

UTILISER LES MODÈLES INDIVIDUS-CENTRÉS COMME LABORATOIRES VIRTUELS POUR IDENTIFIER LES PARAMÈTRES D’UN MODÈLE AGRÉGÉ

Raphaël Duboz, Éric Ramat, Philippe Preux

Laboratoire d'Informatique du Littoral (LIL)
Maison de la Recherche Blaise Pascal
50, rue Ferdinand Buisson - BP 719
62228 Calais Cedex, France
Mél : duboz@lil.univ-littoral.fr

Frédéric Amblard, Guillaume Deffuant

Laboratoire d'Ingénierie pour les Systèmes Complexes
(LISC) Cemagref
24, avenue des Landais BP 50085
63172 Aubière Cedex, France
Mél : frederic.amblard@cemagref.fr

RÉSUMÉ : Nous proposons d'utiliser des modèles individus-centrés comme laboratoires virtuels pour identifier des paramètres de modèles agrégés. Nous illustrons cette approche par un exemple simple : la diffusion de particules animées d'un mouvement brownien. Dans ce cadre, nous disposons donc d'un modèle numérique agrégé et d'un modèle individus-centré de ce phénomène. Nous considérons différentes variantes de cet exemple qui conduisent à des utilisations différentes du laboratoire virtuel. Nous considérons notamment le couplage entre les deux modèles dans lequel le modèle agrégé demande des calculs de paramètres à la volée au modèle individus-centré. Nous évoquons enfin le cas du changement d'échelles dans les modèles et la manière dont il peut être géré par ce type d'approche.

MOTS-CLÉS : *couplage, multi-formalisme, équations différentielles, modèles individus-centrés*

1. INTRODUCTION

Pour décrire un système physique ou biologique, les équations différentielles sont l'outil fondamental des scientifiques. Cette approche tend à décrire un système constitué de nombreux éléments par des variables agrégées (des concentrations, des densités de populations) et s'intéresse à la dynamique de ces variables. Il s'ensuit une représentation très compacte et très générale de l'évolution du système. Cependant, ces modèles font intervenir des paramètres agrégés dont l'identification peut s'avérer délicate.

Au contraire les modèles individus-centrés représentent explicitement l'ensemble des éléments du système ainsi que leurs interactions. Ils sont utilisés de plus en plus largement en particulier en écologie (Grimm, 1999). Là aussi, les équations différentielles sont de mises en général pour exprimer le comportement global résultant de l'ensemble des comportements individuels des éléments du système. Cependant, ces modèles trouvent une expression plus « naturelle » dans des travaux en informatique, depuis certains pionniers de la programmation orientée-objet pour la simulation des systèmes (Dahl, 1966), jusqu'aux travaux plus récents dans le domaine des systèmes multi-agents (Ferber, 1995). Aussi les représentations des dynamiques individuelles peuvent-elles parfois s'écarter d'un formalisme différentiel pur. Les modèles individus-

centrés considérés dans ce texte ne sont pas nécessairement des modèles multi-agents, si nous adoptons la définition stricte de Wooldridge et Jennings (Wooldridge and Jennings, 1995). Nous utilisons donc le terme individus-centré dès lors que la modélisation se focalise sur les entités composant le système, par opposition à une approche considérant des variables agrégées représentant l'ensemble des entités.

En fait, ces deux types de modèles peuvent être vus comme des représentations du même système à deux niveaux d'abstraction différents. Dans ce papier, nous tentons de démontrer l'intérêt de développer les deux niveaux d'abstraction ensemble. Notre argument principal est que ces deux niveaux peuvent s'enrichir mutuellement en ayant recours l'un à l'autre. Dans notre proposition, le modèle individus-centré et le modèle agrégé simulent le même système avec un point de vue différent i.e. chacun d'entre eux ne modélise pas seulement une partie du système, comme c'est le cas généralement lorsqu'on définit des multi-modèles (Fishwick, 1995). Cependant, ce concept de multi-modélisation est apparenté à notre démarche. En effet, on y retrouve cette idée de décrire un phénomène à l'aide de plusieurs modèles, à des niveaux d'abstraction ou de raffinement différents. Nous sommes également proches de l'idée d'abstraction décroissante de Lindenberg (Lindenberg, 1992) qui préconise le développement de modèles de plus en plus fins, dont la compréhension est

nourrie de celle des modèles de granularité plus large, éventuellement dans des formalismes différents.

