Integrative Biology: Modelling and Simulation of the Complexity of Natural Systems

Alain Pavé[1] and Claudine Schmidt-Lainé[2]

[1]Centre national de la recherche scientifique, UPS CNRS-Guyane,
Campus agronomique, BP 316, 97379 Kourou Cedex, French Guyana.
Mail : alain.pave@cnrs-dir.fr

[2]Cemagref, 3 bis quai Chaveau, CP 220 69336 Lyon cedex 9, France.
Mail : claudine.schmidt-laine@cemagref.fr

1. Complexity of natural systems

The concept of complexity has been largely discussed in the literature since the beginning of 1970's. The word "complexity" corresponds to 4246 references in the data base of Nature (from 1995 to 2003), and "complex" to 4404 (but this word includes other meanings such as complex of two molecules, which is far from our context). But a precise definition is not shared by all scientists, and can lead to misunderstandings: "If a concept is not well defined, it can be abused. This is particularly true of complexity." (Viscsek[1])

Here we have used this term within an "intuitive" context. Natural systems are complex in the original sense of the word (at least, one of the following criteria must apply):

- Elementary entities are numerous and diverse, particularly biological ones. These entities are interconnected by many weak or strong relationships (constituting networks or webs).
- The dynamics of these systems can exhibit complex trajectories, mainly resulting from the non-linearity of involved mechanisms (e.g. chaotic dynamics[2, 3]).
- Often, their spatial structures cannot be described by using figures of classical geometry. This is one example for the successful use of fractal geometry. The complexity of a fractal object can be related to the logical complexity of the algorithm which generates it[4].
- The laws which describe the behaviour of a complex system are qualitatively different from those that govern its units (Vicsek, *op. cit*)

Finally, following G. Nicolis[5], complexity is "evidence" for, but also a key concept of the actual epistemology. The advantage or disadvantage of complexity, particularly ecosystem complexity, is still debated in the literature. For example, using a theoretical approach, R. May deduced that complexity could destabilize ecological systems[6], counter to the intuition and understanding of most ecologists. Conversely, McCann *et al*[7] more recently demonstrated that weak interactions in a trophic web lead to stability (see also [8]). In other fields, for example technical systems, complexity is not necessarily synonymous with reliability (the multiplication of components can be a source of failure).

Complexity is often linked to diversity, it is often true in natural systems (because in most cases, entities in nature establish relationships spontaneously; without such connections, the system is simple and reducible to a list of independent entities and dynamics). It seems that diversity (and thus, complexity) in an ecological system increases its productivity (Hulot et al.[9]).

Our purpose is to present some “simple” examples of “complexity” and mainly to show how modelling approaches can help us in the analysis and representation of the complexity of systems[10], particularly natural systems, as well as in conceptual experimentation (cf. May, McCann *et al.,* cited above). We also discuss the interest of “simplifying complexity” by aggregating data, which occasionally hides the richness of structures and phenomena, but which makes it possible to exhibit major tendencies and use simple models to describe them.

2. Analysing and integrating complexity

Usually, the basic question in analysing complexity is to find an order in an apparent disorder. The paradigm is provided by the works of zoologists and botanists during the three last centuries. They show how taxonomy, a systematic approach, leads to classifications of animals and plants (a model of life organisation), which appeared before extremely complex, not immediately understandable.

The problem is often to find a good representation[11], an efficient model, which explains natural complexity. For example, the integrative view of living systems attempts to take into account their organization in hierarchical levels, an embedded organization of interacting entities from macromolecules, cells, organisms, populations… “from genes to ecosystems and biosphere”[12] resulting from 4 billions years of evolution. At each level there are :

- characteristic time and space parameters (cf. Figs. 1 et 2);

