Individual-based model to enrich an aggregate model

Raphaël Duboz* ✉, Frédéric Amblard**, Eric Ramat*, Guillaume Deffuant**, Philippe Preux*

*Laboratoire d'Informatique du Littoral (LIL)	**Laboratoire d'Ingénierie pour les Systèmes
UPRES-JE 2335	Complexes (LISC)
Maison de la Recherche Blaise Pascal	Cemagref
50, rue Ferdinand Buisson - BP 719	24, avenue des Landais BP 50085
62228 Calais Cedex	63172 Aubière Cedex
France	France

✉ Email: duboz@lil.univ-littoral.fr
Phone: +33 321 465 679

Abstract: We propose to use individual-based models as virtual laboratories to identify parameters of aggregate models. We illustrate this approach with a simple example: the diffusion of particles animated by a Brownian movement. We thus have an aggregated numerical model, as well as an individual-based model of this process. We then use the individual-based model as a virtual laboratory. In particular, we consider the coupling between both models in which the aggregate model asks for computations of parameters to the individual-based model. Finally, we evoke the case of scale shifts in the models and the way it can be handled by this type of approach.

Keywords: coupling, multi-formalism, differential equations, individual-based models.

1   Introduction

To describe a social, a physical, or a biological system, the differential equations are the fundamental tool of the scientists. This approach tends to describe a system made of many elements by aggregate variables (densities of populations for instance) and the dynamics of these variables, given by their derivatives. This provides a very compact and general representation of the system's evolution. However, these models use aggregate parameters which can be delicate to identify.

On the contrary, individual-based models (IBM) explicitly represent all the elements of the system, as well as their interactions. They have been increasingly used in particular in ecology, for several years (Grimm, 1999) and also to model social systems (Epstein and Axtell, 1996). Of course, differential equations can be used within the context of IBM to express the global behavior of the system. Nevertheless, IBM find a "natural" expression in recent work in computer sciences, like the object-oriented programming and multi-agents systems (Ferber, 1995). Then, the representations of individual dynamics can sometimes fruitfully deviate from a pure differential formalism. The individual-based models considered in this article are not necessarily multi-agents models if we adopt the strict definition of Wooldridge and Jennings (Wooldridge and Jennings, 1995). We thus use the term "individual-based model" since the model focuses on the entities composing the system, in opposition to an approach considering the aggregate variables representing the whole system.

In fact, these two types of models can be seen like representations of the same system at two different levels of abstraction. In this paper, we try to show the interest to develop both levels of abstraction together. Our principal argument is that the two levels can be coupled, each one being

enriched and getting advantage from the other one. Coupling a traditional numerical model with an individual-based model was already performed in several examples of practical applications. For example, Coquillard and Hill, (1997), use an individual-based model to account for the front of propagation of a forest fire, which is coupled to a traditional model. However, in our proposal, the individual-based and the aggregate models both relate to the same process, while this is not totally the case in the preceding examples (because the individual-based model is used only to model a part of the problem). Our view is closer to Lindenberg's idea of decreasing abstraction (Lindenberg, 1992), who recommends the development of increasingly precise models, which comprehension is nourished of that of the higher abstraction models, in different formalisms. The concept of multi-modeling (Fishwick, 1995) is also related with our approach. Multi-modeling aims at describing a process, or a system, using several models on different levels of abstraction or refinement.

Regarding these different works, the main originality of our proposal concerns the use of a particular type of cooperation between both levels of abstraction: the use of the individual-based model as a virtual laboratory, allowing us to perform experiments and measurements, in order to identify parameters of the aggregate model.

To illustrate this idea concretely, we consider a simple example, the diffusion in a fluid of particles animated by a Brownian movement. The aggregate model is reduced to the equation of evolution of the particle concentration in space. The individual-based model represents the whole set of the particles, and their movements. We consider several variants of the diffusion process, and corresponding dynamical coupling between the individual-based and aggregate models. In some cases, we show that these operations can be related to scale shifts. Finally, we discuss the interest of such couplings compared to an individual-based model alone, or the interest to carry out such an on-line coupling.

2 Aggregate and individual-based models of particle diffusion

The diffusion of particles is a particularly simple example of a process for which microscopic physical description is mathematically connected to the macroscopic description of the system by the use of statistical mechanics. It is thus easy to work out an aggregate model and an individual-based model of this phenomenon, and to establish a link between them. This process thus appears at the same time easy to expose, and rich enough to illustrate our point.