Nous proposons ici un type particulier de coopération entre deux niveaux d'abstraction : l'utilisation du modèle individu-centré comme un laboratoire virtuel, permettant de faire des expérimentations et des mesures, afin d'identifier des paramètres d'un modèle agrégé.

Pour illustrer concrètement cette idée, nous considérons un exemple simple : la diffusion dans un fluide de particules animées d'un mouvement brownien. Le modèle agrégé se réduit à l'équation d'évolution de la concentration des particules dans l'espace. Le modèle individus-centré représente l'ensemble des particules et leurs mouvements aléatoires. Nous considérons différentes variantes de ce phénomène dans lesquelles le modèle individus-centré peut être appelé à la volée par le modèle agrégé pour identifier certains paramètres ou variables au cours du temps. Nous montrons que ces opérations peuvent dans certains cas nécessiter le recours à des changements d'échelles. Enfin, nous proposons une discussion sur l'intérêt de tels couplages par rapport à un modèle individus-centré seul, ou l'intérêt de réaliser un tel couplage en ligne, par rapport à la préparation préalable des mesures dans le laboratoire virtuel.

2. MODÈLE AGRÉGÉ ET INDIVIDUS-CENTRÉ DE LA DIFFUSION DE PARTICULES

La diffusion de particules est un exemple particulièrement simple de phénomène pour lequel la description physique microscopique est reliée mathématiquement à la description macroscopique du système par la mécanique statistique. Il est donc aisé d'élaborer un modèle agrégé et un modèle individus-centré de ce phénomène et d'établir un lien entre eux. Ce phénomène nous a donc paru à la fois suffisamment simple pour être facile à exposer et suffisamment riche pour illustrer notre démarche.

2.1. Modèle agrégé

L'étude du phénomène de diffusion par la théorie cinétique a établi l'équation de Fick, modèle agrégé de la diffusion brownienne (Fick, 1855). On s'intéresse ici uniquement à la composante le long de l'axe des abscisses (x) :

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \quad (1)$$

où c est la concentration en particules, D le coefficient de diffusion et x l'axe sur lequel a lieu la diffusion.

Deux méthodes existent pour résoudre une équation différentielle. La première consiste à intégrer l'équation analytiquement. La deuxième passe par une résolution numérique utilisant un schéma d'intégration approprié.

Lorsque l'on connaît la concentration initiale et les limites de l'espace, il est possible de calculer la solution de l'équation de Fick analytiquement :

$$c(x, t) = \frac{c_0 e^{-x^2/4Dt}}{2\sqrt{\pi Dt}} \quad (2)$$

où c_0 est la concentration initiale, $c(x, t)$ est la concentration en un point à un instant donné et D le coefficient de diffusion.

Cette solution est une gaussienne (cf. courbe en pointillés de la figure 2).

Cependant, si on suppose que le coefficient de diffusion est variable dans l'espace, la solution analytique n'est pas toujours calculable. Il faut alors utiliser une résolution numérique de l'équation à coefficient variable et à pas constant (équations 3) (Dautray et Lions, 1984), en discrétisant l'espace et le temps :

$$\frac{1}{\Delta t} (C_i^{t+\Delta t} - C_i^t) - \frac{1}{\Delta x} \left[D_{i+1} \frac{C_{i+1}^t - C_i^t}{\Delta x} - D_{i-1} \frac{C_i^t - C_{i-1}^t}{\Delta x} \right] = 0 \quad (3)$$

C_i^t est la concentration en particules au pas d'espace i et au pas de temps t , on suppose que C_i^0 est donné.

Δx est le pas d'espace

Δt est le pas de temps

D_i est le coefficient de diffusion au pas d'espace i

2.2. Modèle individus-centré correspondant

Le modèle individus-centré simulant le phénomène de diffusion est particulièrement simple dans un premier temps, mais nous verrons par la suite que celui-ci peut être complexifié à loisir pour répondre à des problématiques plus précises.

