

Introduction à l'école-chercheurs *Analyse de sensibilité, métamodélisation et optimisation de modèles complexes*

Robert Faivre

INRA - Mathématiques et Informatique Appliquées - Université Fédérale de Toulouse



La Rochelle, 26 mars 2018



École-chercheurs

À l'initiative du
réseau Mexico

Analyse de sensibilité, métamodélisation et optimisation de modèles complexes

26-30 mars 2018 ; La Rochelle

Contexte et enjeux

Pour mieux comprendre les écosystèmes ou les systèmes économiques et sociaux, les chercheurs développent des modèles de plus en plus complexes. Explorer et calibrer ces modèles par simulation est indispensable car les interactions entre de nombreux composants dynamiques engendrent souvent des effets inattendus.

L'analyse de sensibilité globale est une méthode qui permet de (i) quantifier l'influence des différentes variables et paramètres sur la variabilité de la réponse d'un modèle; (ii) détecter des interactions entre les différentes entrées ; (iii) résumer un modèle complexe sous la forme d'un métamodèle plus facilement manipulable. Avec l'analyse d'incertitude, elle constitue ainsi une étape indispensable pour dimensionner un modèle, le valider ou orienter les efforts de recherche.

Son utilisation explicite est devenue incontournable dans de nombreux projets de modélisation mais reste souvent difficile à mettre en œuvre par des modélisateurs. Cette famille de méthodes permet d'identifier et hiérarchiser les paramètres clés qui, dès lors, doivent être estimés (calibrés) avec précision pour garantir un bon contrôle des sorties du modèle.

Au-delà de cette première phase d'analyse,

l'optimisation des modèles est l'étape qui permet d'appréhender la calibration des modèles, la recherche d'une décision optimale selon un ou plusieurs critères prédits par le modèle, etc. La principale difficulté de l'exercice d'optimisation vient du fait que les modèles complexes qui sont développés dans nos domaines sont bien souvent non analytiquement manipulables (non linéaires et non dérivables) et présentent des temps de simulation longs et coûteux.

Cette école-chercheurs a pour objectif de se concentrer autour de l'analyse de sensibilité, la métamodélisation et l'optimisation d'un modèle complexe qui présente au moins une des caractéristiques suivantes : le modèle est vu comme une boîte noire, un code de calcul long ou coûteux, de nombreux paramètres continus ou discrets, de multiples variables de sortie et des processus déterministes ou stochastiques. Les méthodes d'optimisation qui seront présentées viseront à trouver i) les valeurs des paramètres permettant de reproduire au mieux les observations du système modélisé (estimation de paramètres, calibration) ou ii) les valeurs de variables décisionnelles (discrètes ou continues) permettant de satisfaire certains critères, le tout en respectant un ensemble de contraintes préalablement définies. ■

Objectifs de l'école

L'école-chercheurs a pour objectifs de permettre aux participants de :

- avoir une vision globale des différentes méthodes d'analyse de sensibilité, de métamodélisation et d'optimisation de modèles complexes et des outils associés ;
- acquérir une démarche permettant de choisir la méthode la plus adaptée à son besoin ;
- favoriser les échanges d'expériences sur les champs d'interventions des organismes représentés tels que l'océanographie, l'halieutique, l'environnement, la gestion de l'eau, l'écologie, l'agronomie mais aussi les mathématiques et l'informatique appliquées. ■

Public

L'école-chercheurs est destinée prioritairement aux agents Inra (chercheurs ou ingénieurs y compris doctorants) ainsi qu'à nos partenaires du réseau Mexico (Cirad, Ifremer, Inraea ...) et aux participants du GDR Mascot-Num (universitaires et partenaires).

