Project no: 043268

Project acronym: PATRES

PATTERN RESILIENCE

Instrument: STREP

Thematic Priority: New and Emerging Science and Technology

Periodic activity report

Period covered: from 01/08/2008 to 31/01/2010 Date of preparation: 15/02/2010

Start date of project: 01/02/2007 Duration: 36 months

Project coordinator: Guillaume DEFFUANT

Project coordinator organisation: Cemagref Revision: Version 2

Content

Publishable executive summary 3
Section 1 – Project objectives and major achievements during the reporting period 9
Section 2 – Workpackage progress of the period 12
Section 3 – Consortium management 23
Section 4 – Other issues 25
Annex – Plan for using and disseminating the knowledge 26

Publishable executive summary

http://www.patres-project.eu

Scientific problem and objectives

The project chose to found its research on Martin's (2004) formalisation of the concept of resilience, which is based on viability theory. This framework is more general than existing definitions because it requires no assumption about the dynamical properties of the system. Moreover, the approach enables to compute laws of actions on the system in order to keep or restore a desired property, lost after a perturbation.

However, solving a viability problem is practically possible only when the problem is expressed in state space of relatively small dimensionality (up to 7 or 8 dimensions). It is therefore impossible to apply the method on systems described by a large number of interconnected entities, because the state space has too many dimensions. Nevertheless, when the interconnected entities generate statistical regularities or patterns, which can be described with a reasonable number of dimensions, and when the desired properties of the system are related to these patterns, the approach can be adapted. The association of patterns with resilience justifies the title of the project: *"pattern resilience"*.

The main objective of the project derives from this scientific challenge: to elaborate efficient methods and tools for modelling and managing pattern resilience in complex systems. The methods integrate contributions from the research on resilience, more particularly its link with viability theory, and methods for pattern identification in models and data.

The main objective therefore includes two aspects:

- *Defining more powerful and more flexible methods and tools for solving viability problems* than current ones, in particular using recent statistical tools such as Support Vector Machines (SVMs), and therefore increase the range of systems in which the resilience problem can be solved.
- *Providing a set of methods and tools for modelling pattern dynamics*, building on current work on the exploration of models with systematic experimental designs, and on general statistical physics approaches.

The project investigates the efficiency of the developed methods and tools on *four case studies*, in very different domains (ecology, social sciences, cognition).

Participants

Co-ordinator contact details:
Guillaume Deffuant
Cemagref, LISC. 24, rue des Landais, BP50085, 63172, Aubière, France.
Tel: 33 4 73 44 06 14
Email: guillaume.deffuant@cemagref.fr

	Participant organisation name	**short name**	**Country**
1	Cemagref Laboratoire d'Ingénierie des Systèmes Complexes	Cemagref	F
2	University of Surrey Department of Sociology	UniS	UK
3	Universitat Illes Balear IFISC	UIB	S
4	Helmholtz Centre for Environmental Research - UFZ Department of Ecological Modelling	UFZ	D
5	Centre National de la Recherche Scientifique Centre de Recherche en Epistémologie Appliquée	CNRS	F

Work performed

In this second period of the project, we derived more refined versions of the case study models, and we used them to test our tools and methods. Globally, we better identified the main steps of a global approach to compute viability based resilience on complex systems. Moreover, for some case studies, new tools for collecting data were developed.

During the period, we held two plenary meetings and three visits between partners took place.

Moreover, we organised two knowledge dissemination workshops:

- a one day presentation in a satellite workshop of the European Conference on Complex Systems in Warwick,
- a three day course for young scientists, which was held in Madeira.

The main dissemination vector for the future is the book produced by the consortium. It describes in details the methods and tools, as well as their applications on case studies. It will be published by the end of 2010.

Results achieved

Extending current definitions of resilience toward an action oriented approach

During the project, we improved collectively the formal definition of resilience based on viability, and we identified better its links with existing definitions. In particular, we considered the main existing formal definitions of resilience: one based on a linearization of the dynamics in the vicinity of an attractor of the dynamics, often called the "engineering resilience", and another one based on regime shifts and attraction basins size. We showed that these definitions can be seen as particular cases of the viability based definition, which is therefore more general. We illustrated this claim on an example from the literature (Anderies et al. 2003). Moreover, we showed that the viability based resilience is more action oriented, because one can compute policies of action to maintain or drive back a system into a desired state set in its framework.

Pattern resilience: a general approach tested on a set of case studies

The main problem of the practical application of the viability-based resilience is that the computation of viable or resilient states is very heavy computationally. In particular, it is impossible to apply it directly to systems defined in a state space with many dimensions. For instance, the models including interacting agents include a lot of variables defining their states, and hence the computation of viability-based resilience in such systems appears out of current computational means.

However, we prove on several case studies that it is sometimes possible to approximate the main pattern dynamics of the system, using more synthetic representations, and then to compute the viability and resilience on the approximated dynamics. The general approach includes the following steps:

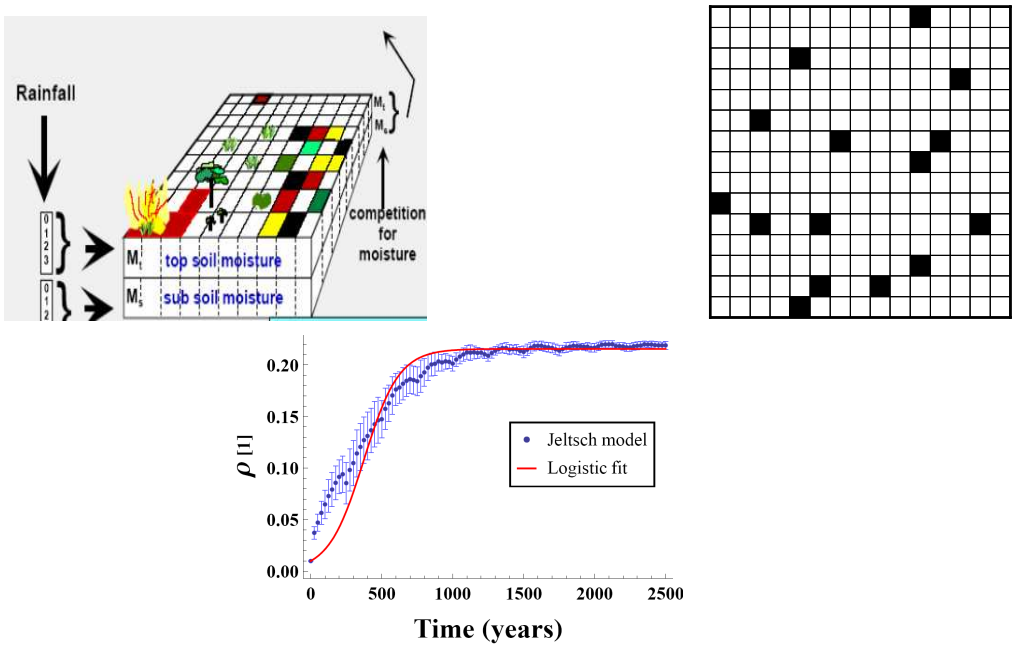

Fig. 1: Savanna case study, first step: the complex model (top left) is simplified into a cellular automaton (top right, black squares are trees, white squares are grass), with rules derived from simple functions fitting the simulations of the complex model (bottom).

