RESEARCH CENTER

FIELD
Computational Sciences for Biology, Medicine and the Environment

Activity Report 2012

Section Application Domains

Edition: 2013-04-24

ABS Project-Team (section vide)

AMIB Project-Team (section vide)

BAMBOO Project-Team

4. Application Domains

4.1. Biology with a focus on symbiosis

The main area of application of BAMBOO is biology, with a special focus on symbiosis (ERC project) and on intracellular interactions.

BEAGLE Team

4. Application Domains

4.1. Application Domains

- Molecular and cellular biology
- Genome evolution

BONSAI Project-Team

4. Application Domains

4.1. Sequence processing for Next Generation Sequencing

As said in the introduction of this document, biological sequence analysis is a foundation subject for the team. In the last years, sequencing techniques experienced remarkable advances with NGS, that allows for fast and low-cost acquisition of huge amounts of sequence data, and outperforms conventional sequencing methods. These technologies can apply to genomics, with DNA sequencing, as well as to transcriptomics, with RNA sequencing allowing to gene expression analysis. They promise to address a broad range of applications including: Comparative genomics, individual genomics, high-throughput SNP detection, identifying small RNAs, identifying mutant genes in disease pathways, profiling transcriptomes for organisms where little information is available, researching lowly expressed genes, studying the biodiversity in metagenomics. From a computational point of view, NGS gives rise to new problems and gives new insight on old problems by revisiting them: Accurate and efficient remapping, pre-assembling, fast and accurate search of non exact but quality labelled reads, functional annotation of reads, ...

4.2. Noncoding RNA

Our expertise in sequence analysis also applies to noncodingRNA analysis. Noncoding RNA genes play a key role in many cellular processes. First examples were given by microRNAs (miRNAs) that were initially found to regulate development in *C. elegans*, or small nucleolar RNAs (snoRNAs) that guide chemical modifications of other RNAs in mammals. Hundreds of miRNAs are estimated to be present in the human genome, and computational analysis suggests that more than 20% of human genes are regulated by miRNAs. To go further in this direction, the 2007 ENCODE Pilot Project provides convincing evidence that the Human genome is pervasively transcribed, and that a large part of this transcriptional output does not appear to encode proteins. All those observations open a universe of “RNA dark matter” that must be explored. From a combinatorial point of view, noncoding RNAs are complex objects. They are single stranded nucleic acids sequences that can fold forming long-range base pairings.This implies that RNA structures are usually modelled by complex combinatorial objects, such as ordered labeled trees, graphs or arc-annotated sequences.

4.3. Genome structures

Our third application domain is concerned with the structural organization of genomes. Genome rearrangements are able to change genome architecture by modifying the order of genes or genomic fragments. The first studies were based onto linkage maps and mathematical models appeared fifteen years ago. But the usage of computational tools was still limited because of lack of data. The increasing availability of complete and partial genomes now offers an unprecedented opportunity to analyse genome rearrangements in a systematic way and gives rise to a wide spectrum of problems: Taking into account several kinds of evolutionary events, looking for evolutionary paths conserving common structure of genomes, dealing with duplicated content, being able to analyse large sets of genomes even at the intraspecific level, computing ancestral genomes and paths transforming these genomes into several descendant genomes.

4.4. Nonribosomal peptides

Lastly, the team has been developing for several year a tight collaboration with Probiogem lab on nonribosomal peptides, and has became a leader on that topic. Nonribosomal peptide synthesis produces small peptides not going through the central dogma. As the name suggests, this synthesis uses neither messenger RNA nor ribosome but huge enzymatic complexes called nonribosomal peptide synthetases (NRPSs). This alternative pathway is found typically in bacteria and fungi. It has been described for the first time in the 70’s [21]. For the last decade, the interest in nonribosomal peptides and their synthetases has considerably increased, as witnessed by the growing number of publications in this field. These peptides are or can be used in many biotechnological and pharmaceutical applications (e.g. anti-tumors, antibiotics, immuno-modulators).

DYLISS Team

4. Application Domains

4.1. Formal models in molecular biology

As mentioned before, our main goal in biology is to characterize groups of genetic actors that control the response of living species capable of facing extreme environments. To focus our developments, applications and collaborations, we have identified three biological questions which deserve integrative studies. Each axis may be considered independently from the others although their combination, a mid-term challenge, will have the best impact in practice towards the long-term perspective of identifying proteins controlling the production of a metabolite of industrial interest. It is illustrated in our presentation for a major algae product: polyunsaturated fatty acids (PUFAs) and their derivatives.

4.2. Biological data integration

Axis 1 (data integration) aims at identifying **who** is involved in the specific response of a biological system to an environmental stress. Targeted actors will mainly consist in groups of genetic products or biological pathways. For instance, which pathways are implied in the specific production of PUFAs in brown algae? The main work is to represent in a system of logical constraints the full knowledge at hand concerning the genetic or metabolic actors, the available observations and the effects of the system dynamics. To this aim, we focus on the use of Answer Set Programming as we are experienced in modeling with this paradigm and we have a strong partnership with a computer science team leader in the development of dedicated grounders and solvers (Potsdam university).

4.3. Asymptotic dynamics of a biological system

Once a model is built and its main actors are identified, the next step is to clarify **how** they combine to control the system (**Axis 2**). Roughly, the fine tuning of the system response may be of two types. Either it results from the discrete combinatorics of the actors, as the result of a genetic adaptation to extreme environmental conditions or the difference between species is rather at the enzyme-efficiency level. For instance, if Pufa's are found to be produced using a set of pathways specific to brown algae, the work in axis 2 will consist to apply constraint-based combinatorial approaches to select consistent combinations of pathways controlling the metabolite production. Otherwise, if enzymes controlling the production of Pufa's are found to be expressed in other algaes, it suggests that the response of the system is rather governed by a fine quantitative tuning of pathways. In this case, we use symbolic dynamics and average-case analysis of algorithms to weight the respective importance of interactions in observed phenotypes (see Fig. 2). This specific approach is motivated by the quite restricted spectrum of available physiological observations over the asymptotic dynamics of the biological system.

