Validation and Fidelity in Numerical Simulations

I. Rosiello[1]*, S. Vaienti[3]†
(1) 6 BD Michelet, 13008 Marseille
(3) UMR-CPT 6207, Universités d'Aix-Marseille I,II
et du Sud Toulon-Var, Marseille, France
and Fédération de Recherche des Unités de Mathématique
de Marseille, Marseille, France

February 20, 2011

Abstract

We discuss the validity of numerical simulations by addressing the question of how the models implemented on a computer will give outcomes possibly different from what is expected, since such models will be perturbed in different unavoidable manners. The question of their stability, global and local in time is discussed.

1 Introduction

The object of this note is to discuss and to investigate the *fidelity* of numerical simulations. This concept is part of the more general issue of the *validity* of numerical simulations and it addresses the fundamental question of how the result of a simulation, or the simulation itself, are faithful to the object they are simulating. This question could be posed at different ontological levels. Let us explain what we mean with the famous example of the well celebrated Schelling model [15]. The opponents of this model wonder if it is possible and meaningful to describe the urban segregation with two dimensional cellular automata. The reduction of the *true* social behavior to the simple rules which govern the automaton is so drastic that the complexity and the variety of the former are inevitably lost and therefore any conclusion that we could get from such a model are either purely academic or ideological. The tenants of the model, who consider it as the *paradigmatic* example of numerical simulations applied to social phenomena, would answer that despite the crude approximations, it produces knowledge since:

- it attempts to reproduce the urban society as a *network* with the underlying assumption, or hope, that simple local rules (e.g. individual and well recognized actions), could produce global and emerging patterns:

*Electronic mail:immacolata.rosiello@gmail.com
†Electronic mail:vaienti@cpt.univ-mrs.fr

> Thus the model continues to be influential, although it has little or no empirical support, because it remains a fruitful source for theorising and for developing new models.[2]

- even if the same phenomenon could be explained otherwise [1], simulations with cellular automata establishes (explain?) this causality link between the individual interactions and the appearance of macroscopic behaviors;

> And a simulation model which generates a macro structure which resembles real-world macro structures from simulated micro structures which resemble micro structures observable in the real world might be accepted as a provisional explanation of real-world macro structures. [7]

- summarizing, we are in presence of an *exploratory simulation* in the sense of Conte and Gilbert:

> Our stress, instead, is on a new experimental methodology consisting of observing theoretical models performing on some testbed, such a new methodology could be defined as *exploratory simulation*. The exploratory aim synthesizes both the prescriptive and descriptive objectives : on the one hand, as with the testing of existing theories, the aim is to increase our knowledge ; but on the other, as happens with studies oriented to the optimization of real life processes, the aim is not to reproduce the social world, but to create new, although not necessarily *better*, or more desirable, systems. Here lies the difference from optimization research. Exploratory research based on social simulation can contribute typically in any of the following ways :
> (a) implicit but unknown effects can be identified. Computer simulations allow effects analytically derivable from the model but as yet unforeseen to be detected;
> (b) possible alternatives to a performance observed in nature can be found;
> (c) The functions of given social phenomena can be carefully observed;
> (d) *sociality*, that is, agent hood oriented to other agents, can be modeled explicitly. [6]

If we accept this approach to simulations as virtual laboratories where different scenarii[2] are proposed and tested on computers, we are therefore tempted to move our investigation to

[1] In the PhD thesis [13] there are examples of such simulations which admit a better and easy explanation: the model of gender discrimination in German school quoted in [7], the example of the fish market in Marseille discussed in [10], etc. At this regard P. Livet wrote

> Nous pourrons donc admettre comme valides des simulations dont les résultats simulés ne sont pas en correspondance directe avec les phénomènes sociaux, ou des programmations qui ne sont pas en correspondance directe avec les processus formels, ou encore des processus psychologiques et sociologiques de constitution du social qui ne sont pas en correspondance directe avec les relations de programmation-simulation. Il faudra seulement que ces correspondances puissent être rétablies par des révisions sur un autre axe de similarité que celui qui est en cause de maniére obvie.[10]