Elle s'adresse (i) aux agents qui développent des modèles, qui les analysent et qui souhaitent consolider leur maîtrise des méthodes et outils d'analyse de sensibilité et d'optimisation ; (ii) aux personnes souhaitant les acquérir. ■

Pré-requis :

- avoir une connaissance de l'utilisation du logiciel R est fortement recommandé, les TP de l'école utiliseront R.
- avoir des notions élémentaires de statistiques et de mathématiques.
- avoir déjà utilisé ou développé des modèles. ■

Modèles de culture

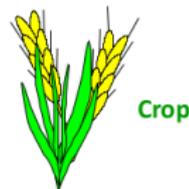
Conceptual modularity



Climate

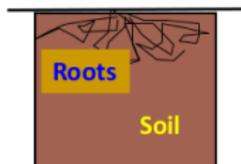
Structure

Development
Leaf area index
Radiation absorption
Biomass growth
Harvested organs



Crop

Irrigation
Water demand
Soil evaporation
Plant transpiration



Water & nitrate transfer

Crop & soil temperature

N fertilizer Organic Manure

N uptake **N₂ Fixation**

SOM N Mineralization

Residue decomposition

- All the processes simulated at a day-time scale
- In interaction with technical acts at the rotation scale

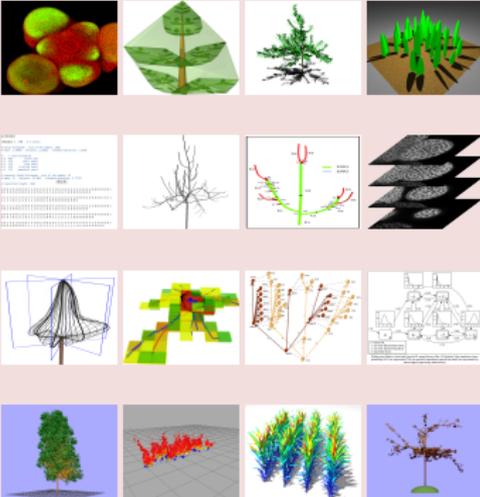
► STICS - Simulateur multidisciplinaire pour les cultures standard

Modèles structure-fonction d'architecture de plantes

OpenAlea in a nutshell Applications Conclusions

Background: plant modelling at a glance

- Different scales:
 - cell
 - branch
 - tree
 - forest
- Heterogeneous data:
 - raw data
 - digitised data
 - tree databases
 - 3D images
- Many tools required:
 - topology
 - Geometry
 - simulation
- Many models possible:
 - theoretical
 - mechanical
 - probabilist



Thomas Cokelaer IHC 2010, Lisbon, 25 August

- ▶ OpenAlea - Software Environment for Plant Modeling (Pradal et al.)

- ▶ Un grand nombre de processus donc de paramètres (Stics ≈ 400), certains corrélés
- ▶ Des processus dynamiques (équations différentielles ou aux différences) ou spatiaux
- ▶ Des entrées temporelles (météo par exemple) ou spatiales (paysage)
- ▶ Des pratiques culturales adaptatives (optimisées parfois)
- ▶ Des fonctions à seuil (stress, phénologie)
- ▶ Des processus parfois stochastiques (diffusion de spores, position/orientation d'organes)
- ▶ Temps de calcul modèle-dépendant ($< 1s$ à $1h$ voire plus)
- ▶ Un grand nombre de variables de sortie

- ▶ Estimation - Calibration des paramètres (optimisation)
- ▶ Qualité prédictive (analyse d'incertitude)
- ▶ Simplification du modèle (analyse de sensibilité, criblage, métamodélisation)
- ▶ Conception de pratiques (optimisation)
- ▶ Couplage de modèles (sensibilité, criblage)

MEXICO et Mascot-Num

- ▶ Méthodes d'EXploration Informatique de modèles COMplexes
- ▶ Méthodes d'Analyse Stochastique pour les COdes et Traitements NUMériques

Réseau Mexico

Méthodes pour l'Exploration Informatique des modèles Complexes



- Réseau proposé en 2006, à l'initiative forte de Vincent Ginot, comme réseau méthodologique du département MIA de l'INRA (Mathématiques et Informatique Appliquées)
- Animé par des chercheurs du département MIA (Hervé Richard, Avignon; Hervé Monod, Jouy-en-Josas; Robert Faivre, Toulouse)
- Regroupe des scientifiques (une vingtaine) de divers organismes : INRA (MIA, EA), Irstea (LISC), Ifremer (EMH) et Code Lutin, Université du Littoral Côte d'Opale (ULCO), Cirad, ...

