

## **STOCHASTREE, UN MODELE DE TRANSITION DE CULTURE UTILISANT DES ARBRES DE DECISION STOCHASTIQUES POUR INTEGRER DES CONTRAINTES AGRONOMIQUES.**

---

**STOCHASTREE, A CROP TRANSITION MODEL BASED ON STOCHASTIC DECISION TREES, WHICH INTEGRATES AGRONOMIC CONSTRAINTS.**

Luc SOREL, Christian WALTER, Patrick DURAND, Valérie VIAUD

INRA, UMR SAS, 65 rue de Saint-Brieuc, CS-84215, 35042 Rennes cedex,  
luc.sorel@rennes.inra.fr

### **RESUME**

L'évaluation des impacts environnementaux des pratiques agricoles nécessite l'utilisation d'approches de modélisation pour simuler les flux hydrologiques à l'échelle du paysage, dont le temps de réponse peut s'étendre sur plusieurs décennies. L'allocation des cultures sur le parcellaire spatiale les pratiques de fertilisation impliquées dans les problèmes de qualité des sols et de l'eau. Afin de faciliter la construction de scénarios agricoles, notre objectif était de développer un modèle de transition de culture tenant compte de déterminants agronomiques et spatiaux: objectifs de production végétale, distribution spatiale des cultures autour des corps de ferme, adéquation des cultures à l'hydromorphie du sol. Nous avons développé un modèle innovant basé sur des arbres de décision stochastiques, Stochastree, pour intégrer des caractéristiques des types de ferme et des parcelles (surface, distance au corps de ferme, hydromorphie, culture en cours) dans le processus de simulation des transitions de culture sans expertise préalable, mais qui s'appuie sur une approche de data-mining. Les résultats de simulation ont été comparés avec des données de terrain et à un modèle markovien basé sur des matrices de transition, Rotomatrix, pour tester ses capacités à suivre les contraintes agronomiques présumées. La structure générale des arbres de décisions ainsi construits avait une structure similaire aux matrices de transitions. Stochastree et Rotomatrix ont montré des performances similaires concernant la prédiction des transitions de culture et les objectifs de production. Stochastree a été significativement meilleur dans sa capacité à maintenir la distribution spatiale des cultures autour des corps de ferme, et légèrement meilleur dans l'allocation préférentielle des cultures aux classes d'hydromorphie du sol. La construction facile de tels arbres de décision ouvre des perspectives de couplage de Stochastree avec des modèles écologiques, comme des modèles de transfert de solutés de diffusion de gènes.

### **ABSTRACT**

Evaluating the environmental impacts of agricultural practices requires the use of simulation approaches to model hydrological fluxes at the landscape scale and with response time spanning over decades. The crop allocation over the field pattern determines the spatial distribution of fertilization practices, which are involved in soil and water quality issues. To facilitate the construction of agricultural scenarios, our objective was to develop a crop transition model able to account for agronomic and spatial driving factors: crop production objectives, spatial distribution of the crops around the farmsteads, and preferential allocation of crops on soil waterlogging classes. We developed an innovative model based on stochastic decision trees, Stochastree, to integrate farm-type and field characteristics (area, distance to farmstead, waterlogging, current crop) in the crop transition simulation process without prior expert knowledge, but relying on a data-mining approach. Simulation results of Stochastree were compared to field data and to a Markovian model based on transition matrices, Rotomatrix, to test its abilities to follow presumed agronomic constraints. Learned decision trees had a general structure similar to transition matrices. Stochastree and Rotomatrix exhibited similar performances in predicting crop transitions and outputting expected crop productions. Stochastree proved to be significantly superior in maintaining the spatial distribution of crops around the farmsteads and slightly better in allocating crops to the proper soil waterlogging class. The ease to construct decision trees suggest many potential couplings of Stochastree to various ecological models, like nutrient diffuse transfer models or gene transmission models.

### **REFERENCES**

- Anderson T. W., L. A. Goodman. 1957. Statistical inference about Markov chains. *Annals of mathematical statistics* 28 (1):89-110.
- Bachinger J., P. Zander. 2007. ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. *European journal of agronomy* 26 (2):130-143.
- Baker W.L. 1989. A review of models of landscape change. *Landscape ecology* 2 (2):111-133.
- Baudry J., C. Dupont, C. Thenail, V. Viaud. 2006. Aménagement du paysage et pratiques agricoles : quelles combinaisons dans la gestion des bassins versants ? In Mérot, P. *Qualité de l'eau en milieu rural : Savoirs et pratiques dans les bassins versants*. Inra Eds, Paris.