[2] We would like to stress the importance of the concept of *scenario* in the understanding of simulation. In the paper [1], Armatte and Dalmedico consider the simulation as a link (*médiateur*) between the numerical model and the narration which we could extract from its results. They wrote:

other issues, in particular we go from the *first level* where the model has been proposed (in our guiding example the cellular automata) to the *second level* where the model will be critically analyzed according to some standard scientific methodology and practices. This is very well depicted by B. Latané:

> It is often assumed that the true test of a scientific model, whether it is a verbal or mathematical theory or a simulation, is whether it can predict empirical outcomes. Yet, I have found relatively little occasion so far to compare the results of these simulations of dynamic social impact directly to data from the real world. Instead, in addition to making unexpected theoretical discoveries such as the predicted emergence of clustering and correlation and the ability of social systems to maintain stable diversity in the face of strong pressures to uniformity, I have found simulation especially useful for three other non-empirical purposes:
> 1. Determining whether the theory is robust with respect to variations in stochastic and other theoretically uninteresting variations (models that are not robust in this way are useless and probably wrong);
> 2. Identifying which elements of the theory are critical to the resulting dynamics (by narrowing the focus of the model, we can separate the theoretical wheat from the chaff, avoiding controversy about irrelevant details and achieving a maximally simple or parsimonious model). [3]

Let us consider in particular the first item in the previous citation, about the robustness of the theory: Latané rightly points out that a model which is not robust is useless. We see that a critical investigation of formal (mathematical or numerical) nature at the second level could invalidate the model not because of its assumptions (first level), but because these assumptions will not produce a realistic or a controlled systems. Suppose that a model always gives the same asymptotic behavior independently of the initial conditions: this could of course happen, for example in presence of global attractive fixed points, but this is not very interesting (realistic), and it is surely false, if we deal with macroscopic social patterns. Analogously, suppose we have a model which is extremely sensitive to small changes of its parameters: the results would be unreliable (out of control) most of the time and the simulation will be again useless for pratical purposes. This second example, sensitivity on the initial conditions, could not be appreciated by those who consider that in this way we are neglecting all the simulations of chaotic systems. This is actually the starting point of our contribution. We will in fact try to understand how the outcomes of a model change if the model is perturbed and this perturbation could be of two types: a change of the parameters or of the initial conditions of the models, or the errors introduced by the numerical algorithms. Our aim is to understand what kind of characteristics the model should have in such a way its *ideal* outcomes (in absence of errors) are faithful to what the computer will produce at the end of its runs. A general answer to this question does not exist; it is however important to keep it in mind when we evaluate a simulation. There is a sort of undecidability related to this numerical processing of the model which let Zellini says:

> Le scénario cristallise pourtant une articulation de plus en plus étroite entre le temps de l'exploration et celui de l'action, entre la dynamique de la science et celle du politique.[ibid]

On another side, the failure of simulation to take into account micro-behaviors, could be absorbed in the production of a scenario which carries global dynamics [13].

I risultati intermedi di un calcolo automatico sono normalmente nascosti nella memoria del calcolatore e sono quindi sconosciuti alla mente del programmatore. Il calcolo tende per questo a diventare un *processo* in gran parte inaccessibile al controllo umano. Di qui la necessità di una delega, fondata su una valutazione a priori della natura del risultato, in particolare del possibile errore da cui tale risultato sar affetto. Il *rumore* causato da instabilità e mal-condizionamento deriva forse ancora più da questo aspetto di inaccessibilità dei risultati intermedi che dall'aumento del numero di operazioni da eseguire. [19]

