THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :

Mathieu BONNEAU

le 30 novembre 2012

Titre :

Échantillonnage adaptatif optimal dans les champs de Markov, application à l'échantillonnage d'une espèce adventice.

École doctorale et discipline ou spécialité :

ED MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de recherche :

INRA - unité de Biométrie et Intelligence Artificielle (UR875) / UMR1347 Agroécologie

Directeur(s) de Thèse :

Régis SABBADIN - Directeur de thèse
Nathalie PEYRARD - Co-Directrice de thèse
Sabrina GABA - Co-Directrice de thèse

Jury :

Liliane BEL - Rapporteur
Bruno ZANUTTINI - Rapporteur
Emmanuelle CAM - Président du jury
Vincent BRETAGNOLLE - Examinateur
Pierre DRUILHET - Examinateur
Régis SABBADIN - Directeur de thèse
Nathalie PEYRARD - Co-Directrice de thèse
Sabrina GABA - Co-Directrice de thèse

Thèse préparée aux
Département de Mathématiques Informatique Appliquées
UBIA (UR 875)
24 chemin de Borde Rouge
31 326 Castanet-Tolosan
Département Santé des Plantes et Environnement
UMR1347 Agroécologie,
pôle Écologie des Communautés et Systèmes de Cultures Durables
17 rue Sully
21 065 Dijon

Résumé

Ce travail de thèse se décompose en deux points : (i) l'étude, d'un point de vue théorique, du problème d'échantillonnage adaptatif dans les champs de Markov et (ii) la modélisation du problème d'échantillonnage d'une espèce adventice au sein d'une parcelle cultivée et la conception de stratégies d'échantillonnage adaptatives de cette espèce.

Le problème d'échantillonnage que nous considérons est le suivant. On considère un vecteur aléatoire discret $\mathbf{X} = (X(1), \ldots, X(n))$, dont la distribution est de type champ de Markov. La valeur de ce vecteur aléatoire est inconnue et doit être inférée à l'aide d'observations de certaines composantes de $\mathbf{X}$. L'observation d'une variable ayant un coût, le nombre d'observations est limité par un budget initial fixé. Le problème d'échantillonnage est alors de choisir quelles sont les variables à observer afin d'optimiser la qualité de la reconstruction du vecteur $\mathbf{X}$, en respectant les contraintes de budget. Ce problème d'échantillonnage peut être formulé comme un problème d'optimisation de la valeur d'une politique d'échantillonnage sous une contrainte de respect du budget initial. Pour de grandes valeurs de n, rencontrées pour des problèmes concrets, ce problème d'optimisation ne peut être résolu de manière exacte. La première partie de ce travail de thèse a consisté à adapter la littérature classique de l'apprentissage par renforcement afin de proposer des méthodes de résolution approchée de ce problème d'optimisation. Le problème du choix d'une politique d'échantillonnage adaptative optimale a d'abord été formulé comme un Processus Décisionnel de Markov à horizon fini. Ensuite un algorithme combinant programmation dynamique et apprentissage par renforcement a été proposé. Ce nouvel algorithme, nommé Least Square Dynamic Programming (LSDP), permet de concevoir des stratégies d'échantillonnage adaptatives pour tout type de champ de Markov et tout type de coût d'observation.

Ce type de problème d'échantillonnage se pose dans beaucoup de domaines finalisés. Ce travail de thèse a été motivé par la question de l'échantillonnage d'une espèce adventice au sein d'une parcelle cultivée. Dans ce cas, le problème est la conception de cartes de répartition spatiale de l'espèce. La conception de telles cartes repose sur des observations de la densité de l'adventice au sein de la parcelle. La taille de la parcelle ne permettant pas une observation complète de celle-ci, la question de la localisation des observations est importante afin d'optimiser la qualité de la carte reconstruite. La deuxième partie de ce travail de thèse a consisté à modéliser ce problème d'échantillonnage et proposer de nouvelles stratégies d'échantillonnage adaptatives. La modélisation de la répartition spatiale à l'aide des champs de Markov a tout d'abord été étudiée. Un modèle de coût d'échantillonnage d'une adventice a également été proposé. Finalement, de nouvelles stratégies d'échantillonnage adaptatives ont été proposées à l'aide de l'algorithme LSDP. Ces stratégies ont ensuite été comparées à d'autres stratégies d'échantillonnage plus classiques, ainsi qu'à une heuristique simple.

Mots-clefs : échantillonnage adaptatif, champ de Markov, processus décisionnel de Markov, apprentissage par renforcement, programmation dynamique, batch, adventice, coût d'échantillonnage.

Optimal adaptive sampling in Markov random fields, application to weed sampling

Abstract

This work is divided into two parts: (i) the theoretical study of the problem of adaptive sampling in Markov Random Fields (MRF) and (ii) the modelling of the problem of weed sampling in a crop fields and the design of adaptive sampling strategies for this problem.

The sampling problem we consider can be described as follows. Let $\mathbf{X} = (X(1), \ldots, X(n))$ be a discrete random vector with a MRF joint distribution. The complete value of $\mathbf{X}$ is not known and has to be inferred from observations of some of its variables. We assume that observing a variable has a cost and that the number of observations is limited by a sampling budget. Then, the sampling problem consists in selecting the variables to observe in order to optimize the quality of the inferred vector under the constraint that the number of observations respects the sampling budget. This sampling problem can be formulated as a constrained optimization problem where the quantity to optimize is the value of a sampling strategy, which cost is constrained to respect the sampling budget. We explore the case of adaptive sampling, where variables to observe are chosen sequentially on the basis of previous observations' results. For large values of n, encountered in real world problems, exact resolution is out of reach. My first contribution was to adapt reinforcement learning algorithms in order to propose approximate solution methods for this sampling problem. This contribution was based on an original finite horizon Factored Markov Decision Process (FMDP). Then, I proposed a generic algorithm for computing an approximate solution to any finite horizon (factored) MDP with known model. This algorithm, called Least-Squared Dynamic Programming (LSDP), combines the concepts of dynamic programming and reinforcement learning. It was then adapted to compute sub-optimal adaptive sampling strategies for any type of MRF distributions and observations costs. An experimental evaluation of this algorithm was performed on simulated problems.

The problem of designing spatial sampling strategies arises in many real world applications. In particular, my work was motivated by the question of weeds sampling in a crop field. In this application, the sampling problem consists in selecting the sites to observe in order to reconstruct a map of the weed density repartition as reliable as possible . The field is divided into a regular grid of n sites. The random vector $\mathbf{X}$ represents a map of the spatial repartition of the weed density and each variable $X(i)$, represents the weed density class at site i. Given the large number of sites in a crop field, observing all sites is too time-consuming. So, a sampling budget is defined to represent the time available for the observer for sampling.A part of my work has consisted in applying the theoretical tools developed in the first part to the weed sampling problem. First, I have modelled the weed density spatial repartition in the MRF framework. Second, I have built a cost model adapted to the weed sampling problem. Finally, both models were used together to design adaptive sampling strategies with the LSDP algorithm. Based on real world data, these strategies were compared to a simple heuristic and to static sampling strategies classically used for weed sampling.

Keywords : adaptive sampling, Markov random field, Markov decision process, reinforcement learning, dynamic programming, batch, weed, sampling cost.

Remerciements

Je remercie tout d'abord l'ensemble des membres de mon jury de thèse d'avoir accepté de lire et de juger mon travail.

Je souhaite ensuite remercier mes trois encadrants de thèse : Sabrina GABA, Nathalie PEYRARD et Régis SABBADIN. Un grand merci de m'avoir donné l'opportunité d'effectuer ce travail de thèse dans de si bonnes conditions. Merci de m'avoir accordé votre confiance, merci pour tous vos bons conseils, merci pour votre pédagogie et merci de m'avoir permis d'évoluer autant scientifiquement que personnellement. Vous m'avez mis sur une bonne voie, j'espère maintenant pouvoir la poursuivre de mes propres ailes.

Je suis très heureux d'avoir pu évoluer au sein du laboratoire de Biométrie et d'Intelligence Artificielle de l'INRA de Toulouse. En son sein, j'ai pu avoir accès à un grand confort de travail, pour ceci merci à Abdé, Fabienne, Mickael, Nathalie et Pascale. Naturellement, ces remerciements s'élargissent à l'ensemble du personnel de BIA. Merci pour votre accessibilité et votre écoute. Je mesure toute la chance que j'ai eu de me lever le matin pour rejoindre un cadre de travail si agréable.

Mes allers et retours dans la ville de Dijon, au sein du laboratoire de Biologie et Gestion des Adventices de l'INRA, resteront pour moi une expérience très enrichissante et me laisseront bon nombre d'agréables souvenirs. Merci à tous de votre patience vis-à-vis d'un novice en agronomie et écologie. Merci de m'avoir dévoilé quelque secrets sur le monde des adventices, j'ai maintenant de solides bases : je maîtrise à merveille la prononciation du mot adventice et connais la différence entre écologie et agronomie ! Je tiens à remercier particulièrement Dominique MEUNIER, qui fut de loin le plus patient, pour avoir pris le temps de répondre à mes nombreuses questions et de m'avoir emmené sur le terrain pour voir et comprendre les données autrement qu'à travers un écran d'ordinateur. Merci également à Émilie CADET, Bruno CHAUVEL, Fabrice DESSAINTS et Nicolas MUNIER-JOLAIN pour leurs conseils avisés.

Un grand merci aux membres de l'équipe MAIA du LORIA Nancy pour votre accueil et nos discussions fructueuses. Je remercie particulièrement Alain DUTECH, Bruno SCHERRER et Olivier BUFFET sans qui ma fin de thèse aurait certainement été bien plus compliquée.

J'adresse également mes remerciements à tous les membres de mon comité de thèse, qui m'ont permis à chaque fois d'avancer un peu plus. Merci donc à Denis ALLARD, Jean-Noël AUBERTOT, Alain DUTECH, Florence FORBES et Nicolas MUNIER-JOLAIN.

Merci à tous les collègues et amis que j'ai pu rencontrer à Dijon : Benjamin B., Benjamin G., Coraline, Delphine, Émilie, Jeanne, Rémy, Solène, Sébastien... Il faut bien avouer que sans vous, Dijon ne me laisserait pas un si bon souvenir !

Merci à tous les collègues et amis toulousains : Aurélie, Céline, Damien, Éric A., Éric C., Gautier, Mahuna, Patrick, Ronan et Jimmy. Je suis très heureux d'avoir pu évoluer avec vous et partager de bons moments.

Un grand merci à mon coloc' de bureau Jimmy VANDEL. J'ai eu beaucoup de chance d'avoir partagé ces 3 ans avec toi et je garderai de cette cohabitation de très bons souvenirs. Merci également au grand Matthieu VIGNES dont les interventions quotidiennes dans notre bureau représentaient une bouffée d'oxygène vitale. Merci à Maria pour nos traditionnelles pauses et apéros. Merci à Damien pour les innombrables cigarettes et matchs de rugby et pour ton aide précieuse lors de la rédaction de ce manuscrit. Finir sa thèse et partir loin peut finalement laisser quelques regrets... Vous me manquez déjà !

Merci à tous ceux avec qui j'ai pu me changer les idées, relativiser et garder les pieds sur terre. À tous les amis du rugby : Ecirb, Lolo, Moutous, Sylvie, Ju... Aux potos Jean-Max & TBO, mon Coloc' Yohann et toute la diaspora Angevine.

Enfin, merci à mes parents sans qui tout cela n'aurait pas été possible. Merci pour vos encouragements. Merci de m'avoir laissé choisir ma voie et de m'avoir donné les moyens d'y arriver. Une pensée à mon grand-père Maurice et mon oncle Étienne qui auraient sûrement apprécié de me voir finir cette aventure.

Table des matières

Introduction générale

Définition et tour d'horizon du problème d'échantillonnage spatial

La notion d'échantillonnage apparaît dans de nombreux de domaines d'étude (industrie minière, sondage d'opinion, traitement du signal, écologie...). Échantillonner consiste toujours à acquérir des observations pour l'étude d'un phénomène qui ne peut être observé dans sa totalité. Se poser un problème d'échantillonnage, c'est se demander quelles sont les observations qui doivent être faites afin de permettre une étude du phénomène considéré de bonne qualité. Un premier exemple est la surveillance de la qualité de l'air en France (phénomène étudié). Pour l'étudier, des relevés (observations) d'ozone, de dioxyde d'azote et de particules sont effectués par les associations agréées de surveillance de la qualité de l'air. Sur la base de ces observations, des cartes de prévisions quotidiennes de la qualité de l'air sont ensuite construites et diffusées au public par l'INERIS [1]. Dans ce premier exemple, le problème d'échantillonnage consiste à se demander où effectuer les relevés, afin de construire quotidiennement des cartes de prévision les plus fiables possibles.

Dans ce travail de thèse nous aborderons le problème d'échantillonnage lorsque le phénomène étudié est un phénomène spatial. Formellement, on considère que ce phénomène est décrit par un champ aléatoire $\mathbf{X} = \big(X(s)\big)_{s\in S}$, où $S \subsetneq \mathbb{R}^2$ est un ensemble fini de sites. Les variables $X(s)$ peuvent représenter directement ou non le phénomène étudié. Dans l'exemple précédent, le phénomène est la qualité de l'air qui n'est pas une grandeur physique directement mesurable. La description de ce phénomène passe alors par la connaissance de trois indicateurs (ozone, dioxyde d'azote, particules). Ici chaque variable $X(s)$ est multidimensionnelle, ce qui définit le cas le plus général.

Il se pose un problème d'échantillonnage dès lors que toutes les variables ne peuvent pas être observées. Lorsque celles-ci sont corrélées spatialement, l'observation d'une variable apporte de l'information sur les valeurs possibles d'autres variables. L'étude globale du phénomène est alors envisageable, même si le nombre d'observations est limité. Résoudre un problème d'échantillonnage consiste alors à utiliser au mieux ces corrélations, afin de choisir les lieux des relevés pour l'étude globale du phénomène à partir d'un nombre limité d'observations.

On peut diviser les problèmes d'échantillonnage en deux parties suivant que (i) l'on souhaite étudier le phénomène dans le temps ou (ii) à un instant donné. Dans chacun des cas, l'échantillonnage peut être adaptatif ou statique. En échantillonnage adaptatif les sites à observer sont déterminés de manière séquentielle, en fonction des résultats précédents. Dans le cas statique les sites à observer sont choisis une fois pour toute au début de l'échantillonnage.

L'étude d'un phénomène dans le temps a souvent pour pour objectif la surveillance de ce phénomène. On souhaite par exemple s'assurer que le phénomène n'évolue pas de manière critique, ou contrôler l'effet de certaines actions. Le phénomène étudié est par conséquent temporel. Il est alors possible d'ajouter une dimension temporelle à l'échantillonnage et de se demander non seulement où, mais également quand acquérir des observations. Cependant, certains capteurs acquièrent des observations en continu, le problème d'échantillonnage consiste alors uniquement à savoir où placer ces capteurs. Ainsi, dans [59], il s'agit de déterminer les emplacements optimaux de capteurs de pollution (observations en continu) au sein d'un réseau de distribution d'eau potable, afin

1. Institut National de l'Environnement Industriel et des Risques

de minimiser le temps de détection d'une contamination de l'eau. Le phénomène étudié est ici spatio-temporel : la qualité de l'eau au sein du réseau de distribution d'eau potable. Dans l'exemple introductif, la qualité de l'air est également un phénomène spatio-temporel. L'emplacement des capteurs a été déterminé pour permettre la construction quotidienne de cartes de la qualité de l'air. Pour ceci, des relevés sont effectués chaque jour le matin et l'après midi. Ici il s'agit bien de surveillance : on souhaite surveiller la qualité de l'air, pour pouvoir prendre des décisions en conséquence. Notons que même si les capteurs effectuent des relevés de manière répétée dans le temps, ce n'est pas un cas d'échantillonnage adaptatif. En effet la localisation des relevés et l'heure d'observation ne changent pas au cours du temps, en fonction des observations passées.
Au contraire lorsque le but de l'échantillonnage est d'obtenir une représentation du phénomène à un instant donné, il s'agit de se demander où acquérir des observations afin de posséder une "photographie" la plus précise possible du phénomène. Par exemple dans [33], l'échantillonnage a pour but de déterminer les caractéristiques d'un sol, phénomène qui évolue généralement peu. Dans [7], 1239 lacs ont été échantillonnés afin de déterminer leurs niveaux de pollution (phénomène temporel) à un moment précis. Ainsi, l'étude d'un phénomène à un instant donné ne se limite pas qu'aux phénomènes purement spatiaux, mais également à des phénomènes spatio-temporels. Toutefois, lorsque le phénomène évolue dans le temps et que l'échantillonnage est adaptatif, il faut alors veiller à ce que les observations soient effectuées dans une échelle de temps suffisamment courte, pour que le phénomène n'évolue pas entre les observations successives.
Enfin, la dernière caractéristique d'un problème d'échantillonnage provient de la nature des observations, qui peuvent être continues (*e.g.* concentration) ou discrètes (*e.g.* comptage).
Même si la notion d'échantillonnage revient toujours à déterminer où (et quand) acquérir des observations, les caractéristiques d'un problème d'échantillonnage sont multiples, expliquant la diversité des problèmes rencontrés. Ces différentes caractéristiques peuvent être résumées de la manière suivante :

- **Phénomène étudié** : spatio-temporel / spatial.
- **Étude :** surveillance dans le temps / photographie à un instant donné.
- **Variables descriptives corrélées :** continues / discrètes.
- **Échantillonnage :** statique / adaptatif.

Les outils méthodologiques permettant de résoudre chaque type de problème d'échantillonnage doivent ensuite être choisis en fonction de ses caractéristiques.
Enfin, pour tout type de problème d'échantillonnage, l'observation d'une variable possède un coût intrinsèque. Par exemple dans le cas de la surveillance de la qualité de l'air, les relevés sont effectués à l'aide de capteurs dont les coûts d'achat et de maintenance limitent leur nombre sur le territoire. Le nombre d'observations est ainsi limité par le budget d'échantillonnage. Le coût d'échantillonnage est rarement modélisé et pris en compte dans la détermination des lieux d'échantillonnage, ce qui peut diminuer la qualité de l'échantillonnage et/ou conduire à un échantillonnage non réalisable sur le terrain.

Problème d'échantillonnage considéré

Dans ce travail de thèse nous étudions le problème d'échantillonnage pour l'analyse d'un phénomène à un instant donné. Pour ce type de problème d'échantillonnage, le cas des variables à valeurs réelles a déjà été largement étudié à l'aide des outils classiques de la géostatistique, voir par exemple [42]. Pourtant d'intérêt dans beaucoup de domaines (*e.g.* écologie), le cas des variables discrètes n'a que très peu été exploré. C'est le sujet d'étude de cette thèse. Le problème d'échantillonnage considéré se formule de la manière suivante : "Quelles sont les variables $X(A),\ A \subseteq S$, à observer afin d'obtenir une reconstruction du vecteur $\mathbf{X}$ de bonne qualité, tout en respectant des contraintes de budget d'échantillonnage ? ".

Les corrélations entre les variables sont exprimées à l'aide d'une distribution de probabilité $\mathbb{P}(.)$ sur le vecteur $\mathbf{X}$. Dans ce travail nous utilisons des distributions de type champ de Markov [31], bien adaptées au cas des variables discrètes.

L'échantillonnage statique étant un cas particulier d'échantillonnage adaptatif, nous nous consacrons à ce dernier.

En échantillonnage adaptatif, on ne définit plus un échantillon (*i.e.* ensemble des variables à observer) mais une "politique d'échantillonnage", fonction qui détermine les prochaines variables à observer en fonction des résultats précédents. Le problème devient alors de trouver la politique d'échantillonnage permettant une reconstruction du vecteur $\mathbf{X}$ de bonne qualité, tout en respectant le budget d'échantillonnage.

Ce problème de recherche d'une politique optimale peut être formulé comme un problème d'optimisation discrète sous contraintes. Pour des problèmes réels, le nombre élevé de variables considérées (*i.e.* $card(S)$) empêche une résolution exacte de ce problème d'optimisation.

Le premier défi, méthodologique, de ce travail de thèse est de proposer une méthode générique de résolution approchée (*i.e.* pour toute distribution de champs de Markov $\mathbb{P}(.)$ et tout coût d'échantillonnage) pour la conception de politiques d'échantillonnage.

Contributions méthodologiques

La piste explorée durant ce travail est celle des algorithmes d'Apprentissage par Renforcement (AR, [75]) pour le calcul approché de politiques optimales.

Le problème du calcul d'une politique d'échantillonnage optimale a d'abord été modélisé comme un Processus Décisionnel de Markov (PDM, [67]) à horizon fini, les PDM constituant le cadre naturel d'application des algorithmes d'AR.

Lorsque le nombre de variables est faible (expérimentalement, $card(S) \leq 16$) l'algorithme TD(λ) [76] peut être utilisé. Cet algorithme bénéficie d'une preuve de convergence asymptotique de la politique calculée vers la politique optimale. Lorsque le nombre de variables augmente, d'autres algorithmes doivent être utilisés. La littérature sur l'AR fournit également des algorithmes possédant des preuves de convergence sous des conditions plus contraignantes que dans le cas précédent. Une simple application de ces algorithmes n'est pas suffisante pour le calcul de politiques d'échantillonnage dans notre cas. Tout d'abord, peu d'algorithmes sont dédiés au cas des PDM à horizon fini. De plus, dans certains problèmes concrets, notamment celui que nous considérerons par la suite, le calcul de la politique d'échantillonnage adaptative doit être fait avant son utilisation sur le terrain. Or, la plupart des algorithmes classiques proposent de calculer la politique "en ligne". Dans le cas de l'échantillonnage, cela signifie que lorsque la première variable à observer est déterminée, il est nécessaire de spécifier à l'algorithme la valeur

de l'observation pour qu'il détermine le lieu de la prochaine observation et ainsi de suite. Ceci nécessite l'utilisation, sur le terrain, d'un ordinateur suffisamment puissant et d'accepter la contrainte d'attendre que le calcul s'effectue entre deux observations. Afin de dépasser ces contraintes, nous avons proposé l'algorithme Least-Squares Dynamic Programming (LSDP, [11]), permettant de calculer hors-ligne une politique sous-optimale de tout PDM à horizon fini. Dans ce cas, la politique d'échantillonnage est calculée une fois pour toute avant la période d'échantillonnage. Étant adaptative, cette politique doit toutefois être stockée sur un ordinateur. Mais le calcul de la prochaine variable à observer s'effectue de manière instantanée. Cet algorithme a été validé expérimentalement, mais ne possède pour l'heure aucune preuve de convergence.

Problème d'échantillonnage pour la cartographie d'une espèce adventice

Ce travail de thèse a été motivé par l'étude de la structure spatiale d'une espèce adventice[2] au sein d'une parcelle cultivée. Ce type d'étude est basé sur l'analyse de cartes de répartition spatiale de l'espèce adventice. Ces cartes sont construites à partir d'observations directes de la densité de l'espèce adventice. Généralement de grandes tailles, ces parcelles ne peuvent être explorées de manière exhaustive, compte tenu du temps et du nombre de personnes que cela demanderait (coût d'échantillonnage). Ainsi l'étude de la répartition spatiale d'une espèce adventice peut se réaliser uniquement sur la base de cartes reconstruites à partir d'un échantillonnage spatial. Le choix des lieux où acquérir des observations impacte directement la qualité de la reconstruction, c'est à dire la ressemblance de la carte reconstruite à la carte réelle. Le problème d'échantillonnage qui en découle revient à choisir les parties de la parcelle à observer afin de reconstruire une carte de bonne qualité, tout en respectant les contraintes de budget d'échantillonnage.
Malgré le caractère générique de l'algorithme LSDP, son utilisation pour la conception de politiques d'échantillonnage adaptatives d'une espèce adventice, comme pour toute autre application, nécessite un travail important de modélisation. Afin de proposer des politiques d'échantillonnage applicables sur le terrain, les notions de coût d'observation et de budget d'échantillonnage doivent être définies. Un modèle de corrélations entre les variables doit également être défini convenablement, sans quoi aucune politique d'échantillonnage ne pourra fournir de résultats satisfaisants.
Ainsi, le deuxième défi, en modélisation, de ce travail de thèse est de proposer à la fois un modèle de coût d'échantillonnage et une famille de modèles de répartition spatiale d'une espèce adventice.

Contributions en modélisation

Tout d'abord le coût de l'échantillonnage d'une espèce adventice a été défini et modélisé en termes de temps nécessaire à l'application d'une stratégie d'échantillonnage.
La deuxième contribution porte sur la modélisation des corrélations spatiales de la densité d'une espèce adventice. Pour ceci plusieurs distributions de champs de Markov ont été proposées. Pour chaque espèce étudiée, une "meilleure distribution" a été calculée à l'aide d'une approximation du score BIC [38]. Ces distributions et le modèle de coût ont ensuite été utilisés pour la conception de politiques d'échantillonnage à l'aide de

2. Une adventice, communément appelée mauvaise herbe, est une espèce végétale autre que l'espèce cultivée au sein de la parcelle.

l'algorithme LSDP. Les politiques produites ont été comparées à des politiques d'échantillonnage classiques, ainsi qu'à une heuristique simple du problème d'optimisation de la valeur d'une politique d'échantillonnage [66].
Enfin, dans un cas réel, aucune donnée n'est disponible a priori pour sélectionner une meilleure distribution et estimer ses paramètres. De plus, lors du choix d'une "meilleure distribution" pour différentes espèces adventices, nous avons constaté qu'il n'existait pas une distribution universelle, meilleure en terme de score BIC, pour chacune des espèces étudiées. Néanmoins, nous avons montré que la conclusion est différente lorsque la pertinence des différentes distributions est discutée dans un objectif de reconstruction de carte à partir d'un nombre d'observations limité. Dans ce cas, une des distributions proposée (le *modèle de Potts avec champ externe*) semble être adaptée pour la plupart des espèces adventices considérées. Enfin, nous avons proposé un ensemble de règles expertes pour en estimer les paramètres, sans observation initiale. Ces règles ont ensuite été testées pour la conception de politiques d'échantillonnage et la reconstruction de cartes, fournissant des résultats comparables au cas où une meilleure distribution et la valeur de ses paramètres sont appris sur des données.

Organisation générale du manuscrit

Ce manuscrit de thèse est construit en deux parties, regroupant respectivement les contributions méthodologiques et en modélisation.
Les parties sont indépendantes du point de vue des notations. La première partie est rédigée en anglais sous forme d'un article qui a été soumis à la revue *Computational Statistics and Data Analysis*, sous le titre *Simulation-based design of sampling strategies under cost constraints in Markov Random Fields.*
La deuxième partie de ce manuscrit sera utilisée pour la soumission d'un article à la revue *Ecological Modelling*, sous le titre *Model-based Adaptive Sampling Strategies for Weed Mapping in Crop Fields.* Cette partie est également rédigée en anglais.
Enfin, une conclusion générale et une description de quelques perspectives de ce travail complètent ce manuscrit.

Première partie

Conception de stratégies d'échantillonnage optimales

Sommaire

1 INTRODUCTION

The question of building probabilistic models of spatial processes and building plausible reconstructions from the model and observed data is classic and has mobilised several research fields in spatial statistics or probabilistic graphical models communities. Nearly as classical is the question of designing optimal *sampling policies* allowing to build reconstructions of high probability. This question is more complex to solve than the pure reconstruction problem and cannot be solved optimally in general. This sampling design problem has been tackled in spatial statistics [26], [55] and artificial intelligence [44], [43], [65]. 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 explore [79].

Among spatial models, the case of real-valued observations (e.g. temperature or pollution monitoring) has been the most studied, mainly within the geostatistical framework of Gaussian random fields and kriging. Much less attention has been paid to the case of discrete-valued observations. However, this problem arises naturally in many studies about biological systems, where discrete-valued observations can be species abundance classes, disease severity classes, presence/absence values...

Since optimal solutions to sampling problems in discrete-valued random fields are out of reach, one should look for approximate solution algorithms with reasonable/moderate complexity and with satisfying approximation quality. In this article, we propose, similarly to [43], [65], [66], to define the optimal sampling problem within the framework of Markov random fields (MRF, [31]), classically used in image analysis. In [43], the authors considered the sampling problem in a particular subset of MRF, defined on polytrees. They looked for static policies, as in [65]. The work in [66] was the first proposition of a naïve heuristic solution to design an adaptive sampling policy in any MRF model, derived from a strong simplification of the MRF model. We propose here to extend this work by using a heuristic built from simulations of the exact model. For this, we first encode the optimal adaptive sampling problem as a finite-horizon Markov Decision Process (MDP, [67]) with factored state space. Casting the optimal sampling problem within the MDP framework allows to exploit principles from the family of *simulation-based* or *Reinforcement-Learning* (RL, [75]) approaches which have been proposed to solve approximately large state space MDPs. As we will demonstrate, classical RL algorithms cannot be applied to solve the optimal sampling problem without being adapted. Therefore we provide a new generic RL algorithm that can be used to solve approximately any large state-space MDPs in the specific case where the horizon is finite : the *Least Square Dynamic Programming* algorithm (LSDP). LSDP is a model-based method, as opposed to model-free methods where an explicit form of the transition probability function is not required. In LSDP the value function is parameterised and simulated trajectories are computed off-line and stored in a *batch*. The batch of trajectories is used to build systems of linear equations which are solved in order to compute approximate policies. We then show how to specialize this generic algorithm to the sampling problem in MRFs. We show experimentally that this algorithm improves classical "one-step-look-ahead" heuristics and RL approaches, thus providing a reference algorithm for sampling design.

