HAL
open science

Les cartes auto-organisatrices de Kohonen appliquées à l'étude des communautés de micro-algues des cours d'eau

Marius Bottin, Jean-Luc Giraudel, Julie Guéguen, Sébastien Boutry, Isabelle Lavoie, Nicolas Antunes, Juliette Tison-Rosebery

To cite this version:

Marius Bottin, Jean-Luc Giraudel, Julie Guéguen, Sébastien Boutry, Isabelle Lavoie, et al.. Les cartes auto-organisatrices de Kohonen appliquées à l'étude des communautés de micro-algues des cours d'eau. 1ères Rencontres R, Jul 2012, Bordeaux, France. hal-00717538

HAL Id: hal-00717538

https://hal.science/hal-00717538v1

Submitted on 13 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les cartes auto-organisatrices de Kohonen appliquée à l'étude des communautés de micro-algues des cours d'eau

M. Bottin[a], **J.L. Giraudel**[b], **J. Guéguen**[a], **S. Boutry**[a], **I. Lavoie**[c], **N. Antunes**[d] and **J. Tison-Rosebery**[a]

[a]Unité REBX
Irstea
50 avenue de Verdun, 33612 Cestas cedex, France
Marius.Bottin@irstea.fr

[b]EPOC
Université de Bordeaux - UMR CNRS 5805
Rue Doyen Joseph LAJUGIE 24019 PERIGUEUX Cedex, FRANCE

[c]Centre Eau Terre Environnement
Institut National de la Recherche Scientifique
490 rue de la Couronne,Québec, Québec,Canada

[d] UMR PACEA/ALEA.
Université de Bordeaux - UMR CNRS 5199 / INRIA
Université Bordeaux 1 - 351 cours de la libération 33405 Talence Cedex

Mots clefs : Écologie, Diatomées, Communautés, Cartes auto-organisatrices de Kohonen, Bioindication.

Les diatomées sont des algues microscopiques libres ou fixées sur les galets ou végétaux des rivières. Les différentes espèces pouvant composer ces communautés présentent des préférences environnementales différentes, ce qui permet une description des conditions écologiques à partir de l'observation des espèces présentes. Elles sont donc utilisées en routine pour l'évaluation de l'état écologique des cours d'eau depuis une vingtaine d'années [1, 2]. Afin de mettre en œvre la Directive Cadre Européenne sur l'Eau, Irstea a participé au développement de nouvelles méthodes d'évaluation, grâce à l'analyse de grandes bases de données d'échelle nationale. Du fait de leur grande diversité et de la complexité des phénomènes agissant sur les communautés de diatomées, l'analyse de leur répartition requiert l'utilisation de méthodes innovantes telles que la logique floue ou les réseaux de neurones. En effet, les méthodes statistiques classiques s'avèrent souvent incapables de décrire une part acceptable de la variation de ces communautés.

Une cartes auto-organisatrice de Kohonen (Self-Organizing Kohonen Maps : SOM) est un réseau de neurones à apprentissage non supervisé. Cette technique permet à la fois une classification objective des communautés et l'étude simultanée de gradients biologiques et environnementaux. Un de ses avantages principaux est la prise en compte de gradients non-linéaires complexes, qui explique en partie sa grande efficacité pour décrire les communautés biologiques [3]. Aussi, cette technique a été utilisée pour décrire, entre autres, les variations spatiales [4] ou l'impact des pesticides [5] sur les communautés de diatomées. En revanche, comme pour les autres réseaux de neurones, de nombreux paramètres doivent être réglés, de façon adaptée aux différents types de jeux de données (nombre de neurones, mesures de dissimilarité, nombres d'étapes, fonctions de voisinages ...).

Des packages existent sous R pour la réalisation des cartes de Kohonen, mais ils sont à notre connaissance assez peu utilisés pour l'ordination de communautés biologiques. Des solutions existantes sous d'autres logiciels (en particulier MATLAB) paraissent mieux adaptées à ce type de données. Afin de pouvoir adapter plus particulièrement cette technique à nos besoins, nous avons développé un ensemble de fonctions sous R permettant la réalisation et la visualisation de SOM. L'utilisation conjointe de R et du langage C nous a permis d'atteindre une efficacité de calcul mieux adaptée à la grande taille de nos tableaux de données, tout en gardant une relative flexibilité de mise en œuvre. Les possibilités graphiques de ce logiciel nous ont de plus permis de développer des outils de visualisation et d'analyses de SOM, eux aussi mieux adaptés à nos besoins.
Les différentes fonctions ont été compilées dans un nouveau package, contenant aussi d'autres techniques (comme les « fuzzy pattern trees »), que nous espérons soumettre aux CRAN d'ici la fin de l'année 2012.

Nous discuterons plus particulièrement de l'algorithme utilisé dans ce package et des possibilités qu'il offre, notamment dans le domaine de l'écologie. Aussi, nous agrémenterons cet exposé d'exemples dans lesquels ces fonctions sont déjà utilisées pour la typologie des communautés :
- des cours d'eau français de métropole.
- des cours d'eau de la Guadeloupe et de la Martinique pour l'application de la Directive Cadre Européenne sur l'Eau.
- de cours d'eau canadiens dans le cadre de la réalisation d'une nouvelle version de l'« Indice Diatomées de l'Est du Canada ».

Références

[1] J. Prygiel, M. Coste, and J. Bukowska. Review of the major diatom-based techniques for the quality assessment of rivers. state of the art in europe. In *Use of algae for monitoring rivers III*, pages 224–238. Agence de I'Eau Artois-Picardie Press Douai Cedex, 1999.

[2] M. Coste, S. Boutry, J. Tison-Rosebery, and F. Delmas. Improvements of the biological diatom index (BDI) : Description and efficiency of the new version (BDI-2006)mo. *Ecological Indicators*, 9(4) :621–650, 2009.

[3] J. L. Giraudel and S. Lek. A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling*, 146(1-3) :329–339, 2001.

[4] J. Tison, Y. S. Park, M. Coste, J. G. Wasson, L. Ector, F. Rimet, and F. Delmas. Typology of diatom communities and the influence of hydro-ecoregions : A study on the french hydrosystem scale. *Water Research*, 39(14) :3177–3188, 2005.

[5] S. Morin, M. Bottin, N. Mazzella, F. Macary, F. Delmas, P. Winterton, and M. Coste. Linking diatom community structure to pesticide input as evaluated through a spatial contamination potential (phytopixal) : A case study in the neste river system (south-west france). *Aquatic Toxicology*, 94(1) :28–39, 2009.