2 Shadowing

As we said above "*local sensitivity to small errors is the hallmark of a chaotic systems*", [14]. As a consequence one could wonder whether the numerical solutions obtained with a computer are valid. We emphasize here that chaoticity is one of the features of complexity and that, contrarily to the latter, chaotic phenomena could now be studied with a well based and accepted mathematical background. For these reasons, we believe that the (partial) answers to the question addressed above could give a legitimate direction and hints to deal with the much more difficult problem of the simulations of (perturbed) complex systems.
As we said at the end of the previous section, there are intrinsic computational errors, the *floating-point calculations*, or *round-off errors*, which affect the solutions of differential equations or the computation of discrete maps, by simply replacing a real number with a version of it with a finite number of digits. This produces a small error at each computational step. Let us suppose that the dynamical systems we are simulating is hyperbolic. To have an idea of such systems, which carry what is usually considered as *chaotic motion*, take the circle of radius one and a point on it which is initially located at the angle θ_0; consider the orbit of such a point when we double it, namely after n step the point will be located at $2^n\theta_0$-mod 1 (where *mod* 1 means that the point $2^n\theta_0$ is unfolded again on the circle). Now, by choosing with probability one the initial point θ_0, its orbit will spread up uniformly on the circle with an exponential speed, and if we take another initial condition θ_0' (with probability 1) and close to the previous one, the orbits of the two will diverge exponentially fast (sensitivity on initial conditions). More complicated examples of such systems are the strange attractors which could be detected and observed even in physical situations (Lorenz attractor, Hénon attractor, etc.). For these hyperbolic systems the perturbation induced on the original system by round-off errors could be considered similar to the perturbation which affects the original system when we slightly change, for instance, its structural parameters. A deep mathematical result due to R. Bowen [5] says that when the perturbation is not too strong, the perturbed orbit could be recovered with a true one but with a different initial condition which allows to conclude that

> locally sensitive trajectories are often *globally insensitive*, in that there exist true trajectories with adjusted initial conditions, called shadowing trajectories, very close to long computer-generated pseudo trajectories.[14]

This result is not useful if it is applied to a single pseudo trajectory (that generated by a computer), since the theorem does not tell us where is the initial condition whose orbit will closely shadow the true one. The strength of the theorem is in the assertion that the asymptotic global phase portrait of the system which we will see on the screen of the computer will be

undistinguishable of the true one, which will be, however, intrinsically inaccessible at the level of individual trajectories.
Another interesting question addressed by this theorem is what happens when its assumptions are not anymore verified, in particular if the system is not hyperbolic. This could be achieved, for example, by constructing a systems with a geometric structure and admitting a few directions with an hyperbolic behavior and other directions which are not hyperbolic at all. In this case the local sensitity will be experienced only along the former directions, the motions being much more regular along the latter: we could define *fluctuating* the resulting global behavior. The conclusion is (at least on the model tested by authors):

> In the presence of this fluctuating [behavior], global sensitivity may led to trajectory mismatch, in particular when long times are considered. The result is that no trajectory of the theoretical model matches, even approximately, the true system outcome over long time spans. The fundamental conclusion...is that to obtain a long trajectory which is even approximately correct is for some systems virtually impossible.[14]

If we think that most of the complex systems which are the target of numerical simulations are fluctuating in the previous sense, this fluctuation being often a signature of complexity, we realize how much we should be careful when we argue about the reliability of a numerical simulation. In particular the three kind of validity formulated in the textbook [4] "to make impossible to distinguish the model and the system *in the experimental frame of interest*", are seriously questioned:

> The most basic concept, replicative validity, is affirmed if, for all the experiments possible within the experimental frame, the behavior of the model and system agree within acceptable tolerance. Stronger forms of validity are predictive validity and structural validity. In predictive validity we require not only replicative validity, but also the ability to predict as yet unseen system behavior. To do this the model needs to be set in a state corresponding to that of the system. Finally, structural validity means that the model not only is capable of replicating the data observed from the system, but also mimics in step-by-step, component-by-component fashion the way in which the system does its transitions. The term accuracy is often used in place of validity. Another term, fidelity, is often used for a combination of both validity and detail. [*ibid*, p. 30]