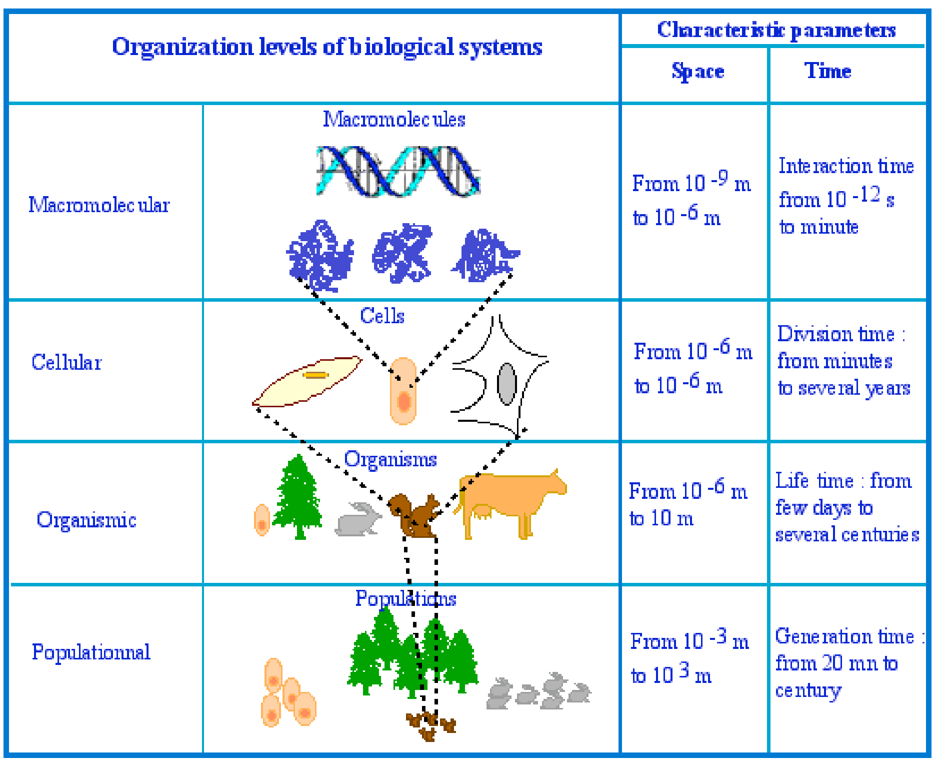


Figure 1. Organization levels of living systems: biological systems

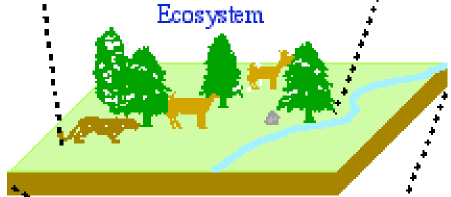
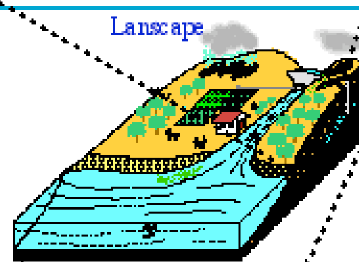
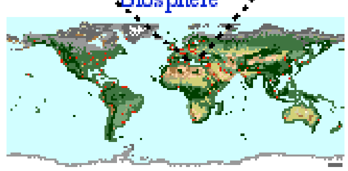
Ecological systems: levels of organization		Characteristic parameters: Space	Characteristic parameters: Time
Community	Community 	From meter to 10^3 m	Renewable time : years to siecle
Ecosystemic	Ecosystem 	From 10^2 m to 10^3 m	Regeneration time : from years to siecle
Lanscape	Lanscape 	From 10^3 m to 10^4 m	Fashionable time : from 10 years to centuries
Ecoregional	Ecoregion 	From 10^5 m to 10^6 m	Change time : centuries
⋮	⋮	⋮	⋮
Biospheric	Biosphere 	About 4 10^7 m	Evolution time : from millenaries to billions of years

Figure 2. Organization levels of living systems: ecological systems (cf. [i])

Note: Living systems are presented in two figures (from André *et al.*[14]), because there is effectively a qualitative break between biological and ecological systems (ecological systems include physical components of the milieu, and more and more take into account anthropogenic actions). This presentation appears more accurate than a previous one (Pavé[15]). In this last reference, a chapter is devoted to complex systems; the most part is always available.

- richness of time and spatial structures, which is not the exact copy of the richness of an elementary basic law, but often results from the iterative application of this law (Nicolis et Prigogine[13], Solbrig and Nicolis, *op. cit.*);
- emergence of “global” or “statistical” laws.

---

i

After two centuries of research oriented toward the study of basic processes and structures, it is really time to develop integrative approaches: to put together the pieces of the puzzle. To achieve this objective, modelling approaches or analysis based on general models (e.g. statistical ones) can help us.

3. Data analysis and modelling of complex systems

Mathematical, statistical or computer based tools are becoming increasingly efficient and more widely used. Some examples are:

- Statistical data analysis (multivariate analysis), which enables detection of complex structures[16]. There are lots of applications to ecosystem and to genome analysis. Based on linear algebra and Euclidian geometry, non-linear extensions are also taken into consideration today[17].
- Mathematical modelling and numerical simulations, which are now largely developed, lead both to basic and practical results (e.g. to answer questions about the stability or the emergence of regularities in complex systems[18], or to provide guidelines to manage living, technical or environmental systems)
- Computer based modelling, quantitative and qualitative simulations (e.g. cellular automata, multi-agents systems).

The extension of these approaches is closely related to advances in computer science and applications. We shall examine more particularly the mathematical and computer based models, theirs differences and complementarities.

4. Mathematical models *vs.* computer based models

Schematically, the principal characteristics of these formal tools are the following:

- Mathematical models are based on a strong theoretical framework, developed over more than 2,000 years. Mathematics offers well-defined concepts which can be successfully used in other fields of science. For example, the theory of dynamic systems introduces the concepts of stability, attractive (or repulsive) sets, chaotic trajectories, etc. Theorems lead to general results. When numerical simulations of these models are used, coherence and consistence of these simulations can be often verified (e.g. the results are not artefacts). But it is difficult to represent some mechanisms or structures, such as individual or social behaviours, or spatial heterogeneity.
- Computer models, such as cellular automata or multi-agent systems[19, 20] can do that, but there are only few theoretical results (mainly for cellular automata[21]). The coherence and consistence of simulations are difficult to verify. Generalisations can only be obtained by multiplying simulation, and even then, we are never sure that the results are "universally true".