Nous supposons que les particules évoluent dans un cube et sont repérées dans l'espace par leurs coordonnées (x, y, z) . L'impact du choix de la topologie de l'espace sur ce type de simulation a été étudié par ailleurs (Duboz *et al.*, 2001). Nous supposons par exemple que les particules sont initialement distribuées aléatoirement dans le plan défini par $x=0$ et séparant le cube en deux moitiés de même volume. La simulation consiste à donner une direction aléatoire (entre 0 et 2π) et une vitesse aléatoire (entre 0 et v_{\max}) à chaque particule pour chaque itération. La condition aux limites du cube est l'impossibilité pour toute particule d'en sortir. On retire donc aléatoirement une direction et une vitesse pour toute particule qui sort du cube jusqu'à obtenir des valeurs correctes. Ce modèle apparaît principalement stochastique. Il est donc nécessaire de simuler un très grand nombre de particules pour respecter l'hypothèse ergodique du modèle agrégé et pour que l'observation du

comportement du modèle ne soit pas un événement particulier mais le reflet d'un comportement moyen de l'ensemble des particules.

Ce modèle donne la position de toutes les particules à chaque instant. Il permet donc de calculer la concentration en particules en toute partie de l'espace, mais aussi éventuellement d'autres grandeurs.

3. LABORATOIRE VIRTUEL POUR LA DÉTERMINATION DE PARAMÈTRES

Le modèle individus-centré peut ainsi nous donner une estimation du paramètre de diffusion D du modèle agrégé, correspondant à une vitesse maximale des particules.

Pour cela, on considère le modèle individus-centré comme un laboratoire virtuel sur lequel on pratique des expérimentations et des mesures (Legay, 1997 ; Grimm, 1999). Ainsi, en partant d'une distribution aléatoire des particules dans le plan $x = 0$, on mesure à chaque instant t , l'écart quadratique moyen \bar{e} de l'ensemble des particules, c'est-à-dire la racine carrée de la somme des carrés des distances entre les particules et l'origine de l'axe des x divisée par le nombre de particules.

On peut en déduire le coefficient de diffusion D de manière expérimentale, à l'aide de la relation :

$$\bar{e} = \sqrt{2Dt} \quad \text{d'où} \quad D = \frac{1}{2t} \bar{e}^2 \quad (4)$$

où \bar{e} est l'écart quadratique moyen, D le coefficient de diffusion et t le temps.

La figure 1 montre l'évolution de la quantité $\frac{1}{2} \bar{e}^2$ en fonction du temps. C'est en calculant la pente de cette droite que nous pouvons obtenir le coefficient de diffusion D .

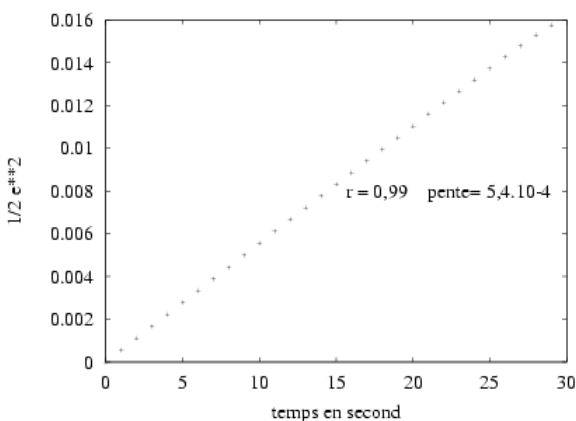


Figure 1. $\frac{1}{2} \bar{e}^2$ en fonction du temps. La pente de cette droite est le coefficient de diffusion D .

On peut ensuite vérifier que le modèle microscopique des particules a le même comportement que l'équation de Fick macroscopique. La figure 2 permet de comparer visuellement les deux modèles.

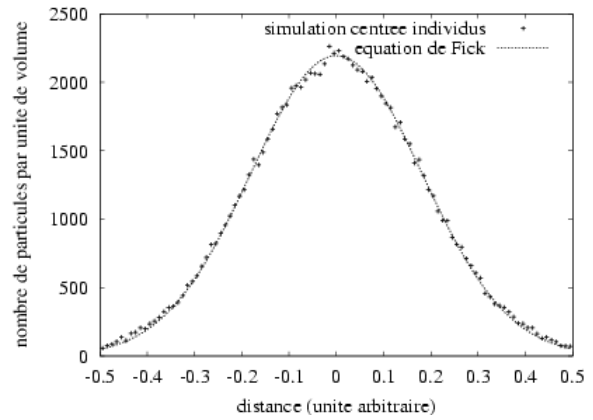


Figure 2. La simulation centrée individus pour 10^5 particules (croix) et l'équation de Fick (ligne pointillée) à $t=30s$. La simulation et la courbe correspondent presque parfaitement. On considère les deux modèles comme équivalents par leur trace d'exécution.