3 Fidelity

We do not know if the concept of fidelity that we are going to use in this section fits the meaning that [4] attributed to it; on the other hand the goal is the same: to give some criteria for declaring satisfactory a numerical simulation. The following technique has been proposed in a series of papers to understand the distribution of errors in randomly perturbed dynamical systems [12, 16, 18]. Contrarily to the shadowing theorem, the attention is now focused on the initial conditions and on the realizations which will produce a random pseudo trajectory close to the true one. Let us consider a given observable Φ defined on the phase space and that we will compute along the true orbit leaving the point x_0, and along a slightly perturbed orbit starting from the *same* point x_0; this second pseudo trajectory could be constructed by adding at the step $n+1$ a small adittive noise to the map at the step n. This noise will be chosen randomly in a small "ball" around 0 (the ball of radius 0 being the *zero noise*): we write $\bar{\epsilon}_n$ for the realization of the particular sequence $\{\epsilon_1, \cdots, \epsilon_n\}$, each member being the small error added at the corresponding step and with all of them being chosen independently in the small ball and with the same distribution. We then consider the difference $\Delta\Phi(x_0, \bar{\epsilon}_n)$ between the values of the observable computed along the true trajectory at time n and along the pseudo trajectory with realization $\bar{\epsilon}_n$. This difference is a random variable depending on the choice of the initial condition x_0 and of the realization $\bar{\epsilon}_n$. In several situations the random process $\Delta\Phi(x_0, \bar{\epsilon}_n)$ will converge in distribution to a random variable $\Delta\Phi_\infty$ which could be considered as the distribution of the error between the true and the perturbed orbit for very large times. The expectation of this asymptotic error will tell us how far is the average of Φ computed along the true orbit with respect to the average computed along the pseudo trajectory. This difference will be in general small, of the order of the size of the noise, but the variance will be large, of the same order of magnitude of the diameter of the phase space, indicating that the values of the observable along the true and the perturbed orbits are uncorrelated. But the interesting feature of this approach is elsewhere, in particular in the *rate of convergence* of the process $\Delta\Phi(x_0, \bar{\epsilon}_n)$ to its limiting value $\Delta\Phi_\infty$. This convergence is ruled out by the decay of a particular integral, which is called *fidelity* [9] and which is, in our context, the characteristic function of the process $\Delta\Phi(x_0, \bar{\epsilon}_n)$. In order to continue we need first to come back to our hyperbolic systems and compare them with another class of systems which we could call *regular* and which exhibit, roughly speaking, a motion which is a non-periodic translation. To give an idea, we still consider, as in the previous section, a point θ_0 on the unit circle, but this time it will evolve by simply adding, at step $n+1$, an irrational number α to the position occupied at time n, so that the location of the initial point after n steps will be $\theta_0 + n\alpha$-mod1. We now return to the decay of the fidelity. When the system is hyperbolic (chaotic), the fidelity shows a sharp transition at a time n^* which is of the order of the logarithm of the reciprocal of the size of the noise. For times n less than n^* the fidelity decays very slowly, which means that *the perturbed system can be considered as equivalent to the unperturbed one*; instead the fidelity will decrease super-exponentially fast for times bigger than the threshold value and with a decay rate which will became insensitive to noise (Fig 1, where we used the chaotic map $3^n\theta_0$-mod 1).

We stress again that this implies that for chaotic systems the true evolution of the system remains close to a generic perturbed version of it, provided the two start with the same initial condition, and moreover they stay close for a time interval proportional to the logarithm of the reciprocal of the size of the noise. We call it a *statistical local stability in time*, when compared with the global stability given by the shadowing theorem, we will return on these concepts in the next section. It is interesting to apply our result to the round-off errors generated in the numerical computations. Although the nature of the roundoff is deterministic (we know at each step how the machine operates on numbers), by increasing the ergodic properties of the system, the *pseudo random* character of roundoff errors becomes more pronounced and the decay law of fidelity becomes equivalent to that of systems perturbed with additive noise. In particular the threshold n^* will grow now like the logarithm of the reciprocal of the accuracy specified by the least significant bit used to represent a real number. We have here a quantitative estimation of the uncertainty associated with the intrinsic approximations given by digital representation

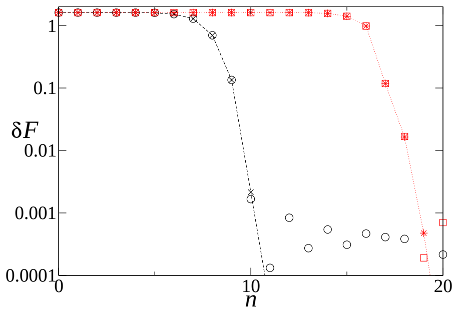

Figure 1: Comparison between analytical results and random orbits for the decay of fidelity for 3θ-mod 1 with $\varepsilon = 10^{-4}$ in black and $\varepsilon = 10^{-8}$ in gray (analytical result: line and stars; random: squares), cf. [12].

