

Résolution de processus décisionnels de Markov à espaces d'état et d'action factorisés

Application en agroécologie

Julia Radoszycki

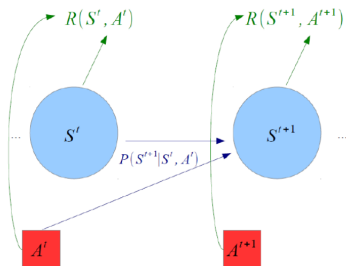
Encadrants : S. Gaba, N. Peyrard, R. Sabbadin

9 Octobre 2015

Motivations

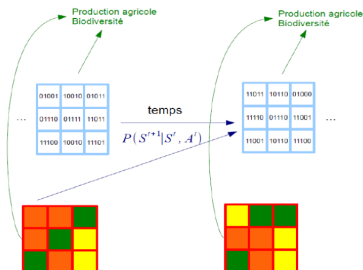
Intelligence artificielle

Comment prendre des décisions au mieux de manière séquentielle pour optimiser le comportement d'un système sur le long terme, dans un environnement incertain?



Agroécologie

Comment agencer les cultures pour faire un compromis entre production et biodiversité dans un paysage agricole?



Le cadre des processus décisionnels de Markov (PDM)

Données :

espace d'état $\mathcal{S}$

espace d'action $\mathcal{A}$

modèle de transition P

modèle de récompense R

horizon de temps T

Politique :

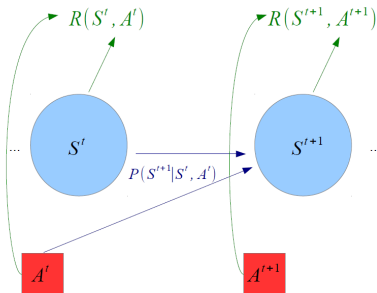
$\delta = \{\delta^t : S^t \rightarrow \mathcal{P}(\mathcal{A}^t), t = 0 \dots T-1\}$

Valeur de la politique : $V_{\delta}^{R,T}(s^0) = \mathbb{E} \left[\sum_{t=0}^{T-1} R(S^t, A^t) + R^T(S^T) \middle| s^0, \delta \right]$

Résoudre le PDM : trouver δ^* telle que $\forall \delta, \forall s^0, V_{\delta^*}^{R,T}(s^0) \geq V_{\delta}^{R,T}(s^0)$

Théorème : Il existe une politique optimale déterministe.

Algorithmes : de complexité polynomiale en $|\mathcal{S}|$ et $|\mathcal{A}|$



PDMs factorisés

- PDMs à espace d'état factorisé :

$$S^t = (S_1^t, \dots, S_n^t), \mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$$

exemple : si $n = 100$ et $\forall i = 1 \dots n \ |\mathcal{S}_i| = 2, |\mathcal{S}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times |\mathcal{A}|$



PDMs factorisés

- PDMs à espace d'état factorisé :

$$S^t = (S_1^t, \dots, S_n^t), \mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$$

exemple : si $n = 100$ et $\forall i = 1 \dots n \ |\mathcal{S}_i| = 2, |\mathcal{S}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times |\mathcal{A}|$

- PDMs à espace d'action factorisé :

$$A^t = (A_1^t, \dots, A_m^t), \mathcal{A} = \prod_{j=1}^m \mathcal{A}_j$$

exemple si $m = 100$ et $\forall j = 1 \dots m \ |\mathcal{A}_j| = 2, |\mathcal{A}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $|\mathcal{S}| \times |\mathcal{S}| \times 2^{100}$



PDMs factorisés

- PDMs à espace d'état factorisé :

$$S^t = (S_1^t, \dots, S_n^t), \mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$$

exemple : si $n = 100$ et $\forall i = 1 \dots n \ |\mathcal{S}_i| = 2, |\mathcal{S}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times |\mathcal{A}|$

- PDMs à espace d'action factorisé :

$$A^t = (A_1^t, \dots, A_m^t), \mathcal{A} = \prod_{j=1}^m \mathcal{A}_j$$

exemple si $m = 100$ et $\forall j = 1 \dots m \ |\mathcal{A}_j| = 2, |\mathcal{A}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $|\mathcal{S}| \times |\mathcal{S}| \times 2^{100}$



- PDMs à espaces d'état et d'action factorisés :

$$\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i, \mathcal{A} = \prod_{j=1}^m \mathcal{A}_j$$

exemple : si $n = m = 100, \forall i = 1 \dots n \ |\mathcal{S}_i| = |\mathcal{A}_i| = 2,$

$|\mathcal{S}| = |\mathcal{A}| = 2^{100}$

$\rightarrow P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times 2^{100}$

$\rightarrow$ besoin d'hypothèses de factorisation sur P et R

PDMs factorisés

- PDMs à espace d'état factorisé :

$$S^t = (S_1^t, \dots, S_n^t), \mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$$

exemple : si $n = 100$ et $\forall i = 1 \dots n \ |\mathcal{S}_i| = 2, |\mathcal{S}| = 2^{100}$

→ $P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times |\mathcal{A}|$

- PDMs à espace d'action factorisé :

$$A^t = (A_1^t, \dots, A_m^t), \mathcal{A} = \prod_{j=1}^m \mathcal{A}_j$$

exemple si $m = 100$ et $\forall j = 1 \dots m \ |\mathcal{A}_j| = 2, |\mathcal{A}| = 2^{100}$

→ $P(S^{t+1}|S^t, A^t)$ serait de taille $|\mathcal{S}| \times |\mathcal{S}| \times 2^{100}$



- PDMs à espaces d'état et d'action factorisés :

$$\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i, \mathcal{A} = \prod_{j=1}^m \mathcal{A}_j$$

exemple : si $n = m = 100, \forall i = 1 \dots n \ |\mathcal{S}_i| = |\mathcal{A}_i| = 2,$

$|\mathcal{S}| = |\mathcal{A}| = 2^{100}$

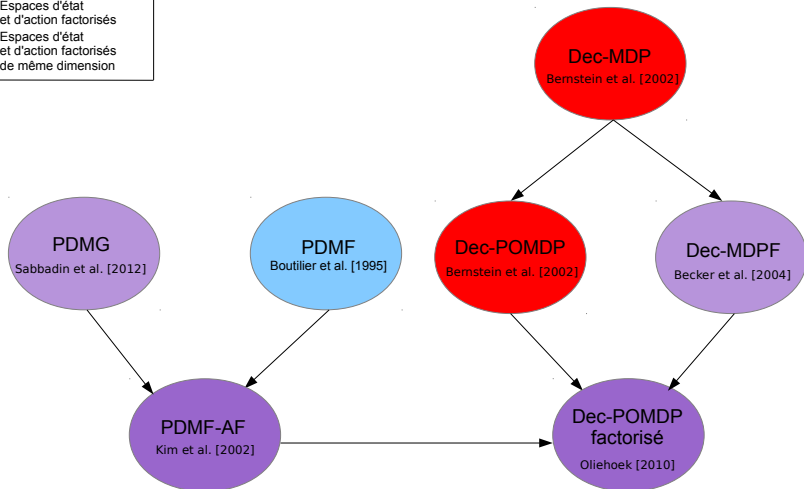
→ $P(S^{t+1}|S^t, A^t)$ serait de taille $2^{100} \times 2^{100} \times 2^{100}$

→ besoin d'hypothèses de factorisation sur P et R

La recherche d'une politique globale optimale, avec les algorithmes pour PDMs, est impossible. → résolution approchée

État de l'art pour les PDM factorisés : cadres

- Espace d'état factorisé
- Espace d'action factorisé
- Espaces d'état et d'action factorisés
- Espaces d'état et d'action factorisés de même dimension



État de l'art pour les PDM factorisés : algorithmes

général mais peu efficace en grande dimension :

- PDMF-AF : Raghavan et al. [2012]
- Dec-POMDP : Pajarinen and Peltonen [2011], Dibangoye et al. [2014]

État de l'art pour les PDM factorisés : algorithmes

général mais peu efficace en grande dimension :

- PDMF-AF : Raghavan et al. [2012]
- Dec-POMDP : Pajarinen and Peltonen [2011], Dibangoye et al. [2014]

grande dimension mais hypothèses restrictives :

- PDMG : Sabbadin et al. [2012], Cheng et al. [2013], Chen et al. [2015]
(même factorisation pour la transition, la politique et la récompense)
- Dec-MDP : Dibangoye et al. [2012] (transitions et observations indépendantes)
- Dec-POMDP : Kumar et al. [2011] (fonction de valeur additive), Oliehoek et al. [2013] (horizon fini petit)

État de l'art pour les PDM factorisés : algorithmes

général mais peu efficace en grande dimension :

- PDMF-AF : Raghavan et al. [2012]
- Dec-POMDP : Pajarinen and Peltonen [2011], Dibangoye et al. [2014]

grande dimension mais hypothèses restrictives :