La figure 2 montre les résultats d'une simulation de la diffusion réalisée avec 10^5 particules pour une durée simulée de 30 secondes dans un cube de 1cm^3 . La vitesse des particules est déterminée par tirage aléatoire sur une loi uniforme entre 0 et $v_{\text{max}}=0.1\text{cm.s}^{-1}$. L'équation 5 donne le coefficient de diffusion $D = 5.10^{-4} \text{ s}^2$. La simulation individus-centrée s'ajuste très bien avec l'équation de Fick. Un test du χ^2 indique que les deux distributions ne sont pas significativement différentes : $p(\chi^2 = 24,76) > 0,5$ avec 30 degrés de liberté. Ce résultat n'est pas surprenant puisque nous ne faisons que reproduire l'expérience réalisée en laboratoire par le physicien J.B. Perrin et ses élèves pour mesurer le coefficient de diffusion (Collectif, 1997).

La première utilisation que nous faisons de la multi-modélisation dans ce cadre permet donc la détermination de valeurs particulières de paramètres ou de fonctions de ces paramètres. Le modèle individus-centré est alors considéré comme un laboratoire virtuel dans lequel nous conduisons des expériences sur le système pour déterminer des paramètres du modèle agrégé, ici D , que nous pouvons par la suite injecter à l'équation différentielle (cf. équation 1) pour une résolution analytique.

4. COUPLAGE DES DEUX MODÈLES

Dans le paragraphe précédent et dans des travaux antérieurs appliqués à l'écologie marine (Duboz et al., 2001), nous avons montré qu'il était possible de paramétrer une équation analytique à l'aide d'un modèle individus-centré. Nous allons montrer ici qu'il est

également possible de paramétrer un schéma numérique en cours de résolution.

Dans la plupart des cas, nous ne disposons pas d'une expression de D en fonction de v_{\max} qui pourrait être directement utilisée dans l'équation 2. Nous simulons donc le phénomène à l'aide du modèle individus-centré pour déterminer à chaque itération et pour chaque tranche du cube la valeur des D_i que nous réintroduisons dans l'équation du modèle numérique pour ensuite déterminer par résolution du modèle numérique les valeurs des différentes concentrations dans chacune des tranches d'espace (cf. Figure 3). Pour identifier chaque D_i , nous utilisons la méthode présentée au paragraphe précédent, en utilisant un modèle individus-centré sur 31 individus seulement pour limiter le temps de calcul. Nous avons vérifié que les résultats sont équivalents à ceux obtenus en simulant 10^5 individus.

Même si les conditions de simulation restent identiques, la grande stochasticité du modèle individus-centré conduit à des valeurs de D_i variables dans un petit intervalle. Pour le calcul d'un tel schéma numérique, il est donc important de connaître le domaine de variation de D_i afin de respecter les conditions de stabilité.

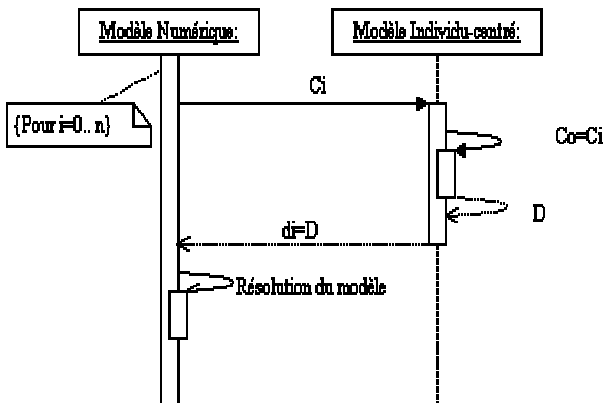


Figure 3. Diagramme de séquences UML pour l'illustration du couplage bidirectionnel entre modèle numérique et modèle individus-centré, pour chaque tranche d'espace i , le modèle numérique va lancer un modèle individus-centré initialisé avec la concentration correspondante pour déterminer le coefficient de diffusion dans la tranche.