R. Faivre

Introduction



MEXICO

Objet

Animation scientifique sur les méthodes d'exploration numérique de modèles complexes et sur leurs applications dans les domaines de la biologie, de l'agronomie, de l'écologie et de l'environnement.

Exploration numérique au sens large : analyse de sensibilité, propagation d'incertitudes, calcul de quantiles associés à de faibles probabilités, aide à la décision par optimisation de la sortie d'un code numérique, métamodélisation, etc.

Compétences, en particulier au sein de MIA : planification, optimisation, processus gaussien, modélisation statistique, etc.

Liens avec le GDR Mascot-Num et le RNSC.

R. Faivre

Introduction



MEXICO

Fonctionnement

- Des actions à vocation nationale voire internationale : conférences, séminaires et écoles-chercheurs
- Des actions à vocation plus opérationnelle :
 - Groupes de travail locaux
 - Réponses ciblées à certains appels d'offre
- Une cellule "développement logiciel" en charge de coordonner les efforts en matière de développement logiciel.
Boîte à outils pour différentes plateformes de modélisation
RECORD, ISIS-FISH, SimExplorer
- La participation à des réseaux à l'échelle nationale voire internationale.
 - RNSC (Réseau National des Systèmes Complexes)
 - GdR MASCOT-NUM (universitaires, industriels, CEA, IFP,...)

Quelques Actions

- ▶ Écoles-chercheurs organisées avec FormaSciences en 2009, 2010 et 2012 et le GdR Mascot-Num, le LJK de l'Université de Grenoble soutenue par MaiMoSINE en 2013 et 2014 et l'École centrale de Lyon en 2015 (~ 60 participants par session).
- ▶ Des journées thématiques : planification expérimentale, apprentissage et modèles, modélisation au Cemagref, aspects spatiaux dans les modèles complexes, système complexe et décision, modélisation intégrative, optimisation de modèles, regroupant plusieurs dizaines de participants.
- ▶ Rencontres Mexico : la dernière à Montpellier en 2017, la prochaine en novembre à Bordeaux.
- ▶ Grammaire XML de description d'une expérience numérique et package R mtk (lien avec le CATI IUMA Informatisation et Utilisation des Modèles dédiés aux Agroécosystèmes)
- ▶ Ouvrage : Faivre R., Iooss B., Mahévas S., Makowski D., Monod H., editors, 2013. Analyse de sensibilité et exploration de modèles. Applications aux modèles environnementaux. Collection *Savoir Faire*, Quae, 2013, 352p.

Afin d'aborder les grands enjeux liés au changement climatique et à la gestion durable de ressources naturelles ou exploitées, des modèles sont développés par les chercheurs en agronomie, écologie, environnement, halieutique, gestion de l'eau, océanographie, etc. Ces modèles intègrent de plus en plus la prise en compte de dynamiques et de processus liés à des systèmes complexes. Pour explorer leurs propriétés et juger de leur pertinence pour assister la décision, il est nécessaire de faire appel à des méthodes d'analyse et d'exploration adaptées. Il est alors souvent fait référence à une grande classe de méthodes, les analyses de sensibilité globales. Forts de leur expérience dans l'organisation d'écoles-chercheurs, les auteurs de cet ouvrage, membres pour la plupart du réseau de recherche interinstitutionnel Mexico (Méthodes pour l'exploration informatique de modèles complexes), ont souhaité transférer par cet ouvrage leur vision globale des différentes méthodes d'analyse de sensibilité et d'exploration et certaines règles d'analyse des modèles développés.

Ce livre s'adresse aux modélisateurs et utilisateurs de modèles qui souhaitent acquérir ou consolider leur maîtrise des méthodes d'analyse et d'exploration de modèles par simulation.

Robert Faivre, docteur en modélisation, calcul scientifique et statistique de l'Université Paris-Sud, Orsay, est directeur de recherche à l'Inra.

Bertrand Iooss, docteur en géostatistique de l'École des Mines de Paris, habilité à diriger des recherches, est chercheur senior au sein d'EDF R&D.