- Beaujouan V., P. Durand, L. Ruiz. 2001. Modelling the effect of the spatial distribution of agricultural practices on nitrogen fluxes in rural catchments. *Ecological modelling* 137:93-105.
- Beaujouan, V., P. Durand, L. Ruiz, P. Aourousseau, G. Cotteret. 2002. A hydrological model dedicated to topography-based simulation of nitrogen transfer and transformation: rationale and application to the geomorphology-denitrification relationship. *Hydrological processes* 16:493-507.
- Bordenave P., F. Ehler, T. Bioteau, N. Turpin, P. Serrand, P. Saint-Cast, E. Le Saos. 2005. Distributed surface-groundwater coupled model applied to long term water management impacts at basin scale. *Ingénieries* 44:13-36.
- Breiman, L., J. H. Friedman, R. A. Olshen, C. J. Stone. 1984. *Classification and regression trees*. Monterey, CA, Wadsworth.
- Castellazzi M. S., G. A. Wood, P. J. Burgess, J. Morris, K. F. Conrad, J. N. Perry. 2007b. A systematic representation of crop rotations. *Agricultural systems*, doi:10.1016/j.agsy.2007.10.006.
- Coppedge B. O., D. M. Engle, S. D. Fuhlendorf. 2007. Markov models of land cover dynamics in a southern Great Plains grassland region. *Landscape ecology* 22:1383-1393.
- Dogliotti S., W. A. H. Rossing, M. K. van Ittersum. 2003. ROTAT, a tool for systematically generating crop rotations. *European journal of agronomy* 19 (2):239-250.
- Gauchere C., N. Giboire, V. Viaud, T. Houet, J. Baudry, F. Burel. 2006. A domain-specific language for patchy landscape modelling: the Brittany agricultural mosaic as a case study. *Ecological modelling* 194:233-243.
- Houet T., L. Hubert-Moy, 2006. Modelling and projecting land-use and land-cover changes with cellular automaton in considering landscape trajectories: an improvement for simulation of plausible future states. *EARSeL eProceedings*, 5 (1):63-76.
- Ladet S., M. Deconchat, C. Monteil, J.-P. Lacombe, G. Balent. 2005. Les chaînes de Markov spatialisées comme outil de simulation. *Géomatique, Cassini* 04:159-173.
- Largouët C., M.-O. Cordier. 2001 Improving the landcover classification using domain knowledge. *AI communication* 1:35-43.
- Lawrence R., A. Bunn, S. Powell, M. Zambon. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote sensing of environment* 90:331-336.
- Lynn H., C. L. Mohler, S. D. DeGloria, C. E. McCulloch. 1995. Error assessment in decision-tree models applied to vegetation analysis. *Landscape ecology* 10 (6):323-335.
- Molénat J., C. Gascuel-Oudou. 2002. Modelling flow and nitrate transport in groundwater for the prediction of water travel times and of consequences of land use evolution on water quality. *Hydrological Processes* 16 (2):479-492.
- Murthy S. K. 1998. Automatic construction of decision trees from data: a multi-disciplinary survey. *Data mining and knowledge discovery* 2:345-389.
- Pal M., P. M. Mather. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote sensing of environment* 86:554-565.
- Landscape Ecology* 23 (2):195-210.
- Provost F., P. Domingos. 2003. Tree induction for probability-based ranking. *Machine Learning* 52 (3):199-215.
- Quinlan, J.R. 1990. Probabilistic decision trees. In *Machine Learning: An Artificial Intelligence Approach - Volume III*, Y. Kodratoff and R.S. Michalski, Eds. Morgan Kaufmann, San Mateo, CA, 140-152.
- Rounsevell M. D. A., J. E. Annetts, E. Audsley, T. Mayr, I. Reginster. 2003. Modelling the spatial distribution of agricultural land use at the regional scale. *Agriculture, ecosystems & environment* 95 (2-3):465-479.
- Thornton P. K., P. G. Jones. 1998. A conceptual approach to dynamic agricultural land-use modelling. *Agricultural System* 7. 505-521.
- Usher M. B. 1992. Statistical models of succession. In Glenn-Lewin DC, Peet RK, Veblen TT (eds) *Plant succession: theory and prediction*. Chapman & Hall, London, pp 215-248.
- Verburg P. H. 2006. Simulating feedbacks in land use and land cover change models. *Landscape ecology* 21:1171-1183.
- Verburg P. H., P. P. Schot, M. J. Dijst, A. Veldkamp. 2004. Land use change modelling: current practice and research priorities. *Geojournal* 61:309-324.
- Walter C., P. Curmi. 1998. Les sols du bassin versant du Coët-Dan : organisation, variabilité spatiale et cartographie. In Cheverry C. (ed.). *Agriculture intensive et qualité des eaux*. Science Update: Paris.
- Walter C., R. A. Viscarra Rossel, A. B. McBratney. 2003. Spatio-temporal simulation of the field-scale evolution of organic carbon over the landscape. *Soil Science Society of America Journal* 67:1477-1486.
- Witten I. H., E. Frank. 2005. *Data Mining: Practical machine learning tools and techniques*, 2nd Edition. Morgan Kaufmann, San Francisco.