The MRF formalization of the optimal adaptive spatial sampling problem is introduced in Section 2, together with a computational complexity study. We show how to

model it as a finite-horizon factored MDP in Section 3 and we discuss classical RL solutions in Section 4. Then, we describe the LSDP algorithm in Section 5 and its application to the problem of sampling in MRF in Section 6. We present an empirical comparison between one-step-look-ahead approaches, classical RL algorithms and LSDP in Section 7. Some methodological and applied perspectives of this work are discussed in Section 8.

2 OPTIMAL ADAPTIVE SAMPLING IN MARKOV RANDOM FIELDS

In this section we formalise the problem of optimal adaptive sampling in a Markov random field and we establish its computational complexity.

2.1 Problem statement

Let $X = (X_1, \ldots, X_n)$ be a vector of discrete random variables taking values in $\Omega^n = \{1, \ldots, K\}^n$. $V = \{1, \ldots, n\}$ is the set of indices of the vector X and an element $i \in V$ will be called a *site*. The distribution $\mathbb{P}$ of X is that of a Markov Random Field (MRF, also named *undirected graphical model*) with associated graph $G = (V, E)$, where $E \subseteq V^2$ is a set of undirected edges. The vector $x = (x_1, \ldots, x_n)$ is a realisation of X and we adopt the following notation : $x_B = \{x_i\}_{i \in B}$, $\forall B \subseteq V$. Then the joint distribution of X is a Gibbs distribution : $\mathbb{P}(X = x) \propto \prod_{c \in \mathcal{C}} \Psi_c(x_c)$, where $\mathcal{C}$ is the set of cliques of V and the $\Psi_c, c \in \mathcal{C}$, are strictly positive potential functions [41].

In order to reconstruct the vector X on a specified subset $R \subseteq V$ of sites of interest, we can acquire a limited number of observations on a subset $O \subseteq V$ of observable sites. We will assume that $R \cup O = V$ and intersection between O and R can be non-empty. The sampling problem is to select a set of sites $A \subseteq O$, named a *sample*, where X will be observed. When sample A is chosen, a *sample output* x_A results, from which the MRF distribution $\mathbb{P}$ is updated. Our objective is, intuitively, to choose A so that the updated distribution $\mathbb{P}(\cdot|x_A)$ becomes as informative as possible (in expectation over all possible sample outputs).

In the following we describe the different elements allowing to formally define the problem of Optimal Adaptive Sampling in a MRF (AOSMRF).

Reconstruction. When a sample output x_A is available, the *Maximum Posterior Marginal* (MPM) criterion is used to derive an estimator x_R^* of the hidden map x_R :

$$x_R^* = \left\{ x_i^* \quad | \quad i \in R, \quad x_i^* = \underset{x_i \in \Omega}{\operatorname{argmax}} \mathbb{P}(x_i \mid x_A) \right\}.$$

Alternately, other reconstruction criteria, such as the *Maximum A Posteriori* (MAP) criterion [65] can also be considered.

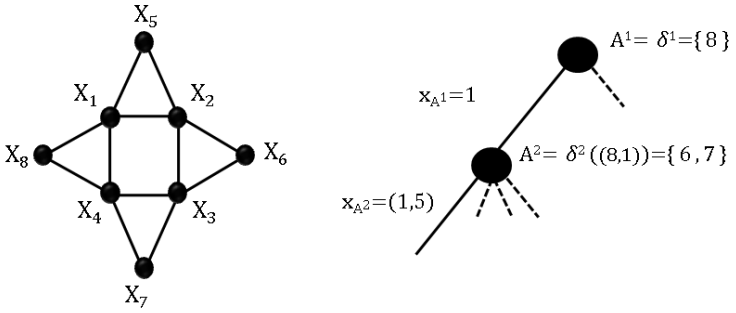


FIGURE 1 – Tree representation of an adaptive sampling policy. Left : graph G associated to the MRF model $(n = 8, K = 5)$; Right : exemple of adaptive sampling policy for $L = 2$, $H = 2$, and $V = R = O$, nodes represent samples and edges represent all possible sample outputs.

Adaptive sampling policy. In adaptive sampling, the sample A is chosen sequentially. The *sampling plan* (the sequence of samples) is divided into H steps. $A^h \subseteq O$ is the sample explored at step $h \in \{1, \ldots, H\}$ and x_{A^h} is the sample output at step h. The sample size is bounded ($|A^h| \leq L$) and Δ_L is the set of all policies satisfying $|A^h| \leq L, \forall h$. The choice of sample A^h depends on the previous samples and outputs. An adaptive sampling policy $\delta = (\delta^1, \ldots, \delta^H)$ is then defined by an initial sample A^1 and functions δ^h specifying the sample chosen at step $h \geq 2$, depending on the results of the previous steps : $\delta^h((A^1, x_{A^1}), \ldots, (A^{h-1}, x_{A^{h-1}})) = A^h$ (see Figure 1 for a graphical representation of an adaptive sampling policy). A *history* is a trajectory $(A^1, x_{A^1}), \ldots, (A^H, x_{A^H})$ followed when applying policy δ. The set of all histories which can be followed by policy δ is τ_δ. We will assume throughout the paper that observations are reliable. As a consequence, we will only consider policies visiting each site at most once ($A^h \cap A^{h'} = \emptyset, \forall h \neq h'$).

Quality of a sampling policy. The quality of a policy δ is measured as the expected quality of the estimator x_R^* that can be obtained from δ. In practice, we first define the quality U of a history $((A_h, x_{A_h}))_{h=1\ldots H}$ as a function of (A, x_A), where $A = \cup_h A_h$:

$$U\big(A, x_A\big) = \sum_{i \in R} \max_{x_i \in \Omega} \Big\{ \mathbb{P}(x_i \mid x_A) \Big\}. \tag{1}$$

The quality of a sampling policy δ is then defined as an expectation over all possible histories :

$$V(\delta) = \sum_{((A_h, x_{A_h}))_h \in \tau_\delta} \mathbb{P}\big(x_A\big) U\big(A, x_A\big).$$

Remark that this definition of the quality of a sampling policy can be adapted to the MAP criterion. In section 5 we will show that our approach can also be adapted to an *entropy* definition of map uncertainty.

Optimal adaptive sampling in MRF (OASMRF). Finally the problem of optimal adaptive sampling amounts to finding the policy of highest quality :

$$\delta^* = \operatorname*{argmax}_{\delta \in \Delta_L} V(\delta). \tag{2}$$

Note that since our definition of the quality of a policy is based on the MPM criterion [3], it does not depend on the order in which observations are received. Therefore, the pair (A, x_A), where $A = \cup_h A^h$, contains exactly the relevant information in a history :

$$A^h = \delta^h\Big((A^1, x_{A^1}), \ldots, (A^{h-1}, x_{A^{h-1}})\Big) = \delta^h\Big(\bigcup_{i<h} A_i, x_{\bigcup_{i<h} A_i}\Big).$$

However this does not imply that the above defined sampling problem is not adaptive. Indeed, sample choices and observations are interleaved, meaning that the optimal policy δ^* is a function, and not a plan (i.e. a constant choice, independent of the successive observations).

2.2 How to represent and handle cost constraints ?

So far, we have considered a sampling budget in terms of a number, H, of allowed sampling steps and a fixed number L of sampled variables per step. This has been defined regardless of any notion of observation costs. In this section, we will discuss more general sampling costs models and discuss how they can be handled within the proposed OASMRF model.

Optimising a trade-off between restoration quality and cost. In order to optimise a global trade-off between restoration quality and cost it is possible to include a measure of sampling cost, which has to give values commensurate with the restoration quality. Then, cost and quality measures can be added in the definition of U, to form the sampling policy quality.
However, one may question the assumption that cost and restoration quality measures be commensurate. One way to avoid this assumption is to consider a sample budget constraint instead of including sample costs into the function U, as we suggest next.

Maximising restoration quality under cost constraint. Instead of considering that $L \times H$ sites can be sampled, we can consider a global (integer-valued) sampling budget B (e.g. a time budget), and assume that each subset $A \subseteq O$ has a different (integer) sampling cost, denoted $S_C(A)$. It may as well be that sample costs depend on the observed values of the variables, x_A : $S_C(A, x_A)$. This cost function S_C is not used in the definition of U, but rather for the definition of the space of allowed trajectories, $\{(A^1, x_{A^1}), (A^2, x_{A^2}) \ldots\}$. In that case the length of a trajectory is no more constant. It depends on the number of samples needed to exhaust the budget and of the corresponding observations. We will see in Section 5 that the simulation-based procedure proposed in this paper can handle this case.

Minimising cost under restoration quality constraint. Conversely, one could also consider that the objective is to minimise the sampling budget, and that we are given a MPM restoration quality threshold. Sampling should be continued until the restoration quality threshold is met, and then stopped. Then, the optimisation problem would be to find the sampling strategy of minimum expected cost, which allows to compute

3. This is also true for MAP and entropy-based criteria.

restored maps which quality is above the fixed threshold. One difficulty to apply the simulation-based procedure proposed in this paper would be that the MPM value has to be computed at every sampling step, in order to check whether the end of a trajectory is reached.

2.3 Computational complexity of optimal adaptive sampling in MRF

In this section we study the computational complexity of the OASMRF problem. More precisely, we study the following, *generalised* OASMRF problem (GOASMRF), expressed in a decision form [4] :
Does there exist δ of depth at most N, such that :

$$\sum_{((A_h, x_{A_h}))_h \in \tau_\delta} \mathbb{P}(x_A) U(A, x_A) \geq G \ ?$$

Here G is a fixed positive threshold, and the definition of quality is more general : $U(A, x_A) = \sum_{i \in R} f_i\big(x_i^*, \mathbb{P}(x_i^* \mid x_A)\big)$, where the functions f_i are non-decreasing functions in their second argument and $x_i^* = \arg\max_{x_i} \mathbb{P}(x_i \mid x_A)$.

This form of quality of a history generalises (1), which is recovered when f_i is a projection on its second argument. The extended form can represent criteria consisting in maximising a weighted expected number of well-restored variables (when some variables are more important than others), or the expected number of variables restored with confidence above a given threshold. The fact that x_i^* is involved and not only its probability, allows to bias restoration towards particular values of x_i. This can be useful, for instance, if we want to build an invasive species map, where we give more weight to restoring invaded sites than non-invaded ones. Finally, the fact that f_i is non-decreasing is not essential for proving the proposition, but reflects the fact that the more certain we are about x_i^*, the better.

Proposition 1 *The GOASMRF problem is Pspace-complete.*

Proof

There is not much difficulty in proving that GOASMRF belongs to *Pspace*. The difficult part is to establish the *Pspace*-hardness of the GOASMRF problem. To prove this, we reduce the *State Disambiguation* problem, which is known to be *Pspace*-hard [22] to it. A detailed proof is given in [49].

The consequence of Proposition 1 is that exact optimisation is intractable. In the next section we present a factored Markov Decision Process (MDP) model of the OASMRF problem [5]. It will allow us to solve OASMRF problems approximately by applying Reinforcement Learning (RL) principles [75].

4. See [60] for a detailed description of computational complexity theory and the decision form of optimisation problems.
5. Which can be easily extended to GOASMRF.

3 FINITE HORIZON MDP MODELLING OF THE OASMRF PROBLEM

A finite-horizon Markov Decision Process model [67] is a 5-tuple $< S, D, T, p, r >$, where S is a finite set of system *states*, D is a finite set of available *decisions*, $T = \{1, \ldots, H\}$ is a finite set of decision steps, termed *horizon*. p is a set of *transition functions* $p^t, t = 1 \ldots H$, where $p^t(s^{t+1}|s^t, d^t)$ indicates the probability that state $s^{t+1} \in S$ results when the system is in state $s^t \in S$ and decision $d^t \in D$ is implemented at time $t \in \{1, \ldots, H\}$. A *terminal state* $s^{H+1} \in S$ results when the last decision is applied, at decision step H. r is a set of *reward functions* : $r^t(s^t, d^t) \in \mathbb{R}$ is obtained when the system is in state s^t at time t and d^t is applied. A *terminal reward* $r^{H+1}(s^{H+1})$ is obtained when state s^{H+1} is reached at time $H + 1$.

A *decision policy* (or *policy*, for short), $\pi = \{\pi^1, \ldots, \pi^H\}$, is a set of decision functions $\pi^t : S \to D$. Once a decision policy is fixed, the MDP dynamic becomes that of a finite Markov chain over S, with transition probability $p^t(s^{t+1}|s^t, \pi^t(s^t))$. The *value function* $V^\pi : S \times T \to \mathbb{R}$ of a policy π is defined as the expectation of the sum of future rewards, obtained from the current state and time step when following the Markov chain defined by π :

$$V^\pi(s,t) = \mathbb{E}_\pi \left[\sum_{t'=t}^{H} r^{t'}(s^{t'}, \pi(s^{t'})) + r^{H+1}(s^{H+1}) \mid s^t = s \right], \forall (s,t) \in S \times T.$$

Solving a MDP amounts to finding an *optimal policy* π^* which value is maximal for all states and decision steps : $V^{\pi^*}(s,t) \geq V^\pi(s,t), \forall \pi, s, t$ (it can be proved that there always exists at least one optimal policy, see [67]).

We now model the OASMRF problem in the MDP framework. It corresponds to the graphical representation of Figure 2.

State space. state $s^t, t = 1, \ldots, H+1$ summarizes current information about variables indexed in O :

$$s^t = \left(\bigcup_{h=1}^{t-1} A^h, \bigcup_{h=1}^{t-1} x_{A^h} \right), \forall t = 2, \ldots, H+1 \text{ and } s^1 = (\emptyset, \emptyset).$$

It may be convenient to model s^t as a vector of length $|O|$, where $s_i^t = -1$ if site i has not yet been sampled, and $s_i^t = k$, $k \in \Omega$ if value k has been observed on site i.

Decision space. A decision d^t is a sample $A^t \subseteq O$ such that $|A^t| \leq L$. In practice, it will never be optimal to sample a site twice (since observations are assumed to be reliable). So, we can restrict the set of decisions to those satisfying $d^t \cap d^{t'} = \emptyset, \forall t' < t$.

Horizon. Decision steps in the MDP correspond to decision steps in the OASMRF problem. Thus, $T = \{1, \ldots, H\}$.

Transition functions. If $s^t = (A, x_A)$ and $d^t = A^t$, the transition function of the MDP can be derived straightforwardly from the original MRF distribution $\mathbb{P}$:

$$p^t\big(s^{t+1} \mid s^t, d^t\big) = \mathbb{P}\big(x_{A^t} \mid x_A\big), \forall t \in T,$$

where x_{A^t} is the realisation of X_{A^t} encoded in s^{t+1}. Note that for all states s^{t+1} corresponding to observations not compatible with state s^t, this transition probability will be zero.

Reward functions. $\forall t = 1, \ldots, H$, rewards are set to zero :

$$r^t(s^t, d^t) = r^t(d^t) = 0, \quad \forall t \in T, s^t, d^t.$$

Note that rewards could be non zero if the objective was to optimise a trade-off between restoration quality and sampling costs (see Section 2.2).

After decision d^H has been applied at decision step H, and state $s^{H+1} = (A, x_A)$ has been reached, the final reward $r^{H+1}(s^{H+1})$, defined as the quality of the MPM reconstruction, is obtained :

$$r^{H+1}(s^{H+1}) = \sum_{i \in R} \left[\max_{x_i \in \Omega} \Big\{ \mathbb{P}(x_i \mid x_A) \Big\} \right].$$

The optimal policy for the above-defined MDP is a set of functions associating samples to unions of past samples outputs. It thus has the same structure as an OASMRF sampling policy. Furthermore, we can establish the following proposition :

Proposition 2 *An optimal policy for the MDP model of an OASMRF problem provides an optimal policy for the initial OASMRF problem (2).*

Proof (Sketched). The proof is only sketched here, the full version is in the Appendix section. The proof follows three steps and uses the fact that the quality of a policy does not depend on the order in which observations are obtained :

(i) We define a function ϕ, transforming any MDP policy π into a valid OASMRF policy $\delta = \phi(\pi)$, which defines decisions independently of the order in which past observations were received, and show that $V(\phi(\pi)) = V^\pi((\emptyset, \emptyset), 1)$.

(ii) We establish that, for any partial history (past observations), the value of an optimal OASMRF policy starting from these observations does not depend on the order in which they were received. As a consequence, we can limit the search for optimal policies of the OASMRF problem to policies prescribing decisions which do not depend on the order of observations.

(iii) We show that any such OASMRF policy δ can be transformed into a MDP policy, through a transformation μ, and that $V(\delta) = V^{\mu(\delta)}((\emptyset, \emptyset), 1)$.

As a result of these three steps, if π^* is an optimal policy for the MDP encoding of the OASMRF problem, then $\phi(\pi^*)$ is optimal for the OASMRF problem.

□

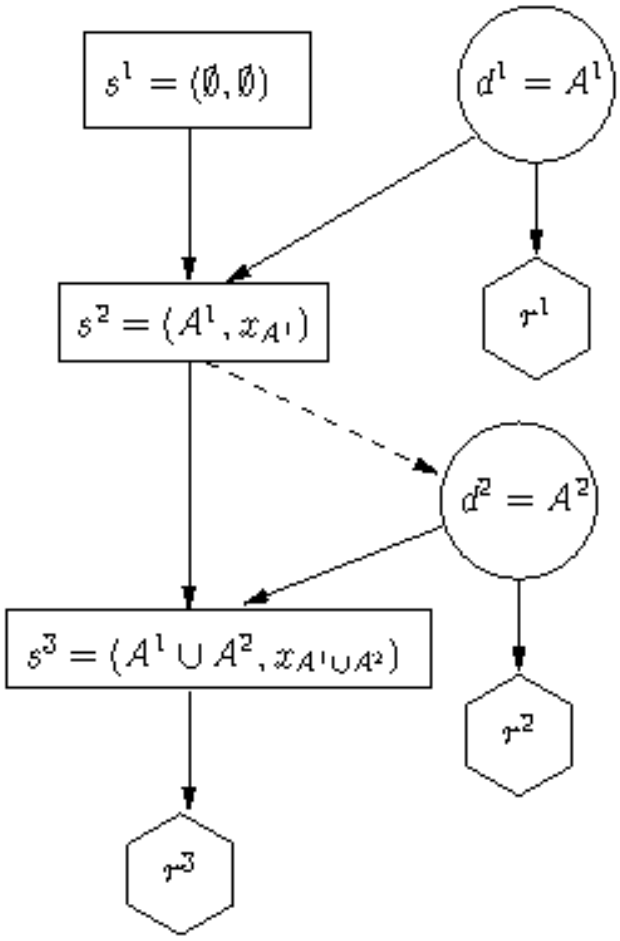


FIGURE 2 – MDP model of a OASMRF problem with horizon $H = 2$.

In the following we will use the same notation δ to represent both OASMRF and MDP policies.

Note that the finite-horizon MDP model of the OASMRF problem has state and action spaces of size exponential in the size of the original problem. This exponential space representation cannot be avoided since MDP are known to be solvable in polynomial time in their representation, while problems of optimal sampling in MRF are *Pspace*-complete (see Section 2). Even though the MDP representation of the OASMRF problem takes exponential space, explicit representation of the problem (and its solution policy) can be avoided, thanks to the use of RL algorithms. We describe the RL approach in the next section.

4 CANDIDATE APPROACHES FOR SOLVING OASMRF

4.1 Exact dynamic programming

Let us define the state-action value function, also called Q-function associated to any finite-horizon MDP problem :

$$Q^{\delta}(s,d,t) = r^t(s,d) + \sum_{s' \in Succ^t(s,d)} p^t(s'|s,d)\mathbb{E}_{\delta}\left[\sum_{t'=t+1}^{H} r^{t'}(s^{t'},\delta(s^{t'})) + r^{H+1}(s^{H+1}) \mid s^{t+1}=s'\right].$$

This function represents the expected reward when applying decision d in state s and thereafter following policy δ. $Succ^t(s,d) = \{s', p^t(s'|s,d) > 0\}$ is the set of possible successors of s when d is applied at time t. In the OASMRF problem the size of $Succ^t(s,d)$ can be small. From the Q-function of a policy, it is straightforward to compute the value of that policy. The *backwards induction* algorithm [67] is based on this property and computes exactly the optimal policy of any finite-horizon MDP. It consists in initialising the value function of the optimal policy at time $H+1$:

$$V^*(s,H+1) = r^{H+1}(s),$$

and then solving iteratively the following equations :

$$\begin{aligned} \forall t = H,\dots,1 \quad \text{and} \quad & \forall s,d \in S \times D, \\ Q^*(s,d,t) &= r^t(s,d) + \sum_{s' \in Succ^t(s,d)} p^t(s'|s,d)V^*(s',t+1), \\ V^*(s,t) &= \max_d Q^*(s,d,t). \end{aligned}$$

At each iteration the optimal policy is progressively build, as follows :

$$\delta^{*,t}(s) = \arg\max_d Q^*(s,d,t)$$

However, since the OASMRF problem typically involve a factored state space, $|S|$ is huge, so the above system has too many equations and variables ($|S| \times |D| \times H$) to be used. Therefore, we have to look for approximate solution methods instead of exact ones. To do this, we can explore two families of approaches for solving OASMRF : *one-step-look-ahead approaches* and *simulation-based* approaches.

4.2 Heuristic approaches

Heuristic approaches are methods for sample selection which provide an arbitrary (most likely suboptimal) sample in reasonable time. These methods can either (i) solve exactly a simpler optimization problem which approximates the original one or (ii) provide policies maximising a function which approximates the optimal Q-function. This approximate Q-function can be, in particular, parameterized and computed through simulations.

One-step-look-ahead heuristics. One-step-look-ahead heuristics provide policies that optimise (exactly or approximatively) the immediate reward. Such heuristics have been proposed, either in Statistics or in Artificial Intelligence, that can be applied to solve the OASMRF problem. In spatial sampling of natural resources, random and regular sampling are classic heuristic approaches [26]. Another classical method to sample $0/1$ variables is Adaptive Cluster Sampling (ACS, [79]). Recently, [66] proposed a heuristic (*BP-max heuristic*), which consists in sampling locations where the marginal probabilities are the least informative (i.e. the closest to $\frac{1}{2}$ in the $0/1$ case), in order to solve (2). It has been shown, experimentally, to outperform random, regular and ACS heuristics. In [44], the authors proposed to optimise a mutual information (MI) criterion to design sampling strategies in Gaussian fields.

Simulation based approaches : reinforcement learning. The main idea of Reinforcement Learning approaches (RL, [75]) is to use repeated simulated *experiences* (s^t, d^t, r^t, s^{t+1}), instead of exact dynamic programming, in order to estimate Q^* or to compute a parameterized approximation of Q^* [75]. They can either estimate Q^* directly (*Q-learning* approach, for example), or interleave estimation steps of a current policy δ ($TD(\lambda)$ [75] can be used) with improvement steps, in a general *policy iteration* scheme.

In general when simulation is used to solve large factored MDP such as in the OASMRF problem, functions Q^δ are too expensive to store in tabular form. A parametric approximation of the optimal Q-function is built as : $Q^w(s,d,t) =< w, \phi(s,d,t) >$, where $w \in \mathbb{R}^m$ is a vector of parameters values and $\phi : (S,D,T) \rightarrow \mathbb{R}^m$ is a mapping from state-action pairs to real-valued m-dimensional vectors (also called features). Simulations are used to compute values of w that give a good approximation of Q^*. In general little can be said about the convergence of such algorithms and no universal properties are given. However in some case performance bounds (e.g. [56],[3]) or convergence guaranties ([77],[50]) can be found. Algorithms for computing w for a specific choice of features are, for example, *LSPI* [46] and *Fitted Q-iteration* ([29],[58]).

5 LEAST-SQUARES DYNAMIC PROGRAMMING (LSDP)

We now present the procedure we proposed to compute an approximate solution to the original OASMRF problem. This procedure is not limited to this problem. It can be used to solve any finite horizon MDP as long as the model (transitions and rewards) is known explicitely. So we first describe it in its general form and then show how it can

be specialised for the OASMRF problem.

5.1 Approximate dynamic programming

The main idea of the algorithm we propose is to combine a parametrized representation of the Q-function with *dynamic programming* (DP) iterations and simulations in order to approximate Q^*. Namely, we consider an approximation Q^w of Q^* as a linear combination of m arbitrary *features* :

$$\begin{aligned} Q^w(s,d,t) &= \sum_{i=1..m} w_i^t \phi_i(s,d,t), \forall s,d,\forall t \in T \text{ and} \\ Q^w(s,H+1) &= r^{H+1}(s), \forall s. \end{aligned}$$

Note that linear approximations of the Q-function have already been proposed for the case of infinite horizon MDP. However, these are less classical in the finite horizon case. The main difference is that when the horizon is finite, the optimal policy needs not be stationnary. Therefore, we define a set of weights $\{w_i^t\}_i$ for every time steps. These weights are computed recursively for $t = H$ to 1, in such a way that equations (3) are approximately satisfied :

$$\begin{aligned} &\forall (s,d,t) \in S \times D \times T, \\ &\sum_{i=1..m} w_i^t \phi_i(s,d,t) = r^t(s,d) + \sum_{s' \in Succ^t(s,d)} p^t(s'|s,d) V^w(s',t+1) \\ &\text{where } V^w(s,t) = \max_d \sum_{i=1..m} w_i^t \phi_i(s,d,t). \end{aligned} \tag{3}$$

Equations (3) form a set of $|S| \times |D|$ linear equations for each time step $t \in T$, with variables $w_i^t, i = 1..m$. In general these systems are clearly over-constrained ($|S| \times |D| \gg m$), therefore we look for *least-squares* approximate solutions, instead of exact ones. The dynamic programming part of the approach comes from the fact that the systems are solved backwards for $t = H$ to 1, each solution vector w^{t+1} being plugged into the system obtained at time t. Then, from the set of weights we can derive the Q-function and thus an approximation of the optimal policy and its value.

In the case where the resources constraints are not defined by a fixed number of sampling steps but by a maximal budget B, the same principle can still be applied. We simply define Q-functions and features as functions of b, the budget used so far, instead of functions of decision steps t performed so far. As a consequence, the sets of weights and features are also indexed by the budget already spent. A trajectory is stopped if all actions have a cost higher than the remaining budget. Since it is more general to define constraints in terms of budget, we will consider this representation in the following. Property 2 still holds with this definition.

5.2 LSDP Algorithm

Systems (3) are too large to build when S is factored, not to mention solving. Therefore, we suggest to *sample* this system, by considering only a subset of equations, corresponding to a subset of triplets $\mathcal{B} = \{(s,d,b)\} \subseteq S \times D \times \{1,\ldots,B\}$, (called *batch* [68]).

The system (3) becomes :

$$\begin{aligned} & \forall (s,d,b) \in \mathcal{B}, \\ \sum_{i=1..m} w_i^b \phi_i(s,d,b) &= r^b(s,d) + \sum_{(s',b') \in Succ^b(s,d)} p^b(s'|s,d) V^w(s',b') \\ \text{where } V^w(s,b) &= \max_d \sum_{i=1..m} w_i^b \phi_i(s,d,b). \end{aligned} \quad (4)$$

Note that now $Succ^b(s,d)$ is the set of pairs (s',b') which can possibly result from the application of d in s when the remaining budget is b.

We propose to build the batch $\mathcal{B}$ from a finite set of simulated trajectories starting in s_1, obtained by simulating successive transitions. Therefore, we have the guarantee that every 4-tuple $(s,d,b,b') \in \mathcal{B}$ effectively corresponds to a reachable configuration. Decisions are chosen randomly according to the ε-greedy method, either maximizing the current estimation Q^w (with probability $1-\varepsilon$) or uniformly (with probability ε) at each time step. Note that ε and the batch size are the only parameters to tune in LSDP.

The restricted system of equations (4) allows to define the *Least-Squares Dynamic Programming* (LSDP) algorithm, a variant of the *policy iteration* algorithm [67]. LSDP is initialised with a set of weights (one for each value $b \leq B$). Then, we iterate steps of batch generation and updates of the weights values from the current batch. Weights updates are performed by applying the approximate dynamic programming procedure described above. The updated values are accepted only if the value of the corresponding policy (estimated by simulation) improves the previous one. If the value is not improved, another batch is built and used. A maximum number of batches to simulate is fixed, and when reached, the current policy is returned. See Figure 3 for a schematic representation of LSDP.