- define a simpler individual based model, with simplified individuals and dynamics (see fig. 1),
- to apply methods of statistical physics (such as pair approximation) on this simplified model, yielding synthetic representations of the system's pattern dynamics,
- to define the desired state set in the state space of the pattern dynamics,
- to compute viability and resilience of the system relatively to this desired state set, and corresponding policies of action (see fig. 2).

$$
\begin{aligned}
\frac{d\rho[1]}{dt} &= \beta\left(z_n q_n[1/0] + z_f q_f[1/0]\right)(1-\rho[1]) \\
&\quad \times P_F^{Surv} e^{-\delta z_n q_n[1/0]} - \alpha\rho[1] \\
\frac{1}{2}\frac{d\rho_n[11]}{dt} &= \beta\left(1+(z_n-1)q_n[1/0] + z_f q_f[1/0]\right)(\rho[1]-\rho_n[11]) \\
&\quad \times P_F^{Surv} e^{-\delta(1+(z_n-1)q_n[1/0])} - \alpha\rho_n[11] \\
\frac{1}{2}\frac{d\rho_f[11]}{dt} &= \beta\left(z_n q_n[1/0] + 1 + (z_f-1)q_f[1/0]\right)(\rho[1]-\rho_f[11]) \\
&\quad \times P_F^{Surv} e^{-\delta z_n q_n[1/0]} - \alpha\rho_f[11],
\end{aligned}
$$

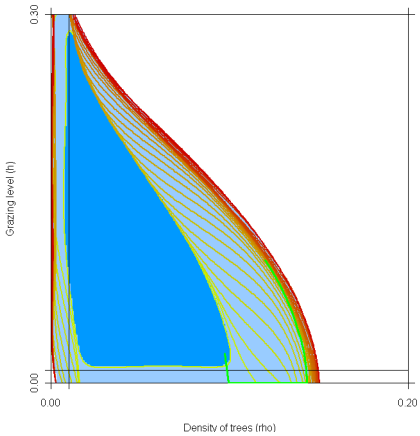

Fig. 2: Left: differential equation defining pattern dynamics from the simple model of savanna. Right: representation of viability kernel (in blue), resilience indices (lines from yellow to red) and trajectory of the system returning to the viability kernel after perturbation (in green).

In this approach, it is necessary to make many computation experiments: first to check that the simpler individual based model (IBM) fits well the more complex one, and then to check that the synthetic pattern dynamics model fits well the simpler IBM. It is also important to perform sensitivity analyses of these models, and to explore their regime shifts.

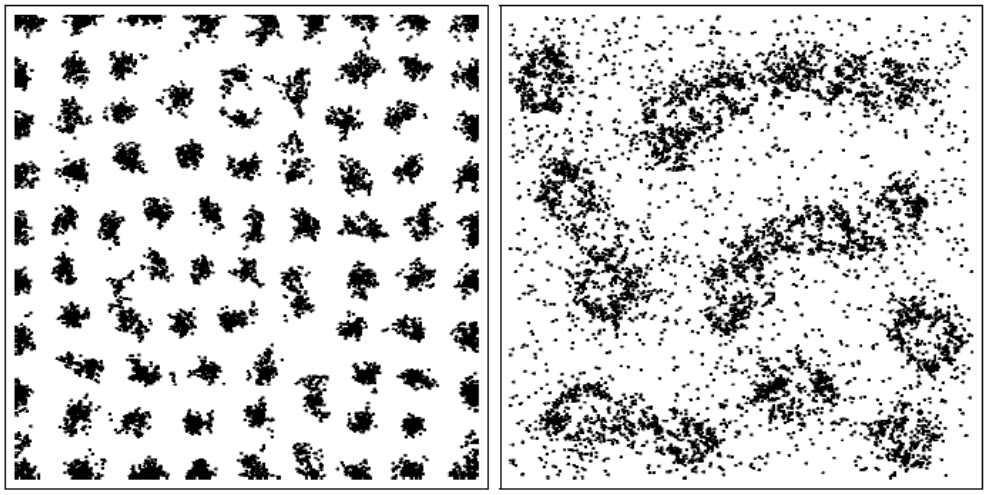

Fig. 3: Examples of patterns obtained from the bacteria model, obtained by changing the value of one parameter. The challenge was to maintain some of their characteristics through adequate actions.

In most cases, it is also interesting to perform a sensitivity analysis of viability and resilience when the parameters of the model vary. This provides the values of the parameters for which the system has the highest viability or resilience. Moreover, we tested the typical optimal policies provided by the software for different values of the parameters.

In addition to this general approach, the work on the case studies achieved some noticeable results, obtained in the second period of the project:

- In the savanna case study, we showed that the interaction between tree-tree establishment competition and fire is fundamental to explaining savanna dynamics. Moreover, we identified specific management strategies for maintaining the savanna in a desired state set.
- In the bacteria case study, we identified a new realistic mechanism for generating observed patterns in bacteria populations (see fig. 3). Moreover, in an idealised setting, we derived policies of actions for maintaining the characteristics of these emergent patterns.
- In the language competition case study, we proposed a more comprehensive analytical model which includes the stochastic part of the dynamics. Moreover, on simpler versions of this model, we computed policies of action on the prestige of languages in order to keep their diversity.
- In the social dilemma case study, we used a toy model of interacting agents with more or less cooperative attitudes, and we derived from it policies of actions for a public body to favour cooperative attitudes.
- We gathered new sets of data about the web 2.0, and observed patterns for the evolutions of wikis and sites for sharing images (Flickr).
- We defined new methods for deriving patterns dynamics from individual-based models defined by two symmetric absorbing states, competitive dynamics and continuous states with local attraction (bounded confidence models) with noise.

Improvement of tools and methods for computing viability and resilience

The main achievement of the project is the development of KAVIAR, a software prototype which includes recent advances on viability kernel approximation with "support vector machines" (SVM). In its simplest version, the algorithm uses a lattice defined on the state space. Then an algorithm defines iteratively a training set separating viable from non-viable points with respect to the current approximation. SVMs are well known for providing efficient and parsimonious solutions to such problems. More sophisticated versions of the algorithm, using only a part of the lattice for learning ("active learning") are also implemented. Then, the software provides the control policies to apply in order to remain in the viability kernel, or to drive the system back into it. A specific variant, optimised for the computation of resilience indices, is also available. We wrote a user guide for this tool, and several participants of the project were trained to use it on their problems. The software is available from the project web site: http://www.patres-project.eu.

In the second period, we extensively tested the software and made it more robust. Moreover, we added different methods of optimisation for computing the controls.

More generally, we proposed a new approach for defining robust control policies to keep a system within its viability kernel. This approach is based on the computation of the distance from the states to the boundary of the viability kernel.

Improvement of SimExplorer, a software tool supporting complex simulation experiments

When one wants to identify patterns in models, repeating heavy simulation experiments is often necessary.

