

Analyse de sensibilité pour les modèles dynamiques: application au modèle AZODYN.

Matieyendou Lamboni¹, David Makowski², Hervé Monod¹

¹MIA-INRA, ² UMR Agronomie-INRA, France.

matieyendou.lamboni@jouy.inra.fr

IHP, 06 Mai 2009



Introduction

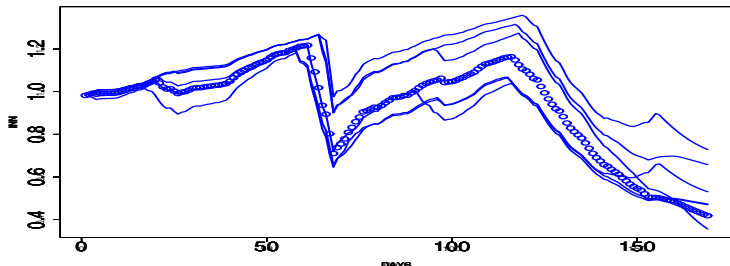
- Contexte : modélisation dynamique en agronomie
 - module biologique (physiologie de la plante)
 - module physique (typologie du sol, transfert d'eau)
 - module environnement (matière organique)
 - module météo (température, rayonnement)
- Modèle :
 - modèle complexe dynamique (pas journalier)
 - nombreux paramètres incertains
 - mais : peu de données disponible
 - AS pour sélectionner les paramètres influents

Introduction

le modèle AZODYN (Jeffroy et al., 1999)

- l'indice de nitrification azotée : INN
- 69 paramètres dont 13 paramètres incertains

Incertitude sur les sorties : 6 simulations du modèle AZODYN



Introduction

- Objectif :

- généraliser l'analyse de sensibilité au cas multivarié
- proposer un indice pour identifier les paramètres à estimer en priorité
 1. facteurs discrets (Campbell et al., 2006)
 2. facteurs continus

- Plan :

- rappels ANOVA et AS
- décomposition de l'inertie (indices)
- nouvelles fonctions de réponse
- application au modèle AZODYN

Méthode : ANOVA et AS

Facteurs (discrets) :

$$z_j \in \{1, \dots, n_j\}, \forall j \in \{1, 2, \dots, d\}$$

Modèle d'ANOVA

$$y_{i_1, \dots, i_d} = \eta_{\emptyset} + \sum_{l=1}^d \eta_{\{l\}, (i_l)} + \sum_{l < r} \eta_{\{l, r\}, (i_l, i_r)} + \dots + \eta_{\{1, \dots, d\}, (i_1, \dots, i_d)}. \quad (1)$$

Modèle vectoriel d'ANOVA

$$\mathbf{y} = \eta_{\emptyset} \mathbf{1}_{\bullet\bullet} + \sum_{l=1}^d \eta_{\{l\}} + \sum_{l < r} \eta_{\{l, r\}} + \dots + \eta_{\{1, \dots, d\}} \quad (2)$$

$$\mathbf{y} = \sum_{u \subseteq \{1, 2, \dots, d\}} \eta_u \quad (3)$$

Méthode : ANOVA et AS

proposition

Sous l'hypothèse d'un plan orthogonal (Gram-schmidt)

$$\mathbb{R}^N = E_{\emptyset} \oplus^{\perp} E_{\{1\}} \oplus^{\perp} \dots \oplus^{\perp} E_{\{1,2,\dots,d\}} \quad (4)$$

$$\eta_u = P_{E_u} \mathbf{y} \quad (5)$$

Si $SC_{\mathbf{y}} = \|\mathbf{y} - \eta_{\emptyset} \mathbf{1}_{\bullet\bullet}\|^2$ et $SC_u = \|\eta_u\|^2$, alors

$$SC_{\mathbf{y}} = \sum_{u \subseteq \{1,2,\dots,d\}} SC_u \quad (6)$$

Indices

$$IS_u = \frac{SC_u}{SC_{\mathbf{y}}} \quad (7)$$

Méthode : ANOVA et AS

Campbell et al (2006)