Today, a complementary application can be recommended, linking mathematical and computer based modelling[2]. For example, simulations of a well-known system of which a mathematical model exists can help us to verify the functioning of a multi-agent simulator (e.g. the simulation of a Lotka-Volterra predator / prey system). Aggregating results of simulations can also lead to simplified results, which then enable mathematical approaches. For example, despite the complexity of microscopic interactions, the growth of an organism or a population is often well represented by simple, logistic models.

5. Some examples of complexity analysis and modelling

In the literature, there are many examples; some of which are already referenced in the previous sections. Here we mention simple "pedagogical" ones, which we used or on which we have worked more particularly:

- the first and second ones are devoted to theoretical and general questions;
- the third one illustrates how the coupling of different models can help to resolve environmental management problems;
- and the fourth one shows how aggregating data of a complex system (such as the evolution of biodiversity at the geological time scale) leads to a simple model which can provide some new results despite this simplification.

5.1. Theoretical questions

The two questions were:

a How to distinguish a stochastic process from a chaotic one?

b What are the effects of interactions in living systems? More disorder, a better order, or nothing?

a Chaos *vs.* stochasticity

This kind of question is now a frequent one in studies of natural population dynamics, since the discovery of chaotic dynamics and the R. May's paper on discrete logistic model[23] and the experimental work of Costantino *et al.*[24]. In fact, when it is possible to find a representation of such stochastic dynamics exhibiting regularities in an appropriate space, we can conclude that the system is chaotic (as shown in Figure 3). Conversely, if we do not find such a representation, we cannot draw any such conclusion, but statistical tests can be also a solution to find random characteristics of a dynamics.

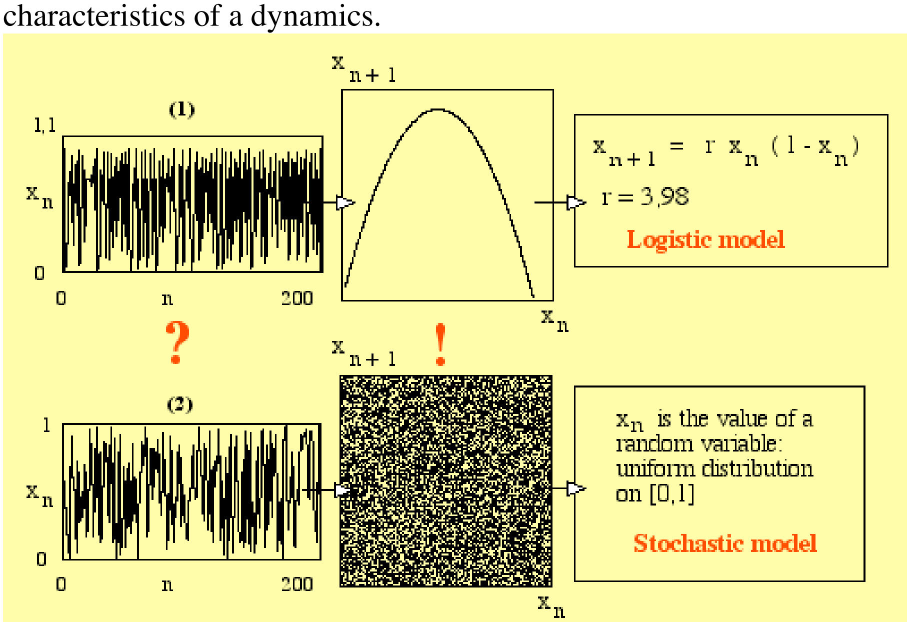


Figure 3. A simple example: the dynamics represented in the left part present small differences, but they exhibit large, non periodic, variations, which look like random processes. But, a simple representation: $x_{n+1}=f(x_n)$, can distinguish the two cases: chaotic (on the top) and random (on the bottom): "A good representation for a good solution" (Winston[10]).

b A « simple-complex » system

To illustrate a theoretical approach, we tried to answer the second simple question, which has been reformulated as follows: can interactions between components of a system lead to regularities? Or, does a complex non-linear system become more regular if entities are interacting?

For this purpose we choose two populations whose dynamics are chaotic, but which are interacting (in competition). Each population is supposed to grow following a discrete time logistic model, where the growth control parameter has the same value, chosen in the chaotic range (i.e. r=3.77). The model is:

$$\left| \begin{array}{l} x_{n+1} = r\, x_n\,(1 - x_n) - a\, x_n\, y_n \\ y_{n+1} = r\, y_n\,(1 - y_n) - a\, x_n\, y_n \end{array} \right.$$

The parameter a represents the intensity of the competition between populations x and y.

