, using tools from the geostatistics domain. In the MRF case with finite state space variables, some heuristics have been proposed for the design problem but no universally accepted solution exists, particularly when considering adaptive policies as opposed to static ones. The problem of designing an adaptive sampling policy in an MRF can be formalized as an optimization problem. By combining tools from the fields of Artificial Intelligence (AI) and Computational Statistics, an original algorithm is then proposed for approximate resolution. This generic procedure, referred to as Least-Squares Dynamic Programming (LSDP), combines an approximation of the value of a sampling policy based on a linear regression, the construction of a batch of MRF realizations and a backwards induction algorithm. Based on an empirical comparison of the performance of LSDP with existing one-step-look-ahead sampling heuristics and solutions provided by classical AI algorithms, the following conclusions can be derived: (i) a naïve heuristic consisting of sampling sites where marginals are the most uncertain is already an efficient sampling approach; (ii) LSDP outperforms all the classical approaches we have tested; and (iii) LSDP outperforms the naïve heuristic approach in cases where sampling costs are not uniform over the set of variables or where sampling actions are constrained.

* Corresponding author. Tel.: +33 0 5 61 28 54 39; fax: +33 0 5 61 28 53 35.
E-mail addresses: bonneau@toulouse.inra.fr (M. Bonneau), sabrina.gaba@dijon.inra.fr (S. Gaba), nathalie.peyrard@toulouse.inra.fr, peyrard@toulouse.inra.fr (N. Peyrard), sabbadin@toulouse.inra.fr (R. Sabbadin).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.10.002

1. Introduction

The issues involved in building probabilistic models of spatial processes as well as plausible maps of the process on the basis of the models and observed data are frequently encountered and have mobilized several research fields in spatial statistics as well as probabilistic graphical model communities. As important is the question of designing optimal *sampling policies* that make it possible to build maps of high probability when the model is known. This optimization problem is more complex to solve than the pure reconstruction problem and cannot generally be solved exactly. This sampling design problem has been tackled in *Spatial Statistics* (de Gruijter et al., 2006; Müller, 2007) and *Artificial Intelligence* (AI) (Krause and Guestrin, 2009; Krause et al., 2008; Peyrard et al., 2010). It is even more complex in the case of adaptive sampling, where the set of sampled sites is chosen sequentially and observations from previous sampling steps are taken into account to select the next sites to be explored (Thompson and Seber, 1996).

The case of sampling real-valued observations (e.g., temperature or pollution monitoring) has been the most frequently studied, mainly within the geostatistical framework of Gaussian random fields and kriging. Much less attention has been paid to the case of sampling variables with finite state space. However, this problem naturally arises in many studies of biological systems where observations may concern species density classes, disease severity classes, presence/absence values, etc. In particular, the generic approach developed here was motivated by the question of optimal weed sampling in a crop field in order to build density class maps. In this article, we focus on the problem of sampling spatial variables with finite state spaces and propose, similarly to Krause and Guestrin (2009) and Peyrard et al. (2010, 2013), to define the corresponding optimal sampling problem within the framework of *Markov Random Fields* (MRF, Geman and Geman (1984)). MRF are well adapted to model variables with finite state space. For example, they are very popular in image analysis to model image segmentation problems.

A sampling policy can be *static* or *adaptive*. In the first case, the set of sampled sites is chosen prior to the sampling (see Evangelou and Zhu (2012) for a recent study on static sampling of count data). With an adaptive policy, the sampling is divided into successive steps and the next set of sampled sites is chosen according to previous observations. Obviously, adaptive policies are more efficient than static ones, but may not always be applicable. In Krause and Guestrin (2009), the authors considered the sampling problem in a particular case of MRF, defined on polytrees. They looked for static sampling policies, as in Peyrard et al. (2010). The work in Peyrard et al. (2013) was the first proposal of a naïve heuristic solution to design an adaptive sampling policy for the general MRF model. The heuristic was derived from a strong simplification of the model. Here we extend the work of Peyrard et al. (2013) by proposing an algorithm based on simulations of the exact MRF model to design a heuristic policy. This algorithm is inspired by tools from the fields of *Operations Research* (OR) and AI: *Dynamic Programming* (DP) and *Reinforcement Learning* (RL, Sutton and Barto (1998)). RL approaches make it possible to approximately solve sequential decision problems by making use of simulations to learn the process dynamics under different decisions. They can be used *on-line* to construct adaptive policies step-by-step, computing only the current action to apply from the set of past observations, or they can be used *off-line* to compute a complete policy before any observation is actually made. Off-line approaches focus their computational effort prior to policy execution, whereas on-line approaches alternate action computation phases and action execution phases. The approach we propose in this paper is an off-line RL algorithm. In particular it is suitable for the weed sampling problem where costly computations during field sampling are not conceivable.

As we will demonstrate, classical RL algorithms cannot be applied to solve the optimal sampling problem without being adapted. By combining AI tools with statistical tools, we were able to derive the *Least-Squares Dynamic Programming* algorithm (LSDP). LSDP relies on three main premises: (1) the value of a policy is approximated using a least-squares linear regression; (2) simulated trajectories of the sampling process are computed off-line using Gibbs sampling (Geman and Geman, 1984) and stored in a *batch*; (3) the weights of the linear approximation are those that minimize the least-squares error evaluated on the simulated trajectories. We show experimentally that this algorithm is an improvement over classical "one-step-look-ahead" heuristics and classical RL approaches, thus providing a reference algorithm for spatial sampling design in the case of finite state space variables.

This paper begins with a description of the case study that motivated this work: weed sampling in a crop field (Section 2). The MRF formalization of the optimal adaptive spatial sampling problem is then introduced in Section 3. Section 4 is devoted to classical OR/AI solutions for computing an optimal policy or an approximation of the optimal policy, and Section 5 contains a description of the LSDP algorithm. An empirical comparison between one-step-look-ahead approaches, classical OR/AI algorithms and LSDP on toy problems is provided in Section 6, and on the weed sampling problem in Section 7. Some methodological and applied perspectives of this work are discussed in Section 8.

2. Case study: weed sampling in a crop field

Weeds are responsible for yield losses in arable fields (Oerke, 2006) because they compete with crops for resources and can be hosts for parasites and diseases, whereas the role of weeds in agro-ecosystem food webs and in providing ecological services has been well established (Gibson et al., 2006). Therefore, innovative weed management policies have to be designed to handle the trade-off between production and regulation services. As a consequence, there has been a growing interest in the study of the spatial distribution of weeds in crops, in particular, for precision agriculture management strategies that target weed patches within fields.