- PDMG : Sabbadin et al. [2012], Cheng et al. [2013], Chen et al. [2015]
(même factorisation pour la transition, la politique et la récompense)
- Dec-MDP : Dibangoye et al. [2012] (transitions et observations indépendantes)
- Dec-POMDP : Kumar et al. [2011] (fonction de valeur additive), Oliehoek et al. [2013] (horizon fini petit)

général, grande dimension (AR) :

- Guestrin et al. [2002], Sallans and Hinton [2004] : politique non explicite, approximation structurelle de la fonction de valeur
- Buffet and Aberdeen [2009] : critère moyen, politique factorisée et paramétrée

Objectifs de ces travaux de thèse

Objectif méthodologique : proposer un cadre et des algorithmes

- génériques
- pour résoudre des PDM dont l'espace d'état et d'action sont factorisés et de grande taille (de l'ordre de 100 variables d'état et d'action, pas forcément binaires)
- dont la structure de décision est connue
- avec un bon compromis efficacité-qualité

Objectifs de ces travaux de thèse

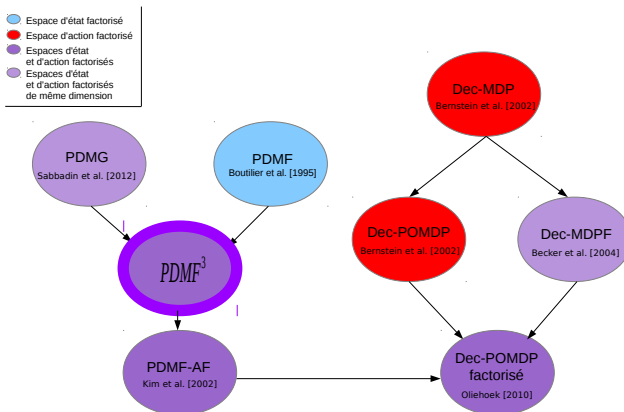
Objectif méthodologique : proposer un cadre et des algorithmes

- génériques
- pour résoudre des PDM dont l'espace d'état et d'action sont factorisés et de grande taille (de l'ordre de 100 variables d'état et d'action, pas forcément binaires)
- dont la structure de décision est connue
- avec un bon compromis efficacité-qualité

Objectif finalisé : sur un cas d'étude, identifier des règles d'allocation spatio-temporelle des cultures permettant de maintenir à la fois le rendement agricole et la biodiversité

Contributions méthodologiques Radoszycki et al. [2014, 2015]

- ❶ le cadre PDMF³
- ❷ une famille d'algorithmes de type itération de la politique (ou *policy search*)



Quel espace des politiques considérer ?

- Un PDM factorisé est un PDM particulier
→ en théorie, on peut trouver une politique déterministe optimale pour tout état initial
- Quand $\mathcal{S}$ et $\mathcal{A}$ sont grands, une politique $\delta : \mathcal{S} \rightarrow \mathcal{P}(\mathcal{A})$ ne peut pas être représentée explicitement
→ nous choisissons de la représenter de manière factorisée, avec une structure donnée, issue de la description du problème
- Mais les politiques optimales d'un PDM factorisé ne sont pas forcément factorisées. De ce fait, les meilleures politiques factorisées
 - peuvent dépendre de l'état initial
 - peuvent être stochastiques (non déterministes)
 - peuvent être non markoviennes

Notre approche

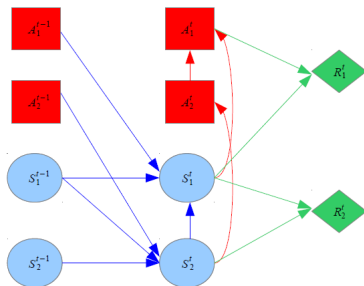
Choix

- considérer une distribution donnée sur l'état initial, représentée sous forme de réseau bayésien
- rechercher dans l'espace des politiques stochastiques factorisées (structure donnée)
- hypothèse de récompense additive

PDMF³ et politiques stochastiques factorisées (PSF)

Le cadre PDMF³

- Distribution initiale P^0 (réseau bayésien)



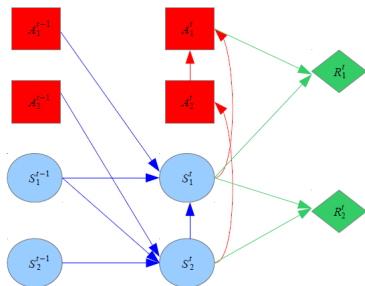
contrainte : pas de circuit avec
les flèches **bleues** et **rouges**

PDMF³ et politiques stochastiques factorisées (PSF)

Le cadre PDMF³

- Distribution initiale P^0 (réseau bayésien)
- Fonction de transition factorisée :

$$P(s^{t+1}|s^t, a^t) = \prod_{i=1}^n P_i(s_i^{t+1}|pa_P(s_i^{t+1}))$$
 où
 $pa_P(s_i^{t+1}) \subseteq \{s_j^t, s_{j'}^{t+1}, a_k^t, j' \neq i\}$



contrainte : pas de circuit avec
les flèches **bleues** et **rouges**

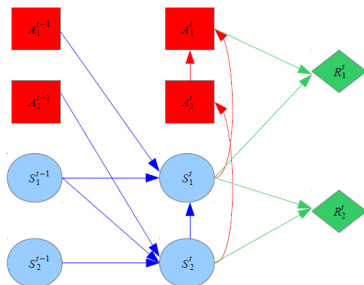
PDMF³ et politiques stochastiques factorisées (PSF)

Le cadre PDMF³

- Distribution initiale P^0 (réseau bayésien)
- Fonction de transition factorisée :

$$P(s^{t+1}|s^t, a^t) = \prod_{i=1}^n P_i(s_i^{t+1}|pa_P(s_i^{t+1}))$$
 où $pa_P(s_i^{t+1}) \subseteq \{s_j^t, s_{j'}^{t+1}, a_k^t, j' \neq i\}$
- Recherche de PSF dont la structure est donnée :

$$\delta(a^t|s^t) = \prod_{j=1}^m \delta_j(a_j^t|pa_\delta(a_j^t))$$
 où $pa_\delta(a_j^t) \subseteq \{s_i^t, a_k^t, k \neq j\}$



contrainte : pas de circuit avec
les flèches **bleues** et **rouges**

PDMF³ et politiques stochastiques factorisées (PSF)

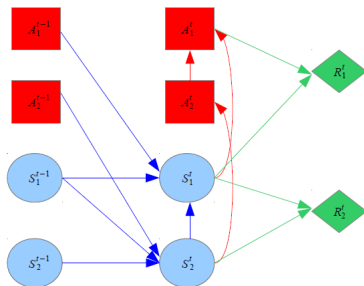
Le cadre PDMF³

- Distribution initiale P^0 (réseau bayésien)
- Fonction de transition factorisée :

$$P(s^{t+1}|s^t, a^t) = \prod_{i=1}^n P_i(s_i^{t+1}|pa_P(s_i^{t+1}))$$
 où $pa_P(s_i^{t+1}) \subseteq \{s_j^t, s_{j'}^{t+1}, a_k^t, j' \neq i\}$
- Recherche de PSF dont la structure est donnée :

$$\delta(a^t|s^t) = \prod_{j=1}^m \delta_j(a_j^t|pa_\delta(a_j^t))$$
 où $pa_\delta(a_j^t) \subseteq \{s_i^t, a_k^t, k \neq j\}$
- Récompense additive :

$$R(s^t, a^t) = \sum_{\alpha=1}^r R_\alpha(pa_R(R_\alpha^t)),$$
 où $pa_R(R_\alpha^t) \subseteq \{s_i^t, a_j^t\}$



contrainte : pas de circuit avec
les flèches **bleues** et **rouges**

PDMF³ et politiques stochastiques factorisées (PSF)

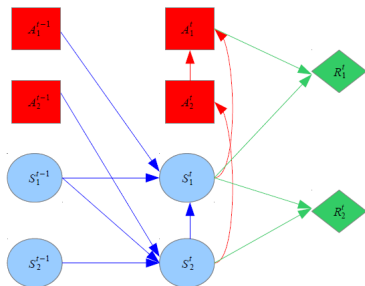
Le cadre PDMF³

- Distribution initiale P^0 (réseau bayésien)
- Fonction de transition factorisée :

$$P(s^{t+1}|s^t, a^t) = \prod_{i=1}^n P_i(s_i^{t+1}|pa_P(s_i^{t+1}))$$
 où $pa_P(s_i^{t+1}) \subseteq \{s_j^t, s_{j'}^{t+1}, a_k^t, j' \neq i\}$
- Recherche de PSF dont la structure est donnée :

$$\delta(a^t|s^t) = \prod_{j=1}^m \delta_j(a_j^t|pa_\delta(a_j^t))$$
 où $pa_\delta(a_j^t) \subseteq \{s_i^t, a_k^t, k \neq j\}$
- Récompense additive :

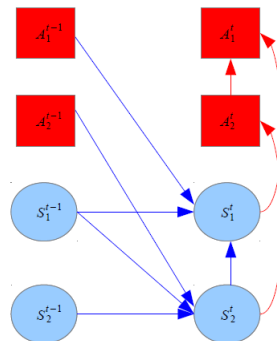
$$R(s^t, a^t) = \sum_{\alpha=1}^r R_\alpha(pa_R(R_\alpha^t)),$$
 où $pa_R(R_\alpha^t) \subseteq \{s_i^t, a_j^t\}$
- Horizon $0 < T \leq +\infty$, facteur d'amortissement $0 < \gamma < 1$



contrainte : pas de circuit avec
les flèches bleues et rouges

PDMF³ et politiques stochastiques factorisées (PSF)

A politique fixée, la distribution de probabilité P_δ^T sur les trajectoires d'horizon fini $(s, a)^{0:T} = \langle s^0, a^0, \dots, s^T, a^T \rangle$ est un **réseau bayésien dynamique**.