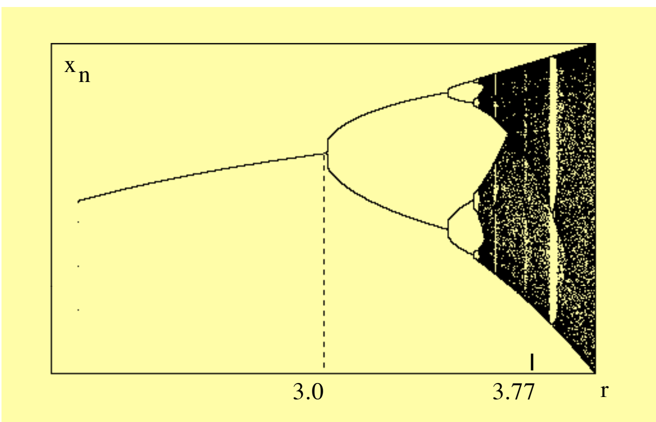


Figure 4. Bifurcation diagram of discrete time logistic model used for the choice of the value of growth parameter r; r = 3.77 is outside the regular domain (in the black part of the graph); then dynamics are chaotic.

The dependence between the simultaneous trajectories is ensured by choosing two different initial values, x(0) and y(0), and verified by examining the results in the sub space {x(t), y(t), t=0, 1,…, n}. For a=0, a very weak structure appears: the dynamics are largely independent (but it is not a random system as represented in Figure 3: (2)). When a increases, the structure becomes more apparent, but changes curiously (a surprising, unanticipated result). The expected result was a fixed point for a sufficiently great value of parameter a (i.e. the dynamics would be completely regular). In fact, the values of x and y do not reach a fixed point but are rapidly arranged along a straight line; the dynamics are synchronized. Then, at least in such a case, an interaction can make a system more regular than it would be without it. We tried other values for r, in the chaotic domain; the results were similar.

Aside from answering our question, the forms of the attractor are quite amazing. For values of a between 0.5 and 0.8, they are as strange as Ediacara fossils, compared to Cambrian and post Cambrian ones. That is why we called this attractor an "Ediacarian type attractor".Such an approach illustrates the interest of models for initial verification of hypotheses (i.e. in this case, the appearance of order in disorder, as the consequence of an interaction which increases the complexity of the system). A formal study of this model has yet to be achieved (providing there is further interest for it), but the results of simulations are already indicative.

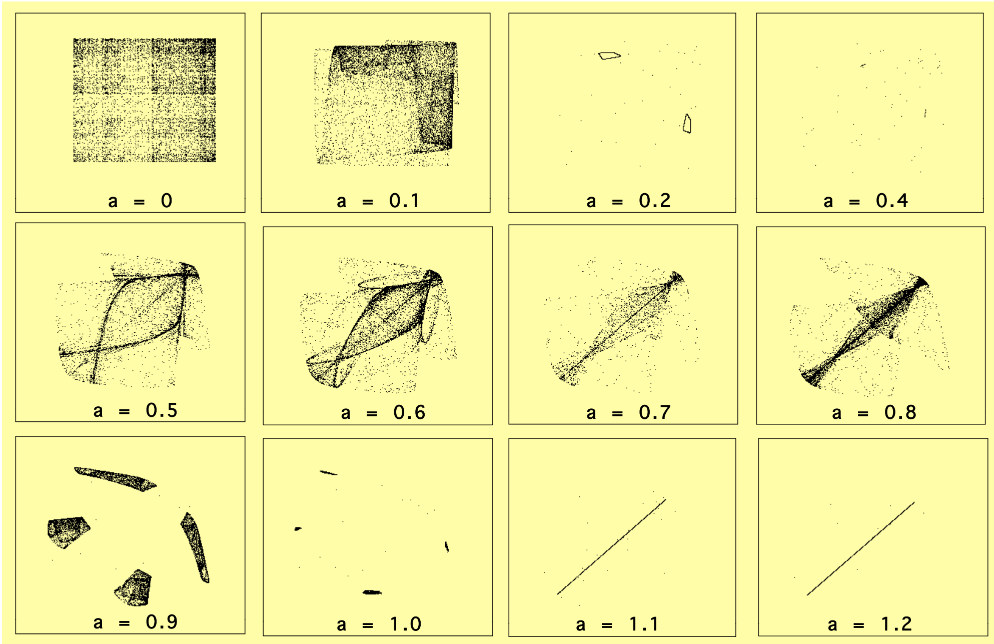


Figure 5. Progressive appearance of regularities in a competitive model of two chaotic populations. Values of parameter a control the intensity of the interaction, which makes the simultaneous dynamics more regular. Note the strange forms of the attractor for values between 0.1 and 1.1.

5.2. Integrated modelling

Traditionally, the modelling of complex systems, in the sense of many elementary interconnected components, was done by developing specific global models of these systems. But the size and the complexity of models themselves soon led to difficulties in interpreting the results and sometimes to strange behaviour in simulations. Today, the developments are based on an association of well-known and verified elementary models, linked together through general or specific interfaces. In fact, the integrated model looks like a network of elementary models. This is the method used by von Dassow *et al.*[25] to study problems of organism development. They demonstrate that the segment polarity network is a robust developmental network. On the basis of this biological example, they show that the many entities and interactions between these entities constitute a stable network and thus contribute to the debate of stability-complexity initiated by ecologists.