2.1 The aggregate model

The study of the diffusion phenomenon by the kinetic theory established the equation of Fick, aggregate model of Brownian diffusion (Fick, 1855). One is interested here only in the component along the x-axis (x):

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

Where c is the concentration in particles, D the diffusion coefficient and x the axis on which takes place the diffusion. c is a function of x and t.

Two methods exist to solve a differential equation. The first consists in integrating the equation analytically; the second uses a numerical resolution using a suitable schema of integration.

When the initial concentration and the limits of space are known, it is analytically possible to calculate the solution of the Fick's equation:

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

Where c_0 is the initial concentration.

This solution is thus a bell-shaped curve (Gaussian) representing the concentration of the particles at a particular time and space coordinates (cf. curve in dotted lines on figure 2).

However, if it is supposed that the diffusion coefficient is variable in space, the analytical solution is not always analytically computable. It is then necessary to use a numerical resolution of the equation with variable coefficient and constant step (equations 3) (Dautray and Lions, 1984), by discretizing space and time.

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

Where C_i^t is the concentration in particles at the space step i and at the time step t. One supposes that C_i^0 is given, Δx is the space step, Δt is the time step, and D_i is the diffusion coefficient at the space step i

2.2 The individual-based model

The individual-based model simulating the process of diffusion is particularly simple initially, but we will see thereafter that it can get more complex to address more detailed settings. We suppose that the particles evolve/move in a cube in three dimensions (x, y, z). We suppose that the particles are initially distributed at random on the plane defined by $x=0$, separating the cube in two halves of equal volume. The simulation of Brownian diffusion merely consists in giving a direction and a random speed (between 0 and v_{max}) to each particle at each iteration. The boundary condition is the impossibility for any particle to leave the cube; one thus draws at random a new direction and a new speed for any particle in the cube. This model is mainly stochastic. It is thus necessary to simulate a huge number of particles so that the observation of the model's behavior is not a particular event but the reflection of an average behavior of the whole set of the particles. At every moment, this model gives the position of every particle. It thus makes it possible to calculate the concentration in particles, or other indicators, in each part of space.

3 Virtual laboratory for the determination of parameters

The aggregate model can give us an estimate of the diffusion parameter D of the aggregate model, corresponding to a maximum speed of the particles.

For that purpose, one considers the individual-based model as a virtual laboratory, on which one performs experiments and measurements (Legay, 1997; Grimm, 1999). Thus, on the basis of an initially random distribution of the particles in the plan $x=0$, one measures at every time step t, the average standard deviation σ of the set of the particles, i.e. the square root of the sum of the square of distance between each particle and the origin of the x-axis divided by the number of particles.

One can deduce the diffusion coefficient D from this figure in an experimental way, using the relation:

$$\sigma = \sqrt{2Dt} \ \textit{ then } D = \frac{1}{2t}\sigma^2 \ \textit{(equation 4)}$$

Where σ stands for the average standard deviation, D is the coefficient of diffusion and t the time. Figure 1 shows the evolution of the quantity $\frac{1}{2}\sigma^2$ according to time. By computing the slope of this line, we obtain an estimation of the diffusion coefficient D.
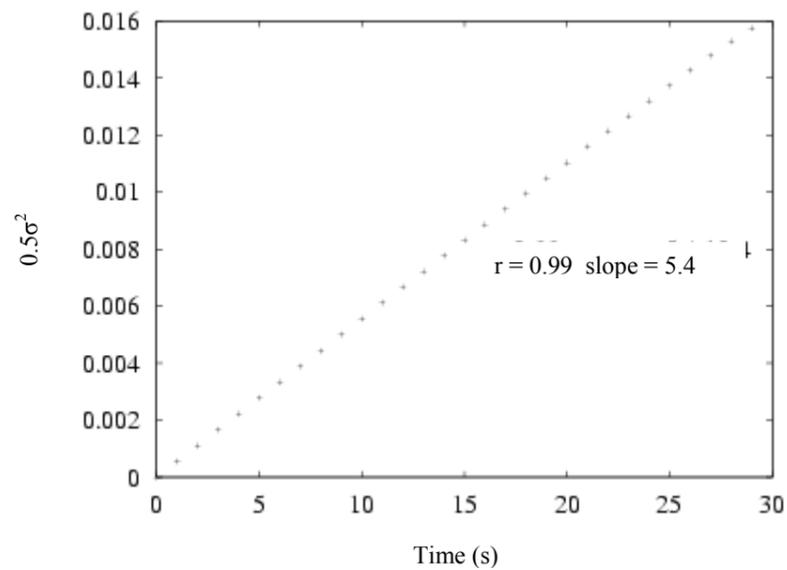

Figure 1: $\frac{1}{2}\sigma^2$ according to time. The slope of this line provides an estimate of the diffusion coefficient D.