La **valeur** d'une politique stochastique factorisée δ pour P^0 donnée est définie par :

$$V_\delta^{R,T}(P^0) = \mathbb{E}_{P_\delta^T} \left[\sum_{t=0}^T R(S^t, A^t) \middle| P^0, \delta \right]$$

Le problème d'optimisation dans un PDMF³

Problème

$$\begin{aligned}
 &\underset{\delta \in (\mathbb{R}^+)^N}{\text{maximiser}} && V_{\delta}^{R,T}(P^0) \\
 &\text{s.c} && \sum_{a_j \in \mathcal{A}_j} \delta_j(a_j | pa_{\delta}(a_j)) = 1 \quad \forall j, \forall pa_{\delta}(a_j)
 \end{aligned}$$

→ problème de maximisation convexe, problème NP^{PP} -difficile

→ nombre de variables d'optimisation N de l'ordre de 1000 ou 10000 sur les problèmes qui vont nous intéresser

Algorithmes d'itération de la politique approchée

Itération de la politique approchée

- **Étape d'évaluation** : étant donnée la PSF courante δ^q , $V_{\delta^q}^{R,T}(P^0)$ est évaluée de manière exacte ou approchée, en utilisant par exemple l'algorithme JT, LBP ou MC.
- **Étape de mise à jour** : δ^q est améliorée en δ^{q+1} , en utilisant soit une descente par coordonnées (CD pour *coordinate descent*) soit une descente de gradient (GD pour *gradient descent*).

→ **Famille d'algorithmes** : GD-LBP, GD-MC, CD-LBP...

Évaluation de politiques stochastiques factorisées (PSF)

- La valeur d'une politique stochastique factorisée δ pour P^0 donnée s'exprime en fonction de **marginales** de la distribution $P_\delta^T((s, a)^{0:T})$:

$$\begin{aligned}
 V_\delta^{R,T}(P^0) &= \mathbb{E}_{P_\delta^T} \left[\sum_{t=0}^T R(S^t, A^t) \middle| P^0, \delta \right] \\
 &= \sum_{t=0}^T \sum_{\alpha=1}^r \mathbb{E}_{P_\delta^T} \left[R_\alpha(pa_R(R_\alpha^t)) \middle| P^0, \delta \right] \\
 &= \sum_{t=0}^T \sum_{\alpha=1}^r \sum_{pa_R(R_\alpha^t)} b_\alpha^t(pa_R(R_\alpha^t)) R_\alpha(pa_R(R_\alpha^t))
 \end{aligned}$$

où $b_\alpha^t(pa_R(R_\alpha^t)) = \sum_{(s,a)^{0:T} \setminus (pa_R(R_\alpha^t))} P_\delta^T((s, a)^{0:T})$

→ calcul de ces marginales par des **méthodes d'inférence** exactes (JT) ou approchées (LBP...)

Évaluation de politiques stochastiques factorisées (PSF)

- La valeur d'une politique stochastique factorisée δ pour P^0 donnée s'exprime en fonction de **marginales** de la distribution $P_\delta^T((s, a)^{0:T})$:

$$\begin{aligned}
 V_\delta^{R,T}(P^0) &= \mathbb{E}_{P_\delta^T} \left[\sum_{t=0}^T R(S^t, A^t) \middle| P^0, \delta \right] \\
 &= \sum_{t=0}^T \sum_{\alpha=1}^r \mathbb{E}_{P_\delta^T} \left[R_\alpha(pa_R(R_\alpha^t)) \middle| P^0, \delta \right] \\
 &= \sum_{t=0}^T \sum_{\alpha=1}^r \sum_{pa_R(R_\alpha^t)} b_\alpha^t(pa_R(R_\alpha^t)) R_\alpha(pa_R(R_\alpha^t))
 \end{aligned}$$

où $b_\alpha^t(pa_R(R_\alpha^t)) = \sum_{(s,a)^{0:T} \setminus (pa_R(R_\alpha^t))} P_\delta^T((s, a)^{0:T})$

→ calcul de ces marginales par des **méthodes d'inférence** exactes (JT) ou approchées (LBP...)

- La valeur d'une PSF peut aussi être approchée par la méthode de Monte-Carlo (MC).

Mise à jour

Algorithmes de descente par coordonnées (CD)

cas de variables d'action binaires

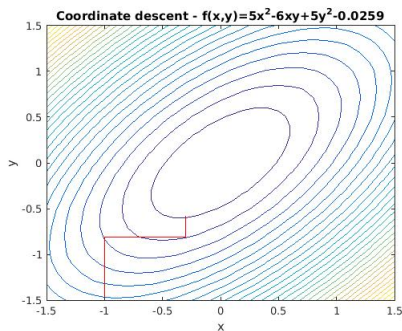
($\forall j = 1 \dots m, |\mathcal{A}_j| = 2$) :

$$\underset{\tilde{\delta} \in [0,1]^{\frac{N}{2}}}{\text{minimiser}} V_{\tilde{\delta}}^{\bar{R}, T}(P^0)$$

$$\tilde{\delta} = \{\delta_j(a_j = 1 | pa_{\delta}(a_j)), \forall j, \forall pa_{\delta}(a_j)\}$$

→ pas optimal

→ garantie de convergence vers un minimum local de l'approximation de la valeur avec la méthode choisie (LBP, MC...)



Mise à jour

Algorithmes de descente de gradient (GD)

cas général :

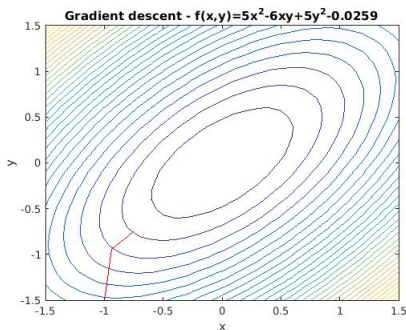
$$\underset{\theta \in (\mathbb{R}^+)^N}{\text{minimiser}} \quad V_{\theta}^{\bar{R}, T}(P^0)$$

$$\delta_j(a_j | p_{a_{\delta}}(a_j)) = \frac{e^{\theta_j(a_j | p_{a_{\delta}}(a_j))}}{\sum_{a'_j \in \mathcal{A}_j} e^{\theta_j(a'_j | p_{a_{\delta}}(a_j))}}$$

→ estimation parallélisée du gradient
par **différences finies** / par calcul de
marginales dans un modèle graphique

→ pas fixe/pas optimal/**conditions
de Wolfe**

→ garantie de convergence vers un point critique de l'approximation de la
valeur avec la méthode choisie (LBP, MC...)