4.4. Biological sequence annotation

In order to check the accuracy of in-silico predictions, a last step (**Axis 3**) is to extract genetic actors responsible of biological pathways of interest in the targeted organism and locate them in the genome. In our guiding example, active proteins implied in Pufa's controlling pathways have to be precisely identified. Actors structures are represented by syntactic models (see figure 4). We use knowledge-based induction on far instances for the recognition of new members of a given sequence family within non-model genomes (see figure 3). A main objective is to model enzyme specificity with highly expressive syntactic structures - context-free model - in order to take into account constraints imposed by local domains or long-distance interactions within a protein sequence.

GENSCALE Team

4. Application Domains

4.1. Sequence comparison

Historically, sequence comparison has been one of the most important topics in bioinformatics. BLAST is a famous software tool particularly designed for solving problems related to sequence comparisons. Initially conceived to perform searches in databases, it has mostly been used as a general-purpose sequence comparison tool. Nowadays, together with the inflation of genomic data, other software comparison tools that are able to provide better quality solutions (w.r.t the ones provided by BLAST) have been developed. They generally target specific comparison demands, such as read mapping, bank-to-bank comparison, meta-genomic sample analysis, etc. Today, sequence comparison algorithms must clearly be revisited to scale up with the very large number of sequence objects that new NGS problems have to handle.

4.2. Genome comparison

This application domain aims at providing a global relationship between genomes. The problem lies in the different structures that genomes can have: segments of genome can be rearranged, duplicated or deleted (the alignment can no longer be done in one piece). Therefore one major aim is the study of chromosomal rearrangements, breaking points, structural variation between individuals of the same species, etc. However, even analyses focused on smaller variations such as Single Nucleotide Polymorphisms (SNP) at the whole genome scale are different from the sequence comparison problem, since one needs first to identify common (orthologous) parts between whole genome sequences and thus obtain this global relationship (or map) between genomes. New challenges in genome comparison are emerging with the evolution of sequencing techniques. Nowadays, they allow for comparing genomes at intra-species level, and to deal simultaneously with hundreds or thousands of complete genomes. New methods are needed to find the sequence and structural variants between such a large number of non-assembled genomes. Even for the comparison of more distant species, classical methods must be revisited to deal with the increasing number of genomes but more importantly their decreasing quality: genomes are no longer fully assembled nor annotated.

4.3. Protein comparison

Comparing protein is important for understanding their evolutionary relationships and for predicting their structures and their functions. While annotating functions for new proteins, such as those solved in structural genomics projects, protein structural alignment methods may be able to identify functionally related proteins when the sequence identity between a given query protein and the related proteins are low (i.e. lower than 20%). Moreover, protein comparison allows for solving the so-called protein family identification problem. Given an unclassified protein structure (query), the comparison of protein structures can be used for assigning a score measuring the "similarity" between the query and the proteins belonging to a set of families. Based on this score, the query is assigned to one of the families of the set. The knowledge acquired by performing such analyses can then be exploited in methods for protein structure prediction that are based on a homology modeling approach.

IBIS Project-Team (section vide)

MAGNOME Project-Team

4. Application Domains

4.1. Function and history of yeast genomes

Yeasts provide an ideal subject matter for the study of eukaryotic microorganisms. From an experimental standpoint, the yeast *Saccharomyces cerevisiae* is a model organism amenable to laboratory use and very widely exploited, resulting in an astonishing array of experimental results. From a genomic standpoint, yeasts from the hemiascomycete class provide a unique tool for studying eukaryotic genome evolution on a large scale. With their relatively small and compact genomes, yeasts offer a unique opportunity to explore eukaryotic genome evolution by comparative analysis of several species.

- Yeasts are widely used as cell factories, for the production of beer, wine and bread and more recently of various metabolic products such as vitamins, ethanol, citric acid, lipids, etc.
- Yeasts can assimilate hydrocarbons (genera *Candida*, *Yarrowia* and *Debaryomyces*), depolymerise tannin extracts (*Zygosaccharomyces rouxii*) and produce hormones and vaccines in industrial quantities through heterologous gene expression.
- Several yeast species are pathogenic for humans, especially *Candida albicans*, *Candida glabrata*, *Candida tropicalis* and the Basidiomycete *Cryptococcus neoformans*.

The hemiascomycetous yeasts represent a homogeneous phylogenetic group of eukaryotes with a relatively large diversity at the physiological and ecological levels. Comparative genomic studies within this group have proved very informative [33], [48], [47], [35], [50], [30], [31], [6].

MAGNOME applies its methods for comparative genomics and knowledge engineering to the yeasts through the ten-year old *Génolevures* program (GDR 2354 CNRS), devoted to large-scale comparisons of yeast genomes with the aim of addressing basic questions of molecular evolution. We developed the software tools used by the CNRS's `genolevures.org` web site. For example, MAGNOME's MAGUS system for simultaneous genome annotation combines semi-supervised classification and rule-based inference in a collaborative web-based system that explicitly uses comparative genomics to simultaneously analyse groups of related genomes.

4.2. Alternative fuels and bioconversion

Oleaginous yeasts are capable of synthesizing lipids from different substrates other than glucose, and current research is attempting to understand this conversions with the goal of optimizing their throughput, production and quality. From a genomic standpoint the objective is to characterize genes involved in the biosynthesis of precursor molecules which will be transformed into fuels, which are thus not derived from petroleum. Biological experimentation by partner laboratories study lipid accumulation the oleaginous yeasts such as *Yarrowia lipolytica* starting from:

- pentoses, produced from lignin cellulose agricultural substrates following a biorefining strategy,
- glycerol, a secondary output of chemical production of biodiesel, and
- industrial residues.

Lipases from *Y. lipolytica* are of particular interest (see [38] for review). Experimental characterization of the lipid bodies produced from these substrates will aid in the identification of target genes which may serve for genetic engineering. This in turn requires the development of molecular tools for this class of yeasts with strong industrial potential. MAGNOME's focus is in acquiring genome sequences, predicting genes using models learned from genome comparison and sequencing of cDNA transcripts, and comparative annotation. Our overall goal is to define dynamic models that can be used to predict the behavior of modified strains and thus drive selection and genetic engineering.

4.3. Winemaking and improved strain selection

Yeasts and bacteria are essential for the winemaking process, and selection of strains based both on their efficiency and on the influence on the quality of wine is a subject of significant effort in the Aquitaine region. Unlike the species studied above, yeast and bacterial starters for winemaking cannot be genetically modified. In order to propose improved and more specialized starters, industrial producers use breeding and selection strategies.