Then, one can check that the microscopic model of the particles has the same behavior as Fick's macroscopic equation. Figure 2 makes it possible to compare the two models.
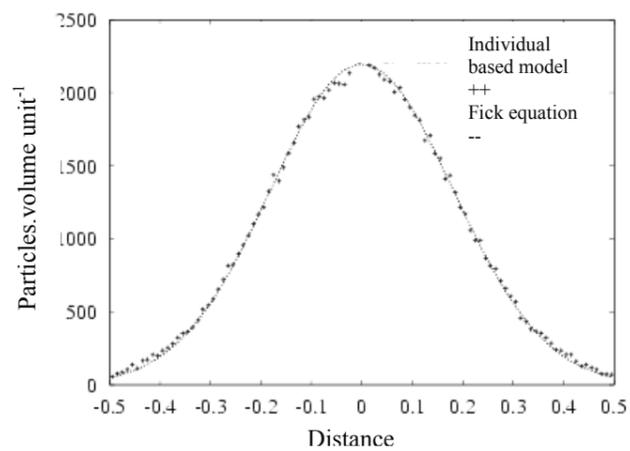

Figure 2: Individual-based simulation of 10^5 particles (cross) and Fick's equation (dotted line) with t=30s. The simulation and the Fick's curve match almost perfectly. One considers the two models as equivalent regarding their execution trace.

Figure 2 shows the results of a simulation of diffusion carried out with 10^5 particles during 30 seconds (simulated time duration) in a cube of $1cm^3$. The speed is determined at random from a uniform distribution between 0 and $v_{max}=0.1cm.s^{-1}$. Equation 4 gives us the diffusion coefficient $D = 5.10^{-4} s^{-2}$. Individual-based simulation fits very well Fick's equation. A χ^2 test indicates that the two distributions are not significantly different: $p(\chi^2 = 24,76) > 0,5$ with 30 degrees of freedom. This result is not surprising since we merely reproduce the experiment carried out in laboratory by physicist J.B. Perrin and his students to measure the diffusion coefficient (Perrin, 1997). The first use that we make of multi-modeling within this framework thus allows the determination of particular values of parameters or functions of these parameters. The individual-based model is then regarded as a virtual laboratory in which we perform experiments on the system to determine some parameters of the aggregate model, here the coefficient of diffusion D. Then, we can integrate its estimate into the analytical model (cf. equation 1) for an analytical resolution.

4 Coupling of both models

In the preceding paragraph as well as in our previous work applied to marine ecology (Duboz et al., 2001), we showed that it is possible to parameterize an analytical equation using an individual-based model. Now, we show that it is also possible to parameterize a numerical integration during its resolution. In most cases, we do not have an expression of D according to v_{max} which could be directly used in equation 2. We thus simulate the process using the individual-based model to determine for each iteration and for each part of the cube, the value of D_i. Then, we reintroduce this value into the equation under resolution in the numerical model. This provides the values of the various concentrations in each section of space (cf Figure 3). To identify each D_i, we use the method presented in the preceding paragraph, by using an individual-based model on 31 individuals only to limit the calculation time. We checked that the results are equivalent to those obtained by simulating 10^5 individuals.