Évaluation expérimentale des algorithmes proposés

Protocole

- 1 petits PDMF³ aléatoires → comparaison avec l'exact
100 problèmes où $n = m = r = 6$

Évaluation expérimentale des algorithmes proposés

Protocole

- ① **petits PDMF³ aléatoires** → comparaison avec l'exact
100 problèmes où $n = m = r = 6$
- ② **gros problèmes aléatoires** → comparaison des valeurs MC
15 problèmes où $n = m = r = 15$
(5 PDMG/5 PDMF³ binaires/5 PDMF³ non binaires)

Évaluation expérimentale des algorithmes proposés

Protocole

- ① **petits PDMF³ aléatoires** → comparaison avec l'exact
100 problèmes où $n = m = r = 6$
- ② **gros problèmes aléatoires** → comparaison des valeurs MC
15 problèmes où $n = m = r = 15$
(5 PDMG/5 PDMF³ binaires/5 PDMF³ non binaires)
- ③ **problème d'épidémiologie agricole**
surveillance à la parcelle (PDMG) :
 $n = m = r = 100, |\mathcal{S}| = |\mathcal{A}| \approx 10^{30}, N = 5184$
surveillance en réseau (PDMF³ binaire) :
 $n = m = r = 25, |\mathcal{S}| = |\mathcal{A}| \approx 3 \times 10^7, N = 800$

Évaluation expérimentale des algorithmes proposés

Protocole

- ① **petits PDMF³ aléatoires** → comparaison avec l'exact
100 problèmes où $n = m = r = 6$
- ② **gros problèmes aléatoires** → comparaison des valeurs MC
15 problèmes où $n = m = r = 15$
(5 PDMG/5 PDMF³ binaires/5 PDMF³ non binaires)
- ③ **problème d'épidémiologie agricole**
surveillance à la parcelle (PDMG) :
 $n = m = r = 100, |\mathcal{S}| = |\mathcal{A}| \approx 10^{30}, N = 5184$
surveillance en réseau (PDMF³ binaire) :
 $n = m = r = 25, |\mathcal{S}| = |\mathcal{A}| \approx 3 \times 10^7, N = 800$
- ④ **problème de conservation d'une espèce** (PDMF³ non binaire)
 $n = 18, m = 9, r = 9, |\mathcal{S}| \approx 2.6 \times 10^5, |\mathcal{A}| \approx 2 \times 10^4, N = 2100$

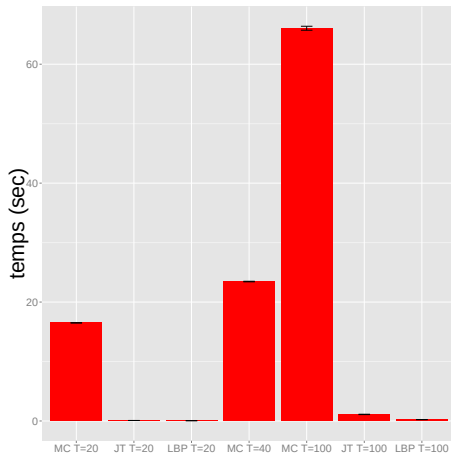
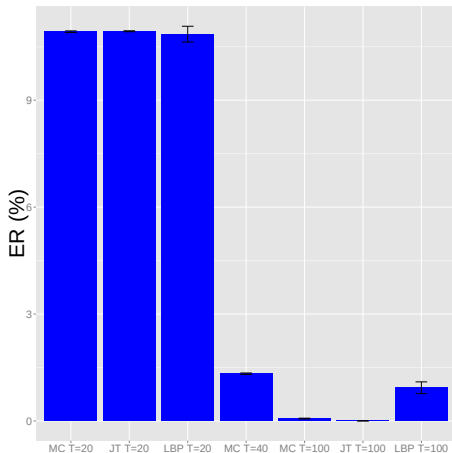
Évaluation expérimentale des algorithmes proposés

Protocole

- ① **petits PDMF³ aléatoires** → comparaison avec l'exact
100 problèmes où $n = m = r = 6$
- ② **gros problèmes aléatoires** → comparaison des valeurs MC
15 problèmes où $n = m = r = 15$
(5 PDMG/5 PDMF³ binaires/5 PDMF³ non binaires)
- ③ **problème d'épidémiologie agricole**
surveillance à la parcelle (PDMG) :
 $n = m = r = 100, |\mathcal{S}| = |\mathcal{A}| \approx 10^{30}, N = 5184$
surveillance en réseau (PDMF³ binaire) :
 $n = m = r = 25, |\mathcal{S}| = |\mathcal{A}| \approx 3 \times 10^7, N = 800$
- ④ **problème de conservation d'une espèce** (PDMF³ non binaire)
 $n = 18, m = 9, r = 9, |\mathcal{S}| \approx 2.6 \times 10^5, |\mathcal{A}| \approx 2 \times 10^4, N = 2100$

Résultats : petits problèmes aléatoires

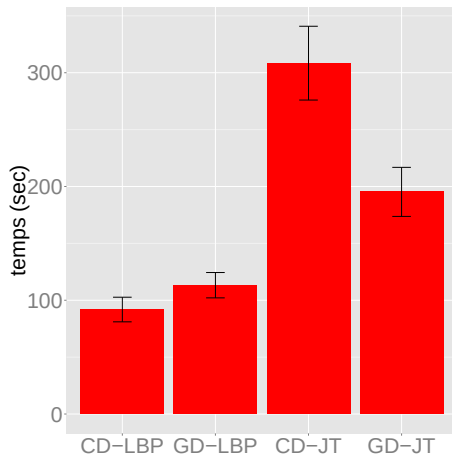
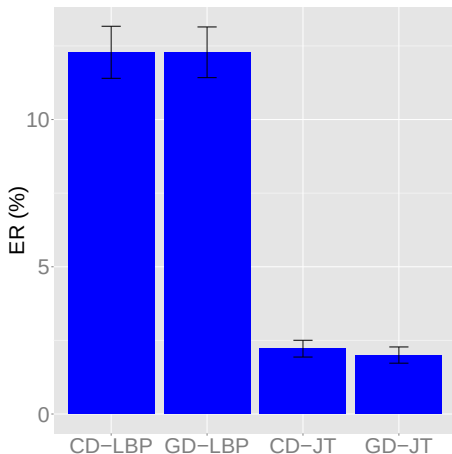
Comparaison des méthodes d'évaluation



→ dans la suite, $T = 20$ pour LBP et JT, $T = 40$ pour MC

Résultats : petits problèmes aléatoires

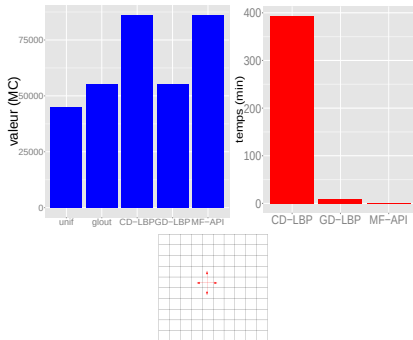
Comparaison des algorithmes de résolution



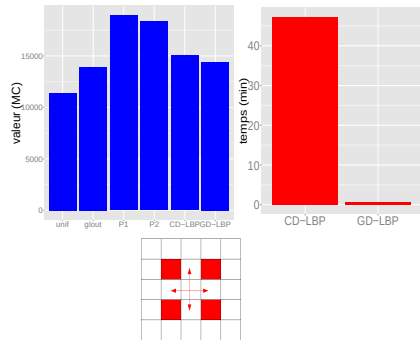
→ GD et CD donnent des résultats du même ordre, avec une ER de l'ordre de 12% si LBP, 2% si JT

Résultats : problème d'épidémiologie agricole

2 actions : ne rien faire, traiter la maladie



(a) Surveillance à la parcelle



(b) Surveillance par un réseau

- GD-LBP et CD-LBP améliorent la politique uniforme
- GD-LBP est beaucoup plus rapide
- MF-API (Sabbadin et al. [2012]) est plus efficace pour les PDMG

Conclusion sur la partie méthodologique

Originalité

- rechercher des politiques stochastiques factorisées de structure donnée
- mêler évaluation par inférence et *policy search*

Résultats

- capacité à traiter des problèmes où $|\mathcal{S}| = |\mathcal{A}| \approx 10^{30}$
- une famille d'algorithmes complémentaires avec de bons résultats en pratique
- famille d'algorithmes disponible dans le f3mdpsolver (Matlab) :
<https://mulcyber.toulouse.inra.fr/projects/f3mdpsolver/>

Perspectives sur la partie méthodologique

Algorithme

- mêler évaluation par simulation et optimisation (apprentissage par renforcement) → comparaison avec FPG (Buffet and Aberdeen [2009])

Perspectives sur la partie méthodologique

Algorithme

- mêler évaluation par simulation et optimisation (apprentissage par renforcement) → comparaison avec FPG (Buffet and Aberdeen [2009])
- varier la méthode d'évaluation au cours de l'optimisation ('multifidélité')

Perspectives sur la partie méthodologique

Algorithme

- mêler évaluation par simulation et optimisation (apprentissage par renforcement) → comparaison avec FPG (Buffet and Aberdeen [2009])
- varier la méthode d'évaluation au cours de l'optimisation ('multifidélité')

Plus généralement

- rechercher des PSF non stationnaires

Perspectives sur la partie méthodologique

Algorithme

- mêler évaluation par simulation et optimisation (apprentissage par renforcement) → comparaison avec FPG (Buffet and Aberdeen [2009])
- varier la méthode d'évaluation au cours de l'optimisation ('multifidélité')

Plus généralement

- rechercher des PSF non stationnaires
- extension au cadre Dec-POMDP factorisé (Oliehoek [2010])

Perspectives sur la partie méthodologique

Algorithme

- mêler évaluation par simulation et optimisation (apprentissage par renforcement) → comparaison avec FPG (Buffet and Aberdeen [2009])
- varier la méthode d'évaluation au cours de l'optimisation ('multifidélité')

Plus généralement

- rechercher des PSF non stationnaires
- extension au cadre Dec-POMDP factorisé (Oliehoek [2010])
- optimisation de la structure de la politique