Yeast starters from the *Saccharomyces* genus are used for primary, alcohol fermentation. Recent advances have made it possible to identify the genetic causes of the different technological differences between strains [55], [54], [53]. Manipulating the genetic causes rather than the industrial consequences is far more amenable to experimental development. An essential tool in identifying these genetic causes is comparative genomics.

Bacterial starters based on *Oenococcus oeni* are used in secondary, malolactic fermentation. Genetically, *O. oeni* presents a surprising level of intra-specific diversity, and clues that it may evolve more rapidly than expected. Studying the diversity of the emphO. oeni genomes has led to genetic tools that can be used to evaluate the predisposition of different strains to respond to oenological stresses. While identifying particular genes has been the leading strategy up to now, recently a new strategy based on comparative genomics has been undertaken to understand the impact and mechanisms of genetic diversity [20], [23], [21], [32], [36], [28] [3].

Starting from historical collaborations by Pascal Durrens and Elisabeth Bon with partners from the Institute for Wine and Vine Sciences in Bordeaux (ISVV), and local industry, and in the framework of an effective partnership, we apply our tools to large-scale comparative genomics of yeast and bacterial starters in winemaking.

4.4. Knowledge bases for molecular tools

Affinity binders are molecular tools for recognizing protein targets, that play a fundamental in proteomics and clinical diagnostics. Large catalogs of binders from competing technologies (antibodies, DNA/RNA aptamers, artificial scaffolds, etc.) and Europe has set itself the ambitious goal of establishing a comprehensive, characterized and standardized collection of specific binders directed against all individual human proteins, including variant forms and modifications. Despite the central importance of binders, they presently cover only a very small fraction of the proteome, and even though there are many antibodies against some targets (for example, >900 antibodies against p53), there are none against the vast majority of proteins. Moreover, widely accepted standards for binder characterization are virtually nonexistent.

Alongside the technical challenges in producing a comprehensive binder resource are significant logistical challenges, related to the variety of producers and the lack of reliable quality control mechanisms. As part of the ProteomeBinders and Affinomics projects, MAGNOME works to develop knowledge engineering techniques for storing, exploring, and exchanging experimental data used in affinity binder characterization. This work involves databases and tools for molecular interaction data [42] [45], standards for data exchange between peers [44], [49], [41] and reporting standards [4] [58].

MORPHEME Team (section vide)

SERPICO Team

4. Application Domains

4.1. Biological pilot models: Birbeck granule and Melanosome biogenesis

In the past recent years, research carried at UMR 144 CNRS Institut Curie contributed to a better understanding of the intracellular compartimentation of specialized model cells such as melanocytes and Langerhans cells, the components and structural events involved in the biogenesis of their specialized organelles: melanosomes and Birbeck granules, respectively. These studies have started to highlight:

- multiple sorting and structural events involved in the biogenesis of these organelles;
- complexity of the endo-melanosomal network of these highly specialized cells;
- complex molecular architecture organizing and coordinating their dynamics;
- intracellular transport steps affected in genetic diseases, among which the Hermansky Pudlak syndrome (HPS) or involved in viral infection (HIV and Langerin in Langerhans cells).

In this context, the central aim of SERPICO is to understand how the different machineries of molecular components involved are interconnected and coordinated to generate such specialized structures. We need to address the following topics:

1. developing new bioimaging approaches to observe and statistically analyze such coordinated dynamics in live material;
2. correlating this statistically relevant spatiotemporal organization of protein networks with the biological architectures and within the overall biological environment as seen at the ultrastructural level;
3. modeling intracellular transport of those reference biological complex systems and proposing new experimental plans in an iterative and virtuous circle;
4. managing and analyzing the workflow of image data obtained along different multidimensional microscopy modalities.

These studies are essential to unravel the complexity of the endomembrane system and how different machineries evolve together (e.g. see Fig. 1). They help to control cell organization and function at different scales through an integrative workflow of methodological and technological developments.

At long term, these studies will shed light on the cellular and molecular mechanisms underlying antigen presentation, viral infection or defense mechanisms, skin pigmentation, the pathogenesis of hereditary genetic disorders (lysosomal diseases, immune disorders) and on the mechanisms underlying cell transformation. Our methodological goal is also to link in a single workflow, dynamics information obtained through diffraction limited light microscopy, eventually at a time regime compatible with live cell imaging. The overview of ultrastructural organization will be achieved by complementary electron microscopical methods. Image visualization and quantitative analysis are of course important and essential issues in this context.

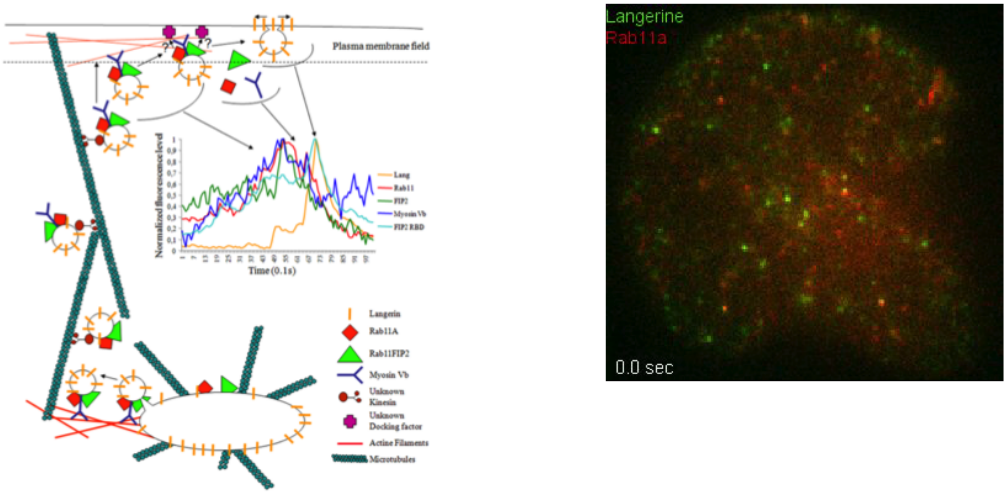

Figure 1. Rab11/Langerin (TIRF) and Birbeck granules: from interactions and trafficking to membrane biogenesis.