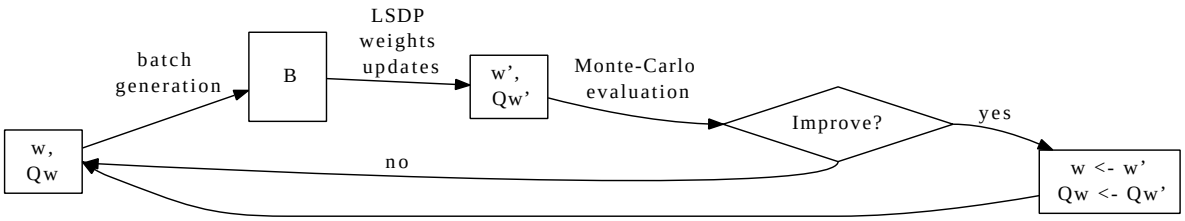


FIGURE 3 – Schematic representation of the LSDP algorithm.

Of course, for a given set of weights values, different batches may be obtained by simulation, leading to different updated weights values and thus to different updated policies. Furthermore, there is no guarantee that the updated policy improves the current policy in state s_1. This is why the value of the updated policy has to be estimated (by simulation) and compared to the value of the previous policy, before being accepted if its estimated value improves that of the current selected policy. This conditional acceptation allows to guarantee that the successive policies returned by the algorithm are of increasing value [6].

6. Since simulation is used to estimate policy values, these estimations may well be incorrect but they hopefully preserve policies values ranking.

6 Application of LSDP to the OASMRF problem

In order to apply the LSDP algorithm to the OASMRF problem, we take into account the problem structure (i) to define features ϕ_i and (ii) to propose a time efficient batch construction method. It also requires to be able to compute efficiently (in terms of time complexity) conditional marginals of the form $\mathbb{P}(x_i \mid x_A)$. These quantities are necessary to compute transition probabilities, to evaluate the final reward (the MPM value) and, as we will see, to compute the features. We also present two other applications of LSDP for the OASMRF problem, based on different feature choices or quality measures.

6.1 LSDP implementation for the OASMRF problem

Features choice. We choose to define one feature per variable in the MRF ($m = n$). The features definition is inspired by the BP-max heuristic (see [66] and section 4). This heuristic consists in selecting for sampling, at each sampling step, the variables which remain the most uncertain. Uncertainty is defined as a conditional max marginal $\max_{x_i \in \Omega} \mathbb{P}(x_i \mid x_A)$ close to 0.5 (in the binary case). The BP-max heuristic can be obtained as the greedy policy with respect to a parametrised Q-function Q^1 with the following features, and all weights equal to $1 : \forall i \in \{1, \ldots, n\}$,

$$\phi_i(s, d, b) \quad = \quad \big(1 - 1_{\{i=d\}}\big) \max_{x_i \in \Omega} \mathbb{P}\big(x_i \mid x_A\big) + 1_{\{i=d\}}, \tag{5}$$

where $A \subseteq O$ is the set of indices of previously observed variables. We adopt definition (5) to apply LSDP to the OASMRF problem. Thus, in practice, we initialise the LSDP algorithm with weights all equal to 1. Then, the LSDP algorithm performs successive updates in order to improve this initial set of weights.

Batch construction. Simulating trajectories in the OASMRF problem is costly since, for each transition, one has to simulate observations $x_{A^{t+1}}$ from the MRF conditional distribution $\mathbb{P}(x_{A^{t+1}} \mid x_{A^t})$. This requires to apply the Gibbs Sampling algorithm a large number of times, which is rather costly, thus severely limiting the size and number of batches that can be constructed. However, larger batches can be constructed if we divide the construction into two phases. First, we simulate, off-line, a batch of maps, $\{x^1, \ldots x^p\}$, from $\mathbb{P}(.)$. It will be used for all iterations of the LSDP algorithm. The construction of this batch is done using Gibbs Sampling, and induces a single overhead cost for the whole algorithm. Then, trajectories are easy to simulate : (i) a map is selected at random in the batch, (ii) actions are chosen following the ε-greedy method with respect to the current policy, and (iii) successor states follow immediately by reading the value of the variables corresponding to the current observation. This second phase of trajectories simulation is fast. Furthermore, simulated trajectories do not have to be stored (only the batch of maps does), thus saving much memory space. In addition, we can establish that it is valid to use this batch procedure to simulate transitions of the MDP encoding of the OASMRF problem. More formally, if we encode a state s^t as a vector of size $|O|$, with $s^t(i) = -1$ if site i has not been visited yet, and $s^t(i)$ equals the observed state x_i otherwise, we establish the following lemma

Lemma 1 *For a given action trajectory* $(d^1, \dots, d^H)$*, a state trajectory* $(s^1, \dots, s^{H+1})$ [7] *simulated according to the following two-steps scheme has the same joint probability distribution as a trajectory simulated according to the OASMRF MDP model transition function.*

1. *Simulate a map* x *according to the joint distribution* $\mathbb{P}(.)$*.*
2. *Deduce iteratively the values* $(s^1, \dots, s^{H+1})$ *according to* $s^1(i) = -1 \ \forall i \in O$ *and :*
$$\forall t \in \{1, \dots, H\}, \ s^{t+1}(i) = s^t(i) \text{ if } d^t(i) = 0 \text{ and } s^{t+1}(i) = d^t(i)x_i \text{ else.}$$

(We recall that a site is visited at most once during a trajectory).

A proof is given in the Appendix.

Approximation of $\mathbb{P}(x_i \mid x_A)$. The *Belief Propagation* (BP) algorithm [61] can be used to compute (approximately) $\mathbb{P}(x_i \mid x_A)$. However since this evaluation has to be performed a huge number of times, BP cannot be applied in practice. So we propose to use the distribution $\widetilde{\mathbb{P}}$ defined below as an approximation of $\mathbb{P}$:

$$\widetilde{\mathbb{P}}(x_i \mid x_A) = \mathbb{P}^{BP}(x_i) + \sum_{j \in A} \Big[\mathbb{P}^{BP}(x_i \mid x_j) - \mathbb{P}^{BP}(x_i) \Big]. \tag{6}$$

This approximation does not necessarily belong to $[0, 1]$ but sums to one. It has the advantage to be fast to compute. Indeed, before running LSDP, all marginals and conditional marginals $\mathbb{P}^{BP}(x_i)$ and $\mathbb{P}^{BP}(x_i|x_j)$ are computed using BP, inducing a fixed overhead computational cost. Then, within an iteration of LSDP, we can compute $\widetilde{\mathbb{P}}(x_i \mid x_A)$ in an incremental way since $\widetilde{\mathbb{P}}(x_i \mid x_A \cup x_j) = \widetilde{\mathbb{P}}(x_i \mid x_A) + \mathbb{P}^{BP}(x_i|x_j) - \mathbb{P}^{BP}(x_i)$. Our approximation is ad-hoc and we could have considered more sound methods to define an approximate distribution of $\mathbb{P}(x_i \mid x_A)$ from the $\mathbb{P}^{BP}(x_i)$ and $\mathbb{P}^{BP}(x_i|x_j)$. Different options are discussed in [1]. In particular the authors pointed out the superiority of methods based on multiplication instead on addition. We did not explore this option since ours provided good empirical results and does not require any extra parameters estimation.

6.2 Two variants to carry out LSDP for OASMRF

Static version of LSDP. It is possible, by changing the features definition, to design a static policy for the OASMRF problem. Here by static we mean that the choice of the next sample does not depend on the value of the variables observed in the previous sampling steps. It depends only on their locations. Therefore the set of sampled sites (a plan) can be computed in advance, before actually sampling the sites. Such a static policy can be obtained by considering the following definition for the features, $\forall i \in \{1, \dots, n\}$,

$$\phi_i(s, d, b) = 1_{\{i=d\} \cup \{s_i \neq -1\}}.$$

The feature is equal to zero for all sites not sampled (at the current step or in previous ones) and 1 otherwise.

7. here H is the length of a given trajectory, and not necessarily a fixed number of sampling steps since under budget constraints trajectories do not have a common length.

Entropy based LSDP. Up to now, the OASMRF problem and the LSDP algorithm have been described for a measure of sampling policy quality based on the MPM criterion (1). This choice is not arbitrary since with this definition the procedure used to restore the MRF state from a sample output and the procedure used to define the sampling policy quality rely on the same criterion. Still, other classical options can be considered to define sampling policy quality. We could, for instance, define the OASMRF problem with an entropy-based criterion. In this case, since entropy has to be minimised, we define :

$$U\big(A, x_A\big) = -H(\mathbb{P}(X_R \mid x_A)) = \sum_{x_R} \mathbb{P}(x_R \mid x_A) \log(\mathbb{P}(x_R \mid x_A)),$$

with r^{H+1} defined accordingly. The steps of the LSDP algorithm would remain roughly unchanged with the entropy criterion but the features definition should be adapted to : $\forall i \in \{1, \ldots, n\}$,

$$\phi_i(s, d, b) = -\big(1 - 1_{\{i=d\}}\big)H(\mathbb{P}(X_i \mid x_A)) + 1_{\{i=d\}},$$

where $A \subseteq O$ is the set of indices of the previously observed variables and $H(\mathbb{P}(X_i \mid x_A)) = -\sum_{x_i \in \Omega} \mathbb{P}(x_i \mid x_A) \log \mathbb{P}(x_i \mid x_A)$ is the marginal entropy of the conditional distribution of X_i given x_A. Evaluating marginal entropy is not simpler than evaluating conditional marginals. In order to approximate these quantities we can again use approximation (6).

Note that the entropy criterion does not provide a rule to estimate the variables X_R from a sample output. This reconstruction step still has to be performed using MPM or MAP methods.

7 EXPERIMENTAL EVALUATION

We present simulated sampling problems and one real problem of weeds sampling to illustrate the gain of using LSDP instead of classical heuristics or RL-based solution algorithms. We compared LSDP to the random heuristic, the LSDP-static policy, the BP-max policy, TD(λ) [76] with tabular representation of the Q-function, and LSPI. LSPI and LSDP were implemented with the same features definition and were run with $\epsilon = 0.9$. We also compared LSDP to a greedy algorithm based on the *Mutual Information* (MI) criterion [44].

The OASMRF problem considered is the following. The graph G is a regular grid and $R = O = V$. One variable is observed at each decision step ($L = 1$) and sampling costs are null on the three first sets of experiments. We considered the following Potts model distribution : $\forall\, x \in \{1, 2\}^n$

$$\mathbb{P}\big(x\big) \propto \exp\left(\frac{1}{2} \sum_{(i,j) \in E} 1_{\{x_i = x_j\}}\right).$$

4×4 grid. This small problem was introduced in the experiments since we were able to compute the corresponding optimal policy, using the backward induction algorithm (see Section 4), and the exact value of any policy. TD(λ) was run with $\lambda = 0.1$, using the

ε-greedy method for action choice ($\epsilon = 0.1$). It was run using 675000 simulated state-action trajectories, in order to reach convergence. To be comparable, we ran LSDP and LSPI with a batch of 100 maps and 6750 iterations (in practice a few hundred iterations are enough). For LSDP the value of the policy obtained at the last iteration of the algorithm was returned, and for LSPI the value of the best policy among all iterations was returned.

The first conclusion is that the absolute difference between the values of all policies is small : an absolute increase of the percentages of 2.2 at most. We also compared the policies in terms of normalised gain compared to the random one δ_R (Figure 4) : the score of a given policy δ is defined as $score1(\delta) = \frac{V(\delta)-V(\delta_R)}{V(\delta^*)-V(\delta_R)}$.

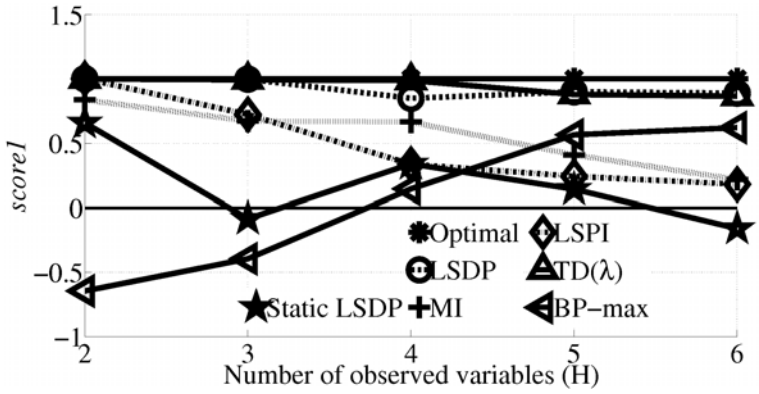


FIGURE 4 – OASMRF problem with 16 variables : $score1$ of LSDP and classical one-step-look-ahead and RL-based heuristic policies. A policy with $score1$ equal to 0 is a policy with the same value as the random policy.

Among RL algorithms, TD(λ) is the best and LSDP gives very similar results. In comparison, LSPI shows a poor behaviour, always returning dominated policies. Surprisingly the relative values of the MI and LSPI policies decrease with the number of observed variables, while the opposite behavior is observed for the BP-max heuristic. The poor performance of the BP-max heuristic with small sample size is explained by the fact that with few observed sites, all sites have similar marginal probabilities. In that situation we arbitrarily choose the site to sample as the one with the lowest index in V.

10×10 grid. For this problem size, only LSDP, LSDP-static, LSPI, BP-max and random policy can be computed. For LSDP, LSDP-static and LSPI we used a batch size of 1000 maps and 1000 iterations. The value of a policy was estimated by Monte Carlo approximation. We modified $score1$ into $score2(\delta) = \frac{V(\delta)-V(\delta_R)}{|V(\delta_{BP-max})-V(\delta_R)|}$: since the value of an optimal policy cannot be computed, δ_{BP-max} serves as a reference. Results are displayed on Figure 5.

We observed again the poor performance of the LSPI algorithm (dominated by the random policy for $H = 10$ to 20). On the contrary, LSDP performs quite better than the BP-max heuristic for small sample sizes. LSDP also performs better than LSPI, in terms of computation time : for $H = 40$, an iteration takes about 7 seconds for LSDP, 77 seconds for LSPI. The LSDP-static policy also leads to an improvement compared to BP-max, but lower than with LSDP : this example and the previous one demonstrate

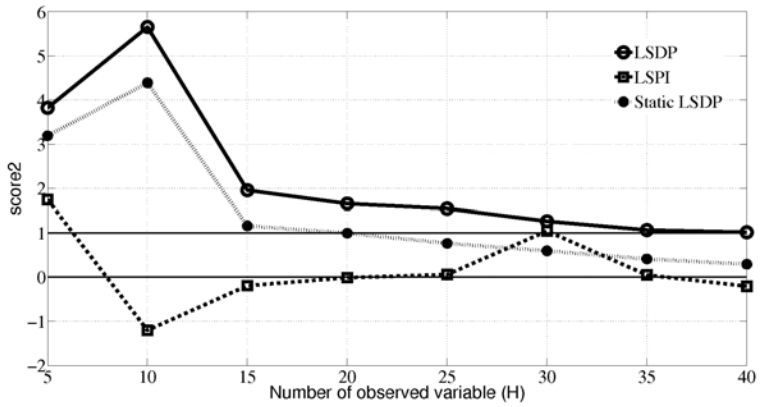


FIGURE 5 – OASMRF problem with 100 variables : *score*2 of LSDP, LSDP-static and LSPI policies. A policy with *score*2 equal to 0 (resp. 1) is a policy with the same value as the random (resp. BP-max) policy.

the interest of looking for adaptive policies.

Constrained moves problem. We compared LSDP, BP-max and random policies on a more realistic sampling problem, involving constrained moves on the grid for observing sites. The agent starts by sampling the site at the top-left corner of the grid.Then, after having observed a site, the agent can only move to distance-2 sites for the next observation.

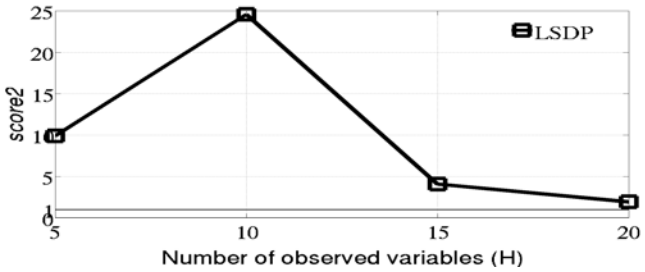


FIGURE 6 – Constrained moves problem with 100 variables : *score*2 of LSDP policy. A policy with *score*2 equal to 0 (resp. 1) is a policy with same value as the random (resp. BP-max) policy.

We again observed that the absolute difference between all policies remained small (for $H = 10$, the value of the LSDP policy is 61.7 while the value of the BP-max policy is 59.4). LSPI showed the same poor behaviour than in the previous experiment. As we expected, the gain provided by LSDP in terms of relative improvement of the random policy ($H \leq 20$, see Figure 6) is significant when the sample size is small (Figure 6).

Sampling under cost constraints With this set of experiments we introduced distincts costs values $S_C(i, x_i)$ and we considered the problem of maximising the restoration quality under the constraint of a fixed allocated budget B. We considered three different cost functions $S_C(A, x_A)$. For the type I and type II, cost depends only on the site location. With cost I, the sampling cost increases with the distance to the grid boundary, while with cost II, we have two different costs in the two diagonals (see Figure 7). With

	Type I cost		Type II cost		Type III cost	
Policy	Value	sampled sites	Value	sampled sites	Value	sampled sites
LSDP	64.80	27.3 (2.5)	63.6	22.8 (0.7)	65.4	25.6 (1.7)
BP-max	61.77	19 (1.9)	60.4	15.8 (1.9)	64.7	25.6 (1.8)
Random	60.27	26.65 (2.8)	59.7	15.6 (2.3)	63.7	25.6 (1.9)

TABLE 1 – Values and mean number of visited sites under different configurations of cost constraints, for the LSDP, BP-max and random policies. Values between parentheses are standard deviations for the mean number of visited sites.

the type III cost, we consider a function S_C which depends only on the value of the observation : $S_C(i, x_i) = 2$ if $x_i = 1$ and 1 otherwise. We ran the LSDP, BP-max and random policy on a 20×10 grid and for a budget $B = 38$. For LSDP we used a batch of size 4000 or 2000, and 1000 iterations. Results in terms of policy values and numbers of sites sampled are presented in Table 1. For the three types of cost function, one can observe that the ranking of the three policies values is always $LSDP > BP - max > random$. The LSDP policy distributes the budget B in a way that enables to sample more sites than with BP-max.

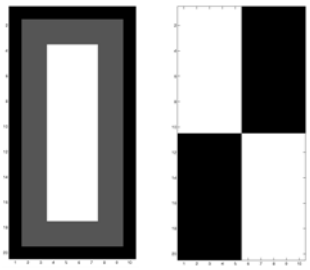


FIGURE 7 – Left : type I repartition of costs, costs are respectively of 1, 2 and 4 for black, grey and white sites. Right : type II repartition of cost, costs are respectively of 1 and 4 for black and white sites.

Weeds sampling in a crop field under time constraint. We applied the LSDP algorithm to a real-life sampling problem : the design of adaptive strategies for weeds sampling in a crop field [10]. A spring barley field has been divided into a grid of 13×13 quadrats and a particular weed species *Galium Aparine* has been sampled in each quadrat. The observation x_i in quadrat i is the weed abundance and belongs to one of the three following classes : 0 (no weeds), 1 (less than one plant per square meter), 2 (between 1 and 3 plants per square meter). We considered different Markov random field models corresponding to different properties and we selected the model with the higher BIC value. This model was an anisotropic Potts model with external field :

$$\mathbb{P}_\beta(x) \propto \exp\left(\sum_{i\in V} \alpha_{x_i} + \beta_t \sum_{(i,j)\in E_t} 1_{\{x_i=x_j\}} + \beta_o \sum_{(i,j)\in E_o} 1_{\{x_i=x_j\}}\right).$$

Subsets E_t and E_s respectively represent the subsets of edges in tillage direction and in the orthogonal direction, since anisotropy was induced by a difference of spatial correlation between these two directions. The estimated parameters where : $(\alpha_1, \alpha_2, \alpha_3) = (0, -0.0357, -3.4401)$ and $(\beta_t, \beta_o) = (0.8813, 0.0876)$.

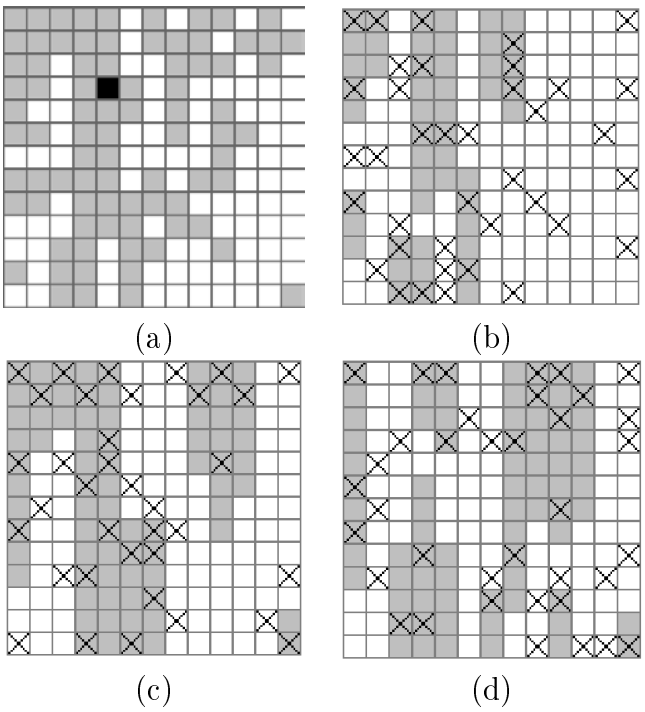


FIGURE 8 – Sampling strategies for weeds mapping. (a) : true abundance map, (b) MPM reconstruction based on the LSDP sampling strategy, (c) MPM reconstruction based on the BP-max sampling strategy, (d) MPM reconstruction based on the random sampling strategy. Sampled quadrats are marked by a cross. White, grey and black quadrats correspond respectively to abundance class 1, 2 and 3.

The cost function $S_C(A, x_A)$ represents the time needed for abundance estimation in a quadrat (a site of the MRF). We used a regression model based on factors identified as the most relevant by experts. Then, we applied LSDP on the MRF model estimated on the weeds data set, with a batch size of 4000 and 1000 iterations. We observed the same ranking of policies (in terms of values, i.e. expected number of quadrats where abundance is well estimated) : LSDP (120) then BP-max (118.7) and then random (117.16). Then we applied the three strategies to sample and reconstruct the original weeds abundance map used to build the MRF model. Figure 8 shows the true abundance map, and maps estimated from sample outputs provided by LSDP, BP-max and the random policy.

8 CONCLUSION

In this article, we have provided a factored Markov Decision Processes (MDP) model to represent problems of optimal adaptive sampling of spatial processes expressed in the Markov random fields framework. Our second contribution is a generic *batch mode reinforcement-learning* algorithm, LSDP, which can be applied to any large state-space finite-horizon MDP problem, as soon as the MDP model is known explicitly. Then, our last contribution is an experimental evaluation of the LSDP approach for solving the OASMRF problem. Our experimental work enables us to draw the following conclusions. First, in small problems where the optimal policy can be computed, we notice that the performance of a purely random strategy is quite close to that of the optimal one. This

seems to also hold for larger problems, where the estimated value of the random policy remains close to that of the LSDP policy. However, in real-life applications of sampling for mapping, small errors in the reconstruction of maps can lead to a significant increase in management costs (think of imperfect mapping and eradication of invasive species, leading to future catastrophic outbreaks). Second, for large problems, non-parametrized RL approaches (such as TD(λ)) are too computationally intensive to apply, and the LSPI approach does not perform well. On the contrary, both BP-max heuristic and the LSDP algorithm provide good results (provided that the sampling budget is large enough, as far as BP-max is concerned). BP-max is less computationally expensive to apply than LSDP. However, its main drawback is that the choice of the sample does not take into account its cost. The budget constraint can only be used to decide when to stop a sampling trajectory. In contrast, LSDP can handle cost functions and our experiments show that when sampling costs are nonhomogeneous the superiority of LSDP over BP-max and random policies is increased.

This work opens several directions for future work : on the problem of sampling in spatial random fields in one hand, and on more general problems of sequential decision under uncertainty.
Regarding the framework and algorithm we proposed for spatial sampling, a first possible extension would be to consider other definitions of sample quality measures. In this paper, the measure used to illustrate the approach is the MPM value. However, the MDP encoding and the application of LSDP do not crucially depend on the quality measure definition. Other criteria, such as MAP, or Entropy should be explored. It would probably require to define new features, as we have illustrated for the entropy case, and belief propagation algorithms could still be used to compute approximately MAP or entropy values.
We largely discussed the different options to introduce cost constraints in the optimal sampling problem. We have modelled our sampling problem as a problem of optimising reconstruction quality, under sampling budget constraint. However, one could, dually, be interested in finding sampling policies achieving a minimum reconstruction quality threshold, while minimising the sampling cost. An MDP encoding of this problem is still possible and the LSDP algorithm could be applied. It would require an MPM evaluation at every sampling step to check if the minimal quality is reached, but this can be evaluated approximately based on our time efficient approximation of the conditional marginals. Other forms of sampling cost could also be discussed : these could be more general than the ones we have considered in the paper. These could be linked, for example, to a maximum sampling trajectory duration, modelled as a sum of transitions (s, a, s') costs.
Finally, even the choice of a Markov Random Field to model map uncertainty can be challenged, while keeping the approach we proposed. One could easily adapt the principles of our approach to continuous space models, provided that the number of potential sampling locations be finite. In a MRF, each variable typically take values in a finite set of small size. We could consider applying LSDP to problems with larger (but still finite) domains, when counts data should be modelled. The only requirement would be to be able to efficiently compute conditional marginals and simulate full maps. If the domain of the variable to map is continuous, this rises the more complex question of the definition of MDP on continuous state space.

Then, as we already mentioned, the LSDP algorithm is not specific to the resolution of the optimal sampling problem. One important contribution of this work is a new model-based reinforcement learning algorithm for large size finite-horizon Markov Decision Processes. This means that it can be applied to solve problems of sequential decision under uncertainly where the state and/or decision space are/is large and factored. (eg. invasive species control, biodiversity conservation, weeds management, ...). MDP formalisms already exist to model the control of spatial processes in time : Factored MDP [34] and Graph-based MDP [71], for example. The structure of these MDPs shares numerous common points with the MDP model of the OASMRF problem. Clearly, the LSDP approach could be adapted to approximately solve FMDP or GMDP problems, when the horizon is finite.

Deuxième partie

Modélisation du problème d'échantillonnage d'une espèce adventice

Sommaire

9 Introduction

In arable fields, weeds are responsible for yield loss [57] (i) they are competing with crop for resources [84], (ii) they can be host for parasites [51] and diseases [82]. In order to optimize yield and hence, an economic criterion, herbicides have been widely used. However, the impact of herbicides on soil and groundwater is now recognized [4] and in the meantime, the role of weeds in agro-ecosystem food webs and in providing ecological services has been established [32], [63]. Therefore weed management has to deal both with economic (i.e. maintaining yield at acceptable levels) and ecological (i.e. promoting ecological services) criteria. The design of more sustainable cropping systems to manage weeds requires first to understand the spatio-temporal dynamics of weeds within fields. Such understanding is based on the knowledge of their spatial distribution within fields and on the influence of agricultural practices on this distribution.
One way to improve knowledge on their spatial structure is to map weeds at the field scale. Construction of such maps is based on observations within the field and sampling methods represent the different ways to acquire these observations.
These methods can be divided into two categories : *continuous* and *discrete* weed sampling [6]. For continuous weed sampling, the field is not discretized. A continuous and complete observation of the field is acquired, like for picture. This method became popular for commercial use and site specific management [47], [73]. In this case, weed maps are constructed using satellite [2], aerial [14] or near-ground imagery [36] and some image processing techniques. For discrete weed sampling, the field is divided into a regular grid of quadrats [8] and due to the large size of fields, observations of weeds are performed only on a subset of quadrats.
Due to the high cost of continuous weed sampling techniques and weak performance for precise weed mapping [73], discrete weed sampling is the most popular method for academic research ([53],[16],[39]). In this case an observer moves within the field and measures weed presence in the selected quadrats.
The aim of our study is to discuss and improve current discrete sampling methods for weeds mapping at the field scale.

Before applying any discrete sampling method, several questions have to be addressed : which type of observations (e.g. counting, density classes or occurrence), which quadrat size, which budget for sampling [9] and which extrapolation method used for mapping. All these questions are linked and the answers depend on the aim of mapping. For example in [6], the weeds maps are created for a site-specific weed management. These maps are used to determine the area of the field where herbicides must be sprayed and in which quantity. Thus, 4 weed density classes were proposed according to the potential yield loss risk and the needed herbicide treatment. In this case the map scale does not need to be very precise since the herbicide spraying equipments are not.
Then, based on the observations nature, an adapted extrapolation method has to be identified.
In the following we will consider the problem of mapping the spatial repartition of one weed species of interest. We consider that such maps will be used for spatial structure

8. A rectangle put down on the ground where weeds population was estimated
9. This is the initial resource available for sampling. For example it can be the number of observers or the time available for sampling.

analysis. We will consider observations of type class (e.g. density classes or occurrence) and suppose that a good weed map is a map where a large proportion of quadrats are reconstructed to their true values . We propose an extrapolation method, several models for spatial repartition of weeds density classes, a way to model budget for sampling and a way to choose observations locations that respect this budget and take into account the objectives of sampling.