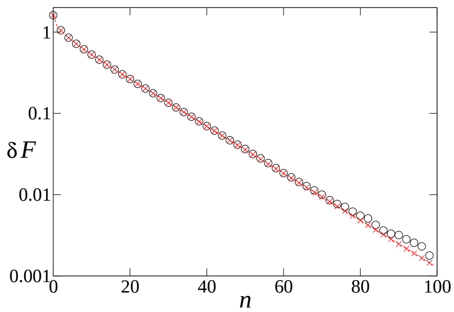

Figure 2: Comparison between analytical results and random orbits for the decay of fidelity in rotations and with $\varepsilon = 0.1$ (analytical result: gray dashed line and crosses; random: circles); cf. [12]

of real numbers.
We introduced above the other class of regular systems. In this case the decay of fidelity is not so straightforward. For some of them the fidelity does not decay at all, for others the decay is much slower and without a sharp transition. On the contrary, this transition is a gradual process; when these systems are simulated on a computer this transition gets longer with the number of bits used to represent real numbers. Regular systems which are affected by a small noise begin to loose memory and to become uncorrelated, but the time they take to do it is much longer than for desordered systems and moreover they do not exhibit precise bifurcation patterns which testify of a sudden change of regime: this seems to happen smoothly in time and in the space (Fig. 2).

4 Conclusions

Complex systems mix chaotic and regular motions with a large number of elements and on different scales. The stability of such systems become therefore a difficult task. Simulations are done to enlighten those systems because simulations offer the only reasonable way, in time, to explore a large number of elements, in space, and of the interactions among them. On the

other hand we have just seen how the simulations themselves suffer of the *complexity* of the object they want to study, since this complexity affects the deterministic equations and rules which transpose the model of simulation in the numerical algorithms. We showed in this note that in particular cases, it is better to assume a probabilistic point of view, which of course will not tell us how and where the system will be in a given future, but for how long we could consider what we compute numerically as faithful to the object that we would like to simulate. We evoked the *stability* at the beginning of this section; actually we explored three kinds of stability up to now and it is probably better to summarize them in a precise way. *Shadowing* gives us the structural stability of the system, as a matter of fact it is the keys of the proof. Structural stability means that, provived we slightly modify the parameters of the system, the equations of motion (the flows) of the perturbed system and of the original one are continuously conjugated, isomorphic in a wide sense [8]. Therefore, if the original system will asymptotically relaxes on a strange attractor, the same will do the perturbed one and the attractor for the latter will be topologically and geometrically *close* to the unperturbed one. In this picture only the *qualitative* global portrait of the system is preserved, we loose the control of each single trajectory.
Our approach using fidelity tries to recover, locally in time, a sort of *statistical stability*, in the sense that the original and the perturbed systems are compared on a probabilistic base by showing that the distribution functions of the observables defined on the systems (perturbed and not), remains close for times growing as the noise goes to zero.
Still in this approach we are not focusing on *individual orbits* and on their evolution: this will lead us to the third kind of stability. It is well known that for chaotic systems (in the precise mathematical sense that we introduced above), the *predictability*, namely the time interval on which one can typically forecast the systems:

> ...is limited up to a time which is related to the first Lyapunov exponent, and the time sequence generated from one of its chaotic trajectories cannot be compressed by an arbitrary factor, i.e. is algorithmically complex. On the contrary, a regular trajectory can be easily compressed (e.g. for a periodic trajectory it is sufficient to have a sequence for a period) so is is "‘simple" [17]

This is again a consequence of the sensitivity on initial conditions which unavoidably affects the accuracy of any future forecasting.
Global, statistical, individual stability. These are three kinds of constraints that each simulation should be aware of, especially when the simulation is built on a model with a mathematical structure. We focused on them in this note, but there are of course other validation criteria. As P. Livet said [11]

> Quand les simulations se bornent à imiter les phénomènes sans se soucier de la comptabilité des opérations théoriques avec les processus effectifs, alors les choses se dégradent. Les problèmes sont encore plus troublants quand les simulations sont les seuls outils de réperage dans les phénomènes, et qu'elles orientent les actions d'acteurs qui font partie des phénomènes.

And Livet also pointed put

> C'est le cas de l'économie et particulièrement de la finance. Les modèles qui sous-estiment la complexité des comportements collectifs ne voient pas arriver les crises,

et les modèles qui tiennent mieux compte de cette complexité ne peuvent pas prévoir quand il n'y aura pas crise.