ASCLEPIOS Project-Team (section vide)

ATHENA Project-Team

4. Application Domains

4.1. Applications of Diffusion MRI

Various examples of CNS diseases as Alzheimer's and Parkinson's diseases and others like multiple sclerosis, traumatic brain injury and schizophrenia have characteristic abnormalities in the micro-structure of brain tissues that are not apparent and cannot be revealed reliably by standard imaging techniques. Diffusion MRI can make visible these co-lateral damages to the fibers of the CNS white matter that connect different brain regions. This is why in our research, Diffusion MRI is the major anatomical imaging modality that will be considered to recover the CNS connectivity.

Clinical domain: Diagnosis of neurological disorder

- *Parkinson's and Alzheimer's diseases* are among the most important CNS diseases. Six million patients (among which 850.000 in France) are suffering from Alzheimer's, making it the most important neurodegenerative disease in Europe. Over 85 years of age, 1 woman in 4 and 1 man in 5 are affected in Europe. In France, the number of Alzheimer's patients is expected to reach at least 2 million in 2025 and will probably double in 2050, with the increasing age of the population. Parkinson's disease is the second most important neurodegenerative disease. There are six and a half million patients in the world and roughly 150.000 patients in France, among which 10% are under 40 and 50% over 58. Together with our partners from NeuroSpin (Saclay), Inserm U678 and CENIR (CHUPS, Paris), we are involved in the ANR project NucleiPark which is about high field MRI of the braintem, the deep nuclei and their connections in the Parkinsonian syndromes.
- *Spinal Cord Injury* (SCI) has a significant impact on the quality of life since it can lead to motor deficits (paralysis) and sensory deficits. In the world, about 2.5 million people live with SCI (http://www.campaignforcure.org). To date, there is no consensus for full rehabilitative cure in SCI, although many therapeutic approaches have shown benefits [69], [71]. It is thus of great importance to develop tools that will improve the characterization of spinal lesions as well as the integrity of remaining spinal tracts to eventually establish better prognosis after spinal injury. We have already started to be active in this domain with our collaborators at Inserm U678 (H. Benali) and CRSN/Faculté de médecine Université de Montréal (Pr. S. Rossignol).

4.2. Applications of M/EEG

Applications of EEG and MEG cover:

- **Clinical domain:** diagnosis of neurological disorders such as
  - Diagnosis of neurological disorders such as epilepsy, schizophrenia, tinnitus, ...
  - Presurgical planning of brain surgery.
- **Cognitive research** aims at better understanding the brain spatio-temporal organisation.
- **Brain Computer Interfaces** look at allowing a direct control of the world using brain signal such as EEG signals. Those can be considered both as an application of EEG processing techniques and as a tool for fundamental and applied research as it opens the way for more dynamical and active brain cognitive protocols.

The dream of all M/EEG researchers is to alleviate the need for invasive recordings (electrocorticograms or intracerebral electrodes), which are often necessary prior to brain surgery, in order to precisely locate both pathological and vital functional areas. We are involved in this quest, particularly through our collaboration with the La Timone hospital in Marseille. M/EEG are also used in **cognitive research**, and we collaborate with the *Laboratory for Neurobiology of Cognition* in order to develop methods that suit their needs for sophisticated data analysis.

CORTEX Project-Team

4. Application Domains

4.1. Overview

Our application domain is twofold:

On one hand, neuro-scientists are end-users of our researches. Data analysis is one issue, but the main outcomes concern modeling, namely the validation of biological assumptions either at a theoretical level or via numerical experiments and simulation of bio-processes. This includes algorithmic expertises and dedicated softwares.

On the other hand, science and technology of information processing is impacted. This concerns embedded systems such as in-silico implementations of bio-inspired processes, focusing on spatial and distributed computing. This also concerns embodied systems such as robotic implementation of sensori-motor loops, the bio-inspiration yielding such interesting properties as adaptivity and robustness.

DEMAR Project-Team

4. Application Domains

4.1. Objective quantification and understanding of movement disorders

One main advantage of developping a model based on a physical description of the system is that the parameters are meaningful. Therefore, these parameters when identified on a given individual (valid or deficient), give objective and quantitative data that characterize the system and thus can be used for diagnosis purposes.
Modelling provides a way to simulate movements for a given patient and therefore based on an identification procedure it becomes possible to analyse and then understand his pathology. In order to describe complex pathology such as spasticity that appears on paraplegic patients, you need not only to model the biomechanics parts - including muscles -, but also parts of the peripheral nervous system - including natural sensors - to assess reflex problems. One important application is then to explore deficiencies globally due to both muscles and peripheral neural nets disorders.

4.2. Palliative solutions for movement deficiencies

Functional electrical stimulation is one possibility to restore or control motor functions in an evolutive and reversible way. Pacemaker, cochlear implants, deep brain stimulation (DBS) are successful examples. DEMAR focuses on movement disorder restoration in paraplegic and quadriplegic patients, enhancements in hemiplegic patients, and some other motor disorders such as bladder and bowel control. Nevertheless, since some advances in neuroprosthetic devices can be exploited for the next generation of cochlear implants, the team also contributes to technological and scientific improvements in this domain.
The possibility to interface the sensory motor system, both activating neural structure with implanted FES, and sensing through implanted neural signal recordings open a wide application area:

- Restoring motor function such as grasping for quadriplegic patient, standing and walking for paraplegic patient, compensating foot drop for hemiplegic patients. These applications can be firstly used in a clinical environment to provide physiotherapist with a new efficient FES based therapy (using mainly surface electrodes) in the rehabilitation process. Secondly, with a more sophisticated technology such as implanted neuroprostheses, systems can be used at home by the patient himself without a clinical staff.
- Modulating motor function such as tremors in Parkinsonian patient using DBS. Techniques are very similar but for the moment, modelling is not achieved because it implies the central nervous system modelling in which we are not implied.
- Sensing the afferent pathways such as muscle's spindles, will be used to provide a closed loop control of FES through natural sensing and then a complete implanted solution. Sensing the neural system is a necessity in some complex motor controls such as the bladder control. Indeed, antagonist muscle's contractions, and sensory feedbacks interfere with FES when applied directly on the sacral root nerve concerned. Thus, enhanced activation waveforms and sensing feedback or feedforward signals are needed to perform a highly selective stimulation.

To achieve such objectives, experimentations in animals and humans are necessary. This research takes therefore a long time in order to go from theoretical results to real applications. This process is a key issue in biomedical research and is based on: i) design of complex experimental protocols and setups both for animals and humans, ii) ethical attitude both for humans and animals, with ethical committee approval for human experiments iii) volunteers and selected, both disabled and healthy, persons to perform experiments with the adequate medical staff.