Even if the conditions of simulation remain identical, the great stochasticity of the individual-based model leads to variable values of D_i in a small interval. For the calculation of such a numerical scheme, it is thus significant to know the field of variation of D_i in order to observe the conditions of stability.
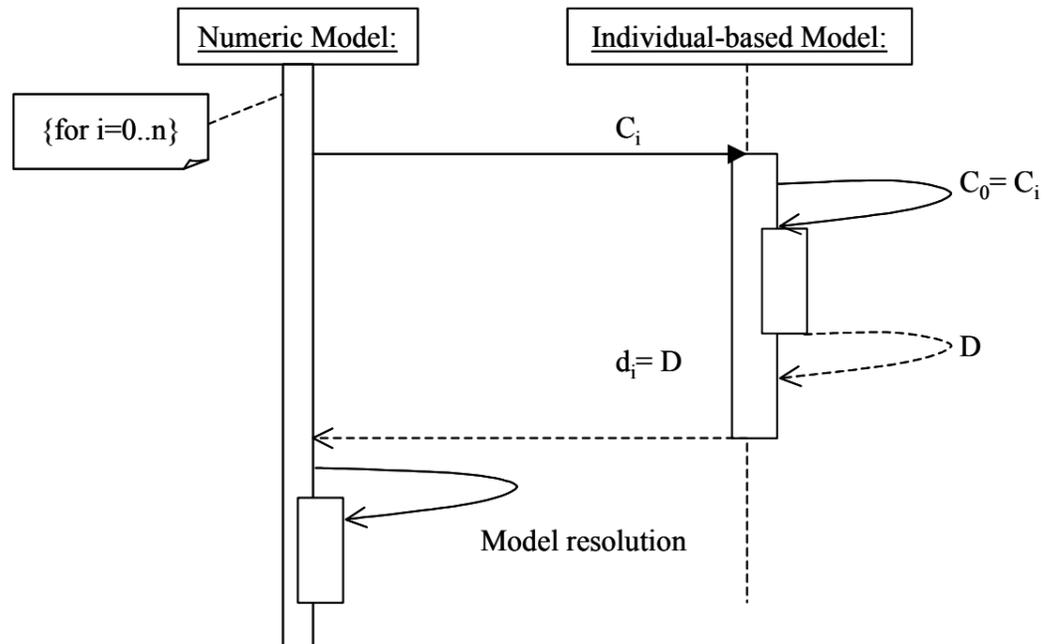
Figure 3: UML Sequence diagram to illustrate the bi-directional coupling between the numerical model and the individual-based model, for each section of space i, the numerical model issues an individual-based model initialized with the corresponding concentration to determine the coefficient of diffusion in the section.

Figure 3 illustrates the coupling of the two models. Such a coupling makes it possible to simulate the space effects of the variation of the diffusion coefficient on the system dynamics. We illustrate this method by considering that the concentration of the particles has an effect on their speed, for example by considering an escape dynamics between particles. This reaction to the concentration can be modeled by a Monod type equation which leads to an increase of speed according to the concentration until a fixed maximum is reached (equation 5).

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

Where v is the speed of the particles, v_{max} their maximum speed and c is the concentration.

The model is the same one as previously seen with $v_{max} = 0.5 \; cm.s^{-1}$ and $b = 10^{-3}$. We obtain the results displayed on figure 4. The difference with the previous case is that the individual-based model requires at every moment the value of the local concentration, which is given by the aggregate model. With such an assumption of modeling, figure 4 shows that the bell form of the equation of Fick is not preserved for the small concentrations. Our intention here is only to illustrate the coupling as a method of possible investigation, and not to elaborate more on this difference.
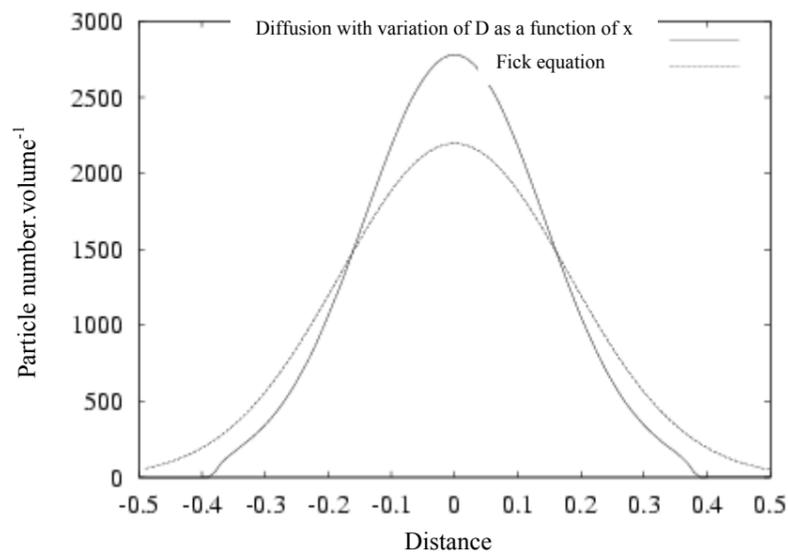
Figure 4: Simulation of the diffusion phenomenon during 30 seconds by considering that the speed of the particles decreases with the concentration (solid line) by comparison with the Fick equation where speed is assumed to be constant(dotted line).