Let us stress another danger which affects the relationship between models and simulations and which happens when simulations are conceived and produced as pure *technology*. This means, roughly speaking, that a *simulateur* is a black box in a computer which receives data to process and gives outputs and where the input devices change accordingly to the intensity and the varieties of the incoming data, but without modifying, eventually, the background on the underlying model. Let us give an example. Recently a colleague working on weather forecast told us that the theoretical models are still those of the sixties and that the only improving in climate predictions came from the much higher performances of the computers and of the numerical tools. He also added that deep conceptual questions deserved to be better investigated, in particular the contribution of the new massive computational techniques could have altered the setting and the conditions where the basic equations of motion were formulated!
Simulations are scientific practices with their own autonomy and protocols; but simulations are only one part of the understanding of the object, this one being the result of different constructions provided by models, theories, experiences, imagination.

References

[1] M. Armatte, A.D. Dalmedico, Modèles et modèlisation, 1950-2000 : Nouvelles pratiques, nouveaux enjeux, *Revue d'histoire des sciences*, **57**-2, Juillet-Décembre 2004, PUF

[2] P. Ahrweiler, N. Gilbert, Caffé Nero: the Evaluation of Social Simulation, *Journal of Artificial Societies and Social Simulation*, **8**, no. 4, (2005)

[3] B. Latané, Dynamic Social Impact, in *Modelling and Simulation in the Social Sciences from the Phylosophy of Science Point of view*, R. Hegselmann, U. Mueller, K.G. Troitzsch eds., Kluwer Academic Publisher, (1996)

[4] B. P. Zeigler, H. Praehofer, T. G. Kim, *Theory of Modeling and Simulation*, Second Edition, Academic Press, (1976)

[5] R. Bowen, Periodic points and measures for Axiom-A diffeomorphisms, *Trans. Am. Math. Soc.*, **154**, 377-397, (1971)

[6] R. Conte, N. Gilbert, Computer simulation for social theory, in N. Gilbert and R. Conte (Eds.), *Artificial societies: The computer simulation of social life*, pp. 1-18, London: UCL Press., (1995)

[7] K. G. Troitzsch, Validating simulation models, Proceedings *18th European Simulation Multiconference Graham Horton (c)*, SCS Europe, (2004)

[8] A. Katok, B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press,(1995)

[9] C. Liverani, Ph. Marie, S. Vaienti, Random Classical fidelity, *Journal of Statistical Physics* **128**, 4, 1079-1091, (2007)

[10] P. Livet, Essai d'pistmologie de la simulation multi-agents en sciences sociales, in *Modèlisation et simulation multi-agents, applications pour les Sciences de l'Homme et de la Société*, Frédéric Amblard et Denis Phan eds., Hermes-Lavoisier (2006)

[11] P. Livet, Rapport sur la thèse *Méthodes de simulation. Analyse de la complexité. Utilisation de modèles. Interdisciplinarité* de I. Rosiello (2009)

[12] Ph. Marie, G. Turchetti, S. Vaienti, F. Zanlungo, Error Distribution in randomly perturbed orbits, *Chaos*, **19**, 043118, (2009)

[13] I. Rosiello, *Méthodes de simulation. Analyse de la complexité. Utilisation de modèles. Interdisciplinarité*, Thèse de Doctorat, Université de Provence, Marseille, (2009)

[14] T. Sauer, C. Grebogi, J. Yorke, How long do numerical chaotic solutions remain valid?, *Phys. Rev. Lett.*, **79**, 59, (1997)

[15] T.S.Schelling, Models of Segregation, *American Economic Review*, **59-2**, (1969), 488-493

[16] G. Turchetti, S. Vaienti, F. Zanlungo, Relaxation to the asymptotic distribution of global errors due to rounf off, *Europhys. Lett.*, **89**, 40006, (2010)

[17] G. Boffetta, M. Cencini, M. Falcioni, A. Vulpiani, Predictability: a way to characterize complexity, *Physics Reports*, **356**, 367-474, (2002)

[18] G. Turchetti, S. Vaienti, F. Zanlungo, Asymptotic distribution of global errors in the numerical computations of dynamical systems, *Physica A*, **389** (2010) pp. 4994-5006

[19] P. Zellini, *Gnomon*, Adelphi, 2ème ed. ,(2007)