La figure 3 illustre le couplage des deux modèles. Un tel couplage permet de simuler les effets spatiaux de la variation du coefficient de diffusion sur la dynamique du système. Nous allons illustrer cette méthode en considérant que la concentration des particules a un effet sur leur vitesse, par exemple en considérant une réaction de fuite des particules les unes par rapport aux autres. Cette réaction à la concentration peut être modélisée par une équation de type Monod qui fait augmenter la vitesse en fonction de la concentration jusqu'à un maximum fixé (équation 5).

$$v = v_{\max} - v_{\max} e^{-bc} \quad (5)$$

où v est la vitesse des particules, v_{\max} la vitesse maximum, c la concentration et b une constante.

Le modèle est le même que précédemment avec, pour l'expression de la vitesse des particules, $v_{\max} = 0.5 \text{ cm.s}^{-1}$ et $b = 10^{-3}$. Nous obtenons les résultats illustrés par la figure 4. La différence avec le cas précédent est que le modèle individus-centré a besoin de la valeur de la concentration locale à chaque instant qui est donnée par le modèle agrégé.

Avec une telle hypothèse de modélisation, la figure 4 montre que les particules n'atteignent pas le bord du cube pour une même durée de simulation et que la forme en cloche de l'équation de Fick n'est pas conservée pour les petites concentrations. Il n'est pas question ici de faire une étude précise de ce phénomène mais d'illustrer le couplage comme une méthode d'investigation possible.

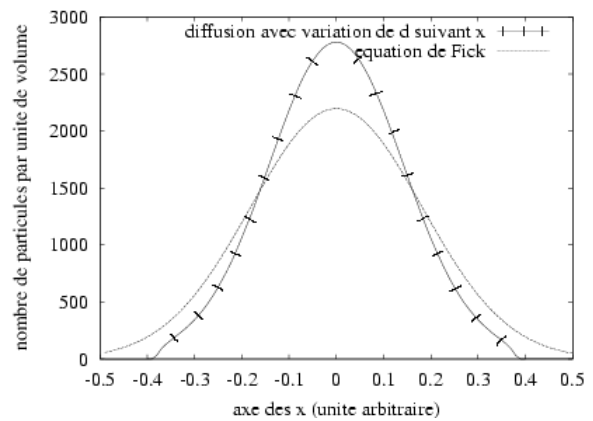


Figure 4. simulation du phénomène de diffusion pendant 30s en considérant que la vitesse des particules diminue avec la concentration (ligne pleine) par comparaison à l'équation de Fick ou la vitesse est supposée constante (voir le texte pour les détails).

5. DISCUSSION

Dans les deux exemples donnés précédemment (paramétrisation et couplage), nous avons simulé une même durée dans un même espace pour une même concentration. Cependant, les mêmes méthodes peuvent être appliquées à des échelles différentes. Dans notre exemple jouet, le changement d'échelle ne modifie pas a priori les dynamiques donc la méthode ne présente pas d'intérêt dans ce cas précis. Nous décrivons cependant comment un changement d'échelle pourrait avoir lieu, à des fins uniquement illustratives.

Considérons le modèle individus-centré simulant la diffusion d'un petit nombre de particules dans un petit volume et, en utilisant un couplage bidirectionnel et les mêmes hypothèses que précédemment, on fournit au schéma d'intégration la valeur du coefficient de diffusion à chaque pas de temps, ce dernier étant supérieur au

temps simulé par le modèle individus-centré, et pour chaque pas d'espace, lui aussi plus grand que le cube simulé par le modèle individus-centré. A l'aide d'un opérateur simple qui fait changer d'unité le coefficient de diffusion pour le donner au schéma d'intégration, nous simulons un transfert d'échelle.

Ce type de couplage peut être particulièrement intéressant pour la simulation des migrations d'animaux par exemple, à l'échelle régionale modélisées à l'aide d'équation différentielle, sachant que le comportement individuel et les réactions à l'environnement influencent fortement la dynamique de l'ensemble du troupeau. Le modèle individus-centré de « boules d'eau » proposé par Servat (Servat 2000), pourrait faire l'objet d'un couplage avec différents modèles agrégés (différent selon la quantité d'eau) utilement inspiré de cette approche. On peut supposer que ce soit également le cas de nombreuses autres applications.