GALEN Team

4. Application Domains

4.1. Clinical Projects

- **MR & Muscular Diseases**: The use of MR and Diffusion Tensor Imaging are investigated in collaboration with the Henri Mondor University Hospital and Institut of Myology towards automatic quantification of muscular mass loss and non-invassive biopsy. The aim is to provide tools that could be used to automatically analyze MR imaging and extract useful clinical measurements (Insitut of Myology), and assess the potential impact of diffusion tensor imaging towards automatic quantification either of muscular diseases progression.
- **Image-driven Radiotherapy Treatment & Surgery Guidance** : The use of CT and MR imaging for cancer guidance treatment in collaboration with the Oscar Lambert Center. The aim is to provide tools for automatic dose estimation as well as off-line and on-line positioning guidance through deformable fusion between imaging data corresponding to perioding patient treatment. The same concept will be explored in collaboration with the Saint-Antoine University Hospital towards image-driven surgery guidance through 2D to 3D registration between interventional and pre-operative annotated data.
- **MR Brain Imaging towards Low-Gliomas Tumor Brain Understanding**: The use of contrast enhanced imaging is investigated in collaboration with the Montpellier University Hospital towards better understanding of low-gliomas positioning, automatic tumor segmentation/identification and longitudinal (tumor) growth modeling.
- **CT/MR Perfusion Imaging**: The use of perfusion imaging is investigated in collaboration with the Georges Pompidou European Hospital towards compartmental analysis and measuring tissue perfusion and capillary permeability in liver tumors.

MNEMOSYNE Team

4. Application Domains

4.1. Overview

One of the most original specificity of our team is that it is part of a laboratory in Neuroscience (with a large spectrum of activity from the molecule to the behavior), focused on neurodegenerative diseases and consequently working in tight collaboration with the medical domain. As a consequence, neuroscientists and the medical world are considered as the primary end-users of our researches. Beyond data and signal analysis where our expertise in machine learning may be possibly useful, our interactions are mainly centered on the exploitation of our models. They will be classically regarded as a way to validate biological assumptions and to generate new hypotheses to be investigated in the living. Our macroscopic models and their implementation in autonomous robots will allow an analysis at the behavioral level and will propose a systemic framework, the interpretation of which will meet aetiological analysis in the medical domain and interpretation of intelligent behavior in cognitive neuroscience.

The study of neurodegenerative diseases is targeted because they match the phenomena we model. Particularly, the Parkinson disease results from the death of dopaminergic cells in the basal ganglia, one of the main systems that we are modeling. The Alzheimer disease also results from the loss of neurons, in several cortical and subcortical regions. The variety of these regions, together with large mnesic and cognitive deficits, require a systemic view of the cerebral architecture and associated functions, very consistent with our approach.

Of course, numerical sciences are also impacted by our researches, at several levels. At a global level, we will propose new control architectures aimed at providing a higher degree of autonomy to robots, as well as machine learning algorithms working in more realistic environment. More specifically, our focus on some cognitive functions in closed loop with a real environment will address currently open problems. This is obviously the case for planning and decision making; this is particularly the case for the domain of affective computing, since motivational characteristics arising from the design of an artificial physiology allow to consider not only cold rational cognition but also hot emotional cognition. The association of both kinds of cognition is undoublty an innovative way to create more realistic intelligent systems but also to elaborate more natural interfaces between these systems and human users.

At last, we think that our activities in well-founded distributed computations and high performance computing are not just intended to help us design large scale systems. We also think that we are working here at the core of informatics and, accordingly, that we could transfer some fundamental results in this domain.

NEUROMATHCOMP Project-Team (section vide)

PARIETAL Project-Team

4. Application Domains

4.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, the neuroimaging *inverse problem* consists in predicting individual information (behavior, phenotype) from neuroimaging data, while an important the *forward problem* consists in fitting neuroimaging data with high-dimensional (e.g. genetic) variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data reasonably well ?), and a regularization schemes that represents a prior on the expected solution of the problem. In particular some priors enforce some properties of the solutions, such as sparsity, smoothness or being piecewise constant.
Let us detail the model used in the inverse problem: Let X be a neuroimaging dataset as an (n_{subj}, n_{voxels}) matrix, where n_{subj} and n_{voxels} are the number of subjects under study, and the image size respectively, Y an array of values that represent characteristics of interest in the observed population, written as (n_{subj}, n_f) matrix, where n_f is the number of characteristics that are tested, and β an array of shape (n_{voxels}, n_f) that represents a set of pattern-specific maps. In the first place, we may consider the columns $Y_1, .., Y_{n_f}$ of Y independently, yielding n_f problems to be solved in parallel:

$$Y_i = X\beta_i + \epsilon_i, \forall i \in \{1, .., n_f\},$$

where the vector contains β_i is the i^{th} line of β. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{\beta}_i = \text{argmin}_{\beta_i} \|Y_i - X\beta_i\|^2 + \Psi(\beta_i), \tag{1}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\beta; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 + \eta_1 \|\nabla\beta\|_1 + \eta_2 \|\nabla\beta\|_2^2 \tag{2}$$

with $\lambda_1,\ \lambda_2,\ \eta_1,\ \eta_2 \geq 0$. In general, only one or two of these constraints will have a non-zero weighting:

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution β is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piecewise constant solution.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

The performance of the predictive model can simply be evaluated as the amount of variance in Y_i fitted by the model, for each $i \in \{1, .., n_f\}$. This can be computed through cross-validation, by *learning* $\widehat{\beta}_i$ on some part of the dataset, and then estimating $(Y_i - X\widehat{\beta}_i)$ using the remainder of the dataset.