In order to have a relevant representation of weeds repartition, locations of observations are crucial [15],[83], [24]. This requires that observations locations are spatially dependent [15]. Several articles have demonstrated the aggregated repartition of weeds into fields [81],[16], [15], [21],[19]. As mentionned in [15] :

> *Most weeds occur at various densities in clumps or patches of various sizes and shapes, and few individual plants occur between patches.*

In addition, weeds density and species vary greatly across field in space and time [25], [52], [81], [19],[21], which makes unsuitable the use of previous sampling results. Literature gives a large range of sampling strategies [10] for weeds mapping [6],[20],[23],[64],[69],[80]. Observations locations are generally chosen randomly or in order to cover a large area within the fields, like for regular sampling. These strategies are generally static. Observations locations are chosen once and for all, whatever the value and locations of previous observations which can lead to reconstructed maps of poor quality. Indeed, when observations locations are fixed in advance, probability that this location falls on places where weeds occur can be small. Shape and size of patches can also be highly different from reality. On the contrary, for adaptive sampling strategies observations locations are determined step by step, taking into account previous observations. When adaptive, the sampling strategy can be modified if a patch has been discovered to enhance the quality of the reconstructed map, as it has been already demonstrated in some other sampling problems [66],[79].

The extrapolation method will also determine the reconstructed map quality [37],[35]. All extrapolation methods are based on a probabilistic model which describes the spatial repartition of the phenomenon under study. In the case of count data, a wide range of models has been proposed [62],[48],[13],[12]. The extrapolation method is generally kriging [28],[16],[39],[37],[24]. Some of these models have been adapted to the case of classes data, but combined with kriging they lead to over-smoothed maps [69]. In this analysis we propose the use of Markov Random Fields (MRF, [31]), a natural model for data of type class, widely used in image analysis. Adapted MRF distributions are selected for different weed species and quadrat sizes. Then a common extrapolation method based on the Maximum Posterior Marginals (MPM, [31]) is used.

Generally, the number of observations is chosen in advance for practical reasons or according to expert rules. However the achievement of a sampling strategy can be compromised if it takes too much time to apply in a field. Therefore in this work we propose to model the time needed to observe one quadrat as a function of the field and quadrat configurations. Thus the budget constraint can also be expressed in terms of available

10. The word strategies is used instead of methods to focus on observations location. The sampling method focus on the quadrat sizes, the extrapolation method and the observation nature.

time for sampling. We also propose an adaptive strategy which selects the next quadrat to observe according to the remaining time.

Finally, the design of adaptive sampling strategies is defined as an optimization problem. The value of any adaptive sampling strategy is based on the expected reconstructed map quality. The optimal adaptive sampling strategy is the one with highest value, respecting the budget constraints. Since this optimization problem is too complex for exact resolution, we used an approximate resolution algorithm, named Least-Squares Dynamic Programming LSDP defined in [11]. A heuristic solution [66] was also tested and these model-based sampling strategies were compared to classical static strategies.

The rest of the paper is organized as follows. In Section 10 we describe all the material necessary to define the problem of finding an optimal adaptive sampling strategy that realizes a trade-off between quality of the reconstructed map and sampling cost. In Section 11 the weeds sampling cost model and the selected MRF models for different weeds species and quadrat sizes are presented. Adaptive sampling strategies are then compared to some classical static strategies for different budget constraints. Concluding remarks are presented in Section 12.

10 Method

In Section 10.1 we present candidate MRF models of the spatial repartition of weeds density classes. A model selection procedure and the extrapolation method are also presented. Then the weed sampling cost model is defined in Section 10.2. Optimal adaptive sampling strategies are defined in Section 10.3. Two datasets are presented in Section 10.4. Finally in Section 10.5 we present the method used to compare the sampling strategies.

First, a mathematical formulation of a weed map is given.

10.1 Weed Map Model

We suppose that the field of interest is divided into n contiguous quadrats, numbered 1,...,n. On each quadrat $q \in \{1, \ldots, n\}$, we define a discrete random variable $X(q)$, representing the density class of the weed species of interest on this quadrat. These variables take values in $\Omega = \{0, \ldots, K\}$, where each $k \in \Omega$ is a density class (or represent presence/absence of the weed species when $K = 1$). Notation $x(q)$ represents a realization of the random variable $X(q)$. This notation is extended to a set of quadrats $A = \{q_1, \ldots, q_H\}$, $x(A) := \{x(q_i)\}_{i=1\ldots H}$. Finally, the random vector $\mathbf{X} = \big(X(1), \ldots, X(n)\big)$ represents a map of spatial repartition of the weed's density classes within the field.
We write $\mathbb{P}(.)$, the probability distribution over all possible maps. For any map $\mathbf{x} = \big(x(1), \ldots, x(n)\big) \in \Omega^n$, $\mathbb{P}(\mathbf{X} = \mathbf{x})$ is the probability that the density classes on quadrat 1 to n are simultaneously $x(1)$,...,$x(n)$.

Possible distributions $\mathbb{P}(.)$ adapted to the spatial repartition of weeds density classes are presented in the next section.

10.1.1 Proposed models for the spatial repartition of weeds density classes

For the choice of $\mathbb{P}(.)$, we propose to use a *pairwise* MRF distribution [11]. The probability of any possible map, $\mathbf{x} = (x(1), \ldots, x(n)) \in \Omega^n$, is written as follow :

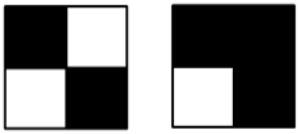
$$\mathbb{P}\big(x(1), \ldots, x(n)\big) \propto \prod_{q=1}^{n} \exp\big(\psi_q(x(q))\big) \prod_{(q,p)\in E} \exp\big(\psi_{qp}(x(q), x(p))\big),$$

where E is the set of pairs of neighboring quadrats, $\psi_q : \Omega \mapsto \mathbb{R}$ and $\psi_{qp} : \Omega^2 \mapsto \mathbb{R}$ are real-valued functions called *potential functions*. Note that the joint distribution is only proportional to the exponential term. In order to get a normalized probability distribution, the exponential term has to be divided by a normalizing constant Z :

$$Z = \sum_{\mathbf{x}\in\Omega^n} \exp\left(\sum_{q=1}^{n} \psi_q(x(q)) + \sum_{(q,p)\in E} \psi_{qp}(x(q), x(p))\right).$$

In practice, exact computation of Z is out of reach due to the large number of terms in the sum. Therefore only approximations of the joint distribution are reachable.
When using a pairwise MRF distribution, the probability value of a weed map $\mathbf{x} = \big(x(1), \ldots, x(n)\big)$, is based on the value of each quadrat independently of each other, $\big\{\exp\big(\psi_q(x(q))\big)\big\}_{q=1}^{n}$ and the value of neighboring quadrats, $\big\{\exp\big(\psi_{qp}(x(q), x(p))\big)\big\}_{(q,p)\in E}$. Then the choice of an appropriate MRF distribution for mapping weeds density classes repartition amounts to the choice of adapted potential functions.
In this paper, we study 8 different models detailed in Table 2. These models are extensions of the model M1, the classical *Potts model* with external field, where $\psi_q(x(q)) = \alpha_{x(q)}$ and $\psi_{qp}(x(q), x(p)) = \beta 1_{\{x(q)=x(p)\}}$. In this model the real-valued vector $\alpha = (\alpha_0, \ldots, \alpha_K)$, called *external field*, influences the relative proportions of the different density classes in the map. For example when β is small, a large value of α_0 (compared to the others), leads to a distribution $\mathbb{P}$ giving more weight to maps with a large proportion of variables in state 0. Then the real valued parameter β governs the proportion of nearby quadrats in the same state (spatial correlation). For example when $\beta = 0$, variables $\big(X(q)\big)_q$ are independent. And when β increases, patches progressively appear.
In order to understand the contribution of each parameter, let's compare the probability of the two maps presented in Figure 9. In this example $\Omega = \{0, 1\}$ and so α is a two dimensional vector. First, if α is equal to $(0, 0)$ and β is positive, the joint probability of the map (b) is greater than that of map (a) because a larger number of neighboring quadrats are in the same state (2 against 0). When $\alpha_1 > \alpha_0$ and β is again positive, the same conclusion arises. And the difference between the joint probabilities of each map is higher than in the first example because both the number of variables in state 1 and the number of neighboring quadrats in the same state are larger for map (b).

11. More complex MRF than pairwise can be defined. However estimation becomes more complex and in practice they are rarely used.



(a) (b)

FIGURE 9 – Two examples of maps configurations for $\Omega = \{0, 1\}$ and n=4. Variable in state 0 (resp. 1) are in white (resp. black).

We now describe the different definitions of potential functions we used.

First we proposed models where $\psi_q(x(q)) = 0$ for all $q = 1, \ldots, n$ and all $x(q) \in \Omega$. This type of models is very simple because β is the only parameter to estimate. These models are denoted M5 to M8 in Table 2.
We also proposed anisotropic models since patches of weeds are known to be more extended in the tillage direction [39], [70], [83]. In these models the spatial parameter is now different in tillage direction (β_t) and in the direction perpendicular to tillage (β_o). The set E is divided into E_t and E_o representing pairs of neighboring quadrats respectively along tillage direction and perpendicular to tillage direction. In Table 2, models M2, M4, M6 and M8 are anisotropic models.
Potential functions ψ_{qp} used for the Potts model are of "all or nothing" type. When $\beta > 0$, they attach the same (null) weight to any neighboring quadrats q and p that are not in the same state, whatever the distance between these two states. We proposed more flexible potential functions [17], $\psi_{qp}(x(q), x(p)) = \beta\big(1 - \frac{|x(q)-x(p)|}{K}\big)$. A maximal weight is given when neighboring quadrats q and p are in the same state and this weight decreases when the absolute difference between $x(q)$ and $x(p)$ increases. As for the Potts model, a weight of 0 is given when one of the two variables is in state 0 and the other in state K (models M3, M4, M7 and M8 in Table 2). When $K = 1$, these models are respectively equivalent to models M1, M2, M5 and M6.
These models can be grouped in 4 different types, depending on their properties. Each model type contains two models, one model is isotropic and the other anisotropic.

In the next section we describe the model selection procedure.

10.1.2 Model selection

We propose to compare the ability of each model to describe weeds repartition using the Bayesian Information Criterion (BIC, [72]). The BIC score of each model is computed from complete maps of different weed species and for different quadrat sizes. The BIC scores are then compared for each weed species and quadrat size and the best model is selected.
An exact computation of the BIC score is out of reach for MRF distributions. When a complete map is available a common method to estimate the parameters of all models and their BIC scores [38] is based on the pseudo-likelihood approximation [8].

Model	$\psi_q(x(q))$	$\psi_{qp}(x(q), x(p))$
		Type I
M1	$\alpha_{x(q)}$	$\beta 1_{\{x(q)=x(p)\}}$
M2	$\alpha_{x(q)}$	$\beta_t 1_{\{x(q)=x(p),(p,q)\in E_t\}} + \beta_o 1_{\{x(q)=x(p),(p,q)\in N_o\}}$
		Type II
M3	$\alpha_{x(q)}$	$\beta\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)$
M4	$\alpha_{x(q)}$	$\beta_t\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)1_{\{(p,q)\in E_t\}} + \beta_o\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)1_{\{(p,q)\in N_s\}}$
		Type III
M5	0	$\beta 1_{\{x(q)=x(p)\}}$
M6	0	$\beta_t 1_{\{x(q)=x(p),(p,q)\in E_t\}} + \beta_o 1_{\{x(q)=x(p),(p,q)\in N_o\}}$
		Type IV
M7	0	$\beta\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)$
M8	0	$\beta_t\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)1_{\{(p,q)\in E_t\}} + \beta_o\big(1 - \frac{\lvert x(q)-x(p)\rvert}{K}\big)1_{\{(p,q)\in N_s\}}$

TABLE 2 – Proposed potential functions for modeling spatial repartition of weeds density classes.

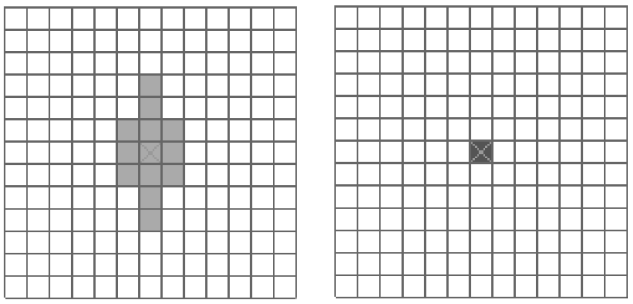
10.1.3 Extrapolation Method

If observations $x(A)$ are available, the reconstructed map, named $\mathbf{x}^{MPM}$, is computed with respect to the *Maximum Posterior Marginals* (MPM) criterion :

$$\forall q \in \{1, \ldots, n\}, \qquad x^{MPM}(q) = \underset{x(q)\in\Omega}{\operatorname{argmax}}\, \mathbb{P}\big(x(q) \mid x(A)\big).$$

Thus, on each quadrat the weed density class is reconstructed at its most probable value. It can be shown that this reconstructed map is the one which maximizes the expected number of well reconstructed quadrats when quadrats A were observed at density classes $x(A)$. The main difficulty when using the MPM criterion is to compute the values of all conditional marginals $\mathbb{P}\big(x(q) \mid x(A)\big)$. In practice, only approximate computation is possible. A common method is the Belief-Propagation (BP,[61]) algorithm. For all the experiments presented in this article, conditional marginals are computed using the BP algorithm.

Logically the reconstructed and true maps differ in the unobserved quadrats. When a quadrat q is far from any observed quadrats, there is no information in order to reconstruct the value of this quadrat. If quadrats A were observed at density classes



(a) (b)

FIGURE 10 – Reconstructed map using model M2 when the centering is observed at value (a) 1 or (b) 2.

$x(A)$, we have :

$$\max_{x(q)\in\Omega} \mathbb{P}(x(q) \mid x(A)) \simeq \max_{x(q)\in\Omega} \mathbb{P}(x(q)).$$

In this case, quadrat q is reconstructed to its most probable value when no other quadrats were observed. On the contrary when the unobserved quadrat q is close to other observed quadrats, the conditional marginal $\mathbb{P}(x(q) \mid x(A))$ is influenced by the observations values $x(A)$. This influence is governed by the model type and its parameters values. In Figure 10 we show two reconstructed maps for model M2 with $\Omega = \{0, 1, 2\}, \alpha = (0, -0.03, -3.58)$ and $(\beta_t, \beta_o) = (0.71, 0.12)$. In Figure 10.a (resp. 10.b) the central quadrat is observed at value 1 (resp. 2). When the central quadrat is observed at density class 1 some closed quadrats are also reconstructed to this density class. This is not the case when it is observed at density class 2 since α_2 has small value. In both cases distant quadrats are reconstructed to density class 0 since α_0 has the highest value. When models of type III are used, the distant quadrats are reconstructed to one of the density classes with the same probability. If models of type IV are used, they are reconstructed to the density classes of median values which is a bias of these models, due to the definition of ψ_{qp} [12].

Finally the best sampling strategy must achieve a good compromise between exploration (all parts of the field are explored) and patches boundaries detection (with small distances between sampled quadrats) while respecting the budget constraint.

10.2 Cost of Weed Sampling

The cost of sampling represents the effort needed to apply a sampling method. In this study, we assume that sampling cost could be restricted to the time needed to sample given a sampling strategy. We proposed and fitted a model on data from a field experiment.

12. When $K = 2k$ is even, they are reconstructed to density class k and when $H = 2k + 1$ is odd, to density class k or $k + 1$ with the same probability.

Most of sampling effort can be summed up by the time needed to observe the weed species within the field. Here we assume that the observations are performed by only one person. The initial budget for sampling is then the available time for the observer to explore the field.

The time needed to follow a sampling strategy can not be known in advance. For example, when there is high weeds species diversity, observations of weeds will be very time consuming. In practice, the time needed to observe a fixed number of quadrats depends on many other parameters than only diversity. From discussion and survey of practitioners we drew up a list (obviously non exhaustive) of these parameters, listed in Table 3 and described hereafter.
The time required to follow a sampling strategy can be decomposed into a sum of observation time, t_{note}, plus the time needed to move to the next quadrat, t_{move}. If the sampling strategy resulted in observing quadrats $\{q_1, \ldots, q_H\}$, the time used is :

$$t_{\text{note}}(q_1, x(q_1)) \; + \; \sum_{i=2}^{H} \left(\; t_{\text{note}}(q_i, x(q_i)) \; + \; t_{\text{move}}(q_{i-1}, q_i) \; \right)$$

Then it is sufficient to model the time $T := t_{\text{note}}(q, x(q)) \; + \; t_{\text{move}}(q, q')$, for all quadrats $q, q' \in \{1, \ldots, n\}$ and observation $x(q) \in \Omega$.
We attempted to explain the relationship between the time T and the parameters of Table 3 by a linear model with interaction between Z_1 and Z_2 and between Z_1 and Z_6. The observation period, Z_1, is a qualitative variable with two states {'unfavorable', 'favorable'}. This variable quantifies the faisability of the observation given at a specific stage. It is set to 'favorable' (resp. 'unfavorable') if the crop coverage [13] is less (resp. more) than 30%. Z_2 is a qualitative variable indicating the cropping system. This variable takes five values $\{1, \ldots, 5\}$, depending on which cropping system the observations are made. These five cropping systems are presented in Section 10.4. Z_3 is a qualitative variable indicating the crop type in the field during the exploration. This variable takes seven values : {'winter wheat', 'winter rape', 'winter barley', 'winter horse bean', 'spring barley', 'sorghum', 'corn', 'no crop' }. Z_4 indicates the number of weed species into the observed quadrat. Z_5 indicates the quantity of weeds in the quadrat. This is a continuous variable equal to the sum of the middle values of density classes of all weeds species present in the quadrat [14].
The support of any variable Z_i is denoted $Sup(Z_i)$. In order to simplify notations, we suppose that the support of categorical variables are positive integers i.e. each possible value of the variable is encoded by a positive integer. For example $Sup(Z_1) = \{0, 1\}$ and $Z_1 = 0$ (resp. $Z_1 = 1$) indicates that the observation period is unfavorable (resp.

13. The crop coverage is the proportion of the observer's field of vision occupied by the crop.
14. The total number of individuals in the quadrat could be more appropriate but only the density class is recorded in the data set. We recall that if a density class is "*between a and b individuals/*m^2 ", the middle of this class is $\frac{b-a}{2}$

Name	Meaning
Z_1	Observation period
Z_2	Cropping system
Z_3	Crop
Z_4	Total number of species into the quadrat
Z_5	Weeds quantity in the quadrat
Z_6	Distance covered, in meters, between two observed quadrats

TABLE 3 – Names and meaning of variables that influence the observation time.

favorable). The model is finally denoted :

$$T \simeq \sum_{i=1}^{3} \sum_{j \in Sup(Z_i)} \rho_{ij} 1_{\{Z_i=j\}} + \sum_{i=4}^{6} \gamma_i Z_i + \sum_{\substack{i \in Supp(Z_1) \\ j \in Supp(Z_2)}} \eta_{ij} 1_{\{Z_1=i, Z_2=j\}} + Z_6 \times \sum_{i \in Supp(Z_1)} \theta_i 1_{\{Z_1=i\}}$$

For the rest of the paper, the available time for sampling will be denoted B. For technical reasons when building our adaptive sampling strategies, the distance between two observed quadrats can not be taken into account when we evaluate the cost of a sampling strategy. This problem is discussed in Section 11.3.2. However we show that the time used to move within the tested adaptive strategies does not vary greatly. Therefore we decided to consider only time needed to observe a quadrat, simply defined as $t_{note}(q, x(q))$. This is computed using the cost model with Z_6 fixed to 0.

We now define the notion of adaptive sampling strategy and we present how to design such a strategy based on our weed map and cost models.

10.3 Design of Adaptive Sampling Strategies

Definition 1 *(adaptive sampling strategy) For any set of observed quadrats $q_1, \ldots, q_t$, at density classes $x(q_1), \ldots, x(q_t)$, an adaptive sampling strategy $\delta = (\delta^1, \ldots, \delta^H)$ is a function which indicates the next quadrat to be sampled given the preceding observed quadrats and the values of the weed plant quantity :*

$$\delta^{t+1}\big((q_1, x(q_1)), \ldots, (q_t, x(q_t))\big) = q_{t+1}.$$

A history is a trajectory $(q_1, x(q_1)), \ldots, (q_H, x(q_H))$ followed when applying a strategy δ. The set of all histories which can be followed by strategy δ is denoted τ_δ.

If a budget B is available for sampling, a strategy δ is said to respect this budget when it stops the exploration as soon as the remaining budget becomes negative or equal to zero. A schematic representation of an adaptive sampling strategy is given in Figure 11.

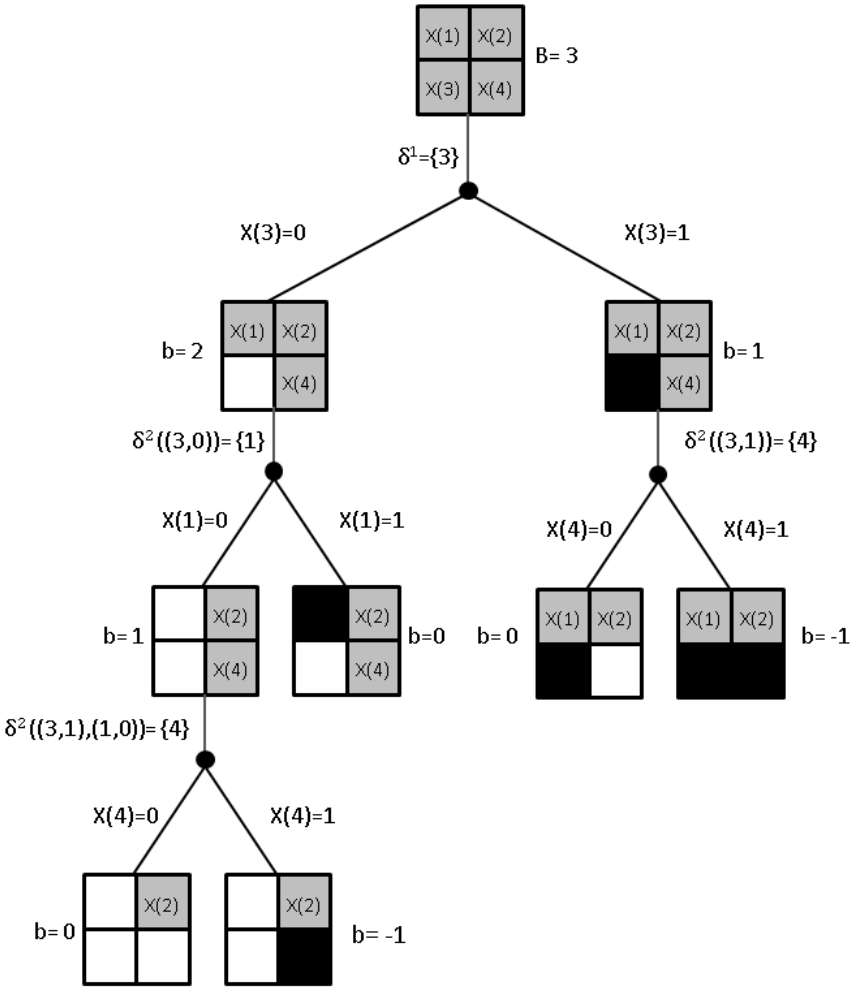


FIGURE 11 – Schematic representation of an adaptive sampling strategy δ. The field is divided into 4 square quadrats and observations can take binary values (e.g. presence/absence of species). The initial budget, B, is 3 and sampling costs only depends on the observation value. For all quadrats $q \in \{1, 2, 3, 4\}$, $t_{note}(q, x(q)) = 1$ when $x(q) = 0$ and 2 when $x(q) = 1$. In this example, quadrat 3 was observed first. Then depending on the observation value a second quadrat is selected. When $X(3)$ is observed at value 0 (left part of the diagram), quadrat 1 is observed next ,$\delta^2((3, 0)) = 1$, and the remaining budget is 2. If $X(1)$ is observed to 1, the exploration is stopped. When $X(1)$ is observed to 0, the remaining budget is still positive and another quadrat can be observed. It is quadrat 4 in this example and exploration is stopped after. If the first observations value is 1 (right part of the diagram) the remaining budget is 1 and quadrat 4 is observed next. Whatever the observations value $x(4)$, the exploration is stopped after.

10.3.1 Quality of an Adaptive Sampling Strategy

First, the value of any history $(q_1, x(q_1)), \ldots, (q_H, x(q_H))$ is defined as the sum over all quadrats of the probability of the most probable density classes. If $A = \cup_{t=1}^{H} q_t$, then

the value of this history, named $U(A, x(A))$, is defined as :

$$\begin{aligned} U\big(A, x(A)\big) &= \sum_{q=1}^{n} \mathbb{P}\big(x^{MPM}(q) \mid x(A)\big) \\ &= \sum_{q=1}^{n} \underset{x(q)\in\Omega}{\operatorname{argmax}}\, \mathbb{P}\big(x(q) \mid x(A)\big) \end{aligned}$$

It can be shown that the value of an history corresponds to the expectation over all possible maps, of the number of well reconstructed quadrats when variables quadrats A were observed to density classes $x(A)$. Thus, the closer $U(A, x(A))$ is to n, the more we can expected a large number of quadrats reconstructed to the true density classes when using the history $(A, x(A))$ for reconstruction.
For weed mapping, it can be more relevant to give a good representation of quadrats with high weed density than quadrats with no or few plants. The value of an history can be adapted accordingly :

$$U\big(A, x(A)\big) = \sum_{q=1}^{n} \mathbb{P}\big(x^{MPM}(q) \mid x(A)\big)\rho(x^{MPM}(q))$$

where $\rho : \Omega \mapsto \mathbb{R}$. In this case the value of an history is a weighted expectation of the number of well reconstructed quadrats. Then if it is more important that quadrats with density class above $l \in \Omega$ are well reconstructed, ρ can be defined as :

$$\forall k \in \Omega, \qquad \rho(k) = \begin{cases} 2 & \text{if } k \geq l \\ 1 & \text{if } k < l \end{cases}$$

Finally, the quality of any adaptive strategy δ is defined as the expected value of possible histories when following δ :

$$Q(\delta) = \sum_{(A,x(A))\in\tau_\delta} \mathbb{P}\big(x(A)\big)U\big(A, x(A)\big). \tag{7}$$

In practice an exact computation of the quality of any strategy is not reachable since, an exact computation of $\mathbb{P}\big(x(A)\big)$ is intractable (see Section 10.1.1). A common approach consists in computing an MCMC estimation of the expectation. First 2000 weeds maps are simulated according to distribution $\mathbb{P}(.)$. On each map, observations are acquired following strategy δ. Then a reconstructed map is computed and compared to the original one. The number of well reconstructed quadrats is then stored. The approximation of the policy value is then the average number of quadrats with good estimates of weed density over the 2000 maps.

10.3.2 Optimal Adaptive Sampling Strategies

An optimal adaptive sampling strategy is defined as follows :

Definition 2 *(optimal adaptive sampling strategy) An optimal adaptive strategy δ^* is a strategy of highest value, respecting the budget constraint :*

$$\delta^* = \arg\max_{\delta} Q(\delta) \text{ such that } \delta \text{ respects the budget constraint .} \tag{8}$$

10.3.3 Approximate Solution Method

In practice, an exact computation of δ^* is out of reach and approximate resolution methods have to be found. This paper focus on two particular strategies, named δ_{LSDP} and δ_{BP-max}, computed from the two known approximate resolution methods. The first method consists in (i) modeling the choice of optimal strategy as a Markov decision process ([78]) and (ii) proposing an adapted reinforcement-learning algorithm ([78],[75]), named LSDP, to solve problem (8). The second method , named *BP-max heuristic*, consists in (i) solving (8) greedily, that is each sampling plan is computed on-line according to previous sampling plans and observations and (ii) making the hypothesis that given the preceding observations, variables $X(1), \ldots, X(n)$ are independent. This strategy consists in observing the quadrat with highest uncertainty in term of marginal probability value. In other words, if quadrats $q_1, \ldots, q_t$ were already observed at density classes $x(q_1), \ldots, x(q_t)$, the next quadrat that will be observed is :

$$\begin{aligned} q_{t+1} &= \delta_{BP-max}\big((q_1, x(q_1)), \ldots, (q_t, x(q_t))\big), \\ &= \underset{q \in \{1,\ldots,n\}}{\operatorname{argmin}} \left\{ \max_{x(q)\in\Omega} \mathbb{P}\big(x(q) \mid x(q_1), \ldots, x(q_t)\big) \right\}. \end{aligned}$$

More details on the strategies δ_{BP-max} and δ_{LSDP} are available respectively in [66] and [11].
Finally note that these adaptive strategies are both model-based. Then to apply such strategies a weed map model has to be known or estimated. For the other sampling strategies the weed map model is only needed for the extrapolation method.