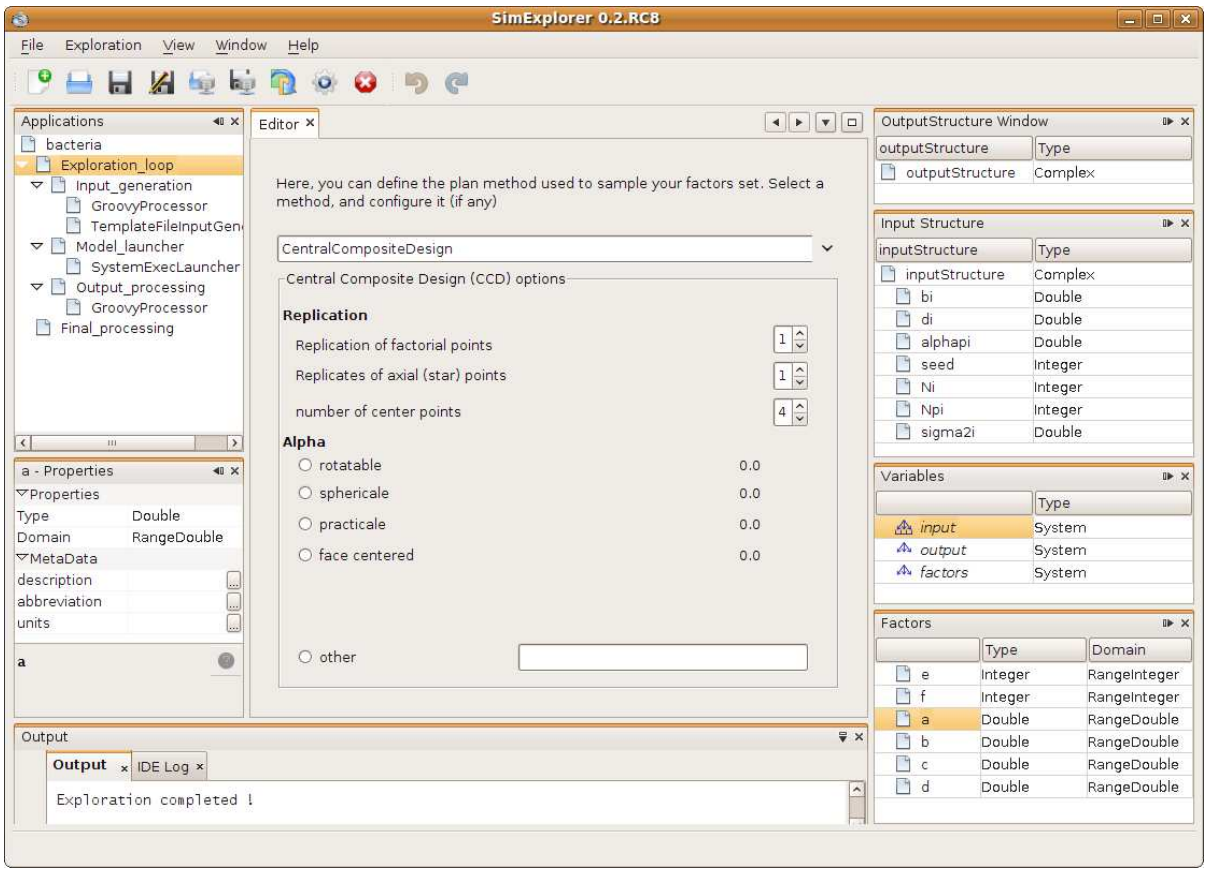

Fig. 5: User interface of SimExplorer.

The prototype software of SimExplorer, developed during the project, helps the user to define a workflow for such experiments: experimental design, generation of model inputs, model launching, model outputs treatments (see fig.5). It includes a link to commonly used libraries to define experimental designs and statistical treatments (R packages), facilities to trace the experiments and manage their quality, facilities to launch experiments on computer clusters or grids. The development of this tool has also been financially supported by LifeGrid program (until September 2008). We used the software on some of the case studies (in particular savanna and bacteria).

The progress during the second period concerns mainly the facilities for launching sets of simulations on clusters and computer grids.

Section 1 – Project objectives and major achievements during the reporting period

Project general objectives and relation to the state of the art

The main objective of the project is *to elaborate an efficient method for modelling and managing pattern resilience in complex systems.* The method will integrate contributions from the research on resilience, more particularly its link with Viability Theory, and methods for pattern identification in models and data.

The efficiency of the developed method depends on the efficiency of its parts. The main objective therefore includes the following ones:

- *Defining more powerful and more flexible methods for solving viability problems* than current ones, in particular using recent statistical tools such as SVMs, and therefore increase the range of systems in which the resilience problem can be solved.
- *Providing a set of methods for modelling pattern dynamics*, building on current work on the exploration of models with systematic experimental designs, and on general statistical physics approaches.

The project will develop two software prototypes which implement the two main parts of the global method, and will be designed to be used together in order to compute pattern resilience:

- A tool *for identifying and modelling pattern dynamics in complex models*, implementing the methodological innovations of the project, and taking into account the requirements of the general method for modelling and managing pattern resilience.
- A tool *for computing resilience properties of a system*, and *the sets of actions to carry out* in order *to maintain or recover a desired pattern dynamics*. The tool will implement the methodological innovations of the project.

Much attention will be paid to a joint use of these tools in the context of pattern resilience computation, but each part will be usable independently.

The project will investigate the efficiency of the developed methods and tools on *five case studies*, in very different domains (ecology, social sciences, cognition). For each case study, the objectives are:

- *to provide an implementation of a micro-level model of the system, and of the general method applied to it, including the pattern dynamics model, and the related resilience computation,*
- *to bring new knowledge about the resilience of the considered complex system.*

Objectives of the period

The period corresponds mainly to development phase of the project, coming after the pilot phase. It included the following objectives:

- To develop a refined model for each case study and then test the refined methods for pattern dynamics identification and resilience computing.
- To design a refined set of methods for pattern identification, to be tested on the case studies. This first version will include a set of already existing tools and methods (SimExplorer, moment approximation…).
- To design a refined general method for computing resilience and related action policies. This first version will include existing components for viability kernel computation and their adaptation to compute resilient action policies (existing viability algorithms).

Work performed during the period

In this second period of the project, we derived more refined versions of the case study models, and we used them to test our tools and methods. Globally, we better identified the main steps of a global approach to compute viability based resilience on complex systems.

In addition to this general approach, the work on the case studies achieved some noticeable results, obtained in the second period of the project:

- In the savanna case study, we showed that the interaction between tree-tree establishment competition and fire is fundamental to explaining savanna dynamics. Moreover, we identified specific management strategies for maintaining the savanna in a desired state set.
- In the bacteria case study, we identified a new realistic mechanism for generating observed patterns in bacteria populations (see fig. 3). Moreover, in an idealised setting, we derived policies of actions for maintaining the characteristics of these emergent patterns.
- In the language competition case study, we proposed a more comprehensive analytical model which includes the stochastic part of the dynamics. Moreover, on simpler versions of this model, we computed policies of action on the prestige of languages in order to keep their diversity.
- In the social dilemma case study, we improved the representation of the problem by including a variable corresponding to the budget of the institution, to take the cost of the action into account.
- We gathered new sets of data about the web 2.0, and observed patterns for the evolutions of wikis and sites for sharing images (Flickr).