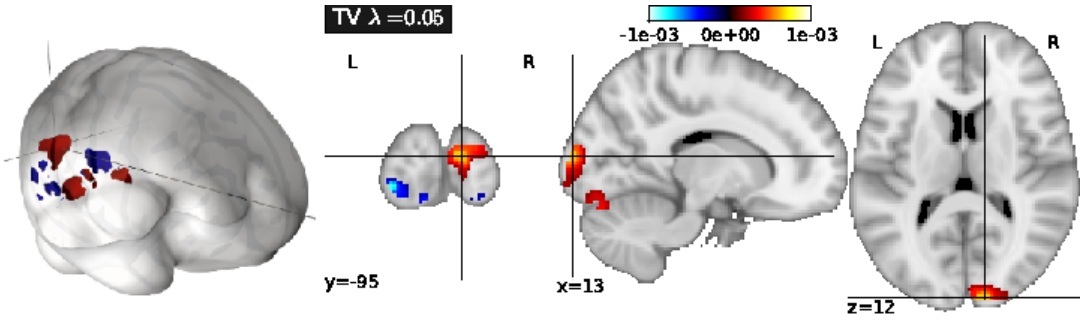

Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problems here consists in predicting the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the observation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that enjoys particularly high accuracy.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This is particularly important to include external anatomical priors on the relevant solution.
- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.
- *Logistic regression*, where a sigmoid non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.
- *Robustness to between-subject variability* is an important question, as it makes little sense that a learned model depends dramatically on the particular observations used for learning. This is an important issue, as this kind of robustness is somewhat opposite to sparsity requirements.
- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector β to have a shared support across all these variables.
  For instance, when one of the variables Y_i is not well fitted by the model, the estimation of other variables $Y_j, j \neq i$ may provide constraints on the support of β_i and thus, improve the prediction of Y_i. Yet this does not impose constraints on the non-zero parameters of the parameters β_i.

$$Y = X\beta + \epsilon, \tag{3}$$

then

$$\widehat{\beta} = \text{argmin}_{\beta=(\beta_i), i=1..n_f} \sum_{i=1}^{n_f} \|Y_i - X\beta_i\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \beta_{i,j}^2} \tag{4}$$

4.2. Multivariate decompositions

Multivariate decompositions are an important tool to model complex data such as brain activation images: for instance, one might be interested in extracting an atlas of brain regions from a given dataset, such as regions depicting similar activities during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.
Let X be a neuroimaging dataset written as an (n_{subj}, n_{voxels}) matrix, after proper centering; the model reads

$$X = AD + \epsilon, \tag{5}$$

where D represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and A the associated subject-wise loadings. While traditional PCA and independent components analysis are limited to reconstruct components D within the space spanned by the column of X, it seems desirable to add some constraints on the rows of D, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging.
This yields the following estimation problem:

$$\min_{D,A} \|X - AD\|^2 + \Psi(D) \text{ s.t. } \|A_i\| = 1 \; \forall i \in \{1..n_f\}, \tag{6}$$

where $(A_i),\ i \in \{1..n_f\}$ represents the columns of A. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.
The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on D for A fixed, and conversely. This readily suggests an alternate optimization scheme, where D and A are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

4.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to assess whether an observation is aberrant or not or in classification problems. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.
The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.
- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is an important question that needs to be addressed.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-samples likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we will consider different means of parametrizing covariance distributions and their impact on the network. Our current work on post-stroke patients (see e.g. Fig. 2) suggests indeed that modeling may prove essential to perform sensitive inference.

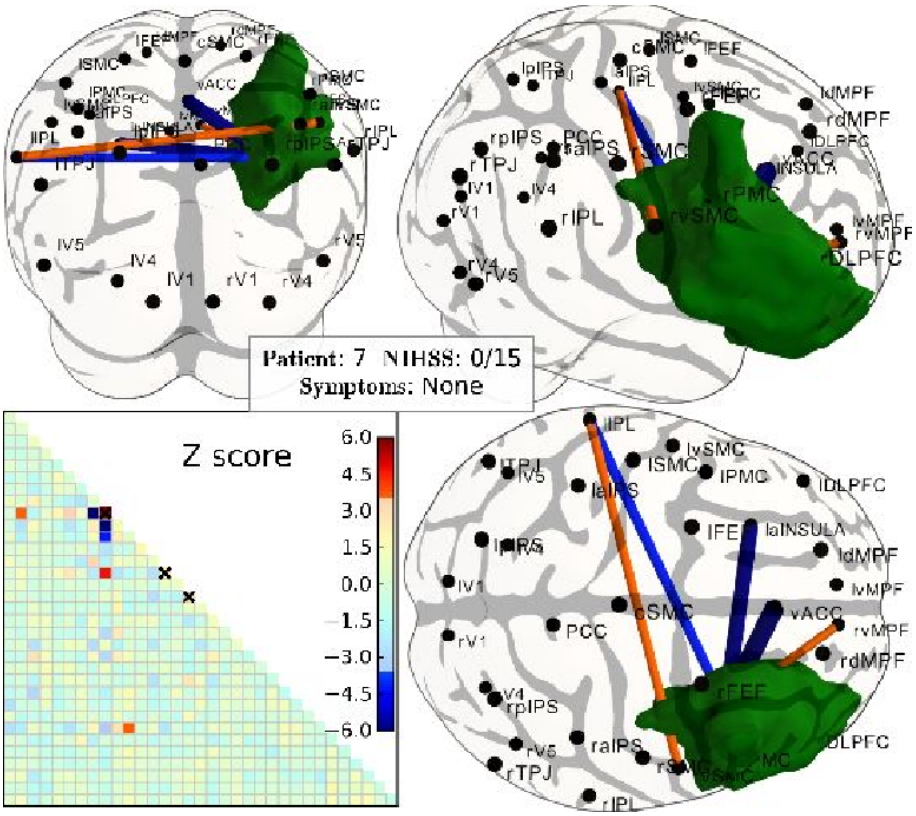

Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion outlined in green) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

SHACRA Project-Team

4. Application Domains

4.1. Clinical Applications

Some of the scientific challenges described previously can be seen in a general context (such as solving constraints between different types of objects, parallel computing for interactive simulations, etc.) but often it is necessary to define a clinical context for the problem. This is required in particular for defining the appropriate assumptions in various stages of the biophysical modeling. It is also necessary to validate the results. This clinical context is a combination of two elements: the procedure we attempt to simulate and the objective of the simulation: training, planning or per-operative guidance. Below are a series of applications we plan to develop. The choice of these applications is not random: the clinical procedures we target are all technically challenging, they highlight various parts of our research, and often they represent an ideal testbed for transitioning from training to planning to guidance. It is important also to note that developing these applications raises many challenges and as such this step should be seen as an integral part of our research. It is also through the development of these applications that we can communicate with physicians, and validate our results. SOFA will be used as a backbone for the integration of our research into clinical applications.