Plan $\xrightarrow{f(\mathbf{z},t)}$ Y \xleftrightarrow{ACP} H

Effets factoriels	Composantes principales				Marginales Complète
	PC_1	PC_2	...	PC_T	
z_1	$SC_{z_1,1}$	$SC_{z_1,2}$...	$SC_{z_1,T}$	SC_{z_1}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
z_1, z_2	$SC_{z_1, z_2,1}$	$SC_{z_1, z_2,2}$...	$SC_{z_1, z_2,T}$	SC_{z_1, z_2}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
Inertie	λ_1	λ_2	...	λ_T	\mathbb{I}

Méthode : ANOVA et AS

proposition

Sous l'hypothèse d'un plan orthogonal (Gram-schmidt)

$$\mathbb{I} = \text{Trace}\{\mathbf{y}'\mathbf{y}\} = \sum_{u \subseteq \{1,2,\dots,d\}} \text{Trace}\{\mathbf{y}'P_{E_u}\mathbf{y}\} \quad (8)$$

$$\text{GSI}_u = \frac{\text{Trace}\{\mathbf{y}'P_{E_u}\mathbf{y}\}}{\mathbb{I}} \quad (9)$$

$$\text{GSI}_{T.z_j} = \frac{\mathbb{I} - \text{Trace}\{\mathbf{y}'P_{-E_{\{j\}}}\mathbf{y}\}}{\mathbb{I}} \quad (10)$$

Facteurs continus : nouvelles réponses et indices

- Modèles :

- facteurs : $\mathbf{Z} = (Z_1, Z_2, \dots, Z_d)$

$$Z_j \sim \text{Unif}[0, 1]$$

- réponse du modèle :

$$Y(t) = f(\mathbf{Z}, t), \quad t \in \{1, 2, \dots, T\} \quad (11)$$

$$\mathbb{E}[Y(t)] = \mu(t), \quad \text{Cov}[Y(t_i), Y(t_j)] = \Sigma(t_i, t_j)$$

- Σ expansion

$$\Sigma = \sum_{k=1}^T \lambda_k \mathbf{v}_k \mathbf{v}_k' \quad (12)$$

- Output \mathbf{Y} expansion

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{H}\boldsymbol{\nu}' \quad (13)$$

Facteurs continus : nouvelles réponses et indices

$$\mathbf{H} = \{H_k = [\mathbf{Y} - \mu]' \mathbf{v}_k, \quad k \in \{1, 2, \dots, T\}\}$$

⇒ non corrélés et de variance décroissante

Indices de Sensibilité sur H_k (Saltelli et al., 2000) :

$$u \subseteq \{1, 2, \dots, d\}, \quad z_u = \{z_j, j \in u\}$$

$$\mathbb{IS}_{z_u, k} = \frac{\text{Var}[\mathbb{E}(H_k \mid z_u)]}{\text{Var}(H_k)} \quad (14)$$

$$\mathbb{IS}_{T, z_j, k} = \frac{\text{Var}(H_k) - \text{Var}[\mathbb{E}(H_k \mid -z_j)]}{\text{Var}(H_k)}. \quad (15)$$

Facteurs continus : nouvelles réponses et indices

Generalized Sensitivity Indices : GSI

proposition

$$\mathbb{I} = \text{Trace}\{\text{Cov}(\mathbf{H})\} = \sum_{u \subseteq \{1,2,\dots,d\}} \text{Trace}\{\text{Cov}[\mathbb{E}(\mathbf{H} \mid z_u)]\}$$

$$\text{GSI}_{z_u} = \frac{\text{Trace}\{\text{Cov}[\mathbb{E}(\mathbf{H} \mid z_u)]\}}{\mathbb{I}} \quad (16)$$

$$\text{GSI}_{T.z_j} = \frac{\mathbb{I} - \text{Trace}\{\text{Cov}[\mathbb{E}(\mathbf{H} \mid -z_j)]\}}{\mathbb{I}} \quad (17)$$

Application : AZODYN-INN

Modèle AZODYN-INN (Jeffroy et al., 1999)

$$Y(t) = f(\mathbf{Z}, t), \quad t = 1 \dots 169$$

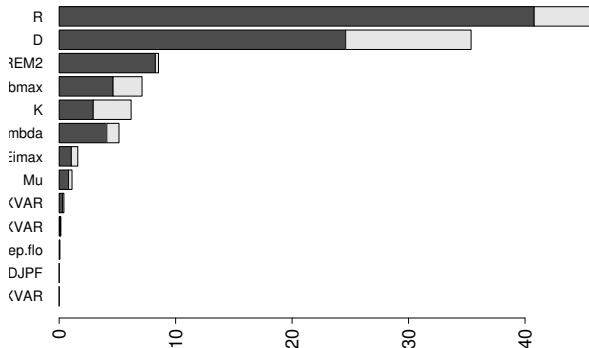
13 paramètres + variables d'entrées (soil, climate)