In terms of methods and tools, this second period has also been devoted to the refinement and development of the achievements of the first phase of the project:

- We improved Kaviar, the software prototype for computing viability kernels and resilience indices, using "support vector machines". The improvements come from the correction of bugs found through intensive testing on the case studies, and the addition of different optimisation methods for computing the controls.
- We defined new methods for deriving patterns dynamics from individual-based models defined by two symmetric absorbing states, competitive dynamics and continuous states with local attraction (bounded confidence models) with noise.

- We proposed a new approach for defining robust control policies to keep a system within its viability kernel. This approach is based on the computation of the distance from the states to the boundary of the viability kernel.

During the period, we held two plenary meetings and three visits between partners took place.

Moreover, we organised two knowledge dissemination workshops:

- A one day presentation in a satellite workshop of the European Conference on Complex Systems in Warwick,
- A three day course for young scientists, which was held in Madeira.

The main dissemination vector for the future is the book produced by the consortium. It describes in details the methods and tools, as well as their applications on case studies. It will be published by the end of 2010.

Adjustments

For the CNRS case study, we explored some models of dynamics of scientific domains, based on data bases of citations. However, we did not manage to express a proper dynamic model beyond the observed data, and hence we could not apply the global approach defined in the project. Therefore, we focused on the social dilemma case study.

In the Surrey case study, we also had difficulties to develop a relevant model for the dynamics of Web 2.0. This model was not advanced enough in the end of the project to apply the global approach.

Section 2 – Workpackage progress of the period

Workpackage 1: Case studies

Objectives

During the period, the main objectives were:

1. To define and implement refined models of the micro level dynamics of each case study, confronted with different sources of data (Pattern Oriented Modelling)
2. To test the version of methods and tools developed in WP2 and WP3 on these refined dynamics, and provide guidance for their improvement

Progress toward objectives

Global progress

The work fits the objectives for 4 case studies (Savanna, bacteria, languages and social dilemma). In this second period of the project, we developed refined dynamics for these systems, and tested the tools for pattern dynamics and viability – resilience computation. In the other case studies (web communities, pattern dynamics in scientific literature) we mainly focused on data driven descriptions of these systems pattern dynamics.

Case studies

Savanna case study (UFZ)

The work in the savnna case study progressed in the following three directions:

- UFZ and IFISC studied the effects of tree-tree establishment competition and fire on savanna dynamics. The mechanisms regulating savanna tree populations are still not well understood. Recent empirical work suggests that both tree-tree competition and fire are key factors in semi-arid to mesic savannas. We developed a minimalistic and analytically tractable stochastic cellular automaton to study the individual and combined effects of competition and fire on savannas. We found that while competition often strongly depresses tree density, fire generally has little impact, but can drive tree extinction in extreme scenarios. When combined, competition and fire interact nonlinearly, magnifying each others negative effects on tree density, a novel result that may help explain several observed phenomena in savannas (Calabrese et al. 2010).
- IFISC and UFZ studied tree-grass coexistence: the "savanna problem". The phenomenon of the coexistence of trees and grasses in savanna ecosystems has been extensively studied during the last decades, and yet it is not well understood. Recent experimental observations show that external forcing like varying weather conditions (rainfalls) is a key factor that regulates savanna dynamics. To tackle the effects of this mechanism from a Statistical Physics point of view, we proposed and studied a simple stochastic model, which is a variant of the Contact Process. We then introduce the concept of "dynamical phase coexistence" to provide a simple solution for the ``savanna problem". The model exhibits a transition from grassland to tree-grass coexistence, whose nature depends on

the values of the internal parameters. We also study the spatial structure of tree distribution and characterize the mechanisms leading to clustering (Vasquez et al. 2010).

- UFZ and Cemagref studied the viability and resilience of the mean-field and pair approximation models of the stochastic cellular automata model. We studied first the case of a desired set including the attractor of the dynamics for a density of trees around 0.22. Then we considered desired state sets where the density of trees should be over a minimum value of 0.01 or 0.03, and should be below 0.2. These desired sets exclude the attractors of the mean-field dynamics. Hence, maintaining the system in these state sets requires to increase or decrease grazing periodically. We observed that the same desired set includes an attractor (for a value of the tree density around 0.18), and leads to qualitatively different viability kernels. More details are provided in deliverable (3.1) and in the book in which a chapter is devoted to this case study.

Publications:

Calabrese, Justin; Vazquez, Federico; López, Cristóbal; San Miguel, Maxi; Grimm, Volker. 2010. The individual and interactive effects of tree-tree establishment competition and fire on savanna structure and dynamics. *American Naturalist.* 175: pp. E44–E65.

Calabrese JM, Deffuant G, Grimm V. Bridging the gap between computational models and viability theory in savanna ecosystems. In: Deffuant G, Gilbert N (eds) *Pattern Resilience: Computing Resilient Action Policies for Social and Ecological Research*. Springer (to appear).

Vazquez, Federico; López, Cristobal; Calabrese, Justin and Muñoz, Miguel Angel. 2010. Dynamical phase coexistence: a simple solution to the savanna problem. *Journal of Theoretical Biology* (in press).

Bacteria case study (Cemagref)

First, we developed a detailed model of bacteria dynamics which was inspired by observations from different experiments. The model includes bacteria with various sizes, depending on the nutrient they consume. The bacteria are motile, and they produce a polymer which affects their ability to move. The diffusion of the substrate is ruled by usual diffusion dynamics. We studied this model, especially considering different values of the influence of the polymer on motility. In some cases, we noticed that a specific spatial organisation emerges, with connected colonies, corresponding to some observation in real settings. This work led to a publication (Mabrouk et al. 2010).

To study the pattern dynamics of this model, we derived a simpler version which exhibits similar patterns, but is easier to study both experimentally and theoretically. In this model, the bacteria are represented by simple points, and we considered that the concentration of substrate is constant in the whole considered volume. In the simplest version, we also eliminated the polymer, because we showed that the model without polymer exhibits the same set of patterns. These patterns go from a globally uniform distribution, to stripes and labyrinth like organisations and also regularly distributed round shape colonies. These patterns have been observed in real biofilms.

We studied these pattern dynamics using second moment approach. The idea is to capture the essential part of the spatial dynamics using by approximating them with the dynamics of spatial correlations. We showed that for most cases this approach is adequate to provide a synthesis of

the average spatial dynamics. However, in the cases where spatial fluctuations of the distribution play an important role in the dynamics, the approach is not appropriate, because these fluctuations are erased by an averaging process.

The moment model could not be directly used in the viability – resilience tools, because it represents a 2D function. Considering only two points of it like in the Savanna case was not possible. However, we managed to define other variables summarizing the pattern observed in the moment function, and to express the problem in a small number of dimensions. This work is the object of a chapter of the book (Mabrouk et al. 2010).

Mabrouk N, Deffuant G, Lobry C. Tolker-Nielsen T. 2010: Bacteria can aggregate into interconnected microcolonies when a self-excreted product reduces their surface motility: evidence from individual-based model simulations, *Journal of Theoretical Biology* – In press.