Contributions finalisées

- un **modèle** des relations spatio-temporelles entre Cultures, Adventices et Pollinisateurs dans le cadre PDMF³
- comparaison de différentes **politiques** d'allocation spatiale des cultures, obtenues par optimisation et évaluation-comparaison

© 2006 The Authors



Gestion durable des adventices



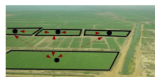
Service de production

Vs



Services de régulation
(pollinisation, prédation)

- A l'échelle de la mosaïque paysagère

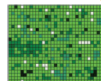


- Deux stratégies évoquées dans la littérature :

land sparing

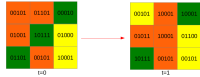
VS

land sharing

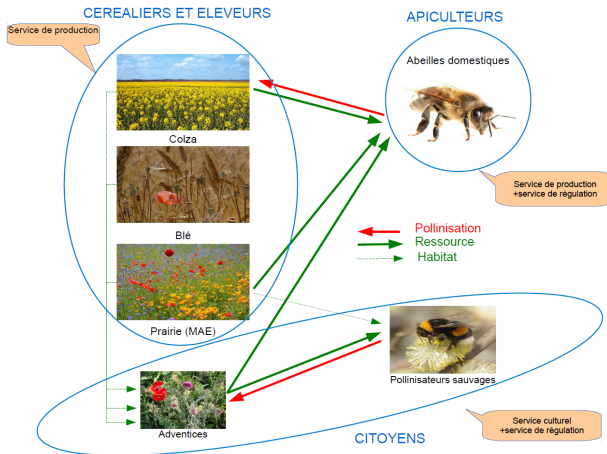


- Besoin de **conception de stratégies par optimisation** plutôt que d'évaluation-comparaison de stratégies expertes
- Besoin d'exploration → pas d'*a priori* sur l'organisation spatiale

Approches pour la gestion des services écosystémiques à l'échelle du paysage

	cartographie + scénarios	modèle + optimisation	PDMF ³
exemples	Chan et al. [2006] Bateman et al. [2013]	Polasky et al. [2005] Brosi et al. [2008]	
avantages	situation particulière - utilisation de données - précis	général - optimisation exacte le plus souvent	général - prise en compte du temps
inconvénients	non prise en compte du temps le plus souvent - exploration partielle 	non prise en compte du temps le plus souvent 	optimisation approchée 

Cas d'étude : problème Cultures-Adventices-Pollinisateurs



Est-il possible d'agencer les cultures dans le paysage de manière à satisfaire au mieux tous les acteurs ?

État de l'art pour la modélisation de ce cas d'étude

- description qualitative du problème dans Bretagnolle and Gaba [2015]
→ problème de décision séquentielle sous incertitude
 - **à l'échelle du paysage** → cadre PDMF³
 - multi-objectifs (multi-services)

État de l'art pour la modélisation de ce cas d'étude

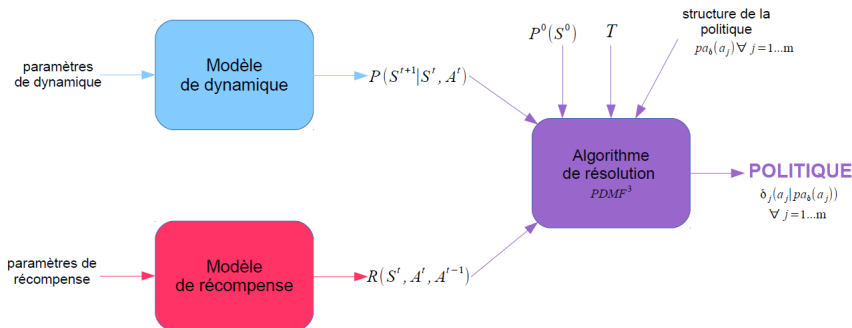
- description qualitative du problème dans Bretagnolle and Gaba [2015]
 - problème de décision séquentielle sous incertitude
 - **à l'échelle du paysage** → cadre PDMF³
 - multi-objectifs (multi-services)
 - dynamique des adventices dans les cultures
 - à l'échelle de la parcelle : systèmes de culture
 - à l'échelle du paysage : flux de gènes (Colbach [2009])
 - paramètres du cycle de vie : peu d'informations
- utilisation d'un modèle de métacommunautés (Dorazio et al. [2010])

État de l'art pour la modélisation de ce cas d'étude

- description qualitative du problème dans Bretagnolle and Gaba [2015]
 - problème de décision séquentielle sous incertitude
 - **à l'échelle du paysage** → cadre PDMF³
 - multi-objectifs (multi-services)
- dynamique des adventices dans les cultures
 - à l'échelle de la parcelle : systèmes de culture
 - à l'échelle du paysage : flux de gènes (Colbach [2009])
 - paramètres du cycle de vie : peu d'informations
- utilisation d'un modèle de métacommunautés (Dorazio et al. [2010])
- évaluation des services écosystémiques (modèle de récompense)
 - quantification des pollinisateurs : Lonsdorf et al. [2009]
 - marge économique colza/blé/prairie : BDD+modèle
 - marge économique du miel : peu d'informations

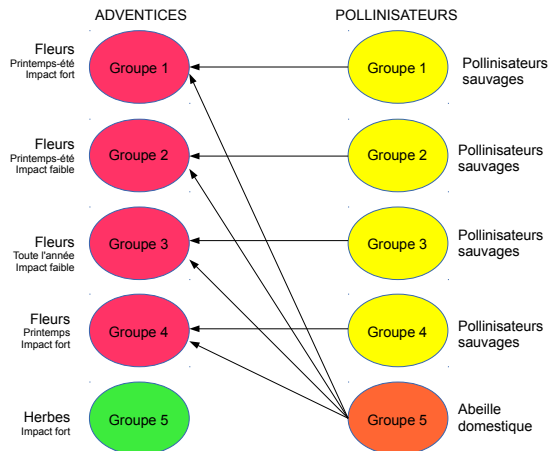
Objectifs de la modélisation

- **proposer un modèle qui combine cultures, adventices et pollinisateurs**, qui soit suffisamment simple pour arriver à faire de l'optimisation tout en prenant en compte toutes les relations
- **optimisation multi-objectif** dans un cadre mono-objectif



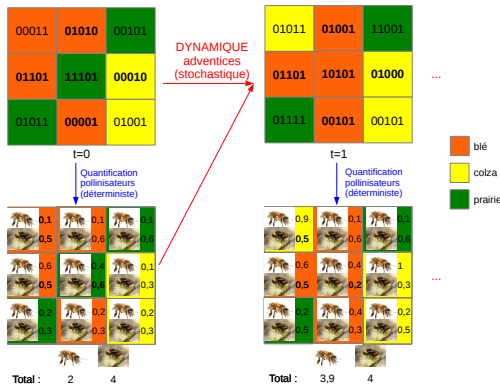
Modélisation sous forme de PDMF³

Réseau trophique



Modélisation sous forme de PDMF³

Modèle de dynamique



inspiré de Dorazio et al. [2010] et Lonsdorf et al. [2009]

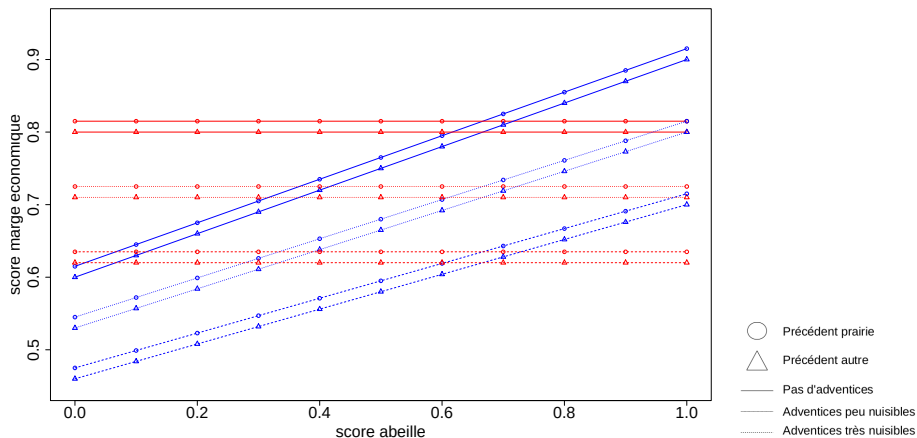
$$|\mathcal{S}| = 2^{500} \approx 10^{150}$$

$$|\mathcal{A}| = 3^{100} \approx 10^{47}$$

Modélisation sous forme de PDMF³

Modèles de récompense

● marge en blé-colza (céréaliers)