David Makowski, statisticien, agronome, habilité à diriger des recherches en sciences de la vie (Université Paris-Sud), est directeur de recherche à l'Inra.

Stéphanie Mahévas, docteur en mathématiques appliquées de l'Université Rennes 1, habilitée à diriger des recherches, est chercheur à l'Ifremer.

Hervé Monod, ingénieur agronome, docteur en statistique, est directeur de recherche à l'Inra. Il dirige l'unité MIA du centre de Jouy-en-Josas.

éditions
Quæ
Éditions Girad, Ifremer, Inra, Instea
www.quæ.com

55 €

ISBN : 978-2-7592-1906-3



9 782759 219063

ISSN : 1952-1251
Ref. : 02370Savoir
faire

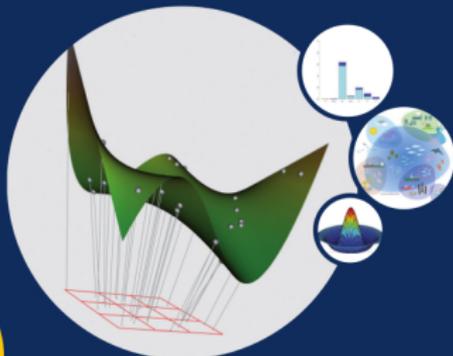
Analyse de sensibilité et exploration de modèles

Savoir
faire

Analyse de sensibilité et exploration de modèles

Application aux sciences de la nature
et de l'environnement

R. Faivre, B. Iooss, S. Mahévas, D. Makowski,
H. Monod, éd.

éditions
Quæ

Coordinateurs : Bertrand Iooss (EDF R&D), Hervé Monod (directeur, MIA Inra),
Clémentine Prieur (UJF Grenoble)

Nouvelles du GdR

Projet de renouvellement

Thématiques du GDR

MASCOT-NUM a pour objectifs:

- d'animer et coordonner les efforts de recherche
- pour développer des outils de planification, modélisation et analyse d'expérimentations *in silico*
- en interagissant avec la R&D industrielle et la recherche publique
- en recourant à toutes les techniques de mathématique appliquée : statistique, probabilité, informatique, analyse numérique, recherche opérationnelle, ...

Soutien 2016-2020

- GDR n° 3179
- Institut de rattachement : Institut national des sciences mathématiques et de leurs interactions (INSMI)
- Institut secondaire : Institut des sciences de l'information et de leurs interactions (INS2I)



MASCOT NUM

Groupes de travail

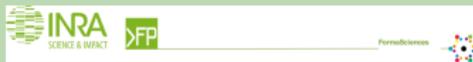
- ▶ **Implémentation Numérique** - Yann Richet, Laurence Viry
- ▶ **Planification d'expériences numériques** - Clémentine Prieur, Luc Pronzato
- ▶ **Métamodélisation** - Anthony Nouy
- ▶ **Sensibilité, calibration, validation** - Sébastien da Veiga, Nabil Rachdi
- ▶ **Risque et incertitudes** - Nicolas Bousquet, Bruno Sudret
- ▶ **Optimisation** - Emmanuel Vasquez
- ▶ **Industrial applications** - Fabien Mangeant
- ▶ **Application en environnement, biologie, agriculture** - Robert Faivre

Réunions

- ▶ Des réunions régulières selon les animations des groupes
- ▶ Des journées annuelles du GdR (la semaine dernière à Nantes)

liste des soutiens

- ▶ FormaSciences



- ▶ Réseau Mexico



- ▶ Départements INRA



- ▶ et le GdR Mascot-Num



comité d'organisation

- ▶ FormaSciences : Geneviève Aubin-Houzelstein et Nathalie Frelat (Inra, FPN Paris)
- ▶ Réseau Mexico : H. Drouineau (Irstea), R. Faivre, V. Picheny et S. Roux (Inra)
- ▶ GdR Mascot-Num : J. Bect (Supelec), R. Faivre et V. Picheny (Inra)

intervenants

- ▶ FormaSciences : G. Aubin-Houzelstein (Inra, FPN Paris)
- ▶ Réseau Mexico : R. Faivre, V. Picheny et S. Roux (Inra), S. Léhuta et S. Mahévas (Ifremer), H. Drouineau et N. Dumoulin (Irstea),
- ▶ GdR Mascot-Num : J. Bect (Supelec), N. Durrande (Prower.io), R. Faivre et V. Picheny (Inra), R. Le Riche (Cnrs),
- ▶ autre : J. Papaïx (Inra, BioSP Avignon)