Mabrouk et al. (2010): Viability and resilience of a stylised biofilm pattern dynamics. G. Deffuant N. Gilbert (Eds.), Pattern Resilience, Springer, 2010. (forthcoming).

Language Competition case study (UIB - IFISC)

This case study had been already well advanced in the first period of the project. IFISC continued the work on pattern dynamics, with the following main stepd:

i) Study of the effects of social network communities on the dynamics of language competition

ii) Characterization of the role of prestige and volatility in language dynamics through simulations of the IBMs of language competition

iii) Derivation and analysis of macroscopic descriptions of language dynamics models in different complex networks.

We investigate the dynamics of the Abrams-Strogatz and Minett-Wang models for the competition between two languages in a bilingual society. In the first model, each individual can be in one of two possible states, either using language A or language B, while the second model incorporates an intermediate state AB representing individuals that use both languages. We analyze both models on fully-connected, complex networks and two-dimensional lattices by analytical and numerical methods. We find that, contrarily to what we were expecting, language coexistence is more difficult to achieve in the Minett-Wang model, where the presence of bilinguals in use facilitates the ultimate dominance of one of the two languages. Also, a stability analysis reveals that language coexistence is more unlikely to happen in poorly-connected than in fully connected networks, and that the dominance of only one language is enhanced as the connectivity decreases.

Moreover, IFISC collaborated with Cemagref on the study of the viability and resilience of two languages in competition represented by the Abram – Strogatz model. In particular, we computed the viability kernel and resilience indices of the system for the different dynamical regimes. In one case, there is a set of attractors located in the desired state set, in the other case, there is not. The viability kernel and resilience sets are very different. Moreover, they lead to policies of action on the prestige of the languages which are also different. This study led to a publication, and a chapter of the book is devoted to it (with more details than in the publication).

Toivonen, R.; Castelló, X.; Eguíluz, V.; Saramäki, J.; Kaski, K; San Miguel,M. Broad lifetime distributions for ordering dynamics in complex networks *Physical Review* E 79, 016109 (1-8) (2009)

Vazquez, F; Castello, X.; San Miguel, M.. 2010. Agent Based Model of Language Competition: Macroscopic descriptions and order-disorder transitions. Submitted to *J. Stat. Mech*.

Chapel, L, Castelló, X, Bernard, C, Deffuant, G, Eguíluz, V, Martin, S, San Miguel, M. Viability and Resilience of Languages in Competition. *Public Library of Science*. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2811194

Web 2.0 case studies (UniS)

UniS activity during the second period focused on the study of the dynamics of two classes of collaborative online communities: peer production systems (wikis) and social media sharing websites (Flickr groups). With respect to the initially planned activity, the second phase was substantially devoted to the development of tools to collect and analyse data on the temporal dynamics of these systems. Indeed, despite a widespread interest in online social networking services as a testbed for social network analysis, to date we lack substantial empirical evidence on the dynamics of *peer production systems* and *collaborative communities*, i.e. online systems (often built on top of social networking services) geared towards the decentralised production of content by large collectives of users. The availability of rich datasets on the evolution of such systems (at the micro-level of individual behaviour, the meso-level of community dynamics and the macro-level of global system dynamics) and their understanding is an essential condition for the development of realistic models of their dynamics. With one sole exception (the case of Wikipedia, which has been extensively studied in the literature but unfortunately represents an outlier of little relevance for the purpose of the present analysis), data about the evolution of an "ecosystem" of collaborative communities was not available before the initial phase of the project and therefore had to be obtained as part of this project. Work conducted in the second phase of the project can be summarised as follows:

- Pilot large-scale study of wiki dynamics. We analyzed the social dynamics of a large datasets of wikis and described the results in a pilot study, showing the effects of governance and demographic properties on the evolution in content and population of wiki-based communities.
- Extension of wiki data collection: WikiTracer. We designed and started implementing a system called Wikitracer[1] allowing the automatic extraction of data from a variety of distributed sources. Our pilot analysis of wiki dynamics demonstrated the empirical predictions that can be made about the evolution of wikis by analyzing their structure and social dynamics over time.
- Case study on growth drivers in social media sharing communities. We analysed several months worth of data extracted via Flickr Group Trackr, a tool that we designed and implemented in the first part of the project. Supplementary data were provided thanks to collaborations with partners external to the consortium (University of Rome la Sapienza and University of Paris 7). The results of this study show the major discrepancies in the dynamics of these systems when compared to peer production systems, and the effects of

[1] http://wikitracer.com

social- and affiliation-network dependent factors on the growth and population turnover of these communities.

- Extension of Flickr group data collection. In collaboration with IFISC we implemented a system to extend the Flickr data collection to collect frequent snapshots of individual social link creation and affiliation behaviour as a condition for a study in preparation between IFISC and UniS.
- Exploring a model of collaborative content production. In collaboration with IFISC we first analyze the properties of the groups formed at Flickr (e.g., distribution of the number of members per group) and second, we propose a model that captures the observed features of the group formation. We observe a lognormal distribution of group sizes for Flickr both in terms of number of photos and number of users. Thus, we were interested in finding a microscopic level model giving rise to lognormal distribution. We show that it can be achieved by means of word-of-mouth spreading processes, closely related to epidemic processes. Finally we analyze the role played by homophily, that is, the tendency to interact with individuals sharing attributes, in the time evolution of group affiliation, and found a large correlation between the network of members' contacts and the overlap in group affiliation. In collaboration with an external partner (University of Southampton), we devised and implemented a general model of collaborative content production with the goal of exploring a range of hypotheses on group-joining/group-leaving behaviour, validating them with empirical results discussed in the literature and with our own results obtained from the analysis of Flickr group dynamics.

Roth, C., Taraborelli, D., and Gilbert, N. (2008) Measuring wiki viability. An empirical assessment of the social dynamics of a large sample of wikis. In WikiSym '08: *Proceedings of the 4th International Symposium on Wikis* (New York, NY, USA, September 2008), ACM. http://nitens.org/docs/wikidyn.pdf

Roth, C., Taraborelli, D., and Gilbert, N. (2008) Démographie des communautés en ligne: Le cas des wikis. *Réseaux* 26, 152 (2008), 205–240. http://dx.doi.org/10.3166/Reseaux.152.205-240

Taraborelli, D., Roth, C., and Gilbert, N. (2008) Measuring wiki viability (II). Towards a standard framework for tracking content-based online communities. Tech. rep., 2008. http://wikitracer.com/refs/wikitrack.pdf

Baldassarri, A., Barrat, A., Cappocci, A., Halpin, H., Lehner, U., Ramasco, J., Robu, V., Taraborelli, D. (2008) The Berners-Lee Hypothesis: Power laws and Group Structure in Flickr. *Proc. of Dagstuhl Seminar*, November 2008, http://drops.dagstuhl.de/opus/volltexte/2008/1789/pdf/08391.SWM.Paper.1789.pdf

Taraborelli, D., Roth, C. (2010) Circles and ties: On the drivers of group dynamics in collaborative media sharing. *Unpublished Ms*, in preparation.