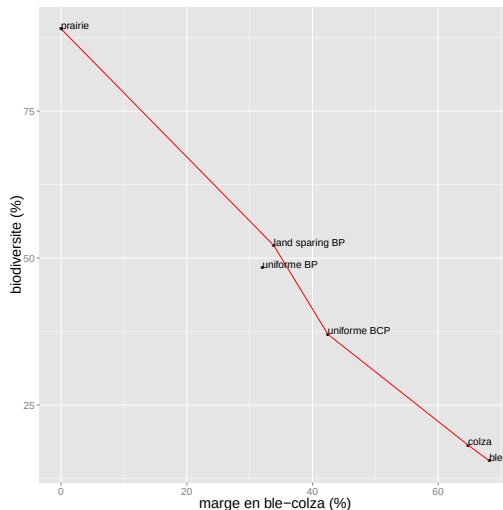
Modélisation sous forme de PDMF³

Modèles de récompense

- marge en blé-colza (céréaliers)
- marge en miel (apiculteurs)
 - proportionnelle à l'abondance des abeilles domestiques dans le paysage
- biodiversité (citoyens)
 - somme des scores de pollinisateurs et des présences-absences de groupes adventices dans le paysage

Recherche de politiques de compromis par optimisation

1. Évaluation de politiques 'simples'



Recherche de politiques de compromis par optimisation

2. Formulation d'un objectif unique (modèle de récompense)

- le compromis ne peut pas se faire à la parcelle
- le compromis ne peut pas se faire à l'échelle du paysage (verrou méthodologique)

→ nous allons rechercher un compromis à l'échelle de l'exploitation

→ cela permet aussi un traitement équitable des agriculteurs

objectif unique : nombre moyen d'années et d'exploitations étant au-dessus de seuils prédéfinis pour chacun des services écosystémiques considérés

$$R(S^t, A^t, A^{t-1}) = \sum_{e=1}^E \mathbb{1} \left\{ \sum_{p \in P_e} \text{marge}_p \geq \beta \right\} \mathbb{1} \left\{ \sum_{p \in P_e} \text{biodiv}_p \geq \zeta \right\}$$

Recherche de politiques de compromis par optimisation

2. Formulation d'un objectif unique (modèle de récompense)

- le compromis ne peut pas se faire à la parcelle
- le compromis ne peut pas se faire à l'échelle du paysage (verrou méthodologique)

→ nous allons rechercher un compromis à l'échelle de l'exploitation

→ cela permet aussi un traitement équitable des agriculteurs

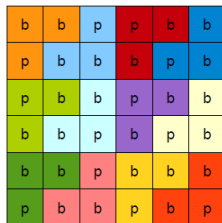
objectif unique : nombre moyen d'années et d'exploitations étant au-dessus de seuils prédéfinis pour chacun des services écosystémiques considérés

$$\begin{aligned}
 R(S^t, A^t, A^{t-1}) &= \sum_{e=1}^E \mathbb{1} \left\{ \sum_{p \in P_e} ble_p(S_p^t, A_p^t, A_p^{t-1}) \geq \beta \right\} \mathbb{1} \left\{ \sum_{p \in P_e} adv_p(S_p^t) \geq \zeta \right\} \\
 &= \sum_{e=1}^E f(S_{P_e}^t, A_{P_e}^t, A_{P_e}^{t-1})
 \end{aligned}$$

Recherche de politiques de compromis par optimisation

3. Utilisation de l'algorithme GD-LBP

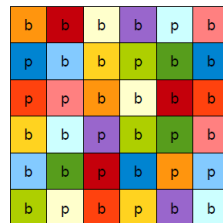
maximiser le nombre moyen d'années et d'exploitations pour lesquelles la marge en blé est supérieure à 30% du maximum et le nombre de groupes adventices est supérieur à 44% du maximum



(a) Très agrégé



(b) Moyennement agrégé

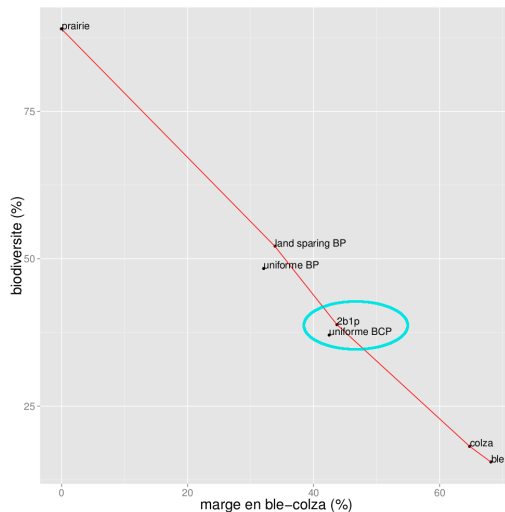


(c) Peu agrégé

→ quelle que soit la structure des exploitations : **deux blés et une prairie par exploitation**

Recherche de politiques de compromis par optimisation

3. Utilisation de l'algorithme GD-LBP



Limites de l'approche par optimisation

Pratiques

- on ne peut pas prendre en compte dans l'objectif la marge économique du miel ou du colza
- pour une grille 6×6 , il faut se limiter à 3 groupes d'adventices et de pollinisateurs

Politiques obtenues

- elles sont déterministes
 - elles ne dépendent pas de l'état de la parcelle
- les cultures ne changent pas au cours du temps

Étude du *continuum land sparing-land sharing* par évaluation-comparaison [Radoszycki et al., in prep.]

Plan d'expériences :

paysage		...		
AI blé-colza	1	...	0.1765	0
AI prairie	1	...	0.2235	0
AI global	1	...	0.2	0

→ 3 compositions : 30% prairie, 50% prairie, 70% prairie

→ 8 paysages par composition

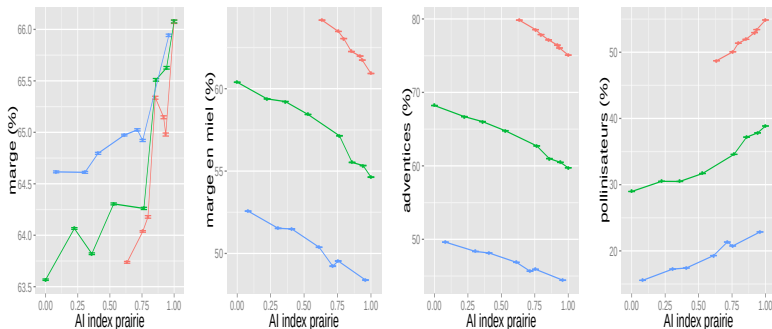
→ évaluation MC de chaque paysage : 4000 simulations sur 20 ans, analyse sur les 10 dernières années

→ étude de la moyenne et de la variabilité dans le temps de la marge et de la biodiversité

→ étude de différents scénarios pour les valeurs de paramètres

Étude du *continuum land sparing-land sharing* par évaluation-comparaison [Radoszycki et al., in prep.]

Quel est l'effet de la composition, quel est l'effet de la configuration (mesurée par l'AI index de He et al. [2000]) ?

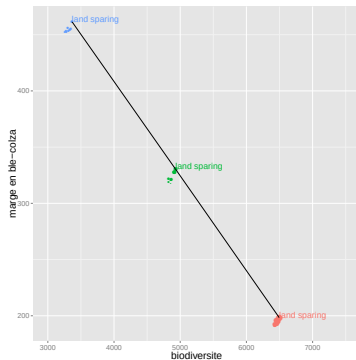


70% prairie, 50% prairie, 30% prairie

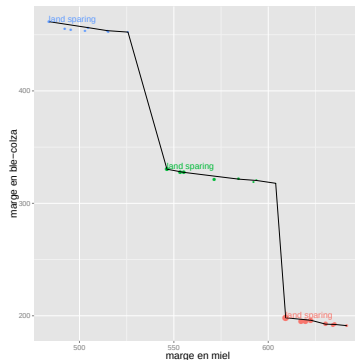
→ effet principal de la composition, effet linéaire de la configuration

Étude du *continuum land sparing-land sharing* par évaluation-comparaison [Radoszycki et al., in prep.]

Pour une composition donnée, est-il possible de trouver des politiques satisfaisantes pour tous les acteurs ? ($\nu = 0.1$)



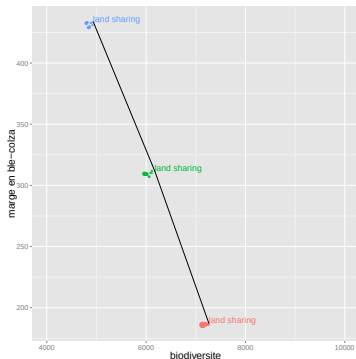
70% prairie, 50% prairie, 30% prairie



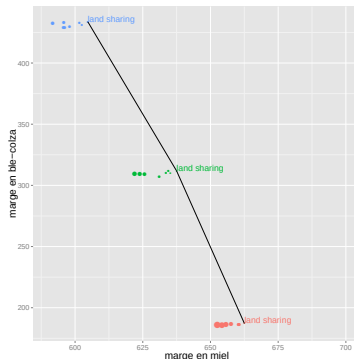
→ il semblerait que non...

Étude du *continuum land sparing-land sharing* par évaluation-comparaison [Radoszycki et al., in prep.]