Étapes pour estimer les IS (facteurs continus)

- Monte-Carlo de taille 10000 (valeurs des paramètres)
- plan (10000 × 14) (Package Sensitivity, Saltelli (2002))
- simulation du modèle ($\mathbf{Y}_{(140000, 169)}$)
- ACP sur \mathbf{Y} pour estimer les vecteurs propres (Hall, 2006)
bootstrap pour intervalles de confiance
- simulation des nouvelles réponses
- calcul des indices de sensibilité (IS) (package sensitivity)

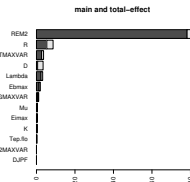
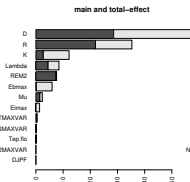
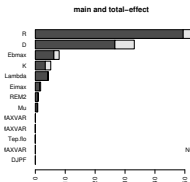
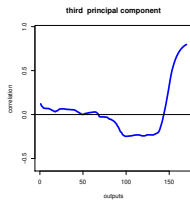
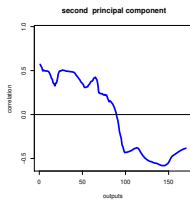
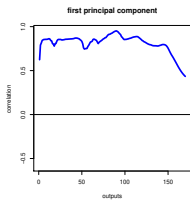
Application : indices (facteurs discrets)

GSI



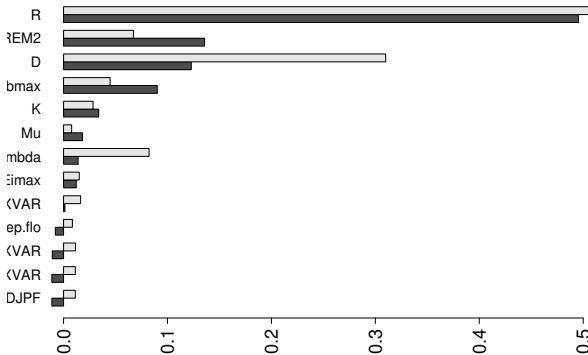
Application : indices (facteurs discrets)

IS sur les composantes



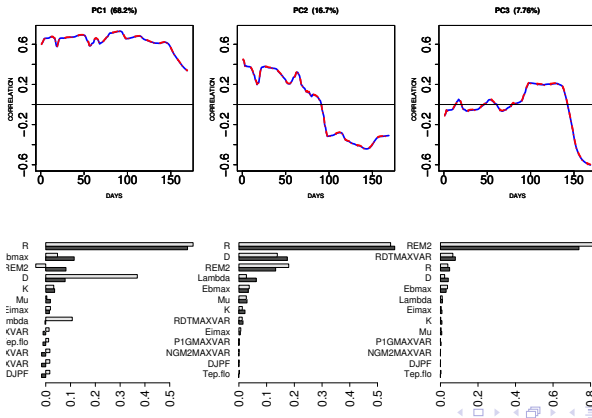
Application : indices (facteurs continus)

GSI



Application : indices (facteurs continus)

IS sur les composantes



Conclusion

- AS multivariée :
 - méthode d'AS pour les modèles dynamiques
 - identification des paramètres influents
 - l'importance des modules
- Perspectives
 - comparer les indices issus des facteurs discrets et continus
 - comparer les plans d'expériences (MC, LHS, factoriels)

Merci de votre attention

- [1] Campbell K., Mckay D. M, Williams J. B, 2006 : Sensitivity analysis when outputs are functions. Reliability Engineering & System Safety 91, 1468-1472
- [2] Campolongo F., Cariboni J., Saltelli A., 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software 22, 1509-1518.
- [3] Gabrielle B. , Laville P., Hénault C., Germon C. J., 2006. Simulation of nitrous oxide emissions from wheat -cropped soils using CERES. Nutrient Cycling in Agroecosystems 74, 133-146.
- [4] A., Tarantola S., Campolongo F., 2000. Sensitivity analysis as an ingredient of modeling. Statistical Science 15, 377-395