This way of integrated modelling is to describe in detail microscopic mechanisms and to simulate entire systems. There is also an approach centred on individuals, which can be implemented by using multiagents systems. This technique makes it possible to represent the behaviour of each individual and to simulate the global behaviour of an ecological system, for example the dynamics of a host-parasite system. But such an approach needs high performance computing. For instance, the simulation of the host-macroparasite system recently studied by G. Latu[26] needs from 100 teraflops to 1.45 petaflops, as well as the use parallel computers whose programming is still the domain of specialists. Such approaches make it possible to take into account the precise knowledge of biologists and to compile a representation of complexity of reality: the 1:1 model (in reference to Borgese's 1:1 map). However, one might question the interest of such maps and models...

The integrated modelling philosophy enables also to envisage coupling or associating models from different scientific areas (e.g. population dynamics, soil processes, climate, etc) in the same computer system. It will be interesting to study environmental problems, which are often dependent upon interdisciplinary approaches. We proposed such a way in a recent paper[27]. Figure 6 shows the conceptual architecture of such an integrated model, based on data fluxes between elementary models of each process. We suggested also that apart its operational interest, integrated modelling can be a dialogue tool in interdisciplinary studies.

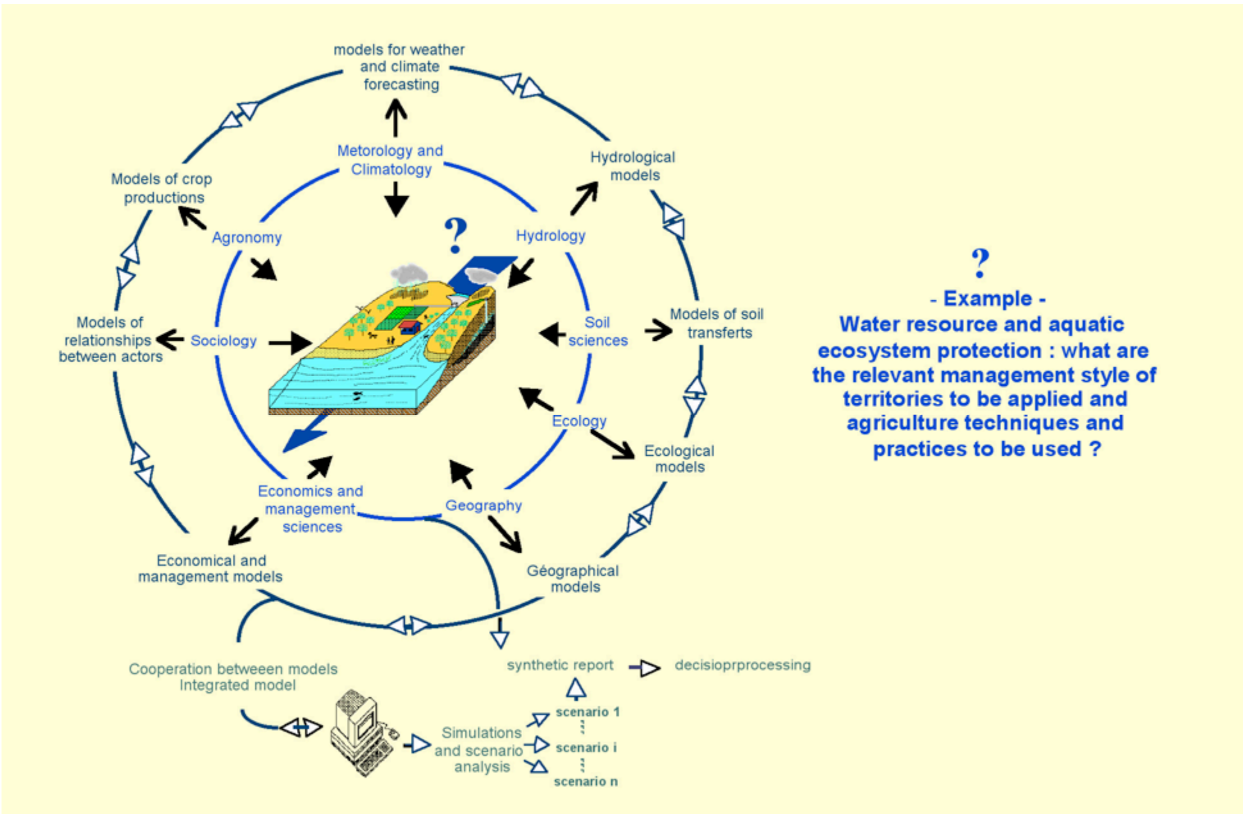


Figure 6. A conceptual example of an integrated modelling approach to solve an environmental problem.

5.3. Aggregation of data and simple modelling: the example of global biodiversity changes at the geological scale

Palaeontology is devoted to the study of fossils and thus biological evolution. Following the classic scientific approach, its trend is towards increasing details: precision of time scale, specialization on certain groups, etc. In this way, the history of life on earth is becoming increasingly well documented[28]. However, sometimes one needs a more synthetic and global point of view. In this line of thought, the use of the data base developed by Sepkoski is illustrative[29]. Despite the resulting simplifications, it enables quantitative analysis of large variations of biodiversity since the end of the Precambrian (Vendian). We use this example to show how aggregation of data, while it may sometimes conceal the complexity of the system, can nevertheless yield some interesting results. The global data concern the number of families during the last 600 M years.