Pour une composition donnée, est-il possible de trouver des politiques satisfaisantes pour tous les acteurs ? ($\nu = 0.25$)



70% prairie, 50% prairie, 30% prairie



→ cela dépend du niveau de dispersion adventice

Conclusion sur la partie finalisée (1/2)

Contributions de la thèse

- un modèle des relations spatio-temporelles entre Cultures, Adventices et Pollinisateurs dans le cadre PDMF³
- une illustration de l'utilisation de GD-LBP pour un problème où $|S| \approx 10^{150}$ et $|\mathcal{A}| \approx 10^{47}$
- comparaison de différentes politiques d'allocation spatiale des cultures, obtenues par optimisation et évaluation-comparaison

Conclusion sur la partie finalisée (1/2)

Contributions de la thèse

- un modèle des relations spatio-temporelles entre Cultures, Adventices et Pollinisateurs dans le cadre PDMF³
- une illustration de l'utilisation de GD-LBP pour un problème où $|S| \approx 10^{150}$ et $|\mathcal{A}| \approx 10^{47}$
- comparaison de différentes politiques d'allocation spatiale des cultures, obtenues par optimisation et évaluation-comparaison

En cours

- conception participative dans la zone atelier *Plaine et Val de Sèvre* [projet AgroBioSE]
- étude de la pertinence d'un modèle non spatialisé, plus facile à optimiser [stage M2 R. Alexandre]

Conclusion sur la partie finalisée (2/2)

Plus généralement

- démarche et modèle originaux pour les problèmes de gestion à l'échelle du paysage, qui peuvent s'appliquer à d'autres cas d'étude :
 - problèmes d'épidémio-surveillance agricole
 - problèmes de gestion d'espèces invasives
 - problèmes de gestion de services écosystémiques
 - ...

Perspectives sur la partie finalisée

- rendre le modèle plus réaliste :
 - prendre en compte l'abondance des adventices
 - prendre en compte la dynamique des pollinisateurs
 - prendre en compte les bordures et les haies
 - prendre en compte les systèmes de production (biologique/conventionnel)
 - structure de décision basée sur la culture précédente en plus de l'état de la parcelle
 - ...

Perspectives sur la partie finalisée

- rendre le modèle plus réaliste :
 - prendre en compte l'abondance des adventices
 - prendre en compte la dynamique des pollinisateurs
 - prendre en compte les bordures et les haies
 - prendre en compte les systèmes de production (biologique/conventionnel)
 - structure de décision basée sur la culture précédente en plus de l'état de la parcelle
 - ...
- 'analyse de sensibilité' sur les politiques obtenues

Perspectives sur la partie finalisée

- rendre le modèle plus réaliste :
 - prendre en compte l'abondance des adventices
 - prendre en compte la dynamique des pollinisateurs
 - prendre en compte les bordures et les haies
 - prendre en compte les systèmes de production (biologique/conventionnel)
 - structure de décision basée sur la culture précédente en plus de l'état de la parcelle
 - ...
- 'analyse de sensibilité' sur les politiques obtenues
- article en préparation : utiliser MULTILAND, un générateur de paysages-grilles de composition et d'agrégation données (Roques [2014])

Perspectives sur la partie finalisée

- rendre le modèle plus réaliste :
 - prendre en compte l'abondance des adventices
 - prendre en compte la dynamique des pollinisateurs
 - prendre en compte les bordures et les haies
 - prendre en compte les systèmes de production (biologique/conventionnel)
 - structure de décision basée sur la culture précédente en plus de l'état de la parcelle
 - ...
- 'analyse de sensibilité' sur les politiques obtenues
- article en préparation : utiliser MULTILAND, un générateur de paysages-grilles de composition et d'agrégation données (Roques [2014])
- prendre en compte explicitement le caractère multi-objectif du problème (Porto et al. [2014], Wang [2014])

Conclusion générale

Nous avons proposé des algorithmes de résolution approchée de PDMF³ qui sont satisfaisants... même si on arrive parfois à trouver des politiques expertes de meilleure qualité... mais :

Conclusion générale

Nous avons proposé des algorithmes de résolution approchée de PDMF³ qui sont satisfaisants... même si on arrive parfois à trouver des politiques expertes de meilleure qualité... mais :

« *L'IA est la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains.* » (Marvin Lee Minsky)

Conclusion générale

Nous avons proposé des algorithmes de résolution approchée de PDMF³ qui sont satisfaisants... même si on arrive parfois à trouver des politiques expertes de meilleure qualité... mais :

« *L'IA est la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains.* » (Marvin Lee Minsky)

Solutions :

- boucle itérative Politique experte → Optimisation → Interprétation experte
- utilisation complémentaire de l'approche par évaluation-comparaison

Remerciements

- A tous les membres du jury : Alain DUTECH, Hervé MONOD, Pierre-Olivier CHEPTOU
- A mes directeurs de thèse : Sabrina GABA, Nathalie PEYRARD, Régis SABBADIN
- Aux membres de mon comité de thèse
- Aux membres de l'unité MIAT et de l'UMR Agroécologie
- Aux membres du projet AgroBioSE
- A Romain ALEXANDRE

- I. J. Bateman, A. R. Harwood, G. M. Mace, R. T. Watson, D. J. Abson, B. Andrews, A. Binner, A. Crowe, B. H. Day, S. Dugdale, C. Fezzi, J. Foden, D. Hadley, R. Haines-Young, M. Hulme, A. Kontoleon, A. A. Lovett, P. Munday, U. Pascual, J. Paterson, G. Perino, A. Sen, G. Siriwardena, D. van Soest, and M. Termansen. Bringing Ecosystem Services into Economic Decision-Making : Land Use in the United Kingdom. *Science*, 341(6141) :45–50, 2013.
- V. Bretagnolle and S. Gaba. Weeds for bees ? A review. *Agronomy for a Sustainable Development*, pages 1–19, 2015.
- B. J. Brosi, P. R. Armsworth, and G. C. Daily. Optimal design of agricultural landscapes for pollination services. *Conservation letters*, 1 : 27–36, 2008.
- O. Buffet and D. Aberdeen. The factored policy-gradient planner. *Artificial Intelligence*, 173 :722–747, 2009.

- K. M. A. Chan, M. R. Shaw, D. R. Cameron, E. C. Underwood, and G. C. Daily. Conservation planning for ecosystem services. *Plos Biology*, 4 (11) :2138–2152, 2006.
- F. Chen, Q. Cheng, J. Dong, Z. Yu, G. Wang, and W. Xu. Efficient approximate linear programming for factored MDPs. *International Journal of Approximate Reasoning*, 2015.
- Q. Cheng, Q. Liu, F. Chen, and A. Ihler. Variational planning for graph-based MDPs. In *Advances in Neural Information Processing Systems 26*, pages 2976–2984, 2013.
- N. Colbach. How to model and simulate the effects of cropping systems on population dynamics and gene flow at the landscape level : example of oilseed rape volunteers and their role for co-existence of GM and non-GM crops. *Environmental Science and Pollution Research*, 16 : 348–360, 2009.

- J. S. Dibangoye, C. Amato, and A. Doniec. Scaling up decentralized MDPs through heuristic search. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence*, pages 217–226, 2012.
- J. S. Dibangoye, C. Amato, O. Buffet, and F. Charpillet. Exploiting separability in multiagent planning with continuous-state MDPs. In *Proceedings of the 13th international conference on Autonomous Agents and Multiagent Systems*, 2014.
- R. M. Dorazio, M. Kery, J. A. Royle, and M. Plattner. Models for inference in dynamic metacommunity systems. *Ecology*, 91(8) :2466–2475, 2010.
- L. J. Gilarranz and J. Bascompte. Spatial network structure and metapopulation persistence. *Journal of Theoretical Biology*, 297 :11–16, 2012.
- C. Guestrin, M. Lagoudakis, and R. Parr. Coordinated reinforcement learning. In *Proceedings of the 19th international conference on Machine learning*, ICML '02, 2002.

- H. S. He, B. E. DeZonia, and D. J. Mladenoff. An aggregation index (AI) to quantify spatial patterns of landscapes. *Landscape Ecology*, 15 : 591–601, 2000.
- A. Kumar, S. Zilberstein, and M. Toussaint. Scalable multiagent planning using probabilistic inference. In *Proceedings of the 22th International Joint Conference on Artificial Intelligence*, 2011.
- E. Lonsdorf, C. Kremen, T. Ricketts, R. Winfree, N. Williams, and S. Greenleaf. Modelling pollination services across agricultural landscapes. *Annals of Botany*, 103 :1589–1600, 2009.
- F. A. Oliehoek. *Value-based planning for teams of agents in stochastic partially observable environments*. PhD thesis, Amsterdam University, 2010.