In the next section we present the dataset used for the model selection procedure, comparing the sampling strategies and estimating the parameters of the weed sampling cost model.

10.4 Presentation of the dataset

10.4.1 Available data set for model selection and comparison of the sampling strategies

The model selection procedure was run independently on complete maps of 22 different weeds species. These maps were obtained on the same subfield of $2500m^2$, located in Dijon-France (47°20'N, 5°2'E) with a semi continental climate. The field was used for barley and observations were recorded between April 11, 2006 and April 18, 2006. This subfield was discretized into contiguous quadrats of $0.36m^2$, which represent around 6950 quadrats. On each quadrat a count of each weed species has been recorded.
The nature of observations was initially count data, so we first converted them into density classes. We defined the six density classes presented in Table 4. These weeds density classes defined in [5] are an adapted version of the Braun-Blanquet cover abundance method [54].
All density classes are not observed for all the weeds species. Then Ω is defined independently for each weed species as the set of observed density classes.

Class number	Definition
0	No weed
1	Less than 1 individual/m^2
2	between 1 and 3 individuals /m^2
3	between 3 and 20 individuals /m^2
4	between 20 and 50 individuals /m^2
5	between 50 and 500 individuals /m^2

TABLE 4 – Definition of the density classes used for model selection

The density classes were defined on square quadrats of different area : $0.36m^2$, $1.44m^2$, $3.24m^2$, $5.76m^2$, $9m^2$, $12.96m^2$ and $17.64m^2$. The model selection procedure was run for each quadrat size. Figures 16 and 17 are examples of weed maps at different spatial scales for *chaenorrhinum minus (L.) lange* and *fallopia convolvulus L.*.

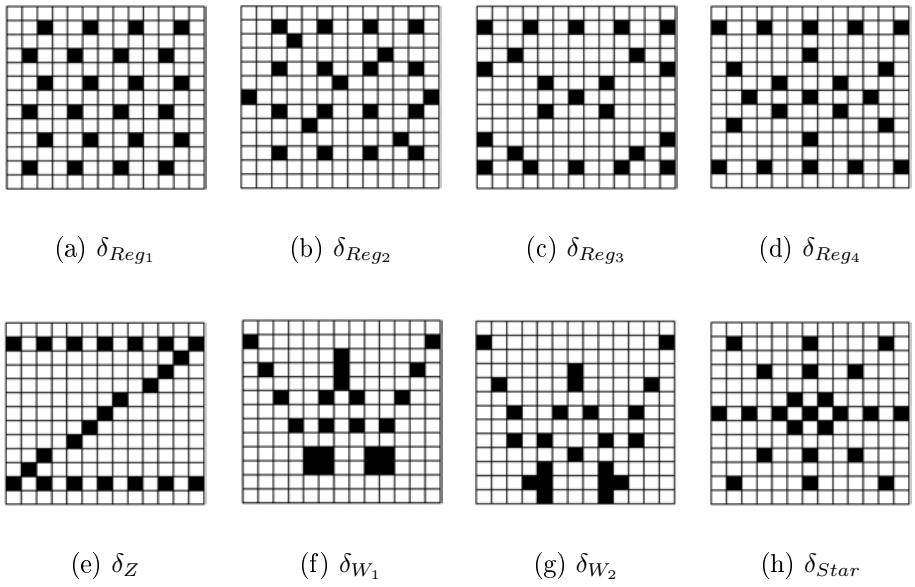
10.4.2 Available dataset for modeling the weed sampling cost

We use a large data set of a six-year experiment located in the same place of the previous dataset. The experiment was designed to compare four different cropping systems based on integrated weed management with a reference standard cropping system. The first cropping system (S1) is a standard reference designed to maximize financial returns. Herbicides are used for weeds control and mouldboard ploughind is carried out each year. The four other cropping systems (S2 to S5) are Integrated Weeds Management (IWM) cropping systems. The use of herbicides gradually decreases from the first cropping system to the fifth one for which no herbicides are used. Over the six-years, no significant differences were observed between the standard and the IWM cropping systems. Weed density varies during the growth of the crop. After weed control, the average weed density was generally lower for the first cropping system than the ITW systems. Before weed control, the average weed density was higher in S2 and S4, intermediate in S1 and lower in S3 and S5. Weed density also varied among weed species. Before weed control, the average density of winter broad-leaved species tended to be lower in IWM cropping systems than in S1. The average density of summer broad-leaved species tended to be higher for IWN cropping systems, both before and after weeds control. For S2, the average density of annual grassy weed species was higher than in the others cropping systems, both before and after weed control. Each of the cropping systems was tested on two different fields. In each field, sampling focused particularly on weed species which were known to be percinious and/or abundant. These weeds species were identified and their density classes estimated using the six density classes presented in Table 4. The density of each species was estimated through visual assessment and georeferenced with a global positioning system (GPS) in 16m^2 square quadrats. The visual assessments were generally done by the same observer. This was repeated two to four times a year in each field. A detailed presentation of this data set is available in [18]. The time of weed sampling per quadrat was automatically recorded by the GPS system. Then the difference between two recording times gives the time needed to asses the density class of a weed species plus the time needed to move from the next observations location (i.e. $t_{note} + t_{move}$). Thus the parameters of the cost model were fitted using about 14 000 observation plus moving time for 80 different weeds species. In general

only few repetitions were available for a given weed species, regarding to the number of model parameters. In order to get a larger data set, no distinctions were made between the observation times of different weed species. The effect of the weed species is then averaged when fitting the parameter of the linear model and it can be used for all types of weed species. Finally note that the model parameters were fitted with observation times for 16m^2 quadrats only. Then the model can be used for prediction only for the same quadrat size.

10.5 Method for the comparison of sampling strategies

We considered two different sampling problems.
In the first problem, we expressed the budget constraint in terms of available time for sampling. Three strategies are then compared : the random strategy δ_{Rand} and the two model-based strategies δ_{LSDP} and δ_{BP-max}. The weed sampling cost model (see Section 10.2) is used to stop the strategies when the available time for sampling is exhausted.
In the second problem, the budget constraint is simply expressed as a number of allowed observations. In this simplified case we will consider 8 additional static strategies : δ_{Reg_1}, δ_{Reg_2}, δ_{Reg_3}, δ_{Reg_4}, δ_Z, δ_{W_1}, δ_{W_2}, δ_{Star}. These strategies are presented in Figure 12. δ_{Reg_1},



(a) δ_{Reg_1} (b) δ_{Reg_2} (c) δ_{Reg_3} (d) δ_{Reg_4}

(e) δ_Z (f) δ_{W_1} (g) δ_{W_2} (h) δ_{Star}

FIGURE 12 – Different static strategies for comparison.

δ_{Reg_2}, δ_{Reg_3} and δ_{Reg_4} are different versions of a regular sampling strategies. For δ_Z the set of observed quadrats forms a Z into the field. For δ_{W_1} and δ_{W_2} it forms a W and for δ_{Star} a star.

For both sampling problem, a density class repartition model is needed in order to compute (i) the adaptive sampling strategies and (ii) the reconstructed maps. In order to compare the strategies, we selected the best models for quadrat of $12.96m^2$, for the weeds species *Lactuca serriola L.*, *Galium aparine L.*, *Picris hieracioides L.*, *Chaenorrhinum minus (L.) lange*, *Cirsium arvense (L.) scopoli*, and *Sinapsis arvensis L.*. The

model parameters were computed using the observed maps of each species, again using the pseudo-likelihood approximation [8]. The models parameters will be detailed in Section 11.3.1.

We compared the different sampling strategies based (i) on simulated maps and (ii) observed maps. The two comparison procedures are summarized in Figure 13 and

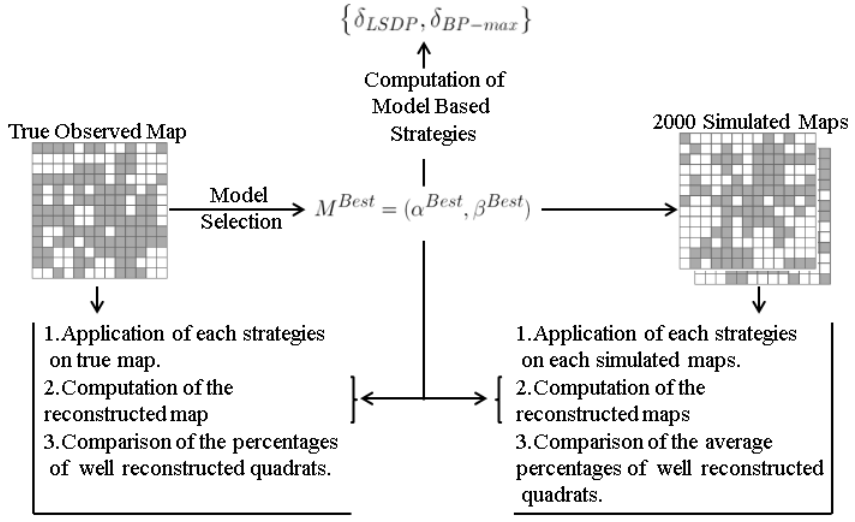


FIGURE 13 – Schematic representation of the procedure for comparing the sampling strategies.

organized as follows.

Simulated maps In this case the model used for computation of the adaptive strategies and the extrapolation method is also used to simulate 2000 weeds maps. All strategies are applied on each map and a reconstructed map is computed. Then the sampling strategies are compared according to their average number of well reconstructed quadrats. It corresponds to a MCMC approximation of the strategy value defined by equation (7). Values are presented in terms of average percentage of well reconstructed quadrats, which we will denoted WR.

Observed maps The sampling strategies are also applied to the six observed maps of the weeds species introduced previously. The strategies are still compared according to their percentage of well reconstructed quadrats. We also compare the strategies in terms of percentage of well reconstructed quadrats in the different density classes. We denote WRi this percentage for density class i.
For each weed species, the repartition model used for computing the adaptive strategies and the extrapolation method corresponds to the best model for the corresponding weed species.

For the first sampling problem (budget defined as a time constraint), strategies are compared on simulated and observed maps. Results are presented in Section 11.3.2. For the second sampling problem (budget defined as a number of observation), the strategies

are only compared on observed maps. Results are presented in Section 11.3.3.

As illustrated in Figure 13, the use of a repartition model is central. In a practical application the model parameters are generally not known and have to be estimated. Parameter estimation could be based on few initial observations but the estimated parameters can be widely different from those estimated on the complete map. Then the extrapolation method and the model-based strategies can performed really poorly. In this work we propose the use of empirical rules in order to estimate the model parameters. The sampling strategies are then compared on observed maps. Results are presented in Section 11.3.4.

11 Results

11.1 Cost model

Parameters of the cost model were fitted using the software R, results are summarized in Table 5.
We first note that all the model parameters have a significant influence. The adjusted R-squared of the fitted model is 0.8925.

ρ	γ
$\rho_{1.} = (-6.61\ ;\ -3.24)$ $\rho_{2.} = (0\ ;\ 17.3\ ;\ 8.13\ ;\ 9.44\ ;\ 9.19)$ $\rho_{3.} = (0\ ;\ -1.69\ ;\ -5.17\ ;\ 6.07\ ;\ 12.89\ ;\ 1.7\ ;\ 12.8\ ;\ 6.54)$	$\gamma_4 = 20.19$ $\gamma_5 = 27.36.10^{-4}$ $\gamma_6 = 2.37$
η	θ
$\eta_{0.} = (0\ ;\ 0\ ;\ 0\ ;\ 0\ ;\ 0)$ $\eta_{1.} = (0\ ;\ -20.34\ ;\ -20.44\ ;\ -18.77\ ;\ -14.09)$	$\theta_0 = 0$ $\theta_1 = 0.49$

TABLE 5 – Values of the fitted parameters for the cost model.

We first discuss the influence of the cropping system and observation period on the observation time t_{note}. Note that the cropping system does not influence the moving time t_{move}. The observation time is longer in the unfavorable observation period than in the favorable one, only for the first cropping system.
In the favorable observation period the first (resp. second) cropping system corresponds to the shortest (resp. longest) observation time and the three other cropping systems are associated approximately to the same observation time. In the unfavorable observation period the third (resp. the first) cropping system corresponds to the shortest (resp. longest) observation time. There is a significant variation of the observation time among cropping system. Figure 14 shows the observation time for all cropping systems in favorable and unfavorable observation periods. In this example the field is used for winter wheat. The observation time is given when the studied weed species is in density

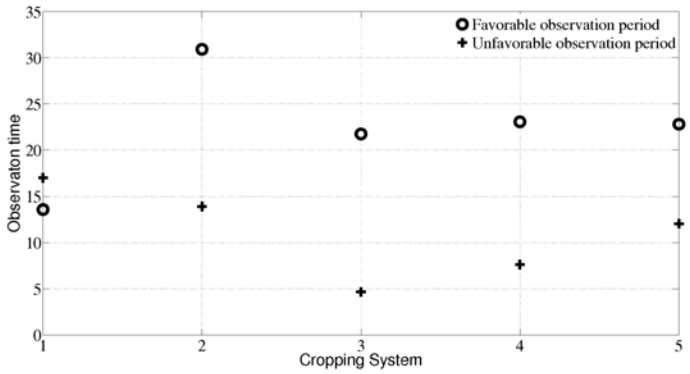


FIGURE 14 – Example of observation time values of one quadrat with a single species in density class 1, for the five cropping systems.

class 1 and no other weed species are present into the quadrat.
In Figure 15 is presented the observation times for all the density classes defined Table 4. The observation times are given for both observation periods, in a field used for winter wheat, managed using the fourth cropping system and where the studied weed species is the only one present in the quadrat. For other configurations the observation times will be different but the shape of the graph will not change. We can note that the observation time increases exponentially with the density class value. This is due to the density class definition.

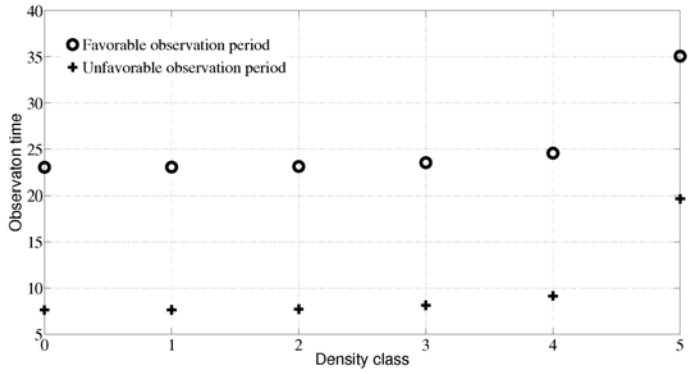


FIGURE 15 – Observation times for the density classes defined in Table 4 for both observation periods. The observation time is given when a single weed species is present into the quadrat, the field is used for winter wheat and managed with the fourth cropping system.

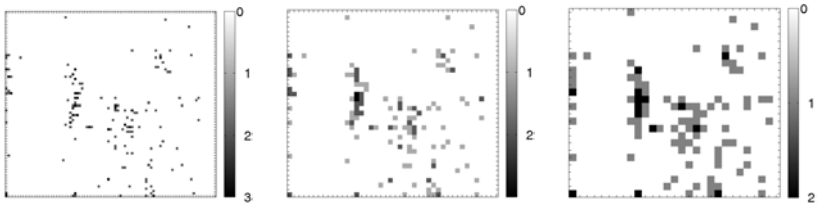
The crop nature also impacts the observation time. Logically, the observation time is shortest when there is no crop into the field. The sorting, in ascending order of observation time, of the crop is the following : no crop, winter barley, winter wheat, winter horse bean, winter rape, sorghum, spring barley and corn.
The presence of an additional weed species into the quadrat increases the observation time by 20.19 seconds. The walking speed of the observer is estimated to 1.52 km/h^{-1} in the favorable observation period and to 1.26 km/h^{-1} in the unfavorable observation period.

11.2 Model Selection

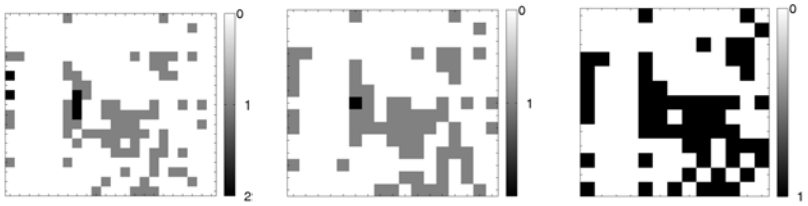
The selection model procedure was run independently for the 22 weeds species of the dataset presented in Section 10.4.1. As expected, there is no universal model describing weeds density repartition for all weeds species and/or quadrat sizes. The selected models for the different quadrat sizes and species are presented in Tables 6 and 7. When the selected model is not given, it corresponds to a technical limit of the simplex search method [45] we used to maximize the pseudo-likelihood. In these cases the BIC scores could not be calculated.

For all weeds species, the selected models change with quadrat size. This evolution of the best model can, most of the time, be explained by the effect of the quadrat size on the weed map. As illustrated in Figures 16 and 17, when the quadrat sizes increases the weeds map is smoother as the information on each quadrat is averaged. In Figure 17 we can see that the proportion of the field in density class 0 decreases with the quadrat size, replaced by density class 1 (or more) quadrats. As a consequence, models with smooth spatial variation (i.e. models of type II and IV) are progressively preferred. This is true for all weed species. For quadrats of $0.36m^2$ only 10% of the selected models are of type II or IV , over all weeds species. For the other quadrat sizes they are selected in 78%, 100%, 100%, 85%, 90% and 95% of the cases. These results are given in ascending order of quadrat size. For the rest of the paper, if no confusion is possible, the results will always be given in ascending order of quadrat size. Indeed, as long as the quadrat size increases the number of density classes present on the map decrease. When only two density classes are present, models of type II and IV are respectively equivalent to models of type I and III. This is illustrated in Figure 16. For quadrats of $12.96m^2$ and $17.64m^2$ area, only density classes O and 1 are present. In both cases, model M5 is selected which is equivalent to model M7. Thus for large quadrat sizes the percentage of selected models with "real"smooth spatial variation is lower. Models with "real"smooth spatial variation are selected in 10%, 73%, 78%, 77%, 59% and 51% of the cases. We can also note that models with external field and with or without smooth spatial variation (i.e. models of type II or I) are practically equivalent when two density classes are present in large proportion and a third one in small proportion (i.e. observed in one or two quadrats). Here by practically equivalent we mean that probabilities of each possible map are very closed. If we consider that this models are also equivalent the number of selected models with "real"smooth spatial variation again decrease. They will be selected in 10%, 73%, 72%, 76%, 54%, 37% and 36% of the cases.

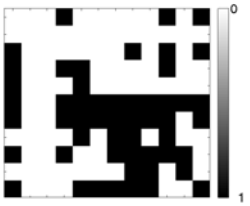
Finally when the quadrat size increases the weed maps are simpler to describe, because there is less spatial variation and fewer density classes are present. Thus the simplest models with no external field (i.e. models of type III and IV) are progressively chosen. They are selected in 18%, 14%, 51%, 82%, 68%, 55%, 68% of the cases over all weeds species. Finally we can notice that the proportion of selected models which are non-isotropic increases slightly with the quadrat size. They are selected in 19%, 18%, 19%, 32%, 23%, 23%, 18% of the cases over all weeds species. This conclusion can vary for different weeds species and it seems that no general interpretation can be made. For example for *fallopia convolvulus* (see Figure 17), the selected models are alternatively non-isotropic and isotropic.



(a) M2, $0.36m^2$ (b) M3, $1.44m^2$ (c) M3, $3.24m^2$



(d) M7, $5.76m^2$ (e) M8, $9m^2$ (f) M5=M7, $12.96m^2$

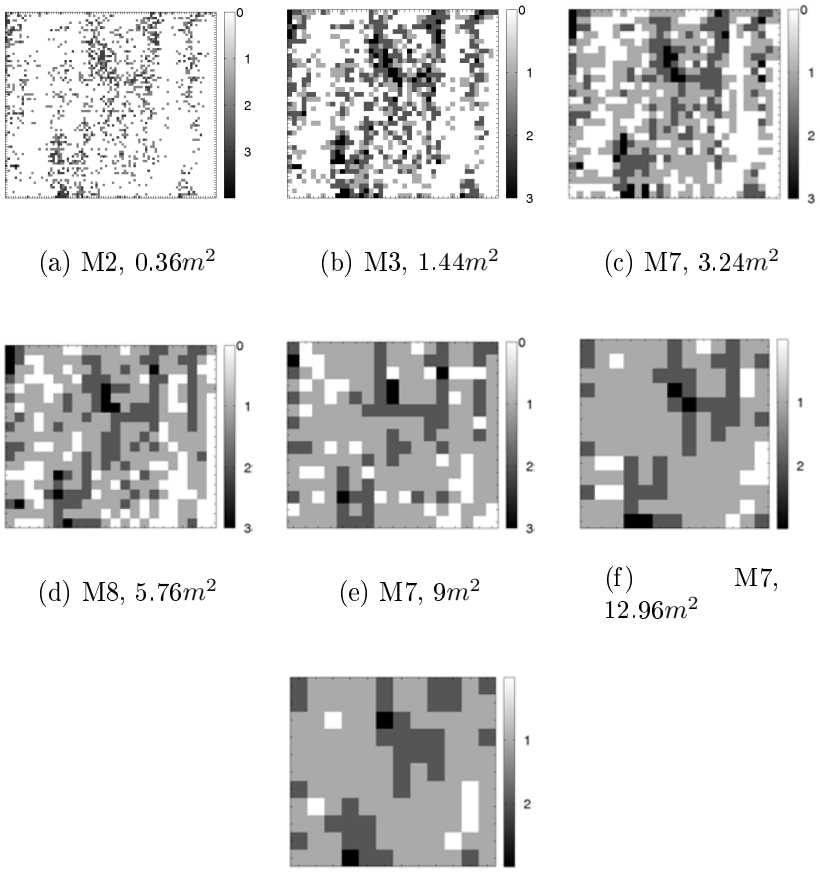


(g) M5=M7, $17.64m^2$

FIGURE 16 – Weed maps for *chaenorrhinum minus (L.) lange* and selected models for all quadrat sizes.

Quadrat size	0.6	1.2	1.8	2.4	3	3.6	4.2
Sonchus spp.	M1	M4	M3	M7	M7		
Thlaspi arvense L.	M5	M5	M5*	M7	M5	M5*	M5*
Viola arvensis murray	M1	M3	M3	M1*	M1*	M1*	M5*
Valerianelle locusta (L.) laterrade	M5	M1	M1*	M5*	M1*	M1*	M1*

TABLE 6 – Selected models for the first four weeds species depending on quadrat size. A star indicates that there are only two different density classes on the weed map, so models M1 and M5 are respectively equivalent to models M3 and M7.



(a) M2, $0.36m^2$ (b) M3, $1.44m^2$ (c) M7, $3.24m^2$

(d) M8, $5.76m^2$ (e) M7, $9m^2$ (f) M7, $12.96m^2$

(g) M8, $17.64m^2$

FIGURE 17 – Weed maps for *fallopia convolvulus L.* and selected models for all quadrat sizes.

Quadrat size	0.6	1.2	1.8	2.4	3	3.6	4.2
Aethusa cynapium L. Parley	M1	M3	M3	M3	M1	M4	M3
Alopecurus agrestis L.	M1	M3	M7	M7	M7	M7	M7
Anagallis arvensis L.	M4	M3	M7	M8	M6		
Brassica napus L. spp. napus	M1	M4	M4	M8	M8		
Chaenorrhinum minus (L.) lange	M2	M3	M3	M7	M8	M5*	M5*
Cirsium arvense (L.) scopoli	M3	M3	M7	M7	M7	M7	M7
Convolvulus arvensis L.	M1	M3	M7	M7	M7	M7	M7
Euphorbia exigua L.	M1	M3	M3	M8	M3	M7	M5
Fumaria officinalis L.	M5	M5*	M5*	M5*	M5*	M5*	M5*
Galium aparine L.	M1	M3	M3	M3	M3	M4	M3
Geranium spp.	M5	M5	M5*	M5*	M5*	M5*	M5*
Lactuca serriola L.	M1	M1	M1*	M1*	M1*	M1*	M1*
Silene latifolia poiret	M1	M1	M7	M7	M5*	M5	M5*
Myosotis arvensis (L.) hill	M1	M3	M8	M8	M7	M3	M7
Picris hieracioides L.	M2	M4	M4	M7	M3	M4	M8
Polygonum aviculare L.	M1	M4	M4	M8	M8	M8	M8
Fallopia convolvulus L.	M2	M3	M7	M8	M7	M7	M8
Sinapsis arvensis L.	M1	M3	M7	M8	M8	M8	M8

TABLE 7 – Selected models for the last eighteen weeds species depending on quadrat size. A star indicates that there are only two different density classes on the weed map, so models M1 and M5 are respectively equivalent to models M3 and M7.

11.3 Comparison of sampling strategies

11.3.1 Selected Weed for the comparison

In this section we present the density repartition models and weed maps used for the comparison of the sampling strategies.

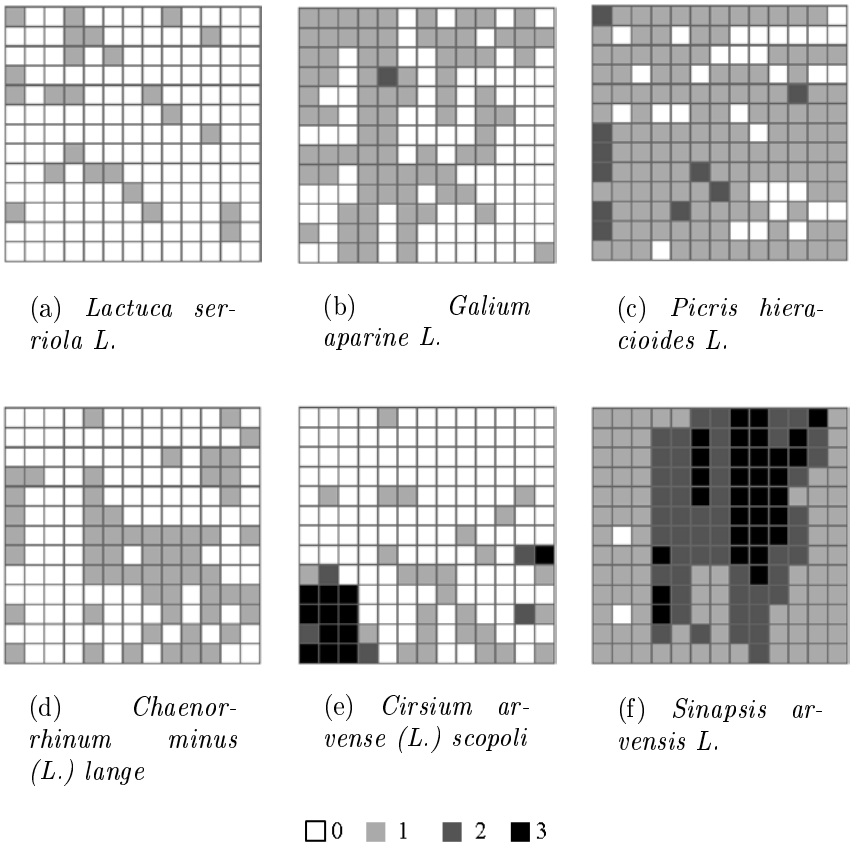
First the comparison was achieved for quadrat of $12.96m^2$ area in order to be able to use the cost model proposed in Section 10.2 and define the cost constraint in term of available time for sampling. We recall that the cost model gives the observation time for quadrat of $16m^2$ area. Here we used smaller quadrats but we supposed that the observation time would be similar.
The models we have chosen correspond to the best models for this quadrat size, for the weeds species *Lactuca serriola L.*, *Galium aparine L.*, *Picris hieracioides L.*, *Chaenorrhinum minus (L.) lange*, *Cirsium arvense (L.) scopoli*, and *Sinapsis arvensis L.*. These weeds species were selected because (i) their density repartition maps exhibit a variety of spatial patterns and (ii) the selected models are all different.
The models selected for each weed species and their parameters values are summarized in Table 8. The complete maps used for parameters estimation and comparison are presented in Figure 18.
The weeds species *Picris hieracioides L.* and *Sinapsis arvensis L.* are present practically in every quadrats of the field. On the contrary *Lactuca serriola L.* is rare and seems to be distributed randomly over the field. No patches are observed. *Cirsium arvense (L.) scopoli* is also rare but a patch of density class 3 is located on the bottom left part of the field. For the maps of *Galium aparine L.* and *Chaenorrhinum minus (L.) lange* density class 0 and 1 are present in close proportion. *Sinapsis arvensis L.* is the weed species present in the larger proportion. A large patch of density class 3 is located in the centre of the field.