Sepkoski used the global data quantitatively by applying statistics, principally to analyse the frequencies of extinction events[30]. The first global modelling approach was attempted by Benton[31]. He showed that the average exponential tendency of biodiversity increases. Courtillot and Gaudemer[32] proposed a more precise model: the logistic one, which fits well over long periods (each fitting was obtained independently for 4 periods). In recent "variations on a same theme," we published a twofold refinement[33]: (1) an interpretation of the logistic model by explicitly introducing a new variable which represents the number of ecological niches, and (2) by chaining logistical model by a global fitting. It is not our purpose to present the results in detail here. Just remember that the logistic model depends of two parameters r (the growth parameter, characteristic of the diversification process) and K (the plateau, interpreted as the maximum number of ecological niches). The main results are as follows (cf. Figure7) :

- The growth rate of biodiversity is the same apart from the Jurassic-Cretaceous period (interpreted as a succession of "minor" environmental perturbations which makes this rate lower).
- The increase of the plateau is interpreted following the Michod's theory on Darwinian dynamics[34]: the progressive appearance of new ecological relationships, such as cooperative ones, which enable « niche sharing » and thus enable species to coexist, even to establish cooperative interactions. The importance of these kinds of cooperative relationships in ecosystem functioning is becoming increasingly clear[35].
- A simple logistic model chained along major periods provides a good description of the global dynamics of a complex system: the residual information obtained after the aggregation of data permits global interpretation. This is one way to simplify complex system analysis: first to detect large tendencies, and progressively to detail involved process, to refine models.

It is a further exercise to develop global ecology (in both senses: spatial, at the planetary level, and temporal, at the geological scale) toward a *macroecology*.

6. Conclusion

Modelling is certainly an efficient methodology for integrative approaches, particularly in life sciences[36]. Remember also that ecology, and population dynamics are the disciplines in these sciences where the modelling approach was first developed[37] (it is also true for bioinformatics, despite the better known current developments in genomics[38]). Sometimes the modelling approach permitted the emergence of new paradigms. For example, the studies on HIV infections changed after the papers of Nowak *et al.*[39], based on a population dynamics approach to the viral infection at the level of the individual; another example is chaos and complexity concepts themselves[40].

Models can help us either in speculative thinking, conceptual approaches or practical needs. Today, we could say: there is no scientific work without modelling, although that would certainly be a "provocative" or an "imperialist" position. In any case, the debate is still open around the following questions:

- What about complex models *vs*. simple models? "Small is beautiful", but sometimes we have to work with "big models". But small is not synonymous with simple, nor big with complex. For example, mathematical models of a limited number of equations may, in fact, be highly complex (e.g. Navier-Stoke equations of fluids dynamics or equations of combustion[41], or even the model proposed above of two interacting populations). Any modelling is preferable to working without a model.