5 Discussion

The method requires that an aggregate and an individual based model of the same phenomenon are available. Moreover, a clear link relating some results of the IBM to parameter values of the aggregate model is necessary. This leads to distinguish two main cases for using this method: on the one hand, starting from an aggregate model and deriving an IBM from it; on the other hand, starting from an individual-based model and deriving an aggregate model from it.

The first case could be particularly relevant for the simulation of animal migrations. For instance, starting from differential equations modeling the regional scale, one could derive an individual-based approach locally knowing that the individual behavior and the reactions to the environment strongly influence the dynamics of the whole herd. The individual-based model of "water balls" proposed by Servat (Servat 2000), could also be the subject of a coupling with various aggregate models (different models according to the water quantity), and with the aggregate model derived from the IBM or vice-versa.

In fact, the illustrative example taken here could have hidden some difficulties of this method, because in the case of particles diffusion, the aggregative process that enables to link both models is clearly identified. Such a link between the models is a prerequisite for the method, and it unfortunately, it not available in general. The first task is therefore to build either the aggregate or the individual based model, with the adequate link. This task can be difficult. Consider, for instance, a problem of modeling a population migration at a large scale, as in Bogota (Vanbergue et al., 2000). The individual-based model could be the only way to apprehend the variety of individual behaviors leading to intra-urban migration, but it it would lead to intractable simulation times with several millions of individuals. Then the recourse to differential equations could be one of the mean to avoid this pitfall. The point is, now that we have identify one of the possible solution, how to build the link between a light (considering the size of the population modeled) individual-based model focusing on the individual motivations and behaviors to migrate and an aggregated model that take into account the whole population? Both aggregation variables and aggregation process chosen matter, and this can be a difficult problem.

Nevertheless, we think that this approach makes it possible in some cases to use the advantages of both types of modeling, and to obtain a real improvement compared to the use of only one of the models. Indeed, the aggregate model makes it possible to obtain a compact and easily usable representation of the whole system dynamics. The individual-based model makes it possible to simplify the use of the aggregate model by identifying its parameters locally.

6　References

Coquillard, P. and Hill, D.R.C., 1997. *Modélisation et Simulation des Ecosystèmes,* Masson.

Dautray R. and Lions J.L., 1984. *Analyse mathématique et calcul numérique pour les sciences et les techniques*. Evolution : numérique, transport, vol 9, *INSTN CEA*, 1984.

Duboz R., Ramat E. and Preux P., "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, 2001

Epstein, J. and Axtell, R., 1996. *Growing artificial societies, Social Science from the bottom-up,* MIT Press.

Ferber, J., 1995. *Les systèmes Multi-agents, vers une intelligence collective*. Inter-editions, 1995.

Fick, A., 1855. „Über diffusion", *Annalen der Physik und Chemie,* 94, p.59-86. 1855.

Fishwick P., 1995. *Simulation Model design and Execution. Building Digital Worlds*. Prentice Hall, 1995.

Forrest S., "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, 1990.

Grimm, V., 1999. "Ten years of individual-based modelling in ecology: what we have learned and what could we learn in the future?", *Ecological Modelling*, 115, p.129-148, 1999.

Legay, J.-M., 1997. *L'expérience et le modèle, Un discours sur la méthode*. INRA éditions, 1997.

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, 1992.

Perrin, J.-B., 1997. *L'ordre du chaos*. Bibliothèque Pour La Science, Pour La Science eds, 1997.

Servat, D. Modélisation de dynamique de flux par agents. Application aux processus de ruissellement, infiltration et érosion. *PhD Thesis* Paris VI, 162p, 2000.

Vanbergue, D., Treuil, J.-P. and Drogoul, A., 2000. "Modelling urban phenomena with cellular automata", in Ballot, G. and Weisbuch, G. (eds.) *Applications of Simulation to Social Science,* Hermès.

Wooldridge, M. and Jennings, N., 1995. "Intelligent Agents: Theory and Practice", *The Knowledge Engineering Review*, 10(2), pp. 115-152, 1995.