Weed species	Ω	Model	α	β or (β_t, β_o)
Lactuca serriola L.	$\{0,1\}$	M1	(0 ;-1.66)	0.05
Galium aparine L.	$\{0,1,2\}$	M2	(0,-0.03,-3.58)	(0.71,0.12)
Picris hieracioides L.	$\{0,1,2\}$	M4	(0,0.89,-0.75)	(-0.47,4.01)
Chaenorrhinum minus (L.) lange	$\{0,1\}$	M5	$\emptyset$	0.378
Cirsium arvense (L.) scopoli	$\{0,1,2,3\}$	M7	$\emptyset$	3.86
Sinapsis arvensis L.	$\{0,1,2,3\}$	M8	$\emptyset$	(9.42,2.39)

TABLE 8 – Density class repartition models for the six weeds species used to compare the sampling strategies

11.3.2 Time constraint

Simulated maps We first compare the two adaptive sampling strategies δ_{LSDP}, δ_{BP-max} and the random policy δ_{Rand} using a fixed sampling budget defined by the available time for sampling. For all strategies we fix it to 2 hours and 30 minutes. This is only the time available for quadrats observations, the moving time is not taken into account (see Section 10.2). Using the cost model described in Section 10.2, the mov-



(a) *Lactuca serriola L.*

(b) *Galium aparine L.*

(c) *Picris hieracioides L.*

(d) *Chaenorrhinum minus (L.) lange*

(e) *Cirsium arvense (L.) scopoli*

(f) *Sinapsis arvensis L.*

FIGURE 18 – Complete weed map for the weed species used to compare the different sampling strategies.

ing times can also be computed a posteriori for all strategies. Results are presented in Table 9. The moving time for the random strategy is around 35 minutes, it is logically approximately constant for all models. For δ_{LSDP} it is also around 35 minutes. We can observed that for model M1 the strategy δ_{LSDP} spends exactly 39 minutes for moving on each of the 2000 simulated maps. On simulated map this strategy acts as a static strategy. δ_{BP-max} generally requires fewer move into the field. The moving time varies greatly among the models. For model M4 it is estimated to 18 minutes and 34 minutes for model M1. Thus when the moving time is taking into account, the time needed to apply these strategies is between 2H50 and 3H05.

Model	δ_{LSDP}	δ_{BP-max}	δ_{Rand}
M1	39 (0)	34.2 (6.3)	34.8 (2.9)
M2	36.3 (2.3)	22.29 (2.8)	35 (2.9)
M4	37.1 (3)	17.9 (3.2)	35 (2.9)
M5	38 (0.5)	29.4 (2.6)	35 (2.9)
M7	38.6 (3)	31 (3.1)	35 (2.9)
M8	30 (2.8)	25.9 (4.1)	34.9 (2.9)

TABLE 9 – Mean moving times in minutes (over 2000 simulated maps) for the strategies δ_{LSDP}, δ_{BP-max} and δ_{Rand} and all models. The standard deviation is given in brackets.

The average percentage of well reconstructed quadrats for the different models are presented in Table 10.
First we can note that there are no big differences between the strategies values. The largest absolute difference is 1.6%, observed for the model M8 between the random policy and δ_{BP-max}. Apart for model M1 the largest difference between adaptive and random strategies is about 1%. For model M1 all strategies have approximately the same values.
Even if the absolute difference between the strategy value is small, the random strategy is always dominated by the adaptive strategy δ_{BP-max}. And except for model M8 it is also the case with δ_{LSDP}.

We can also compare the average number of observations of all sampling strategies. Results are presented in Table 11. As expected the average number of observations is stable for the random strategy. We can see that δ_{LSDP} is the sampling strategy which generally has the highest number of observations. This is also logical since the other

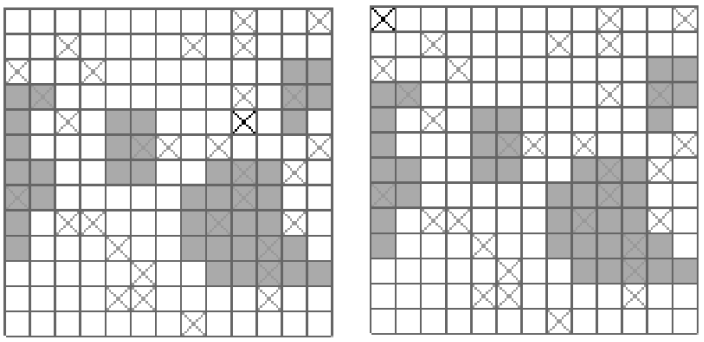
Model	δ_{LSDP}	δ_{BP-max}	δ_{Rand}
M1	**89.2%**	89%	89%
M2	**69.1%**	68.8%	67.6%
M4	81.5%	**82.2%**	80.9%
M5	**67.9%**	67.8%	66.8%
M7	**65.5%**	64.5%	64.4%
M8	77.5%	**80%**	78.4%

TABLE 10 – Average percentage of well reconstructed quadrats computed on 2000 simulated maps for sampling strategies δ_{LSDP}, δ_{BP-max} and δ_{Rand}.

Weed species	δ_{LSDP}	δ_{BP-max}	δ_{Rand}
Lactuca serriola L.	40 (0)	35.3 (0.6)	37.1 (0.8)
Galium aparine L.	38.7 (0.6)	36.8 (0.6)	37.1 (0.8)
Picris hieracioides L.	36.2 (0.9)	34.1 (0.6)	37.1 (0.8)
Chaenorrhinum minus (L.) lange	38.8 (0.5)	36.5 (0.7)	37.1 (0.8)
Cirsium arvense (L.) scopoli	37.3 (0.9)	35.4 (0.8)	37.1 (0.8)
Sinapsis arvensis L.	36.9 (0.9)	36.2 (0.8)	37.1 (0.8)

TABLE 11 – Average number of observations (over 2000 simulated maps) for the strategies δ_{LSDP}, δ_{BP-max} and δ_{Rand} and all models. The standard deviation is given in brackets.

sampling strategies does not take into account the possible cost of the future observations. This is illustrated in Figure 19. The strategy δ_{LSDP} was applied on the real map of *Chaenorrhinum minus (L.) lange* and based on the corresponding best model (see Table 8). Observation was stopped when the remaining time for sampling was about 30 minutes. For this budget, the next quadrat to observe is located line 5 and row 10. This quadrat is marked with a black cross on Figure 19.a while the quadrats observed previously are marked with gray crosses. Then we voluntarily reduced the remaining budget to 15 minutes. For this new budget, the next quadrat to observe is located line 1 and row 1, see Figure 19.b. For the other strategies, the next quadrat to observe will be the same whatever the time remaining for sampling.



(a) (b)

FIGURE 19 – Next observed quadrat with sampling strategy δ_{LSDP} when the remaining time for sampling is about 30 minutes in case (a) and 15 minutes in case (b). Previously observed quadrats are marked with gray crosses and the next quadrat to observe by a black cross. In the background is the reconstructed map. The strategy and the reconstructed map are computed using model M5 with parameters fitted on the observed map of *Chaenorrhinum minus (L.) lange*. The model parameters are presented in Table 8.

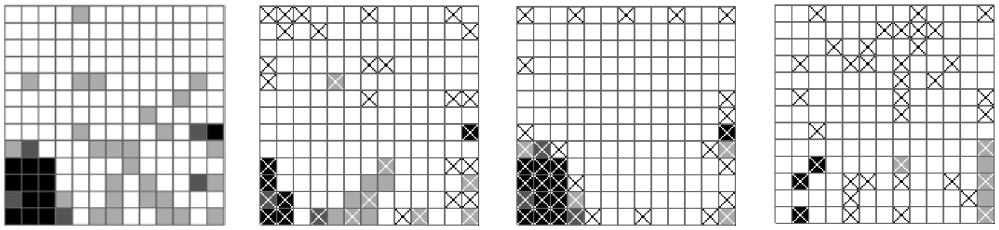
Observed maps We now compare these strategies on the observed weed maps. The number of well reconstructed quadrats for all weeds species and density classes are presented in Table 12. The observed maps, locations of observations and reconstructed

maps for all sampling strategies are presented in Figures 21 and 22.

First we can observe that the absolute difference between the percentage of well reconstructed quadrats of the adaptive strategies and the random strategy is higher than with comparison using simulated maps. The largest difference is now of 10.7%, between δ_{LSDP} and δ_{BP-max} for the *Sinapsis arvensis L.* map.
Generally the relative performances of the three strategies are the same that with simulated maps. This is not true for *Sinapsis arvensis L.* where the order of performance is reversed.
For the *Lactuca serriola L.* map the three strategies give similar results, even if for the random strategy WR1 is a beat higher. We can note that using model M1 with parameters described in Table 8, the reconstruction method is very simple. It consists in reconstructing the quadrats where no observations are available at the density class 0. Then for such weeds repartitions, the simplest strategy δ_{Rand} can be preferred and only the quadrats where the weed is observed have to be stored.
For the *Galium aparine L.* map, adaptive strategies become more efficient. We can see that the left part of the map was more explored by the adaptive strategies. This is beneficial since the weed species is more present in this area. And both adaptive strategies allow to find approximately correct spatial patterns in this area. For δ_{BP-max} when a quadrat is observed in density class 1, the sampling strategy consists in finding the boundaries of patches of density 1. For δ_{LSDP} this is similar but softened. This strategy also tries to keep some time to explore other areas of the map. This explains why δ_{LSDP} leads to higher performance. For the random strategy, the reconstructed map looks really different from the observed map. Unlike for the adaptive strategies the boundaries between quadrats in density classes 0 and 1 are never estimated. So the reconstructed map over estimates the number of quadrats in density class 1 around an observed quadrat at this density class.
The same conclusion generally holds for the other weeds species. For *Picris hieracioides L.* δ_{BP-max} allows the largest number of well reconstructed quadrats. For this strategy the observed quadrats are grouped in different areas of the field, which increases WR0 and WR1.
For *Chaenorrhinum minus (L.) lange* the reconstructed maps of both adaptive strategies precisely located the regions where *Chaenorrhinum minus (L.) lange* is present, even though these regions are better delimited by δ_{LSDP}. The better performance of the adaptive strategy is due to a larger number of well reconstructed quadrats in density class 1. WR0 is approximately constant for all sampling strategies.
For *Cirsium arvense (L.) scopoli* only the two adaptive strategies recover the patch of density class 3. Unlike the other maps where a large proportion of quadrats are in density class 0 (e.g. *Lactuca serriola L.*, *Galium aparine L.*, *Chaenorrhinum minus (L.) lange*), WR0 is surprisingly small (about 55%). The main reason is the biased behaviour of model M7. In this case this simple model exhibits its limit for reconstruction. We saw in Section 10.1.3 that density classes 1 and 2 are the most probable when no observation are available. The number of quadrats in density class 1 is thus largely overestimated whatever the strategy. For comparison, we ran the same reconstruction experiment with model M1 ($\alpha = (0, -0.327, -1.54, -1.07)$ and $\beta = 0.68$) instead of model M7. We also computed the three strategies with model M1. WR is respectively equal to 82.25%, 85.27% and 76.92% for δ_{LSDP}, δ_{BP-max} and δ_{Rand}. All strategies give better results. This appears logical since model M1 and more generally model with external field (i.e.

of type I and II), include informations about the proportions of the different density classes within the field. The reconstructed maps and the locations of observations are presented Figure 20.
In certain cases type IV models also perform well, like for *Sinapsis arvensis L.* where δ_{LSDP} leads to the largest number of well reconstructed quadrats. In addition this is the second larger number of well reconstructed quadrats over all weeds species. δ_{Rand} is the second best strategy and the general spatial pattern of the map is respected for these two strategies. Surprisingly δ_{BP-max} is the worst strategy. It is the only strategy which does not find the patch of density class 3.

Cirsium arvense (L.) scopoli with model M1



(a) Observed Map (b) δ_{LSDP} (c) δ_{BP-max} (d) δ_{Rand}

FIGURE 20 – Location of observations and reconstructed map when following δ_{LSDP}, δ_{BP-max} or δ_{Rand} on the observed map of *Cirsium arvense (L.) scopoli*. The adaptive strategies and the reconstructed map are computed using model M1 with $\alpha = (0, -0.327, -1.54, -1.07)$ and $\beta = 0.68$.

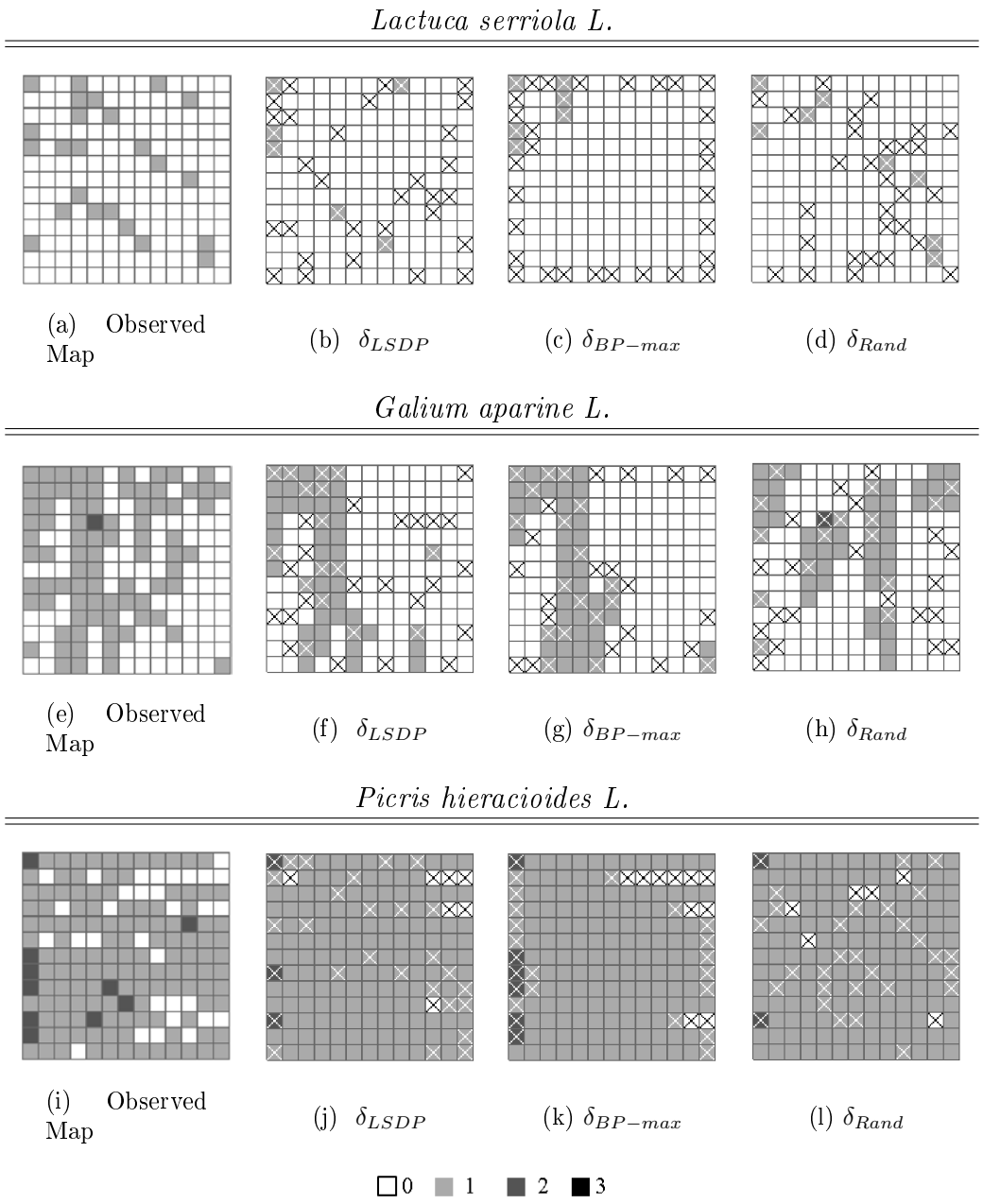


FIGURE 21 – Locations of observations (quadrats with crosses) and reconstructed maps when following δ_{LSDP}, δ_{BP-max} or δ_{Rand} on the observed map of the first three selected weed species. The cost constraint is expressed in terms of time.

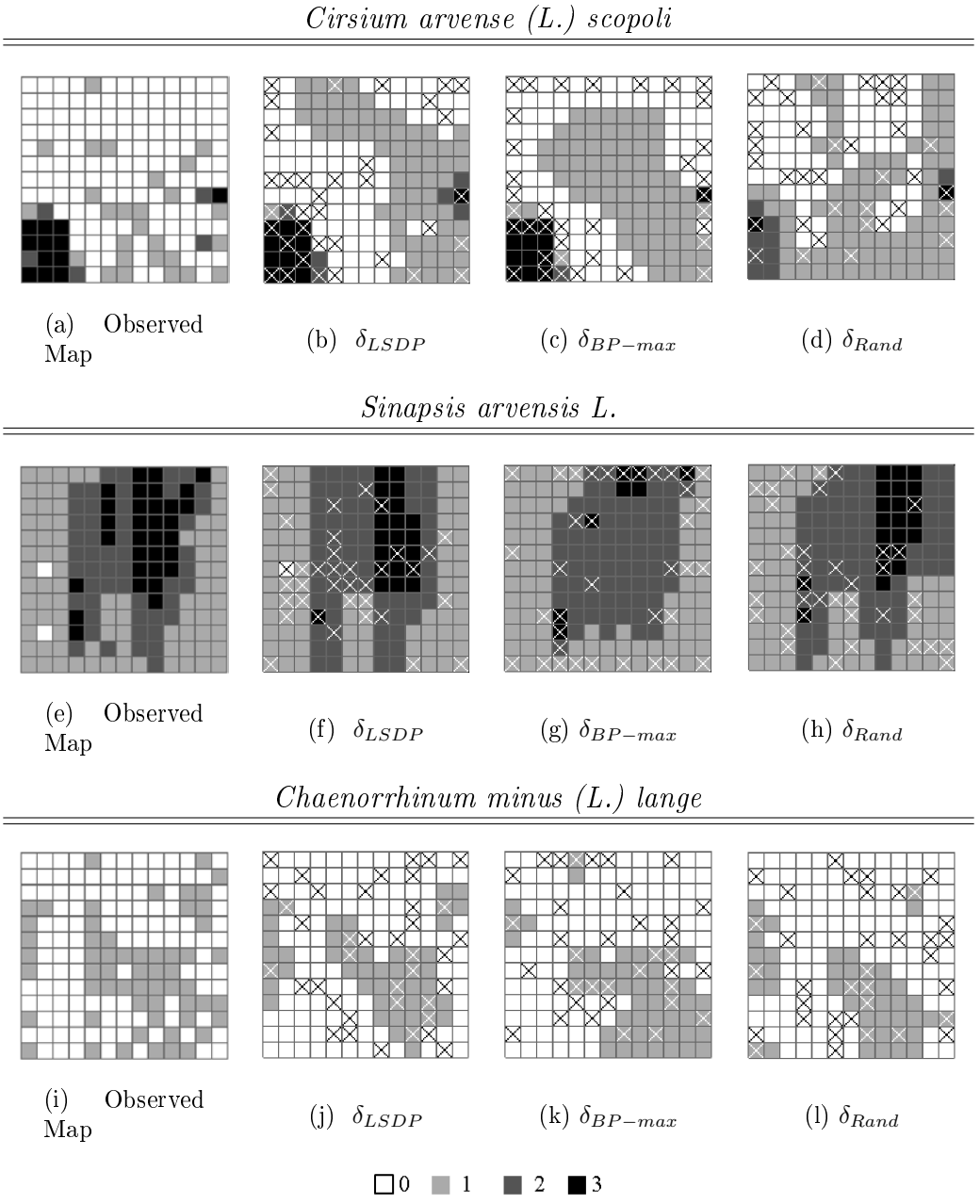


FIGURE 22 – Locations of observations (quadrats with crosses) and reconstructed maps when following δ_{LSDP}, δ_{BP-max} or δ_{Rand} on the observed map of the three last selected weed species. The cost constraint is expressed in terms of time.

Weeds Species	δ_{LSDP}					δ_{BP-max}					δ_{Rand}				
	WR	WR0	WR1	WR2	WR 3	WR	WR0	WR1	WR2	WR3	WR	WR0	WR1	WR 2	WR 3
L. serriola L.	89.3	100	25			89.3	100	25			**90.5**	**100**	**33.3**		
G. aparine L.	**74**	**95.3**	52.1	0		73.4	94.2	**52.4**	0		65.7	83.7	46.3	**100**	
P. hieracioides L.	81.1	21.9	**100**	30		**84.6**	**31.25**	**100**	**60**		79.9	18.7	**100**	20	
C. minus (L.) lange	**78.1**	**87.4**	60.3			77.5	85.6	**62.1**			74	86.5	50		
C. arvense (L.) scopoli	59.2	**57.6**	48.1	**60**	**100**	**60.3**	56.8	66.7	20	**100**	56.2	56	**81.5**	20	16.7
S. arvensis L.	**85.8**	**50**	90.5	**92.1**	**65.6**	75.1	0	**91.7**	82.3	25	76.3	0	82.1	84.3	53.1

TABLE 12 – Number of well reconstructed quadrats of each density class when applying δ_{LSDP}, δ_{BP-max} or δ_{Rand} on the observed map of the six selected weed species with time constraint.

11.3.3 Fixed number of observations

In this section we compare the three strategies δ_{LSDP}, δ_{BP-max} and δ_{Rand} with 8 static strategies named, δ_{Reg_1}, δ_{Reg_2}, δ_{Reg_3}, δ_{Reg_4}, δ_Z, δ_{W_1}, δ_{W_2}, δ_{Star} presented in Figure 12. The budget constraint is now expressed in terms of possible number of observations, which is fixed to 23 (about 14% of the quadrats are observed).
The percentage of well reconstructed quadrats are presented in Table 13 for all strategies. The locations of observations and the reconstructed maps of the two adaptive and best static strategies are presented in Figures 24 and 25. The percentage of well reconstructed quadrats in all density classes are presented in Table 14 for these strategies.
In Figure 23 we presented the box-plot of the absolute differences between the percentage of well reconstructed quadrats of the best strategy and that of all other strategies. We can note that δ_{Star} is generally the best strategy to follow. δ_{Reg_2} and δ_{LSDP} also provide good results. Unlike the other static strategies, δ_{Star} and δ_{Reg_2} realize a good compromise between exploration of the field, with distant observed quadrats, and delimitation of the boundaries between patches of different density classes, with closed observed quadrats. The adaptive strategies provide good results only for the three first weeds species, when models with external field are used.

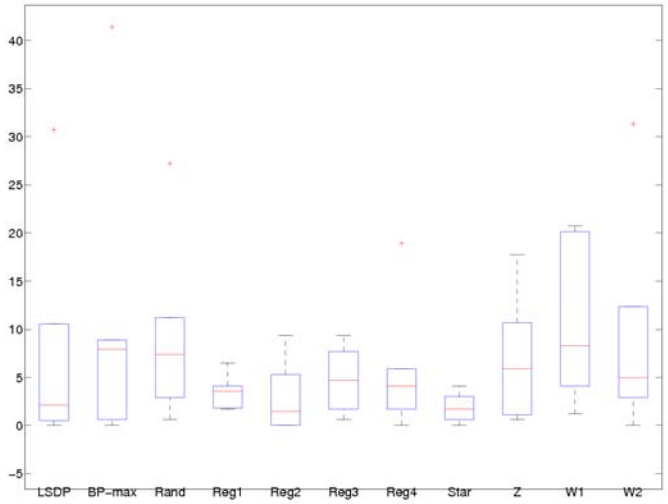


FIGURE 23 – Box-plot of the differences between the percentage of well reconstructed quadrats of the best strategy and that of each strategy, for all weed species.

The advantage of the adaptive strategies is not demonstrated in this case. δ_{Star} provides good performances over all weeds species. Except for *Chaenorrhinum minus (L.) lange* and *Cirsium arvense (L.) scopoli* the strategy δ_{LSDP} also provides good results, close to δ_{Star} and often better. We can expect that the use of models with external field can improve the efficiency of δ_{LSDP} for *Chaenorrhinum minus (L.) lange* and *Cirsium arvense (L.) scopoli*. But this will certainly also improves the results of δ_{Star} and the other strategies.
We can also note that the efficiency of the other static strategies vary greatly among weeds maps. δ_{Reg_2} is in general the best regular sampling strategy. The other strategies δ_{Rand}, δ_Z, δ_{W_1} and δ_{W_2} seem to be less appropriate.
For the best sampling strategies, the general spatial pattern of the observed map is globaly respected. For the map of *Lactuca serriola L.* and *Picris hieracioides L.*, this is easy because a density class dominate the other. This information is carried by the

external field α, making the extrapolation method efficient [15].
For the map of *Sinapsis arvensis L.* model M8 particularly well fitted the observed map and thus reconstruction was good. For the other maps the task is more complex and the reconstructed maps generally look different from the observed maps for all strategies.

	δ_{LSDP}	δ_{BP-max}	δ_{Rand}	δ_{Reg_1}	δ_{Reg_2}	δ_{Reg_3}
L. serriola L.	87.6	87.6	87.6	86.4	**88.2**	87.6
G. aparine L.	**72.2**	65.1	63.3	69.2	66.9	65.7
P. hieracioides L.	81.1	**81.6**	78.7	77.5	78.7	78.7
C. minus (L.) lange	63.9	66.9	68.6	70.4	**74.5**	72.8
C. arvense (L.) scopoli	45	34.3	48.5	74	66.3	66.3
S. arvensis L.	72.8	67.4	65.1	69.8	**76.3**	68.6

	δ_{Reg_4}	δ_{Star}	δ_Z	δ_{W_1}	δ_{W_2}
L. serriola L.	**88.2**	87.6	87.6	87	**88.2**
G. aparine L.	66.3	69.2	61.5	62.1	59.8
P. hieracioides L.	78.1	79.9	80.5	77.5	78.7
C. minus (L.) lange	72.8	72.8	68	68	68
C. arvense (L.) scopoli	56.8	**75.7**	58	55.6	44.4
S. arvensis L.	71.6	72.2	71	55.6	72.8

TABLE 13 – Percentage of well reconstructed quadrats on observed maps, when the number of observations is fixed to 23 for all strategies.

15. We can expect that if type III or IV models are used instead, the efficiency of all the sampling strategies will decrease.

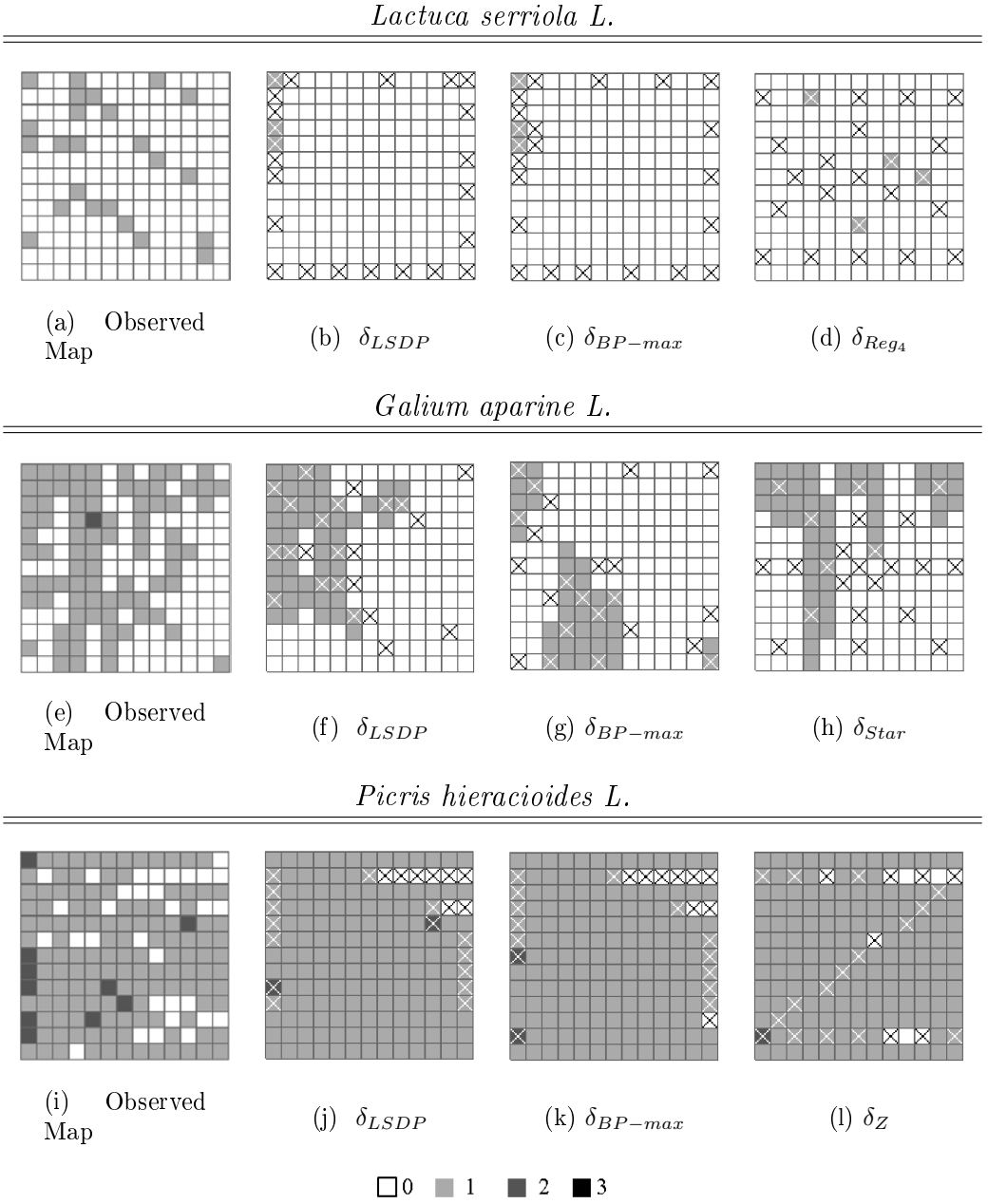


FIGURE 24 – Locations of observations and reconstructed maps when following δ_{LSDP}, δ_{BP-max} and the best static strategy on the observed maps of the first three selected weeds species. All strategies were allowed 23 observations.