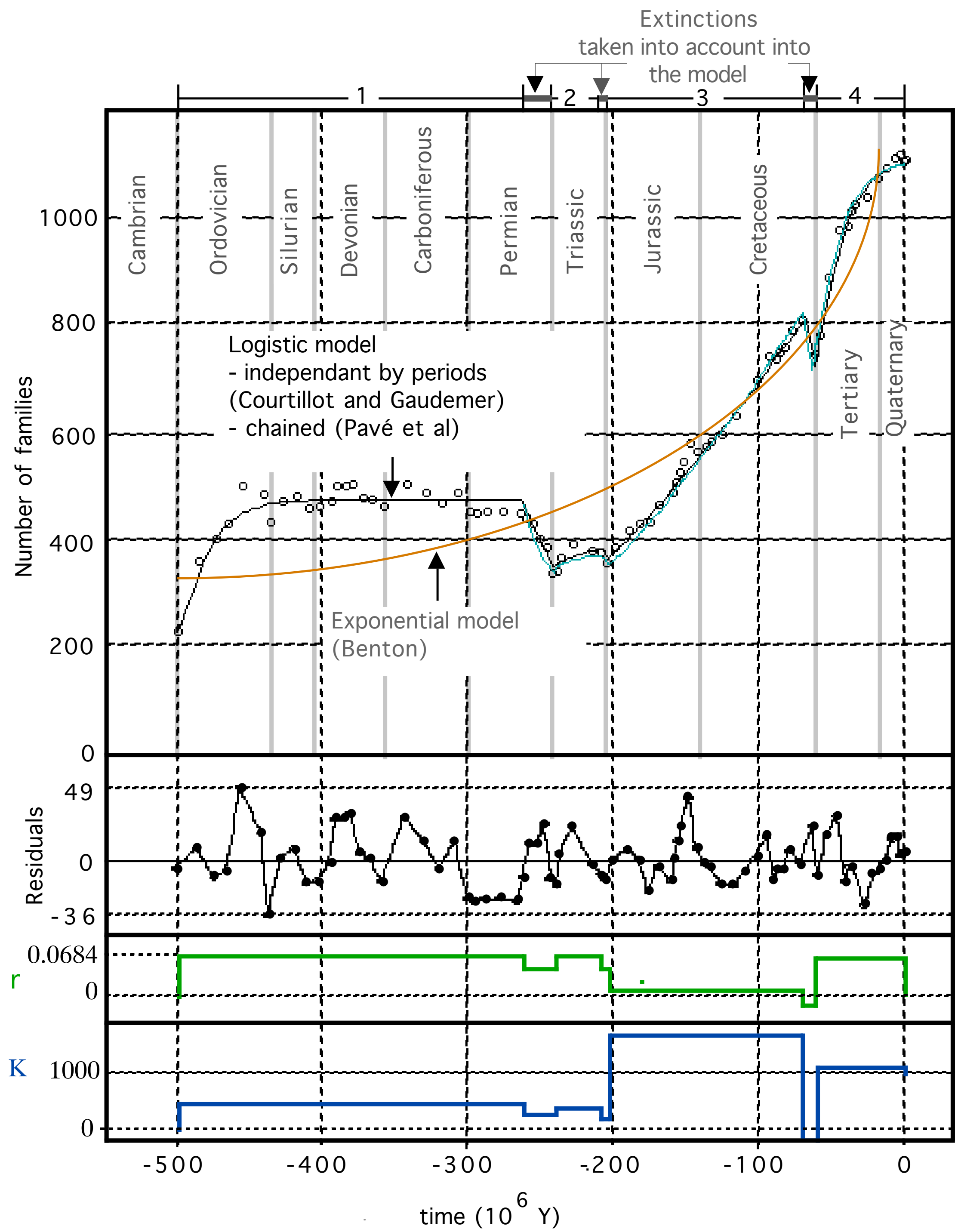


Figure 7. Biodiversity changes at the geological time scale (only data after the Cambrian are taken into account, i.e. after 500 MYA). Data are represented by circles (margins of error are not represented to facilitate interpretation of the graph). The chained logistic model dN/dt = r N (1-N/K) provides a good representation of the major tendencies, both during growth periods and also during extinctions. “Chained” model (Pavé *et al.*[33]) means that parameters r and K are estimated for each periods, but the initial condition is only estimated at the beginning (t=-500 MA). In fact it constitutes a continuous solution. “Piece by piece” model (Courtillot and Gaudemer[34]) means that, apart r and K, initial conditions are re-estimated at the beginning of each period.

- For complex systems (such as an "environment"), what is the better way: to associate models of different kinds of processes (e.g. climatic models, ecological models[42], even economic models), or to construct new "global" models? There is no obvious answer to this question.
- Aggregate data could simplify models, but do the higher scales or upper levels always have something to teach us? Only the development of global approaches (such as macroecology[43]) will enable us to answer to this question.

One of the main practical questions we have to answer in the near future is how to deal with complex systems. It is now evident that most natural processes lead to diversity and complexity. This is also true for many human activities, including technological developments and social structures. In the past, the tendency was to simplify, with the result that most management rules and techniques were elaborated for simplified systems (for example, for mono-specific crops in agriculture). Now we know that was not always a good solution: since we are dealing with complexity, we must therefore find new rules adapted to the management of such systems. Modelling and simulation can provide effective methodologies for developing new and appropriate management tools.

References

1-Vicsek T. The bigger picture. *Nature*, 418, 2002, 131.

2-Lewin R. - *Complexity : Life and the Edge of Chaos.* Macmillan,1993, 208 p

3-Schesinger M. From chaos to complexity (a review of books). *Nature*, 394, 1998, 538.

4-Delahaye J.P. - Complexités, la profondeur logique selon C. Bennett , *Pour La Science*, 166, 102-104. (1991).

5-Nicolis G. Le vocabulaire de la complexité. *La Science au Présent*. Ed. Encyclopædia universalis, Paris, Vol. 1, 208-211, 1992.

6-May R.M. - *Stability and Complexity in Model Ecosystems.* Princeton University Press. (1973).

7-McCann K., Hastings A. and Huxel G.R. Weak trophic interactions and the balance of nature. *Nature*, 395, 794-798, 1998.

8-Polis G.A. Stability is woven by complex webs. *Nature*, 395, 1998, 744-745

9-Hulot F.D., Lacroix G., Lesher-Moutoué F., Loreau M. Functional diversity governs ecosystems response to nutrient enrichment. *Nature*, 405, 2000, 340-344.

10-Legay J.-M. - Méthode et modèles dans l'étude des systèmes complexes. In *"Pour une agriculture diversifiée"*. Ed. Jollivet M., l'Harmattan, 1988, 14-24.

11-Winston P.H. *Intelligence artificielle*. InterEdition, Paris, 1988.

12-Solbrig O., Nicolis G. *Perspectives on Biological Complexity.* IUBS Monograph series, n°6, Paris, (1991).

13-Nicolis G. and Prigogine I. *Self-Organization in Nonequilibrium Systems*, J. Wiley & Sons, New York, 1977.