4.1.1. Interventional radiology

Over the past twenty years, interventional methods such as angioplasty, stenting, and catheter-based drug delivery have substantially improved the outcomes for patients with vascular disease. Pathologies that used to require a surgical procedure can now be treated in a much less invasive way. As a consequence, interventional radiology procedures represent an increasing part of the interventions currently performed, with more than 6 million patients treated every year in Europe and about 5 millions the United States. However, these techniques require an intricate combination of tactile and visual feedback, and extensive training periods to attain competency. To reinforce the need to reach and maintain proficiency, the FDA recently required that US physicians go through simulation-based training before using newly developed carotid stents. Besides simulation for training, interventional radiology is a perfect target to illustrate the potential of planning and rehearsal of procedures. As an initial step in this direction, Alcove and Magrit were partners in an ARC project (Simple) to develop a planning tool for the treatment of aneurysms using coils. This collaboration still goes on after the end of the ARC, and led to a series of papers in key conferences [5] [28], [34], [21].

4.1.1.1. Interventional neuro-radiology

We will continue the development of our simulation and planning system for interventional radiology, with two principal clinical partners: Massachusetts General Hospital in Boston and University Hospital in Nancy. We have completed the integration in SOFA of improved versions of algorithms for describing the behavior of catheters, guide-wires, coils, as well as the interactive simulation of fluoroscopic images, the modeling of complex contacts. Our future efforts will focus on the development of an advanced planning system for interventional radiology, in particular for coil embolization. This will require the integration of new methods of reconstruction of vascular anatomy from medical images (in collaboration with the MAGRIT team). We will also add our recent results on blood flow simulation in aneurysms.

4.1.1.2. Interventional cardiology using radio-frequency ablation

Cardiac arrhythmias (or dysrhythmias) are problems that affect the electrical system of the heart muscle, producing abnormal heart rhythms, and causing the heart to pump less effectively. About 5% of people over 40 years old are affected by this pathology, with a rather high morbidity rate. Radio-frequency ablation is a non-surgical procedure that has been used for about 15 years to treat tachyarrhythmias, i.e. rapid, uncoordinated heartbeats. The procedure is performed by guiding a catheter with an electrode at its tip to the area of heart muscle where there is an accessory pathway. The catheter is guided under fluoroscopic imaging. When the catheter is positioned at the site where cells give off the electrical signals that stimulate the abnormal heart

rhythm, a low radio-frequency energy is transmitted to the pathway. This destroys heart muscle cells within a very small area near the tip of the catheter and stops the area from conducting the extra impulses that caused the arrhythmia. In this context, a simulation system would be able to provide added value in two main areas: 1) to train physicians in the early stages of their apprenticeship and 2) to provide quantitative information during the planning phase of a complex procedure, using patient-specific data. Most aspects of this simulation will rely on components developed during our research program but we will also extend our collaboration with the ASCLEPIOS team and the CardioSense3D project on the modeling of the heart the Cadiosense3D project. This involves an important integration task, and it will also validate the reusability aspects of the code developed within SOFA.

4.1.2. Minimally-invasive surgery

4.1.2.1. Laparoscopic hepatic resection

The liver is one of the major organs in the human body. It is in charge of more than 100 vital functions. Because of its many functions, its pathologies are also varied, numerous and unfortunately often lethal. This is for instance the case of hepatitides which today affect about 300,000 people in France for hepatitis B and 600,000 people for hepatitis C. The most advanced state of evolution of these pathologies is generally cirrhosis followed by cancer, which represents the third cause of cancer related death. In 2005, 14,267 liver cancer cases and 20,497 cirrhosis cases have been diagnosed in France. The surgical solution remains the option offering the best success rate for these pathologies. More than 7,000 surgical interventions have been carried out on the liver in 2005 and partial resection of the liver remains the most common approach. In this context, the ability to train surgeons, and to be able to plan complex procedures using computer-based simulations, would be a formidable help to the current apprenticeship model: “See One, Do One, Teach One”. Right now, only a few commercial systems are available to the medical community, and they are limited to basic skills training. Developing a realistic simulation system that could be used to plan and rehearse procedures would be a very important step in the introduction of new training paradigms in medicine. This is the main objective of the PASSPORT european project in which we are actively contributing at two levels. First, our research results on biomechanical modeling of solid organs and on coupling will be used to propose a realistic model of the deformation of the liver and its vascular network. Second, SOFA has been chosen in this project as the software for integrating all results from the different partners. Both aspects will help validate our models, test SOFA and obtain feedback from the clinicians.

4.1.2.2. Ophthalmology and cataract surgery

A cataract is an opacity in the natural lens of the eye. It represents an important cause of visual impairment and, if not treated, can lead to blindness. It is actually the leading cause of blindness worldwide, and its development is related to aging, sunlight exposure, smoking, poor nutrition, eye trauma, and certain medications. The best treatment for this pathology remains surgery. Cataract surgery has made important advances over the past twenty years, and in 2005, more than 5 million people in the United States and in Europe underwent cataract surgery. Most cataract surgeries are performed using microscopic size incisions, advanced ultrasonic equipment to fragment cataracts into tiny fragments, and foldable intraocular lenses to minimize the size of the incision. All these advances benefit the patient, but increase training requirements for eye surgeons. At the end of 2007, we started the development of a new training system for cataract surgery. The main objectives of this simulation are to reproduce with great accuracy the three main steps of cataract surgery: 1) capsulorhexis 2) phacoemulsification and 3) implantation of an intraocular lens. We have already started the development of this simulation. The main research effort went in the choice of appropriate deformable models for the lens and lens capsule. An important effort also went into the development of topological changes corresponding to the capsulorhexis and phacoemulsification [20]. The modeling of the intraocular implant and its deployment in the capsule has been published to the major conference in medical simulation [26].