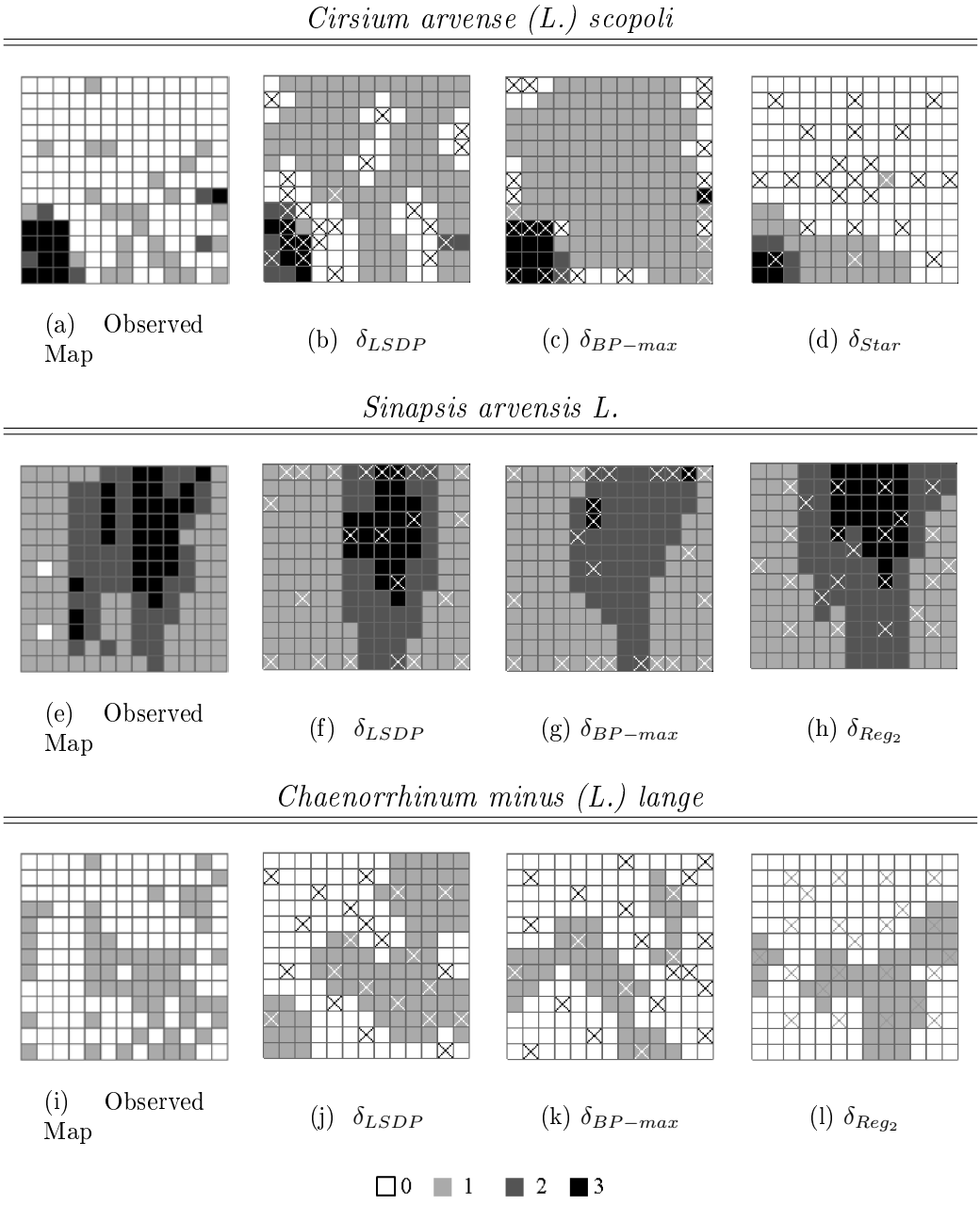


FIGURE 25 – Locations of observations and reconstructed maps when following δ_{LSDP}, δ_{BP-max} and the best static strategy on the observed maps of the three last selected weeds species. All strategies were allowed 23 observations.

Species	δ_{LSDP}					δ_{BP-max}					Best static				
	WR	WR0	WR1	WR2	WR 3	WR	WR0	WR1	WR2	WR3	WR	WR0	WR1	WR 2	WR 3
L. serriola L.	87.6	**100**	12.5			87.6	**100**	12.5			**88.2**	**100**	16.7		
G. aparine L.	**72.2**	84.9	**59.7**	0		65.1	**91.9**	37.8	0		69.2	88.4	50	0	
P. hieracioides L.	81.1	25	**100**	**20**		**81.6**	**28.1**	**100**	**20**		80.5	**28.1**	99.2	10	
C. minus (L.) lange	63.9	62.2	**67.2**			68.6	81.4	44.8			**74.5**	**84.7**	55.2		
C. arvense (L.) scopoli	45	38.4	66.7	**60**	58.3	34.3	17.6	**85.2**	20	**100**	**75.7**	**92**	37	0	25
S. arvensis L.	72.8	0	88.1	51	**71.9**	67.4	0	**94**	62.7	9.4	**76.3**	0	82.1	**74.5**	68.7

TABLE 14 – Number of well reconstructed quadrats of each density class when applying δ_{LSDP}, δ_{BP-max} or the best static strategy on the observed map of the six selected weed species. The sampling cost is expressed in terms of number of observations. It is fixed to 23.

11.3.4 When no data are available to select the model

In previous experiments, a best model was selected for each weed species according to the corresponding observed maps and parameters were fitted on this maps. In practice, when no data are available to estimate and select a best model, methods have to be found in order to select a density class repartition model and fix the values of the model parameters. This is what we will focus on in this section.

We have seen in previous sections that the use of an external field can increases the reconstructed map quality since it indicates which density class is present in largest proportion. For quadrat size 12.96m^2, model M1 is the most often selected model among models with external field. So, we propose to use model M1 whatever the studied weed species.
For parameters estimation, the most common method consists in fitting the parameters using incomplete observations. In practice, we have notice that this can leads to wrong models, highly different from those estimated from complete maps. In addition, when initial observations are needed for parameters estimation, less time is available for sampling in order to optimize the reconstructed map quality. Here we propose to fix the parameters values using the following rules.
First, β is fixed to 0.58, which is the average value of this parameter for all weeds species for model M1.
The value of the external field α is more complicated to fix without any information on the weed species repartition. We propose to use the two following informations : (i) which are the different density classes present within the field and (ii) which is the density class present in largest proportion. Then if K is the highest density class present within the field and $l \in \{0, \dots, K\}$, the density class present in largest proportion, the external field, $\alpha = (\alpha_0, \dots, \alpha_K)$, is equal to :

$$\forall k \in \{0, \dots, K\}, \qquad \alpha_k = \begin{cases} 0 & \text{if } k = l \\ -0.2 & \text{if } k \neq l \end{cases}$$

Thus, in the non explored areas of the field, quadrats are reconstructed to the density class which appears to be the most frequently present. And when a quadrat is observed in density class $k \in \{0, \dots, K\}$, the influence on the neighboring quadrats is averaged among weeds species.

As in Section 11.3.3 we compare the two adaptive strategies, the random strategy and the eights static strategies only on observed maps. Results are presented in Table 15.

First we can note that δ_{BP-max} provides good results for all weeds species. It leads to higher performances in half of the case. Surprisingly, this is not the case for δ_{LSDP}, which in general leads to poor results, like the static strategies.
We can also note that for *Cirsium arvense (L.) scopoli*, all strategies leads to better results than with the selected model and the estimated parameters on the observed map. For the other weeds species, all strategies generally leads to lower performances. In addition, we can see in Figures 26 and 27 that the reconstructed map of the adaptive

	δ_{LSDP}	δ_{BP-max}	δ_{Rand}	δ_{Reg_1}	δ_{Reg_2}	δ_{Reg_3}
L. serriola L.	87.6	88.2	**88.7**	86.4	87	87.6
G. aparine L.	62.7	**64.5**	58.6	63.3	63.3	63.3
P. hieracioides L.	**79.3**	78.7	74.5	76.3	76.9	77.5
C. minus (L.) lange	68.6	70.4	71.6	72.2	69.8	**73.4**
C. arvense (L.) scopoli	81.6	**83.4**	79.3	77.5	75.1	78
S. arvensis L.	66.3	**76.3**	67.4	71	75.1	65.1

	δ_{Reg_4}	δ_{Star}	δ_Z	δ_{W_1}	δ_{W_2}
L. serriola L.	88.2	86.4	86.4	87	87.6
G. aparine L.	59.8	60.9	62.7	59.8	55
P. hieracioides L.	78.1	74	74.5	78.1	76.3
C. minus (L.) lange	71	71	69.2	72.8	71.6
C. arvense (L.) scopoli	72.8	76.9	80.4	73.4	72.2
S. arvensis L.	66.9	68.6	70.4	46.7	63.9

TABLE 15 – Percentage of well reconstructed quadrats on the observed maps, when the number of observations is fixed to 23 for all strategies. Adaptive sampling strategies and the extrapolation method used model M1 with parameters fixed with empirical rules.

strategies look really different from the observed maps.
When the parameters are not estimated form a complete map, a larger number of observations is needed to find the correct spatial pattern of the true map.

Finally, we also try to estimate the parameters from the observed quadrats. After applying each strategy the parameters of model M1 were fitted using variational version of the EM algorithm [30] and the new model is used to compute the reconstructed maps. This decrease the performances of all strategies.

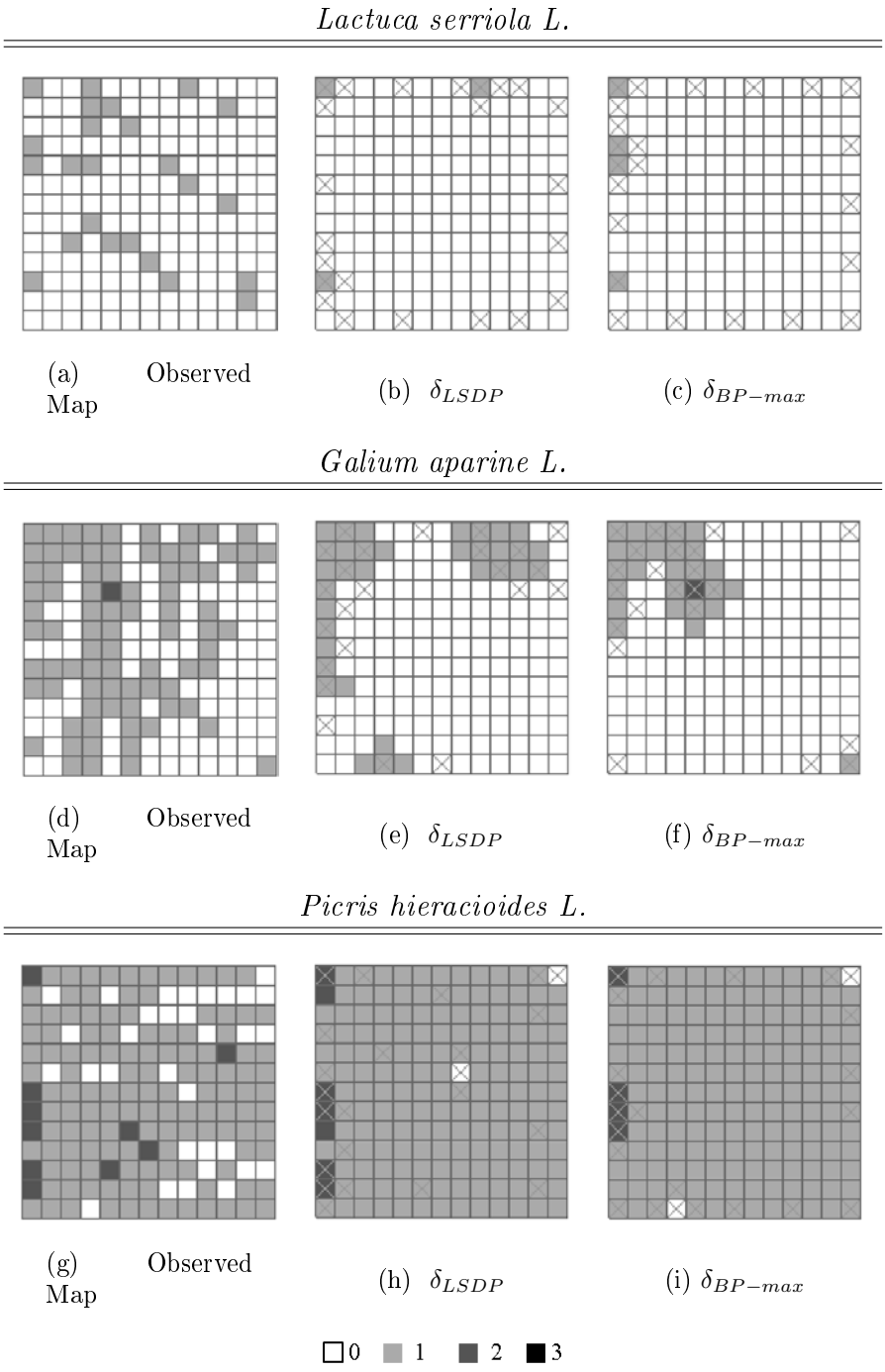


FIGURE 26 – Locations of observations and reconstructed map when following δ_{LSDP}, δ_{BP-max} and the best static strategy on the observed map of the three first selected weed species. All strategies were allowed 23 observations. Model M1 is used for the extrapolation method and computation of the adaptive strategies. Parameters are fixed by empirical rules.

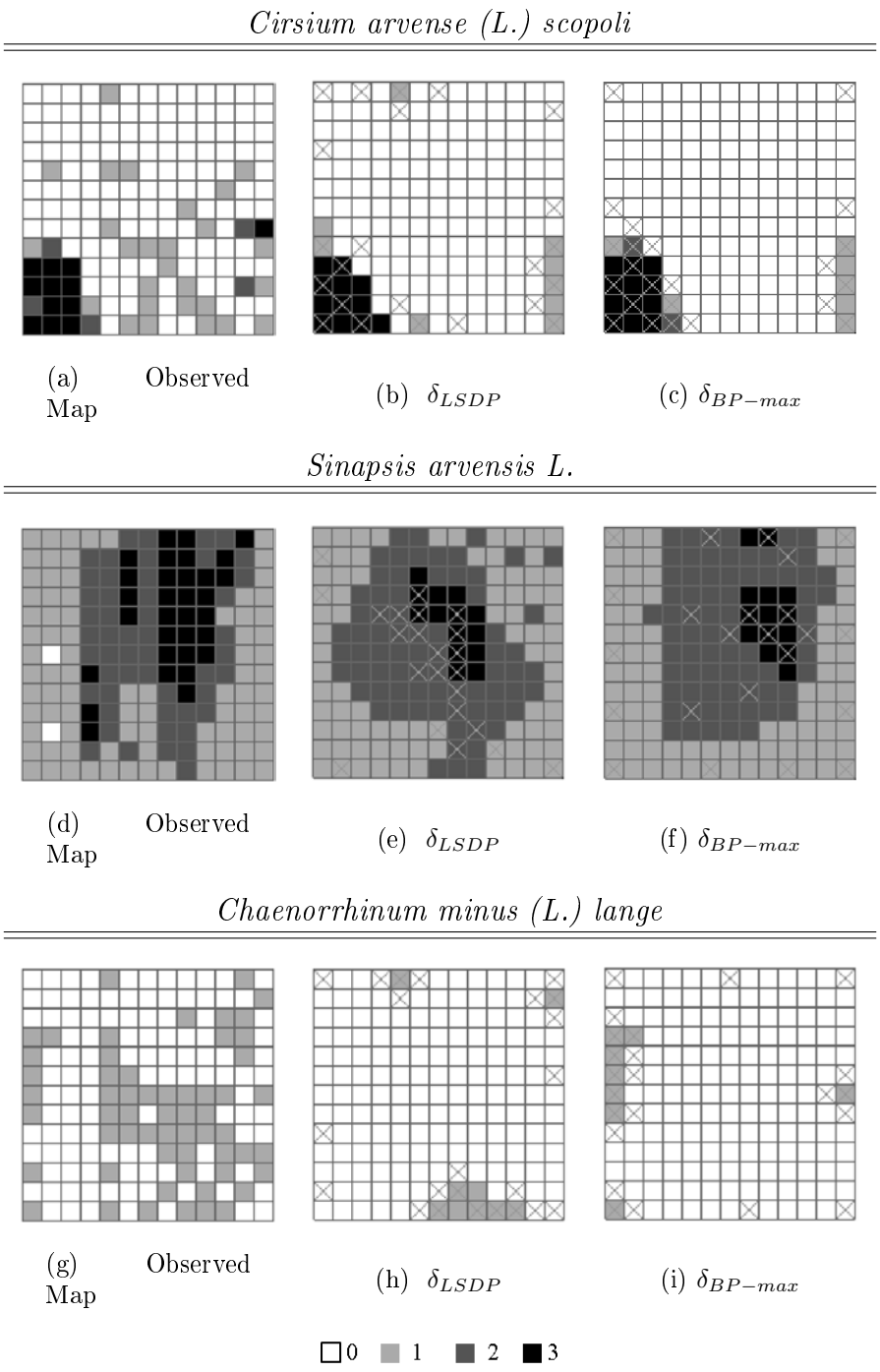


FIGURE 27 – Locations of observations and reconstructed map when following δ_{LSDP}, δ_{BP-max} on the observed map of the three last selected weed species.. All strategies were allowed 23 observations. Model M1 is used for the extrapolation method and computation of the adaptive strategies. Parameters are fixed by empirical rules.

12 Conclusion

The aim of this study was to discuss and propose improvements of current sampling strategies for weed mapping at the field scale. In order to meet this objectives, we have tackled the three following important questions : (i) which weeds sampling cost model, (ii) which models of spatial distribution of weeds density classes and (iii) which method for the design of sampling strategies.

We have proposed to base the weeds sampling cost model on the time needed to apply a sampling strategy. This model represents the time needed to observe one quadrat, plus the time needed to move between quadrats within the field, as a regression involving factors defined as relevant by experts. We have used this model in order to constrain the number of observations a sampling strategy can make. In order to estimate the parameters of this cost model, we have used a data set collected from various sampling strategies under different conditions (i.e. crop type, cropping system, quantity and diversity of weeds into the observed quadrats, crop coverage). Now, the model still has to be validated by dedicated experiments. For instance, by preliminary analysis it seems that the model overestimates the time needed to observe quadrats when the weeds quantity and diversity are high and underestimates it when they are low. However this cost model provides reasonable estimated times and its precision is enough for our purpose of sampling strategies comparison. Having a reliable cost model can also be helpful to compute bounds on the time needed to apply a sampling strategy before actually applying it, which can improve the scheduling of sampling strategies on multiple fields.

We have also proposed several models of spatial distributions of weeds density classes. All of them are based on the Markov random field framework, commonly used for quantitative variables. The models we have proposed consider several features : (i) isotropy/anisotropy, (ii) smooth/abrupt spatial variation and (iii) presence/absence of an external field. We have performed a model selection study over the eight proposed models using the BIC score. The selection was performed using data on the spatial repartition of 22 different weeds species and we have considered 7 different quadrat sizes. We have shown that, among the proposed models, no one model dominates the others. However a general trend can be highlighted, considering quadrat size : When quadrat size increases, the information is “averaged” and thus, spatial variation of weeds density is smoother and the number of density classes decreases. As a consequence, models with smooth spatial variations are progressively chosen more often. Furthermore the weeds maps are simpler to describe, because there is less spatial variation and fewer density classes are present. Thus the simplest models with no external field are progressively chosen. However the objective of map reconstruction is not reflected by the BIC score and the use of an external field brings valuable information when reconstructing maps using few observations. We have notice that the use of such models generally increases the quality of reconstructed maps. Once a model has been chosen, the difficulty is that for practical applications the model parameters have to be estimated using few observations. We have noticde that this leads to poor model quality. So we recommended (i) to preferably use a model with external field and (ii) to apply a set of empirical rules in order to estimate the model parameters. We obtained good results using these two recommendations on real weed sampling and mapping problems.

Finally, we have proposed a mathematical formulation of the weeds sampling problem (using the above cost and spatial repartition models) and we have formulated it as a complex combinatorial optimization problem. In order to solve this problem and to design adaptive sampling strategies, we have considered two different approximate resolution methods from literature : LSDP and BP-max. We have compared these strategies with more classical static sampling strategies on two sampling problems, where the cost constraint is defined either in terms of available time for sampling or in terms of allowed number of observations. We have shown that these strategies offer different compromises between field exploration and patch boundary detection sampling efforts : the strategy δ_{LSDP} realizes more exploration than the strategy δ_{BP-max} in general. The interest of using adaptive strategies is demonstrated when time constraints are considered, even though for some "pathological" problems where spatial correlations are weak, a simple random strategy is as (in)efficient as any more complex strategy. For the second sampling problem (fixed number of observations), adaptive sampling strategies did not clearly dominate static ones. In some cases, the static sampling strategies δ_{Star} and δ_{reg2} gave better results than strategies δ_{LSDP} and δ_{BP-max} in terms of percentage of well reconstructed quadrats. Static strategies appear especially efficient when models without external field are used. The two mentioned static strategies were able to provide a satisfying compromise between field exploration and patch boundary detection, unlike the other static strategies. However, the maps obtained when applying static strategies do not, in general, reflect the spatial pattern of the true map. Even though the best static strategy leads to acceptable spatial pattern, this best static strategy is not the same for all studied weeds species. These strategies are penalized by the fact that they cannot focus sampling on areas with high weeds density variation, which is detrimental for spatial structure analysis. On the contrary, adaptive strategies generally focus on these boundary areas when density variations are detected, thus helping to improve map quality for all weeds species. Finally, when applying all strategies in field condition (i.e. no data to estimate a model) by using our above-mentioned recommendations to build a model, we showed that δ_{BP-max} gives better results, than the other strategies. Following our results, a general sampling scheme can be proposed. First the observer makes a walk over the field (exploration) and acquires observations in order to find rough boundaries of weeds patches (adaptive patch boundaries detection). This first survey can be achieved using the sampling strategy δ_{LSDP}, which allocates most efforts for exploration. Part of this sampling effort should also be devoted to determining which density class is the most present, in order to fix the external field value. Then the strategy δ_{BP-max} can be used for a second survey for patch boundaries detection within the areas of high density variation.

Building optimal sampling strategies certainly consists in finding the optimal compromise between field exploration and patch boundaries detection according to observations values. This can be an other criterion to optimize in future works. Besides, several questions remain open when dealing with weed sampling, like for instance sampling for mapping of multiple weeds species or mapping at the landscape scale.

Conclusion générale et perspectives

Dans ce travail nous avons tout d'abord abordé, dans un cadre purement théorique, le problème d'échantillonnage d'un processus spatial à états discrets dans un but de cartographie. À partir de cette étude, des outils génériques permettant la conception de politiques d'échantillonnage adaptatives ont été proposés. L'utilisation de ces outils dans le cadre de l'échantillonnage d'une espèce adventice a ensuite nécessité un travail de modélisation des corrélations spatiales et du coût d'échantillonnage d'une espèce adventice. Dans cette conclusion, les résultats obtenus seront d'abord résumés et discutés. Ensuite, quelques pistes de réflexion laissées ouvertes par ce travail seront présentées.

Résumé des résultats obtenus et discussions

Tout d'abord, le problème d'échantillonnage considéré a été formulé comme un problème d'optimisation. Pour le résoudre de manière approchée, le problème du choix d'une politique d'échantillonnage a été modélisé à l'aide des PDM à horizon fini, transformant ainsi la résolution du problème initial en celui de la résolution d'un PDM. Le PDM proposé, de taille exponentielle en nombre de variables à reconstruire, ne peut également pas être résolu de manière exacte. Pour le résoudre, nous avons proposé l'algorithme LSDP combinant les principes de l'apprentissage par renforcement en mode *batch* et de la programmation dynamique approchée. Cet algorithme peut être utilisé pour la résolution de tout PDM à horizon fini, mais nécessite un calcul rapide des probabilités de transition du PDM étudié.

Cet algorithme a été validé sur plusieurs expérimentations dans le cadre de la résolution du problème d'échantillonnage spatial.

Une première validation a été effectuée sur des exemples jouets et lorsque le modèle de champ de Markov est connu. Ici, la politique δ_{LSDP} domine toutes les politiques sous-optimales testées, excepté dans un seul cas, où l'algorithme TD(λ) fournit une politique de meilleure qualité. Pour des coûts d'échantillonnage homogènes, les politiques δ_{LSDP} et δ_{BP-max} fournissent des résultats similaires. Toutefois, la politique δ_{LSDP} se démarque plus largement lorsque le nombre d'observations est faible. En revanche, la politique δ_{LSPI}, calculée à partir d'un algorithme d'apprentissage par renforcement classique, fournit des résultats arbitrairement mauvais. Lorsque les coûts d'observations sont inhomogènes, la supériorité de la politique δ_{LSDP} augmente.

Une deuxième validation a été effectuée sur des cartes réelles de répartition spatiale d'espèces adventices. Cette fois les modèles de champ de Markov ont été appris soit sur la carte complète, soit à partir de règles expertes. Les performances des politiques d'échantillonnage testées sont très variables, aussi bien les politiques utilisées classiquement en échantillonnage d'une espèce adventice que δ_{LSDP} et δ_{BP-max}. Par conséquent, aucune politique d'échantillonnage n'est universellement meilleure. Toutefois, la politique δ_{LSDP} permet, sur la majorité des exemples, de reconstruire des cartes de bonnes qualités. Lorsque le modèle de champ de Markov est construit en suivant les règles expertes, la politique δ_{BP-max} fournit de bons résultats sur l'ensemble des cas testés.

Enfin, lors d'une étude récente [74], le problème d'échantillonnage dans les Markov Logic Networks (MLN, [27]) a été abordé. Contrairement aux expérimentations précédentes, les distributions utilisées ne sont plus stationnaires et les graphes représentant les corrélations entre les variables sont non réguliers. Dans ce cas les performances de la politique δ_{LSDP} sont significativement meilleures que celles de l'heuristique δ_{BP-max}.

Au-delà de ces résultats prometteurs, il faut être conscient que les performances de l'algorithme LSDP sont tributaires de certaines contraintes. En effet, sur l'exemple de l'échantillonnage, l'algorithme LSDP fournit de bons résultats lorsque le modèle de corrélation spatiale est connu ou estimé sur des données complètes. En revanche lorsque le modèle utilisé est de qualité moindre et que le nombre d'observations est faible, l'efficacité de l'algorithme LSDP est discutable. Ces conclusions sont sûrement dues au fait que cet algorithme repose sur l'apprentissage par renforcement. En effet, un des principes de l'algorithme LSDP, et de tout algorithme d'apprentissage par renforcement, est de simuler des réalisations du phénomène étudié à partir du modèle de corrélations spatiales, puis "d'apprendre" sur ces réalisations ce que peut être une bonne politique d'échantillonnage. Ainsi, sur des cas réels, où le phénomène n'est pas une réalisation d'un modèle connu, l'algorithme LSDP peut être peu performant dès lors que les réalisations du modèle sur lesquelles l'apprentissage a été effectué sont différentes du phénomène réel. Cette discussion reste vraie pour le modèle de coût : si celui-ci est fiable, l'apprentissage sera bénéfique. Dans les comparaisons effectuées, lorsque le coût d'échantillonnage varie en fonction de la position de la variable et/ou de sa valeur, la politique δ_{LSDP} observe un plus grand nombre de variables que toutes les autres politiques d'échantillonnage testées, ce qui permet d'augmenter ses performances.
Pour le PDM correspondant au problème d'échantillonnage spatial, le calcul des probabilités de transition correspond à un calcul de probabilités conditionnelles de champ de Markov, tout comme les *features* que nous avons définies. Afin d'effectuer ce calcul rapidement, nous avons proposé une approximation ad-hoc de ces probabilités conditionnelles, limitant vraisemblablement les performances de l'algorithme.