Taraborelli, D., Roth, C. (2010) *Viability of collaborative Web communities: Two case studies*. G. Deffuant N. Gilbert (Eds.), Pattern Resilience, Springer, 2010. (forthcoming)

Social dilemma (CNRS)

In this case study, we propose to explore the possibility to apply viability theory to this social dilemma and emergence of cooperation dynamics. The main idea is that there are several ways for an institution to act on the strength of a social dilemma (tax, subsidies, etc) but this action is

often costly. Thus the different measures that an institution can undertake can thus be viewed as control parameters that help the institution to maintain the social system within its domain of normal functioning. Since these measures are costly, the institution faces a viability problem with resilience episodes: how to maintain the social systems in a given domain and what are the optimal control to apply to restore the system when it leaves this domain. In a previous case study (Chavalarias 2006), we focused on a model of social dilemma on a network based on a sequential prisoner's dilemma game.

In the first phase of PATRES, we began to study this model in the viability framework, in collaboration with the Cemagref (Chavalarias & Chapel 2008). In the first setting, we considered only as state variable the strength of the social dilemma, the proportion of altruists, and the proportion of reciprocators. The control was defined as a modification of the strength of the social dilemma, chosen between a minimum and a maximum. We found that below a critical value of this maximum, the viability kernel is restricted to a very tiny strip of the state space. Above the critical value, the viability kernel is a large part of the desired set (more than 80%).

In the second phase of the project, we included the cost of the action in the model. To do this, we considered that the institution has a budget for this type of action at each time step, and we compute the cumulated remaining budget. We fix a constraint that this cumulated remaining budget cannot be below a threshold value. This completes the desired set, and we could study the resilience of the system in this setting.

Pattern dynamics in scientific literature (CNRS)

In this case study, more the 18 Million references of biomedical literature, taken from the database MedLine, were analysed. The aim was to identify different modes of science evolution and the viability constraints of different scientific fields. Methods of co-word and co-occurrence analysis were used to identify scientific fields and their evolution in the course of time. The questions then are how scientific fields emerge, evolve, differentiate, and die. To quantify the proximity between scientific fields, a new proximity measure was developed (Chavalarias & Cointet 2008), as well as a new measure of the consistency of scientific fields (Cointet & Chavalerias 2008). As a new tool for analyzing bibliometric datasets, the algorithm FieldBuilder was developed, that implements the new proximity and consistency measure. These approaches also allow to construct “phylogenies” of scientific fields, i.e. a field can give birth to one or more new branches, or children. A first application, research related to bias in biomedical research was analyzed (Chavalerias & Ioannidis, in press).

Chavalarias D. & Cointet J-P. (2008) Bottom-up scientific field detection for dynamical and hierarchical science mapping - methodology and case study *Scientometrics* Vol. 75 No. 1 , (DOI): 10.1007/s11192-007-1825-6.

Cointet J-P., Chavalarias D. (2008) Multi-level Science mapping with asymmetric co-occurrence analysis: Methodology and case study, *Networks and Heterogeneous Media*, Vol 3 Number 2, june 2008, p267-276

Chavalarias & Ioannidis (in press) Science mapping analysis characterizes 235 biases in biomedical research

Chavalarias & Cointet (submitted) The Reconstruction of Science Phylogeny

Deviations from the workprogramme

The main deviations from the workprogramme are related to the case studies for which the whole approach including pattern dynamics and viability – resilience computation couldn't be completed in the time frame of the project (web 2.0 and scientific patterns). For both of them, the reason is that the dynamic modeling of these systems was not as mature as for the others. Hence a longer period of time had to be devoted to this modeling work.

Deliverables List

Del. no.	Deliverable name	WP no.	Date due	Actual/ Forecast date	Lead contract.
D1.1	A report on the different versions of the models for each case study, and the results of pattern resilience computation	1	31/01/2010	31/01/2010	UFZ
D1.2	Computer programs of the models, pattern dynamics and resilience	1	31/01/2010	31/01/2010	UFZ

Milestones and expected result

Milest. no.	Milestone name	WP no.	Date due	Actual/ Forecast date	Lead contract.
M1.1	Result of the test of the pilot tools and methods on the simplified dynamics	1	01/02/2008	30/06/2008 except web	UFZ
M1.1	Refined set of core models	1	01/08/2008	01/12/2008	UFZ

Workpackage 2: Pattern dynamics

Objectives

In this second part of the project, the objective of the workpackage was:

- Improve progressively the methods and tools, taking into account their performance on the refined models of the case studies.

Progress toward objectives

In this development phase, we firstly refined the methods identified in the pilot phase for computing synthetic dynamics from individual based models in 3 directions:

- From microscopic to macroscopic dynamics for systems with two symmetric absorbing states (Vasquez et al. 2009). The dynamics of agent based models with two symmetric absorbing states (configurations from which the system cannot escape) is very relevant to the so-called voter models, in which the state of a particle in a site evolves according to the density of states in its near neighborhood. This type of dynamics has many applications in different disciplines such as opinion formation, language spreading, genetics, and chemical kinetics. We have proposed a general approach to study these

systems, so that starting from the micro-dynamics we have derived a stochastic equation for the time evolution of the particle-density field that successfully explains many of their macroscopic properties. In particular, we show that the macroscopic behaviour only depends on the first derivatives of the microscopic transition rates.

- Competitive interactions (Hernandez-Garcia et al. 2009, Poglietti et al. 2010). In a set of papers we have studied the macroscopic properties of entities competing for a common finite set of resources. A simple model can be derived for the dynamics of the populations in competition where the crucial information is kept in the competitive interaction function. The form of this competition function determines the equilibrium distribution of entities: clustered, homogenously distributed and/or with exclusion zones.

- Master equation description of noisy dynamics (Pineda et al. 2009). We study the Deffuant et al. model for continuous--opinion dynamics under the influence of noise. In the original version of this model, individuals meet in random pairwise encounters after which they compromise or not depending of a confidence parameter. Free will is introduced in the form of noisy perturbations: individuals are given the opportunity to change their opinion, with a given probability, to a randomly selected opinion inside the whole opinion space. We derive the master equation of this process. One of the main effects of noise is to induce an order-disorder transition. In the disordered state the opinion distribution tends to be uniform, while for the ordered state a set of well defined opinion clusters are formed, although with some opinion spread inside them. Using a linear stability analysis we can derive approximate conditions for the transition between opinion clusters and the disordered state. The master equation analysis is compared with direct Monte-Carlo simulations. We find that the master equation and the Monte-Carlo simulations do not always agree due to finite-size induced fluctuations that we analyze in some detail.

Secondly, Cemagref improved the software tool SimExplorer. This tool aims at facilitating programming and executing experimental designs on simulations. The software is particularly relevant when testing a complex model to identify its relevant patterns of behaviour when parameters of initial conditions are modified. The developed version of the model (which is also financed by another project), includes a client server information system for exchanging methods and results, but also for ensuring the traceability of the experiments. The main R packages for experimental designs, and data treatments are available from the software, together with an evolving library system. Moreover, the software provides facilities to use cluster and grid computing. In this development phase, we improved:

- the user interface for editing the experimental designs,
- the integration of the information system,
- the facilities for launching experiments on clusters and computer grids.

The development phase of the project is well underway with application of these methods to different case studies:

- Language competition: We applied the techniques of derivation of macroscopic equations for the IBM of language competition as mentioned in WP1.