14-André J.C., Mégie G., Schmidt-Lainé Cl. Échelles et changements d'échelles : problématiques et outils. *Rapport Science et Technologie Académie des Sciences et Académie des Technologies* " L'environnement à toutes les échelles: du territoire au continent", 2003.

15-Pavé A. *Modélisation en biologie et en écologie*. Aléas, Lyon, 1994..

16-Thioulouse J., Chessel D., Champély S. 1995. Multivariate analysis of spatial patterns : an unified approach to local and global structures. *Environmental and Ecological Statistics*, 2, 1-14.

17-Gershenfeld N., Schoner B., Metois E. Cluster-weighted modelling for time-series analysis. *Nature*, 397, 1999, 329-332.

18-Cazelle B. *Non-linearity and stochasticity in population dynamics*. HDR dissertation, University of Paris 6, Paris, 2001.

19-Bousquet F. *multi-agent simulations and natural renewable resources management*, HDR dissertation, Université Claude Bernard, Lyon, 2000.

20-Bousquet F., Cambier C., Mullon C., Morand P., Quensière J. and Pavé A., 1993. Simulating the Interaction Between a Society and a Renewable Resource. *Journal of Biological Systems*. 1:2, 199-214.

21-Elmoznino H., 1999. *Influence du cycle de vie individuel sur la dynamique spatiale d'une forêt monospécifique. Analyse à travers un automate cellulaire*. Thèse de l'Université de Nice-Sophia Antipolis.

22-Bakam I., Kodron F., La Page C., Bousquet F., 2000. Formalization of a spatialized multi-agent system using coloured Petri nets for the study of a hunting management system. *Lecture Notes in Computer Systems*, Springer.

23-May R. M. Simple mathematical model with very complicated dynamics. *Nature*, 261, 459-467, 1976

24-Costantino R.F., Desharnais R.A., Cushing J.M., Dennis B. – Chaotic Dynamics in an Insect Population. *Science*, 275, 389-391, 1997.

25-von Dassow G., Meir E., Munro E.M. and Odell G.M., The segment polarity network is a robust developmental module, *Nature* 406, 2000, 188-192.

26-Latu G. *Algorithmique parallèle et calcul haute performance dédiés à la simulation d'un système hôte-macroparasite (Parallel algorithmic and high performance computing dedicated to the simulation of a host-macroparasite system)*. Thèse Université Bordeaux I, France, 2002,231p.

27-Schmidt-Lainé Cl., Pavé A. - Environnement : modélisation et modèles pour comprendre, agir et décider dans un contexte interdisciplinaire. *Natures-Sciences-Société : Sciences pour l'ingénierie de l'environnemen,t* 10 : s.1, 2002, 5-25.

28-Carroll S.B. Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409, 2001, 1102-1109.

29-Sepkoski, J.J., A compendium of fossil marine families. *Milwauk, Public Mus., Contrib. Biol. Geol*. 83 (1992), 1-156.

30-Sepkoski J.J., Is the periodicity of extinctions a taxonomic artefact ? *Nature*, 330 (1987), 251-252.

31-Benton M.J., Diversification and extinction in the history of life, *Science*, 268 (1995)

32-Courtillot V. and Gaudmer Y., Effects of mass extinction on biodiversity. *Nature*, 381 (1996), 146-148.

33-Pavé A., Hervé J.C. et Schmidt-Lainé Cl. Mass extinctions, biodiversity explosions and ecological niches. *Comptes rendus de l'Académie des Sciences-Biologies*, 325 : 7, 2002, 755-765.

34-Michod R.E., *Darwinian Dynamics*. Princeton paperbacks, 2000, Princeton. 262p.

35-Cardinale B.J., Palmer M.A. Collins S. L. species diversity enhance ecosystem functioning through interspecific facilitation. *Nature*, 415, 2002, 426-429.

36-Maddox J. – Polite row about models in biology. *Nature*, 373, 1995, 555.

37-Levin S.A., Grenfell Bryan, Hastings A., Perelson A.S., Mathematical and Computational Challenges in Population Biology and Ecosystems Science. *Science*, 1997, 275, 334-342.

38-Schatz B.R., Levin S.A., Weinstein J.N., *et al* – Bioinformatics. *Science*, 275, 1997, 327-349.

39-Nowack M.A., May R.M., Phillips R.E., Rowland-Jones S., Lalloo D.G., MacAdam S., Klenerman P., Köppe B., Sigmund K., Bangham C. R. M., McMichael A.J., 1995. Antigenic oscillations and shifting immunodominance in HIV-1 infections. *Nature*, 375, 606-611.

40-Waldrop M. - *Complexity : the Emerging Science and the Edge of Order and Chaos.* Simon and Schuster/Viking, 380 p. (1992).

41-Schmidt-Lainé Cl. (Ed.). *Modélisation de la combustion*. CNRS, Paris, 1996.

42-Melillo J.M., Mc Guire A.D., Kicklighter D.W., Moore B., Voromarty C.J., Schloss A.L., 1993. Global climate change and terrestrial net primary production. *Nature,* 353, 234-240.

43-Nee S. Biodiversity: Thinking big in ecology, *Nature*, 417, 2002, 229-230