Pour appliquer l'algorithme LSDP au problème d'échantillonnage d'une espèce adventice, un travail conséquent d'analyse des caractéristiques des corrélations spatiales et du coût d'échantillonnage d'une espèce adventice (enquêtes, tests de permutation) a été réalisé avec des écologues et des techniciens familiers avec l'échantillonnage des espèces adventices. Ce travail n'a pas été détaillé ici, mais a permis la conception des modèles présentés dans cette thèse. Tout d'abord, le coût d'échantillonnage adventice a été défini comme le temps nécessaire à la réalisation d'une politique d'échantillonnage. Un modèle a été proposé pour calculer ce temps, en fonction des caractéristiques de la parcelle.
Ensuite, différents modèles de corrélation spatiale d'une espèce adventice ont été proposés. À l'aide d'une approximation du score BIC, un meilleur modèle a été sélectionné pour plusieurs espèces adventices et différentes échelles spatiales. Pour chaque échelle spatiale, nous avons montré que parmi les modèles proposés, aucun n'était universellement meilleur pour toutes les espèces adventices. Toutefois, nous avons également montré que pour l'objectif de reconstruction de cartes à partir d'un nombre limité d'observations, les modèles de Potts avec champ externe fournissaient, en général, les meilleurs résultats pour l'ensemble des espèces. Des règles expertes ont ensuite été proposées pour fixer la valeur des paramètres de ce modèle dans des cas réels, sur la base d'une première exploration de la parcelle. Ces règles répondent aux contraintes rencontrées en pratique : l'exploration initiale peut être rapide car elle ne nécessite aucun relevé de densité et aucun calcul n'est nécessaire pour fixer la valeur des paramètres.
Différentes politiques d'échantillonnage ont été testées sur des cartes réelles de répartition spatiale d'espèces adventices. Les politiques ayant été testées sont les deux politiques adaptatives δ_{LSDP} et δ_{BP-max}, la politique aléatoire ainsi que huit politiques d'échantillonnage statiques. Comme nous l'avons énoncé plus haut dans cette conclusion, aucune

politique d'échantillonnage n'est universellement meilleure. Nous avons également montré que l'échantillonnage adaptatif permettait, la plupart du temps, de reconstruire des cartes respectant l'apparence générale de la carte réelle. Enfin, à partir des quelques exemples que nous avons traités, la politique d'échantillonnage δ_{BP-max} semble la plus adaptée pour des applications réelles, la politique δ_{LSDP} restant soumise aux contraintes évoquées plus haut.

Perspectives

Tout d'abord, l'algorithme LSDP pourrait être amélioré au niveau de la construction du *batch*. Pour l'instant, seule une méthode très classique (*i.e.* ϵ-greedy) est utilisée. La construction du *batch* pourrait s'appuyer sur des méthodes plus élaborées, en utilisant par exemple le formalisme proposé dans [68].
Ensuite, pour la résolution du problème d'échantillonnage, l'algorithme LSDP pourrait être appliqué en utilisant d'autres approximations des probabilités de transition du PDM. Nous rappelons que cela revient en fait à calculer de manière approchée des probabilités marginales d'un champ de Markov. L'utilisation d'autres méthodes d'approximation pourrait améliorer l'approximation proposée dans ce travail, en s'inspirant par exemple de [1] et ainsi améliorer les performances de LSDP.

Il serait également intéressant de comparer les résultats de l'algorithme LSDP avec ceux des algorithmes d'apprentissage par renforcement "en-ligne", de type Monte-Carlo Tree Search [40]. On peut s'attendre à ce que ces algorithmes fournissent des politiques d'échantillonnage de meilleure qualité que les politiques présentées dans ce manuscrit. De plus, ils permettraient de mieux situer les politiques δ_{LSDP} et δ_{BP-max} par rapport à la politique optimale sur de grands problèmes.

Les travaux réalisés durant cette thèse pourraient également s'étendre naturellement au cas des champs de Markov cachés [31], permettant de prendre en compte d'éventuelles erreurs de mesure rencontrées dans certaines applications.
Comme nous l'avons déjà énoncé plus haut, l'algorithme LSDP a été appliqué au problème d'échantillonnage dans les MLN. Lors d'une étude préliminaire, la politique d'échantillonnage δ_{LSDP} a permis d'améliorer significativement l'heuristique δ_{BP-max}. Dans ce cas les distributions utilisées sont non-stationnaires et le graphe décrivant les corrélations spatiales entre les variables est non régulier. D'autres tests avec des distributions semblables pourraient être effectués afin de montrer, si c'est le cas, l'efficacité de l'algorithme LSDP sur ce type de problème. Plus généralement, cela démontre que le travail réalisé durant cette thèse peut s'appliquer au cas de l'échantillonnage non spatial.

Pour la modélisation de la répartition spatiale d'une espèce adventice, les modèles de champs de Markov proposés pourraient être comparés aux modèles de Cox Log-Gaussien [12]. Ces modèles, issus de la géostatistique, sont plus classiques pour la modélisation de la répartition spatiale d'une espèce adventice. Ils permettent de considérer des observations continues ou discrètes (e.g. classe de densité).
D'autre part, dans ce travail, les modèles proposés ont été comparés grâce à une approximation du score BIC qui ne prend pas en compte l'objectif de cartographie. Prendre en compte cet objectif permettrait de sélectionner les modèles les plus pertinents pour

l'échantillonnage. Pour ceci, d'autres critères doivent être proposés, en s'inspirant par exemple du critère ICL [9].
Enfin, le modèle de coût d'échantillonnage proposé doit être validé par une expérimentation terrain.
D'autres questions restent ouvertes, comme la modélisation des interactions entre les espèces adventices ou l'échantillonnage à l'échelle de la mosaïque paysagère.

Appendice

Proof of Proposition 2 :

First, let us define $h^1 = ((\emptyset, \emptyset))$ and $\forall t = 2, ..., H$, $h^t = ((A^1, x_{A^1}), \ldots, (A^{t-1}, x_{A^{t-1}}))$. From any history h^t, a unique MDP state $s^t(h^t)$ can be defined as $s^1(h^1) = (\emptyset, \emptyset)$ and $\forall t = 2, ..., H$, $s^t(h^t) = \left(\cup_{k=1}^{t-1} A^k, x_{\cup_{k=1}^{t-1} A^k} \right)$. Then, we define the following transformation ϕ from the set of MDP policies to the set of OASMRF policies. For π, a MDP policy, $\delta = \phi(\pi)$ is defined as : for any $t = 1 \ldots H$ and any reachable trajectory h^t, $\delta^t(h^t) = \pi^t(s^t(h^t))$.

(i) We first show that $V^\pi((\emptyset, \emptyset), 1) = V(\phi(\pi))$. Indeed, we recall that

$$\begin{aligned} V^\pi((\emptyset, \emptyset), 1) &= \mathbb{E}_\pi\Big[\sum_{t=1}^{H+1} r^t \mid s^1 = (\emptyset, \emptyset) \Big] \\ &= \sum_{s^2, \ldots, s^{H+1}} \mathbf{P}(s^2, \ldots, s^{H+1} \mid \pi, s^1 = (\emptyset, \emptyset)) \Big[\sum_{t=1}^{H} r^t(\pi^t(s^t)) + r^{H+1}(s^{H+1}) \Big], \end{aligned}$$

where $\mathbf{P}(s^2, \ldots, s^{H+1} \mid \pi, s^1)$ is the probability of the state trajectory $(s^2, \ldots, s^{H+1})$, starting from s^1 and following policy π. Note that for any "feasible" MDP state trajectory $s^1, \ldots, s^{H+1}$ we can define a unique history $h^{H+1} = ((A^1, x_{A^1}), \ldots, (A^H, x_{A^H}))$, where A^t is the set of vertices involved in s^{t+1} and not in s^t. Then :

$$\mathbf{P}(s^2, \ldots, s^{H+1} \mid \pi, s^1) = \begin{cases} 0 \text{ if state trajectory not reachable,} \\ \mathbb{P}(x_A \mid) \text{ otherwise.} \end{cases}$$

with $A = \cup_{t=1}^H A^t$. In addition, we have that : $r^t(\pi(s^t)) = 0$ and $r^{H+1}(s^{H+1}) = \sum_{r \in R} \max_{x_r \in \Omega} \left\{ \mathbb{P}(x_r \mid x_A) \right\}$. Finally

$$\begin{aligned} V^\pi((\emptyset, \emptyset), 1) &= \sum_{h^{H+1} \in \tau_{\phi(\pi)}} \mathbb{P}(x_A \mid) \Big[-\alpha \sum_{t=1}^{H} \sum_{a \in A^t} c_a + \sum_{r \in R} \max_{x_r \in \Omega} \left\{ \mathbb{P}(x_r \mid x_A) \right\} \Big] \\ &= \sum_{h^{H+1} \in \tau_{\phi(\pi)}} \mathbb{P}(x_A) U(A, x_A) \\ &= V(\phi(\pi)). \end{aligned}$$

(ii) Then, we prove by backwards induction that an optimal policy δ^* for the OASMRF problem can be defined, which prescribes successive samples independently of the order of past observations. Let us consider $\delta^{*,H}$ first.

$$\delta^{*,H}((A^1, x_{A^1}), \ldots, (A^{H-1}, x_{A^{H-1}})) = \arg\max_{A^H} \sum_{x_{A^H}} \mathbb{P}(x_{A^H} | x_{A^1}, \ldots, x_{A^{H-1}}) U(A, x_A),$$

where $A = A^1 \cup \ldots \cup A^H$. Both $\mathbb{P}(x_{A^H} | x_{A^1}, \ldots, x_{A^{H-1}})$ and $U(A, x_A)$ do not depend on the order of observations $x_{A^1}, \ldots, x_{A^{H-1}}$. Thus, $\delta^{*,H}$ does not depend on the order of its

arguments.
Now, at time $h = H - 1$:

$$\delta^{*,H-1}((A^1, x_{A^1}), \ldots, (A^{H-2}, x_{A^{H-2}})) = \arg\max_{A^{H-1}} \sum_{x_{A^{H-1}}}$$

$$\sum_{x_{A^H}} \mathbb{P}(x_{A^{H-1}}, x_{A^H} | x_{A^1}, \ldots, x_{A^{H-2}}, \delta^{*,H}(\ldots)) U(A, x_A).$$

Since $\delta^{*,H}$ does not depend on the order of its arguments, $\delta^{*,H-1}$ is also independent of the order of its arguments $x_{A^1}, \ldots, x_{A^{H-2}}$.
Following the same reasoning for $h = H - 2, \ldots, 1$, we prove that an optimal policy δ^* can be computed, which prescribes samples independently of past observations. This result implies that we can limit the search for optimal policies of the OASMRF problem to policies δ prescribing actions which do not depend on the order of observations.

(*iii*) Let us now consider a given policy δ of the OASMRF in our limited search space. We can derive a policy π of the corresponding MDP model. The construction is also by induction : $\pi^1((\emptyset, \emptyset)) = \delta^1$, and for $t = 2$ to H and a reachable state s^t we define a history $((A^1, x_{A^1}), \ldots, (A^{t-1}, x_{A^{t-1}}))$ of size t-1, where the order in which observations are made are choosen arbitrarily, and we set $\pi(s^t) = \delta^t((A^1, x_{A^1}), \ldots, (A^{t-1}, x_{A^{t-1}}))$. With this procedure, π is defined only for states s^t reachable from policy δ. For other states, the policy is set to an arbitrary decision (the value of π will not depend on this choice since the corresponding state will never be reached). Let us call μ this transformation from a OASMRF policy to a MDP policy. Following the same reasoning as in (*i*), we can easily show that $V^{\mu(\delta)}((\emptyset, \emptyset), 1) = V(\delta)$.

(*iv*) Finally, let π^* be the optimal policy of the MDP model of the OASMRF problem :

$$V^{\pi^*}(s, t) \geq V^{\pi}(s, t) \qquad \forall \pi, s, t$$

Therefore the policy $\phi(\pi^*)$ is optimal for the OASMRF problem. Indeed, let δ be a given policy of the OASMRF problem (with the property of independence on the observations order) and $\mu(\delta)$ the corresponding policy of the MDP model. We have that

$$V^{\pi^*}((\emptyset, \emptyset), 1) \geq V^{\mu(\delta)}((\emptyset, \emptyset), 1),$$

and since $V^{\pi^*}((\emptyset, \emptyset), 1) = V(\phi(\pi^*))$ and $V^{\mu(\delta)}((\emptyset, \emptyset), 1) = V(\delta)$, we obtain $V(\phi(\pi^*)) \geq V(\delta)$. This establishes Proposition 1.

Proof of Lemma 1 :
For a given action trajectory $(d^1, \ldots, d^H)$, let us consider a state trajectory $(s^1, \ldots, s^{H+1})$ simulated according to the following two-steps scheme.

1. Simulate a map x according to the joint distribution $\mathbb{P}(.)$.
2. Deduce iteratively the values $(s^1, \ldots, s^{H+1})$ according to $s^1(i) = 0, \ \forall i \in O$ and :

$$\forall t \in \{1, \ldots, H\},\ s^{t+1}(i) = s^t(i) + d^t(i)x_i. \tag{9}$$

We have that

$$\mathbf{P}(s^1, \ldots, s^{H+1} \mid d^1, \ldots, d^H) \ = \ \sum_{x_V \in \Omega^n} \mathbb{P}(x_V)\mathbf{P}(s^1, \ldots, s^{H+1} \mid x_V, d^1, \ldots, d^H).$$

The probability $\mathbf{P}(s^1, \ldots, s^{H+1} \mid x_V, d^1, \ldots, d^H)$ is either equal to zero or to 1, since only one state trajectory can be reached from x_V and $(d^1, \ldots, d^H)$ according to (9). Furthermore, given $(d^1, \ldots, d^H)$, the state trajectory $(s^1, \ldots, s^{H+1})$ can be reached from any configuration x_V which agrees with the observations of this state trajectory on the subset A of sites visited by the action trajectory $(d^1, \ldots, d^H)$. Thus, if x'_A is the set of observations collected on A along the state trajectory $(s^1, \ldots, s^{H+1})$

$$\mathbf{P}(s^1, \ldots, s^{H+1} \mid d^1, \ldots, d^H) = \sum_{x_V \in \Omega^n} \mathbb{P}(x_V)1_{\{x_A = x'_A\}},$$

which by definition is equal to $\mathbb{P}(x'_A)$.

Let us now evaluate the propability to observe the same state trajectory $(s^1, \ldots, s^{H+1})$, given $(d^1, \ldots, d^H)$, when simulating according to the OASMRF MDP transition function :

$$\mathbf{P}(s^1, \ldots, s^{H+1} \mid d^1, \ldots, d^H) = \mathbb{P}(x'_{d^1}) \prod_{t=2}^{H} \mathbb{P}(x'_{d^t} \mid x'_{d^{t-1}}, \ldots, x'_{d^1}).$$

Using the Bayes rule, one can see that $\mathbb{P}(x'_{d^1}) \prod_{t=2}^{H} \mathbb{P}(x'_{d^t} \mid x'_{d^{t-1}}, \ldots x'_{d^1})$ is exactly $\mathbb{P}(x'_{d^1}, \ldots, x'_{d^H})$, which is equal to $\mathbb{P}(x'_A)$.

Therefore, with the two simulation schemes, for a given action trajectory $(d^1, \ldots, d^H)$ the same state trajectories can be reached (those where visited sites are coherent with the actions) and each state trajectory has the same probability in both schemes.

Références

Références

[1] D. Allard, A. Communian, and P. Renard. Probability aggregation methods in geoscience. *Mathematical Geosciences*, 44(5) :545–581, 2012.

[2] G.L. Anderson, J.H. Everitt, A.J. Anderson, and D.E. Escobar. Using satellite data to map false broomweed (*Ericameria austrotexana*) infestations on south texas rangelands. *Weed Technology*, 7 :865–871, 1993.

[3] A. Antos, C. Szepesvarí, and R. Munos. Learning near-optimal policies with bellman residual minimization based fitted policy iteration and a singl sample path. *Machine Learning*, 71(1) :89–129, 2008.

[4] M. Arias-Estevez, E. Lopez-Periago, and E. Martinez-Carballo. The mobility and degradation of pesticides in soils and the pollution of groundwater resources. *Agriculture, Ecosystems and Environment*, 123 :247–260, 2008.

[5] G. Barralis. Méthode d'étude des groupements adventices des cultures annuelles. *5ème colloque International sur l'Ecologie et la biologie des mauvaises herbes. INRA, Dijon*, pages 59–68, 1976.

[6] J. Barroso, D. Ruiz, C. Fernandez-Quintanilla, E.S. Leguizamon, P. Hernaiz, A. Ribeiro, B. Diaz, B.D. Maxwell, and L.J. Rew. Comparison of sampling methodologies for site-specific management of Avena sterilis. *Weed Research*, 45 :165–174, 2005.

[7] C. Bellehumeur, D. Marcotte, and P. Legendre. Estimation of regionalized phenomena by geostatistical methods : lake acidity on the Canadian shield. *Environmental geology*, 39 :211–220, 2000.

[8] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser.*, 6 :192–236, 1974.

[9] C. Biernacki, G. Celeux, and G. Govaert. Assessing a mixture model for clustering with the Integrated Completed Likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 719–725, 2000.

[10] M. Bonneau, S. Gaba, N. Peyrard, and R. Sabbadin. Model-based adaptive sampling strategies for weed mapping in crop fields. in prep.

[11] M. Bonneau, N. Peyrard, and R. Sabbadin. A reinforcement-learning algorithm for sampling design in markov random fields. *European Conference on Artificial Intelligence*, 2012.

[12] A. Bourgeois, S. Gaba, N. Munier-Jolain, B. Borgy, P. Monestiez, and S. Soubeyrand. Inferring weed spatial distribution from multi-type data. *Ecological Modelling*, 226 :92–98, 2012.

[13] A. Brix and J. Chadoeuf. Spatio-temporal modelling of weeds by shot-noise g cox processes. *Biometrical Journal*, 44(1) :83–99, 2002.

[14] R.B. Brown and J.P. Steckler. Weed patch identification in no-till corn using digital imagery. *Canadian Journal of Remote Sensing*, 19 :88–91, 1993.

[15] J. Cardina, G.A. Johnson, and D.H. Sparrow. The nature and consequence of weed spatial distribution. *Weed Science*, 45(3) :364–373, 1997.

[16] J. Cardina, D.H. Sparrow, and E.L. McCoy. Analysis of spatial distribution of common lambsquarters (*Chenopodium album*) in no-till soybean (*Glycine max*). *Weed Science*, 43 :258–268, 1995.

[17] M Charras-Garrido, D Abrial, N Peyrard, and S. Dachian. New classification method for disease mapping based on discrete hidden markov random fields. *Biostatistics*, 13 :241–255, 2012.

[18] R. Chikowo, V. Faloya, S. Petit, and N.M. Munier-Jolain. Integrated weed management systems allow reduced reliance on herbicides and long-term weed control. *Agriculture, Ecosystems and Environment*, 132 :237–242, 2009.

[19] S.A. Clay, B. Kreutner, D.E. Clay, C. Reese, J. Kleinjan, and F. Forcella. Spatial distribution, temporal stability, and yield loss estimates for annual grasses and common ragweed (*Ambrosia artimisiifolia*) in a corn/soybean production field over nine years. *Weed Science*, 54(2), 2006.

[20] N. Colbach, F. Dessaint, and F. Forcella. Evaluating field-scale methods for the estimation of mean plant densities of weeds. *Weed Research*, 40(5) :411–430, 2000.

[21] N. Colbach, F. Forcella, and G.A. Johnson. Spatial and temporal stability of weed populations over five years. *Weed Science*, 48 :366–377, 2000.

[22] V. Conitzer and T. Sandholm. Definition and complexity of some basic metareasoning problems. In *Proc. of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 1099–1106, 2003.

[23] J.S. Conn, C.H. Proctor, and W.A. Skroch. Selection of sampling methods to determine weed abundance in apple (malus domestica) orchards. *Weed Science*, 30(1) :35–40, 1982.

[24] R.D. Cousens, R.W. Brown, A.B. McBratney, B. Whelan, and M. Moerkerk. Sampling strategy is important for producing weeds maps : a case study using kriging. *Weed Science*, 50 :542–546, 2002.

[25] R.D. Cousens and J.L. Woolcock. Spatial dynamics of weeds : an overview. *Proceedings of the Brighton Crop Protection Conference - weeds*, 1997.

[26] J. de Gruijter, D. Brus, M. Bierkens, and K. Knotters. *Sampling for Natural Resource Monitoring*. Springer, 2006.

[27] P. Domingos and D. Lowd. *Markov Logic : An Interface Layer for Artificial Intelligence*. Synthesis Lectures on AI and ML. Morgan and Claypool, San Rafael, CA, 2009.

[28] W.W. Donald. Geostatistics for mapping weeds, with a canada thistle (*Cirsium arvense*) patch as a case study. *Weed Science*, 42(4), 1994.

[29] D. Ernst, P. Geurts, and L. Wehenkel. Tree-based batch mode reinforcement learning. *Journal of Machine Learning Research*, 6 :503–556, 2005.

[30] F. Forbes and N. Peyrard. Hidden markov random field model selection criteria based on mean field-like approximations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 :1089–1101, 2003.

[31] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 :721–741, 1984.

[32] R.H. Gibson, I.L. Nelson, G.W. Hopkins, B.J. Hamlett, and J. Memmott. Pollinator webs, plant communities and the conservation of rare plants : arable weeds as a case study. *Journal of applied ecology*, 43 :246–257, 2006.

[33] J.W. van Groenigen, W. Siderius, and A. Stein. Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma*, 87 :239–259, 1999.

[34] C. Guestrin, D. Koller, R. Parr, and S. Venkataraman. Efficient solution algorithms for factored MDPs. *Journal of Artificial Intelligence Research*, 19 :399–468, 2003.

[35] G. Guillot, N. Loren, and N. Rudemo. Spatial prediction of weed intensities from exact count data and image-analysis estimates. *Journal of Royal Statistics Society Series C : Applied Statistics*, 58 :525–542, 2009.

[36] D.E. Guyer, G.E. Miles, M.M. Schreiber, O.R. Mitchel, and V.C. Van-Derbilt. Machine vision and image processing for plant identification. *Transactions of the ASABE*, 29(6) :1500–1507, 1986.

[37] T. Heisel, C. Andreasen, and A.K. Ersboll. Annual weed distributions can be mapped with kriging. *Weed Science*, 36(4), 1994.

[38] C. Ji and L. Seymour. A consistent model selection procedure for Markov random fields based on penalized pseudolikelihood. *The annals of applied probability*, 6 :423–443, 1996.

[39] G.A. Johnson, D.A. Mortensen, and C.A. Gotway. Spatial and temporal analysis of weed seedling populations using geostatistics. *Weed Science*, 44(3), 1996.

[40] L. Kocsis and C. Szepesvári. Bandit based monte-carlo planning. In *ECML'06*, pages 282–293, 2006.

[41] D. Koller and N. Friedman. *Probabilistic Graphical Models : Principles and Techniques*. MIT Press, 2009.

[42] A. Krause. *Optimizing sensing : theory and applications*. PhD thesis, School of Computer science, Pittsburg, 2008.

[43] A. Krause and C. Guestrin. Optimal value of information in graphical models. *Journal of Artificial Intelligence Research*, 35 :557–591, 2009.

[44] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes : theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9 :235–284, 2008.

[45] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM journal of optimization*, 9 :112–147, 1998.

[46] M. Lagoudakis and R. Parr. Least-squares policy iteration. *Journal of Machine Learning Research*, 2003.

[47] D.W. Lamb and R.B. Brown. Remote-sensing and mapping of weeds in crops. *Journal of agricultural engineering research*, 78 :117–125, 2001.

[48] C. Lavigne, B. Ricci, P. Franck, and R. Senoussi. Spatial analyses of ecological count data : A density map comparaison approach. *Basic and applied Ecology*, 11 :734–742, 2010.

[49] Bonneau M., Peyrard N., and Sabbadin R. A reinforcement-learning algorithm for sampling design in Markov random fields. In *European Conference on Artificial Intelligence (ECAI'12)*.

[50] H Maei, C. Szepesvarí, S. Bhatnagar, and R. Sutton. Toward off-policy learning control with function approximation. *ICM*, 2010.

[51] P.G. Mantle and S. Shaw. A case study of the aetiology of ergot disease of cereals and grasses. *Plant pathology*, 26 :121–126, 1977.

[52] D.A. Mortensen, J.A. Dieleman, and G.A. Johnson. Weed spatial variation and weed management. *Integrated weed and soil management*, 1997.

[53] D.A. Mortensen, G.A. Johnson, and L.J. Young. Weed distributions in agricultural fields. *Proceedings of Soil Specific Crop Management : a workshop on research and development issues*, pages 113–124, 1993.

[54] D. Mueller-Dombois and H. Ellenberg. *Aims and methods in vegetation ecology.* Wiley and Sons, New-York, 1974.

[55] WG Müller. *Collecting spatial Data.* Springer Verlag : Heidelberg, 2007. 3rd ed.

[56] R. Munos and C. Szepesvarí. Finite-time bounds for fitted value iteration. *Journal of Machine Learning Research*, 9 :815–857, 2008.

[57] E.C. Oerke. Crop losses to pests. *Journal of Agricultural Science*, 144 :31–43, 2006.

[58] D. Ormoneit and S. Sen. Kernel-based reinforcement learning. *Machine Learning*, 49 :161–178, 2002.

[59] A. Ostfeld and col. The battle of the water sensor networks : a design challenge for engineers and algorithms. *Journal of Water Resources Planning and Management Division*, 134 :556–568, 2009.

[60] C. H. Papadimitriou. *Computational complexity.* Addison-Wesley, 1994.

[61] J. Pearl. *Probabilistic Reasonning in Intelligent Systems.* Morgan Kaufmann, 1988.

[62] J.N. Perry, E.D. Bell, R.H. Smith, and I.P. Woiwood. Sadie ; software to measure and model spatial pattern. *Aspects of Applied Biology - Modelling in Applied Biology : Spatial Aspects*, 46 :95–102, 1996.

[63] S. Petit, A. Boursault, M. Le Guilloux, N. Munier-Jolain, and X. Reboud. Interactions between weeds and cultivated plants as related to management of plant pathogens. *Agronomy for sustainable developpment*, 31 :309–317, 2011.

[64] A.N. Pettitt and A.B. McBratney. Sampling designs for estimating spatial variance components. *Journal of the royal statistics society. Serie C*, 42(1) :185–209, 1993.

[65] N. Peyrard, R. Sabbadin, and U. F. Niaz. Decision-theoretic optimal sampling with hidden Markov random fields. In *European Conference of Artificial Intelligence (ECAI'10)*, 2010.

[66] N. Peyrard, R. Sabbadin, D. Spring, R. Mac Nally, and B. Brook. Model-based adaptive spatial sampling for occurrence map construction. *Statistics and Computing*, 2012.

[67] M. Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc, 1994.

[68] E. Rachelson, F. Schnitzler, L. Wehenkel, and D. Ernst. Optimal sample selection for batch-mode reinforcement learning. *In Proceedings of ICAART*, pages 41–50, 2011.

[69] L.J. Rew and R.D. Cousens. Spatial distribution of weeds in arable crops : are current sampling and analytical methods appropriate ? *Weed Research*, 41 :1–18, 2001.

[70] L.J. Rew and G.W. Cussans. Horizontal movement of seeds following tine and plough cultivation : implications for spatial dynamics of weed infestations. *Weed Research*, 37 :247–256, 1997.

[71] R. Sabbadin, N. Peyrard, and N. Forsell. A framework and a mean-field algorithm for the local control of spatial processes. *International Journal of Approximate Reasoning*, 53(1), 2012.

[72] G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, 1978.

[73] D.R. Shaw. Remote sensing and site-specific weed management. *Frontiers in Ecology and the Environment*, 3(10) :526–532, 2005.

[74] G. Sirvent, N. Peyrard, and R. Sabbadin. Échantillonnage dans les Markov Logic Networks. Master's thesis, Master IRR, université Paul Sabatier, 2012.

[75] R. S. Sutton and A.G. Barto. *Reinforcement Learning : An Introduction*. MIT Press, 1998.

[76] R.S. Sutton. Learning to predict by the method of temporal differences. *Machine Learning*, 3 :9–44, 1988.

[77] R.S. Sutton, C. Szepesvari, and H.R. Maei. A convergent $o(n)$ temporal difference algorithm for off-policy learning with linear function approximation. In *NIPS*, pages 1609–1616, 2008.

[78] CS. Szepesvári. *Algorithms for Reinforcement Learning*. Morgan and Claypool, 2010.

[79] S. Thompson and G. Seber. *Adaptive sampling*. Series in Probability and Statistics. Wiley, New York, 1996.

[80] L.J. Wiles. Sampling to make maps for site-specific weed management. *Weed science*, 53(2) :228–235, 2005.

[81] L.J. Wiles, G.W. Oliver, A.C. York, H.J. Gold, and G.G. Wilkerson. Spatial distribution of broadleaf weeds in north carolina soybean (glycine max) fields. *Weed Science*, 40 :554–557, 1992.

[82] G.C. Wisler and R.E. Norris. Interactions between weeds and cultivated plants as related to management of plant pathogens. *Weed Science*, 53 :914–917, 2005.

[83] G. Zanin, A. Berti, and L. Riello. Incorporation of weed spatial variability into the weed control decision-making process. *Weed Research*, 38 :107–118, 1998.

[84] R.L. Zimdaht. *Weed-crop competition*. Wiley-Blackwell, 2008.