- Bacteria dynamics: In this development phase we applied the moment approximation approach to get synthetic representations of the spatial dynamics using correlation function. Moreover, we used SimExplorer to explore the IBM and to relate the patterns to the values of the parameters.
- Savanna case study: In the refinement work to relate the abstract model from the Jeltsch savanna model, it appeared finally that the second moment terms of the equations modify significantly the dynamics. Moreover, we used SimExplorer to find different patterns in the Jeltsch model, patterns used then to simplify the dynamics.

Vazquez, Federico; López, Cristobal, Systems with two symmetric absorbing states: relating the microscopic dynamics with the macroscopic behaviour. Physical Review E, 78, 061127 (1-5) (2008)

Hernández-García, Emilio; López, Cristóbal; Pigolotti, Simone; Andersen, Ken H. Species competition: coexistence, exclusion and clustering. Philosophical Transactions of the Royal Society A, 367, 3183-3195 (2009)

Pigolotti, Simone; López, Cristóbal; Hernández-García, Emilio; Andersen, Ken H. How Gaussian competition leads to lumpy or uniform species distributions. Theoretical Ecology, in press, (2010)

M. Pineda, R. Toral and E. Hernandez-Garcia Noisy continuous-opinion dynamics . Journal of Statistical Mechanics: Theory and Experiment, P08001, (1-18) (2009)

Deviations from the workprogramme

None.

Deliverables List

Del. no.	Deliverable name	WP no.	Date due	Actual/ Forecast date	Lead contract.
D2.1	Software prototype of SimExplorer, including facilities for model simulation experimental designs, and pattern detection	2	31/01/2010	31/01/2010	Cemagref
D2.2	Report on methods and tools to detect and define pattern dynamics	2	31/01/2010	31/01/2010	IFISC

Milestones List

Milest. no.	Milestone name	WP no.	Date due	Actual/ Forecast date	Lead contract.
M2.1	Pilot set of methods	2	01/08/2007	30/06/2007	IFISC
M2.1	SimExplorer pilot version	2	01/08/2007	01/06/2008	Cemagref

Workpackage 3: Resilience and Viability

Objectives

In the development phase, the objectives were:

1. To specify and implement an improved version of the methods and tools, building on the results of the pilot tests on the case studies,
2. Improve progressively the methods and tools, taking into account their performance on the refined models of the case studies.

Progress toward objectives

The development phase reached its objectives.

Definition of resilience

The improved definition of the resilience proposed in the pilot phase has been presented in the actions of dissemination of the project. The feed-back from these presentations helped us to improve again the formulation of the definition. Moreover, we used an example of savanna model from the literature on which the more standard mathematical definition of resilience was applied in order to show the main differences with our view. We believe that this clarified the presentation, and we adopted it in the corresponding paper of the book.

KAVIAR

Cemagref issued a new version of the software tool for computing viability kernels and resilience indices. The software also computes viable or resilient action policies. The main improvement of the software developed during the period are:

- We improved the procedure to include a new model into the software, it is now easier.
- We included several possibilities for computing the optimal action at each time steps (conjugate gradient, Newton method, exhaustive search).
- We included a new method for computing the parameters of the SVM, in order to avoid possible trouble due to the geometry of the problem.

Moreover, we already began the development phase and worked on some improvements of the method.

Robustness of control

We investigated some methods to define the robustness of the viable control. The main idea is to keep the system far from the boundaries of the viability kernel. This led us to develop efficient methods for computing the distance from any point of the kernel to this boundary. Then, we developed methods for choosing the actions that maximise this distance. This work is the subject of a paper of the book.

Deviations from the workprogramme

None.

Deliverables List

Del. no.	Deliverable name	WP no.	Date due	Actual/ Forecast date	Lead contract.
D3.1	Final version of the prototype software tool for computing pattern resilient action policies on complex systems	3	31/01/2010	31/01/2010	Cemagref
D3.2	Report on the methods for computing pattern resilient action policies	3	31/01/2010	31/01/2010	CNRS

Milestones List

Milest. no.	Milestone name	WP no.	Date due	Actual/ Forecast date	Lead contract.
M3.1	Pilot version of the prototype software for computing pattern resilient action policies.	3	01/08/2007	31/01/2008	Cemagref

Section 3 – Consortium management

Consortium management tasks and achievements

The main body dealing with the project management is the management committee which meets at each plenary project meeting. During the reporting period, the management committee held two meetings. No particular consortium management problem occurred during the period.

Contractors

We had no major change of contractor responsibilities. The contributions of contractors globally correspond to their commitments.

Project timetable and status

The consortium held two plenary meetings, two meetings involving 2 contractors for more specific collaborations, and two meetings for dissemination purposes (workshop and course).

The time table is globally respected, except for one case studies where we did not manage to complete de development phase, for different reasons (see deviations from the work programme).

Nevertheless, the major main objectives of the project are achieved:

- We have developed two functional prototypes of software tools: one for exploring models and detecting dynamical patterns, the other for computing viability kernels and resilience indices. These software prototypes are downloadable from the internet.
- We have tested these methods on 5 case studies, leading the production of published new knowledge in 4 of them.
- We developed new methods for computing viability and resilience indices, and applied methods of physics (particularly moment approximation) to assess pattern dynamics, leading also to specific publications.
- We carried out dissemination actions towards different audience:
    - satellite workshop of the European Conference on Complex Systems,
    - tutorial workshop organised for PhD students and researchers,
    - Hand-book based on the work of the project, to be published soon by Springer Verlag.

Project barchart and status.

Acronym: PATRES

Contract n°: 043268

	02/07-07/07			08/07-01/08			02/08-07/08			08/08-01/09			02/09-07/09			08/09-01/10		
WP1: Case studies																		
Bacteria test pilot tools																		
Bacteria pilot dynamics																		
Bacteria refinement																		
Bacteria test refined tools																		
Savanna pilot dynamics																		
Savanna test pilot tools																		
Savanna refinement																		
Savanna test refined tools																		
Languages pilot dynamics																		
Languages test pilot tools																		
Languages refinement																		
Languages test refined tools																		
Web 2.0 pilot dynamics																		
Web 2.0 test pilot tools																		
Web 2.0 refinement																		
Web 2.0 test refined tools																		
Social pilot dynamics																		
Social test pilot tools																		
Social refinement																		
Social test refined tools																		
WP2: Pattern dynamics																		
KL & Other pilot methods																		
KL & Other developt																		
KL & Other refinement																		
SimExplorer pilot																		
SimExplorer developt																		
SimExplorer refinement																		
WP3: Resilience																		
Resilience pilot methods																		
Resilience develpt																		
Resilience refinement																		
KAVIAR pilot																		
KAVIAR developt																		
KAVIAR refinement																		
WP4: Dissemination																		
Web site																		
Book																		
User meetings																		

Section 4 – Other issues

Annex – Plan for using and disseminating the knowledge

Exploitable knowledge and its Use

No economically exploitable knowledge was identified at this stage.

Dissemination of knowledge

The following overview table lists all the actions of dissemination from the projects. It includes mainly scientific conferences, seminar and publications.