58 participants

- ▶ Universités et Écoles (8), Inra (27), Cirad (5), Ifsttar (4), Isrtea (3), Ifremer (2), Cnrs, Ird, Mhnh, Arvalis, Cea, Ifpen, Ineris, IRSN, IMR
- ▶ Inra : Bap (2), Cepia (1+1), EA (13), Efpa (2), Mia (2), Phase (2), SA, Sad, Sae2, Spe (2)
- ▶ Doctorants (21), PostDocs (5), CR - MC (19), Ingénieurs (12), Informaticien

PROGRAMME DE L'ECOLE-CHERCHEURS

26 au 30 Mars 2018

Analyse de sensibilité, métamodélisation et optimisation de modèles complexes



Lundi 26 Mars

9h30-11h00	Accueil
11h00-11h30	Introduction générale (Robert Faivre)
11h30-12h30	Introduction à l'exploration et l'optimisation de modèles sur un exemple (Sigrid Léhuta)
12h30-14h00 : Déjeuner	
14h00-15h00	Introduction à l'exploration de modèles : plans d'expérience, métamodélisation. (Nicolas Durrande et Victor Picheny)
15h00-16h30	Cours 1 : Analyse de sensibilité (Robert Faivre)
16h30-17h00 : Pause	
17h00-18h00	Mise en place des TP (Victor Picheny)
19h Pot d'accueil	

Mardi 27 Mars

8h30-10h30	TP Cours 1 : Analyse de sensibilité (Robert Faivre et Sébastien Roux)
10h30-11h00 : Pause	
11h00-12h30	Cours 2 : Processus gaussiens (Nicolas Durrande et Victor Picheny)
12h30-14h00 : Déjeuner	
14h00-15h30	TP Cours 2 : Processus gaussiens (Nicolas Durrande et Victor Picheny)
15h30-16h00 : Pause	
16h00-16h30	Introduction à la calibration de modèles (Stéphanie Mahévas)
16h30-18h00	Cours 3 : approche bayésienne (Julien Papaix)
19h30 : Dîner	
21h15	Modèle au bar

Bien respecter le choix de la visite entre Aquarium et Ville

Mercredi 28 Mars

8h30-10h30	TP Cours 3 (Julien Papaix)
10h30-11h00 : Pause	
11h00-11h45	Etude de la colonisation de bassin-versant par les anguilles par développement et calibration d'un modèle complexe (Hilaire Drouineau)
11h45-12h30	Exposé méthode ABC (Nicolas Dumoulin)
12h30-14h00 : Déjeuner	
14h-19h30: Après-midi détente et visite	
19h30 : Dîner	
21h15	Modèle au bar

Jeudi 29 Mars

8h30-10h00	Cours 4 : optimisation avec métamodèle (Nicolas Durrande et Victor Picheny)
10h00-10h30 : Pause	
10h30-12h30	TP Cours 4 : optimisation avec métamodèle (Nicolas Durrande et Victor Picheny)
12h30-14h00 : Déjeuner	
14h00-15h30	Cours 5 : optimisation sans métamodèle (Rodolphe Le Riche)
15h00-15h30 : Pause	
15h30-17h30	TP Cours 5 : optimisation sans métamodèle (Rodolphe Le Riche)
19h30 : Dîner Repas Régional	

Vendredi 30 Mars

9h00-10h30	Exposé extentions : optimisation multiobjectif, bruitée, ... (Julien Bect)
10h30-11h00 : Pause	
11h00-12h00	Conclusion, Synthèse et Discussions sur les méthodes (Stéphanie Mahévas)
12h00-12h30	Évaluation à chaud de l'école-chercheurs (Geneviève Aubin-Houzelstein)
12h30-14h00 : Déjeuner ou panier repas	