- F. A. Oliehoek, S. Whiteson, and M. T. J. Spaan. Approximate solutions for factored Dec-POMDPs with many agents. In *Proceedings of the 12th international conference on autonomous agents and multiagent systems (AAMAS)*, 2013.
- J. Pajarinen and J. Peltonen. Efficient planning for factored infinite-horizon DEC-POMDPs. In *Proceedings of the 22th international joint conference on Artificial intelligence*, 2011.
- S. Polasky, E. Nelson, E. Lonsdorf, P. Fackler, and A. Starfield. Conserving species in a working landscape : land use with biological and economic objectives. *Ecological applications*, 15(4) :1387–1401, 2005.
- M. Porto, O. Correia, and P. Beja. Optimization of landscape services under uncoordinated management by multiple landowners. *Plos One*, 9 (1), 2014.

- J. Radoszycki, N. Peyrard, and R. Sabbadin. Finding good stochastic factored policies for factored Markov decision processes. In *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic*, pages 1083–1084, 2014.
- J. Radoszycki, N. Peyrard, and R. Sabbadin. Solving F^3 MDPs : Collaborative Multiagent Markov Decision Processes with Factored Transitions, Rewards and Stochastic Policies. In *PRIMA 2015 - Principles and Practice of Multi-Agent Systems, 26-30 October 2015, Bertinoro, Italy*, 2015.
- A. Raghavan, S. Joshi, A. Fern, P. Tadepalli, and R. Khordon. Planning in factored action spaces with symbolic dynamic programming. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, 2012.
- L. Roques. Multiland software : user guide. Technical report, Juin 2014.

- 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) :66–86, 2012.
- B. Sallans and G. E. Hinton. Reinforcement learning with factored states and actions. *Journal of Machine Learning Research*, 5 :1063–1088, 2004.
- W. Wang. *Multi-objective sequential decision making*. PhD thesis, Université Paris-Sud, 2014.

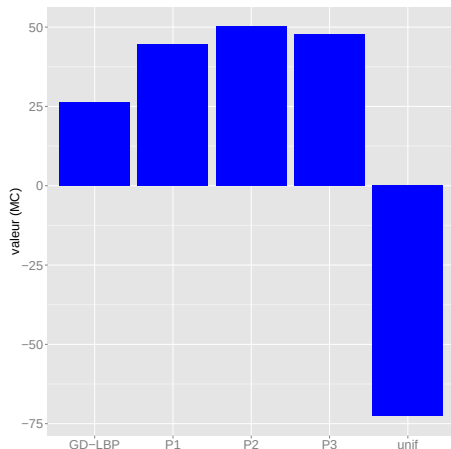
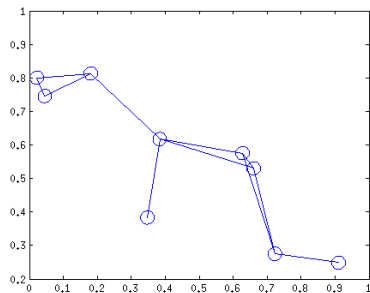
Pour que $\mathbb{P}(|\hat{V}_{\delta,MC}^{R,T}(P^0) - V_{\delta}^{R,T}(P^0)| < \epsilon) \geq \kappa$ il faut que :

$$n_{\text{sim}} \geq \frac{(1 - \gamma^{T+1})^2 (R_{\max} - R_{\min})^2}{2\epsilon^2(1 - \gamma)^2} \ln \frac{2}{1 - \kappa} \quad (1)$$

où $R_{\max} = \max_{s,a} R(s, a)$, $R_{\min} = \min_{s,a} R(s, a)$.

Résultats : problème de conservation d'une espèce

3 actions : ne rien faire, réintroduire, empoisonner



GD-LBP renvoie une politique de valeur correcte, mais on arrive à trouver des politiques expertes de meilleure valeur

Modèle de dynamique des adventices

Formalisme

- approche métacommunautés ; on suit la **présence-absence des 5 groupes adventices** sur chaque parcelle
- réseau bayésien dynamique (modèle markovien)
- proche de modèles utilisés en écologie (Dorazio et al. [2010] ou Gilarranz and Bascompte [2012])
- 4 paramètres (2 de survie et 2 de dispersion)

Présence-absence des 5 groupes d'adventices

00101	01101	00010
01001	10111	01000
01101	00101	10001

t=0

DYNAMIQUE
adventices

Présence-absence des 5 groupes d'adventices

00101	10001	10001
01011	10001	01100
10111	00101	00101

t=1



Modèle de dynamique des adventices

Formalisme

- approche métacommunautés ; on suit la **présence-absence des 5 groupes adventices** sur chaque parcelle
- réseau bayésien dynamique (modèle markovien)
- proche de modèles utilisés en écologie (Dorazio et al. [2010] ou Gilarranz and Bascompte [2012])
- 4 paramètres (2 de survie et 2 de dispersion)

Hypothèses

- indépendance des groupes adventices (pas de compétition)
- paramètres identiques pour tous les groupes adventices
- dispersion avant la compétition avec la culture
- survie supérieure en prairie
- dispersion temporelle > dispersion spatiale
- prise en compte de l'effet des pollinisateurs sauvages sur la reproduction des adventices fleurs

Modèle de dynamique des adventices

$$P(S^{t+1}|S^t, A^t) = \prod_{p=1}^P \prod_{i=1}^I P_{pi} \left(S_{pi}^{t+1} \middle| S_{V(p)i}^t, A_{V(p)}^t \right)$$

$$S_{pi}^{t+1} | S_{V(p)i}^t, A_{V(p)}^t \sim \text{Bernouilli} \left((\phi_i + (1 - \phi_i)\gamma_i) S_{pi}^t + \gamma_i(1 - S_{pi}^t) \right)$$

$$\gamma_i = \gamma_i(S_{V(p)i}^t, A_p^t) = \epsilon + (1 - \epsilon) \left[1 - (1 - \kappa(A_p^t))(1 - \nu)^{\sum_{j \in V(p)} S_{ji}^t} \right]$$

$$\begin{aligned} \forall i = 1 \dots 4, \phi_i &= \phi_i(S_{V(p)i}^t, A_{V(p)}^t) \\ &= \frac{\sum_{b \in RT^{-1}(i)} PO_{pb}(S_{V(p)RT(b)}^t, A_{V(p)}^t) + \eta(A_p^t)}{|RT^{-1}(i)| + \eta(A_p^t)} \\ \phi_5 &= \phi_5(S_{V(p)5}^t, A_{V(p)}^t) = \eta(A_p^t) \end{aligned}$$

$$\epsilon = 0.01, \nu = 0.1, \eta(A_p^t) = \kappa(A_p^t) = 0.6 \text{ (prairie)} / 0.1 \text{ (blé/colza)}$$

Modèle de récompense : quantification des pollinisateurs

Nous proposons de s'inspirer du modèle de Lonsdorf et al. [2009] :

$$\begin{aligned} & \text{score pollinisateur } s \text{ sur parcelle } p \\ &= \text{habitabilité sur parcelle } p \text{ et voisines pour } s \\ &\times \text{ressource florale sur parcelle } p \text{ et voisines pour } s \end{aligned}$$

① score d'habitabilité d'une parcelle p' pour s (entre 0 et 1) :

- pollinisateurs sauvages : maximal pour les prairies, nul pour les autres cultures
- abeilles domestiques : maximal quelle que soit la culture (ruches)

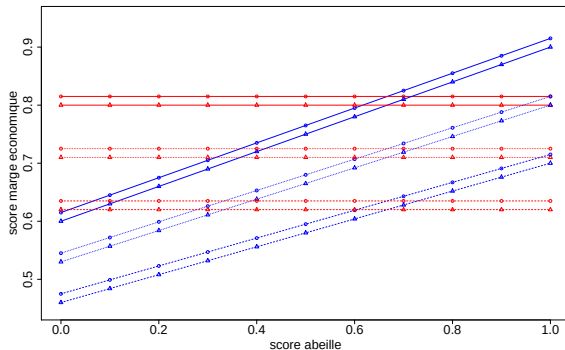
② score de 'ressource florale' d'une parcelle p' pour s (entre 0 et 1) :

- pollinisateurs sauvages : maximal si le groupe adventice que pollinise s est présent, nul sinon
- abeilles domestiques : score pondéré sur les saisons, pour une saison donnée maximal si présence d'une culture à fleurs (colza ou luzerne), moyen si présence d'adventices en fleur, nul si aucun des deux

③ portée des pollinisateurs : poids de $p' \propto e^{-D(p,p')/\alpha}$

- pollinisateurs sauvages : voisinage immédiat
- abeilles domestiques : voisinage plus large

Modèle de récompense : marge économique



1 colza et blé

- Précédent prairie
- △ Précédent autre
- Pas d'adventices
- Adventices peu nuisibles
- Adventices très nuisibles

2 prairie : 0.5 si pas de prairie avant, 0.8 sinon

3 miel : proportionnelle à l'abondance d'abeilles dans le paysage