During the project, we carried out the following dissemination of knowledge actions:

- We presented the project during a workshop in the European Conference on Complex Systems in Warwick. It attracted some of the participants of the conference, and we invited American partners in a panel.
- We organised a more detailed tutorial on the project in Madeira, which attracted 11 students. The course alternated theoretical presentations with practical exercises using the software tools developed in the project.

Moreover, the main actions for dissemination of knowledge are:

- The project website has been running since the start of the project and is gradually expanding to cover the results of our work. It includes the main publications and the links to the software sites.
- The websites with the software tools: SimExplorer and KAVIAR, from where the software and documentation are downloadable.
- The book, to be published in Springer, presents the methods and the results of the project on the case studies (it one of the deliverables).

Planned/actual Dates	**Type**	**Name**	**Type of audience**	**Countries addressed**	**Size of audience**	**Partner responsible /involved**
10.09.2008	conference	ACM WikiSym 2008 (Porto, Portugal)	academic	Int'l	100-200	UniS
26.09.2008	seminar	Dagstuhl Seminar "Social Web Communities" (Dagstuhl, Germany)	academic	Int'l	50-70	UniS
01.10.2008	seminar	London Wiki Wednesdays, British Computer Society	corporate / IT	UK	50-70	UniS
24.02.2009	seminar	Nature Network London	academic / IT	UK	30-50	UniS
12.06.2009	seminar	Centre for Human Computer Interaction Design, City University, London	academic	UK	15-20	UniS
06.10.2009	workshop	Mobile A2K, Rockefeller Foundation (Bellagio, Italy)	academic / NGO	Int'l	50-70	UniS
28.01.2010	workshop	QualityCommons, Sorbonne University (Paris)	academic	Int'l	50-60	UniS
August 18 – 23, 2008	presentation in scientific conference	Discussion Panel. Workshop on Challenges and Visions in the Social Sciences. Zurich (Suitwerland)	Research	International		UIB-IFISC/Maxi San Miguel
August 18 – 23, 2008	presentation in scientific conference	Group Formation. Workshop Challenges and Visions in the Social Sciences. Zurich (Suitwerland)	Research	International		UIB-IFISC/Maxi San Miguel
December 10 – 12, 2008	presentation in scientific conference	Broad lifetime distributions for ordering dynamics in complex networks. BCNetWORKSHOP, Barcelona (Spain)	Research	International		UIB-IFISC/Xavier Castelló
December 1 – 5, 2008	presentation in scientific conference	From microscopic to macroscopic dynamics in systems with two symmetric absorbing states.	Research	International		UIB-IFISC/Federico Vázquez

Planned/actual Dates	**Type**	**Name**	**Type of audience**	**Countries addressed**	**Size of audience**	**Partner responsible /involved**
		MEDYFINOL08, Punta del Este (Uruguay)				
October 21, 2009	presentation in scientific conference	The continuum PATRES tutorial Workshop of agent-based models (part I). PATRES tutorial, Madeira (Portugal)	Research	International		UIB-IFISC/Cristóbal López
October 21, 2009	presentation in scientific conference	The continuum PATRES tutorial Workshop of agent-based models (part II).PATRES tutorial, Madeira (Portugal)	Research	International		UIB-IFISC/Federico Vázquez
November 23 – 27, 2009	presentation in scientific conference	Dynamics of language competition. 150 years after Darwin, Palma de Mallorca (Spain)	Research	International		UIB-IFISC/Maxi San Miguel
August 24, 2009	presentation in scientific conference	Dynamics of language competition: bilingualism and social structure effects. Economic sociology working group, MIT (USA)	Research	International		UIB-IFISC/Xavier Castelló
September 9 – 12, 2009	presentation in scientific conference	Local effects of global languages: evidence from language modelling. SLE, Lisbon (Portugal)	Research	International		UIB-IFISC/Lucía Loureiro

Planned/actual Dates	**Type**	**Name**	**Type of audience**	**Countries addressed**	**Size of audience**	**Partner responsible /involved**
September 10 – 12, 2009	presentation in scientific conference	The savanna problem from a statistical physics point of view. FISES2009, Huelva (Spain)	Research	International		UIB-IFISC/Federico Vázquez
September 21 – 25, 2009	presentation in scientific conference	Viability and resilience in the dynamics of language competition. ECCS '09, Warwick (UK)	Research	International		UIB-IFISC/Xavier Castelló
September 21 – 25, 2009	presentation in scientific conference	The macroscopic description of agent-based models. ECCS '09, Warwick (UK)	Research	International		UIB-IFISC/Federico Vázquez
May 27 – 30, 2009	presentation in scientific conference	Model of group formation for Flickr online community. COST action MP0801, Rome (Italy)	Research	International		UIB-IFISC/Przemyslaw Grabowicz
September 10 – 12, 2009	presentation in scientific conference	From microscopic to macroscopic dynamics in systems with two symmetric absorbing states. FISES2009, Huelva (Spain)	Research	International		UIB-IFISC/Federico Vázquez
September 10 – 12, 2009	presentation in scientific conference	Savanna-Fire Model: Combined effects of tree-tree establishment competition and spatially explicit fire on the spatial pattern of trees in savannas. FISES2009, Huelva (Spain)	Research	International		UIB-IFISC/Flora S. Bacelar
November 23, 2009	presentation in scientific conference	Mesoscale structure effects in language competition: topological traps and broad lifetime distributions. 150 Years After Darwin, Palma de Mallorca (Spain)	Research	International		UIB-IFISC/Xavier Castelló
November	presentation in	Bilingualism, language death and language birth	Research	International		UIB-IFISC/Lucía

Planned/actual Dates	Type	Name	Type of audience	Countries addressed	Size of audience	Partner responsible /involved
23, 2009	scientific conference	in language contact situations.150 Years After Darwin, Palma de Mallorca (Spain)				Loureiro
17/11/08	Seminar	University of Potsdam	Research	Germany	30	UFZ
23/03/09	Campus Talk	CEMAGREF	Research	France	80	UFZ
24/04/09	Workshop Talk	Editorial Board meeting of GAIA	Research	Europe	50	UFZ
24/09/09	Conference Talk	European Conference on Complex Systems	Research	International	120	UFZ
24/09/09	Conference Talk	European Conference on Complex Systems	Research	International	120	UFZ
21/10/09	Tutorial Workshop	PATRES Tutorial Workshop	Research	International	30	UFZ
21/10/09	Tutorial Workshop	PATRES Tutorial Workshop	Research	International	30	UFZ
22/10/2008	Workshop	Mathematical modeling for microbial ecology. INRA, La Grande-Motte	Research	France	30	Cemagref
6/07/2009	Workshop	Introductory lectures on Aspects of Complexity	Research	Europe	50	Cemagref
8/10/2009	Workshop	IXXI seminar	Research	France	30	Cemagref
23/08/2009	Conference	ICCGI 2009	Research	International	200	Cemagref
24/09/09	Conference Talk	European Conference on Complex Systems	Research	International	120	Cemagref
21/10/09	Tutorial Workshop	PATRES Tutorial Workshop	Research	International	30	Cemagref

Publishable results

The consortium has no economically exploitable result to publish.