Depuis les méthodes de Monte-Carlo, il est connu que la simulation ayant recours aux nombres pseudo-aléatoires peut par exemple servir au calcul d'intégrales présentant des limites d'intégration qui ne sont pas aisées à exprimer analytiquement (Press et al., 1994). La simulation vient ici « au secours » des mathématiques, lorsque celles-ci atteignent leurs limites analytiques et même numériques. De même les méthodes d'intégration numérique, bien qu'issues de théorèmes et axiomes démontrés formellement, sont issues du besoin de résoudre numériquement, c'est-à-dire en les simulant, des systèmes d'équations différentielles le plus souvent insolubles analytiquement. Notre approche se situe dans la même lignée, en ajoutant l'idée d'expérimentation virtuelle sur le modèle individus-centré. De plus, la notion de laboratoire virtuel permet d'étendre l'utilisation des modèles individus-centrés à la détermination de variables ou même de fonctions d'un modèle agrégé, comme c'est le cas pour des expériences réelles classiques.

Nous pensons que cette approche permet dans certains cas d'utiliser au mieux les avantages des deux types de modélisation, et d'obtenir une réelle plus-value par rapport à l'utilisation d'un seul des modèles. En effet, le modèle agrégé permet d'obtenir une représentation compacte et facilement utilisable de l'ensemble de la dynamique du modèle. Le modèle individus-centré permet d'assouplir l'utilisation du modèle agrégé en identifiant localement ses paramètres.

On peut par ailleurs se poser la question de l'efficacité du couplage en ligne. Il est clair que dans certains cas il peut être plus avantageux d'établir a priori un protocole de tests dans le laboratoire virtuel et en tirer l'ensemble des valeurs intéressantes pour le modèle agrégé. C'est le cas dans notre exemple si l'ensemble des valeurs de v_{max} atteignables sont connues a priori. En effet, il suffit alors

d'établir des expérimentations pour différentes valeurs de v_{max} dans cet intervalle, afin d'obtenir par interpolation la fonction qui donne D en fonction de v_{max} . Cette fonction est ensuite directement utilisée dans le modèle agrégé. Il s'agit là d'une méthode généralisable à de nombreux autres modèles. En effet beaucoup des modèles phénoménologiques de la physique font appel à des coefficients constants du même type que le coefficient de diffusion dans notre exemple. En élaborant un modèle centré sur les entités équivalent au modèle phénoménologique, notre méthode est alors applicable.

REFERENCES

- Collectif, 1997. *L'ordre du chaos*. Bibliothèque Pour La Science, Pour La Science eds.
- Dahl O. J. et Nygaard K., 1966. *SIMULA – An Algorithmic based simulation language*, communication of the ACM, 9(9), 671-678.
- Dautray R. et Lions J.L., 1984. *Analyse mathématique et calcul numérique pour les sciences et les techniques*. Evolution : numérique, transport, vol 9, INSTN CEA.
- Duboz R., Ramat E. et Preux P., 2001. *Towards a coupling of continuous and discrete formalism in ecological modelling: influences of the choice of algorithms on results*, in proceedings of the ESS'01 conference, Marseille (France), p.481-487.
- Ferber, J., 1995. *Les systèmes Multi-agents, vers une intelligence collective*. Inter-editions.
- Fick, A., 1855. „Über diffusion“, Annalen der Physik und Chemie, 94, p.59-86.
- Fishwick P., 1995. *Simulation Model Design and Execution. Building Digital Worlds*. Prentice Hall.
- Forrest S., 1990. *Emergent computation : self organizing, collective, and cooperative phenomena in natural and artificial networks*, Introduction to the proceedings of the ninth annual CNLS conference, Physica D 42, p1-11.
- Grimm, V., 1999. *Ten years of individual-based modeling in ecology: what we have learned and what could we learn in the future ?*, Ecological Modelling, 115, p.129-148.
- Legay, J.-M., 1997. *L'expérience et le modèle, Un discours sur la méthode*. INRA éditions.
- Lindenberg, S., 1992. *The method of decreasing abstraction*. In Rational choice theory: advocacy and critique. T.Fararo (ed.), vol.7, Sage publications, p.3-20.
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1994. *Numerical Recipes in C, (FORTRAN, PASCAL)*, The art of Scientific Computing, 2nd Ed, Cambridge University Press.
- Servat, D., 2000. *Modélisation de dynamique de flux par agents. Application aux processus de ruissellement, infiltration et érosion*. PhD Thesis Paris VI, 162p.
- Wooldridge, M. and Jennings, N., 1995. *Intelligent Agents: Theory and Practice*, The Knowledge Engineering Review, 10(2), pp. 115-152