4.1.2.3. Neurosurgery and deep brain stimulation

Deep brain stimulation (DBS) is a neurosurgical treatment which stimulates the brain with low electrical signals. The signals reorganize the brain’s electrical impulses (similarly as what was presented above for radio-frequency ablation for cardiac problems). This results in major improvements in several pathologies

such as Parkinson disease. The principle of the procedure is the following: a thin, insulated wire lead with several electrodes at the tip is surgically implanted into the affected area of the brain. A wire runs under the skin to a battery-operated pulse generator implanted near the collarbone. The generator is programmed to send continuous electrical pulses to the brain. To implant the electrodes, a neurosurgeon uses a stereotactic head frame and magnetic resonance or computed tomography imaging to map the brain and pinpoint the problem area. The main difficulty in this procedure comes from the deformation of the brain (small brain shift when the skull is opened, and local deformation of the brain due to the insertion of the electrode) and the deflection of the electrode itself during and after the procedure. This results in a difference between the planned target and the location of the end effector of the electrode. Our main objective is to use our work on soft tissue deformation, vascularized structures, as well as our recent results on constraint solving between soft tissues and flexible devices [29]. This work will be done in collaboration with the VISAGES team and we will dedicate an important effort in validating our results, analyzing post-operative medical images, and interacting with surgeons. This project has a strong potential as DBS is being increasingly used yet most research groups only consider non deformable planning systems (geometrical planning). Our proposal could make a important difference in the accuracy of the planning as it takes into account the biophysics of the brain.

VISAGES Project-Team

4. Application Domains

4.1. Neuroimaging

neuroimaging, clinical neuroscience, multiple sclerosis, multispectral MRI, brain atlas

One research objective in neuroimaging is the construction of anatomical and functional cerebral maps under normal and pathological conditions.

Many researches are currently performed to find correlations between anatomical structures, essentially sulci and gyri, where neuronal activation takes place, and cerebral functions, as assessed by recordings obtained by the means of various neuroimaging modalities, such as PET (Positron Emission Tomography), fMRI (Functional Magnetic Resonance Imaging), EEG (Electro-EncephaloGraphy) and MEG (Magneto-EncephaloGraphy). Then, a central problem inherent to the formation of such maps is to put together recordings obtained from different modalities and from different subjects. This mapping can be greatly facilitated by the use of MR anatomical brain scans with high spatial resolution that allows a proper visualization of fine anatomical structures (sulci and gyri). Recent improvements in image processing techniques, such as segmentation, registration, delineation of the cortical ribbon, modelling of anatomical structures and multi-modality fusion, make possible this ambitious goal in neuroimaging. This problem is very rich in terms of applications since both clinical and neuroscience applications share similar problems. Since this domain is very generic by nature, our major contributions are directed towards clinical needs even though our work can address some specific aspects related to the neuroscience domain.

4.2. Multiple sclerosis

Over the past years, a discrepancy became apparent between clinical Multiple sclerosis (MS) classification describing on the one hand MS according to four different disease courses and, on the other hand, the description of two different disease stages (an early inflammatory and a subsequently neurodegenerative phase). It is to be expected that neuroimaging will play a critical role to define *in vivo* those four different MS lesion patterns. An *in vivo* distinction between the four MS lesion patterns, and also between early and late stages of MS will have an important impact in the future for a better understanding of the natural history of MS and even more for the appropriate selection and monitoring of drug treatment in MS patients. Since MRI has a low specificity for defining in more detail the pathological changes which could discriminate between the different lesion types, but a high sensitivity to detect focal and also widespread, diffuse pathology of the normal appearing white and grey matter, our major objective within this application domain is to define new neuroimaging markers for tracking the evolution of the pathology from high dimensional data (e.g. nD+t MRI). In addition, in order to complement MR neuroimaging data, we ambition to perform also cell labelling neuroimaging (e.g. MRI or PET) and to compare MR and PET data using standard and experimental MR contrast agents and radiolabeled PET tracers for activated microglia (e.g. USPIO or PK 11195). The goal is to define and develop, for routine purposes, cell specific and also quantitative imaging markers for the improved *in vivo* characterization of MS pathology.

4.3. Modelling of anatomical and anatomo-functional neurological patterns

The major objective within this application domain is to build anatomical and functional brain atlases in the context of functional mapping and for the study of developmental, neurodegenerative or even psychiatric brain diseases (Multiple sclerosis, Epilepsy, Parkinson, Dysphasia, Depression or even Alzheimer). This is a very competitive research domain; our contribution is based on our previous works in this field [52], [54], [53], [55], and by continuing our local and wider collaborations.

An additional objective within this application domain is to find new descriptors to study the brain anatomy and/or function (e.g. variation of brain perfusion, evolution in shape and size of an anatomical structure in relation with pathology or functional patterns, computation of asymmetries ...). This is also a very critical research domain, especially for many developmental or neurodegenerative brain diseases.

CLIME Project-Team

4. Application Domains

4.1. Introduction

The central application domain of the project-team is atmospheric chemistry. We develop and maintain the air quality modeling system Polyphemus, which includes several numerical models (Gaussian models, Lagrangian model, two 3D Eulerian models including Polair3D) and their adjoints, and different high level methods: ensemble forecast, sequential and variational data assimilation algorithms. Advanced data assimilation methods, network design, inverse modeling, ensemble forecast are studied in the context of air chemistry. Note that addressing these high level issues requires controlling the full software chain (models and data assimilation algorithms).

The activity on assimilation of satellite data is mainly carried out for meteorology and oceanography. This is addressed in cooperation with external partners who provide numerical models. Concerning oceanography, the aim is to improve the forecast of ocean circulation, by assimilation of fronts and vortices displayed on image data. Concerning meteorology, the focus is on correcting the model location of structures related to high-impact weather events (cyclones, convective storms, *etc*.) by assimilating images.

4.2. Air quality

Air quality modeling implies studying the interactions between meteorology and atmospheric chemistry in the various phases of matter, which leads to the development of highly complex models. The different usages of these models comprise operational forecasting, case studies, impact studies, *etc*, with both societal (e.g., public information on pollution forecast) and economical impacts (e.g., impact studies for dangerous industrial sites). Models lack some appropriate data, for instance better emissions, to perform an accurate forecast and data assimilation techniques are recognized as a major key point for the improvement of forecast's quality.

In this context, Clime is interested in various problems, the following being the crucial ones:

- The development of ensemble forecast methods for estimating the quality of the prediction, in relation with the quality of the model and the observations. Sensitivity analysis with respect to the model's parameters so as to identify physical and chemical processes, whose modeling must be improved.
- The development of methodologies for sequential aggregation of ensemble simulations. What ensembles should be generated for that purpose, how spatialized forecasts can be generated with aggregation, how can the different approaches be coupled with data assimilation?
- The definition of second-order data assimilation methods for the design of optimal observation networks. Management of combinations of sensor types and deployment modes. Dynamic management of mobile sensors' trajectories.
- How to estimate the emission rate of an accidental release of a pollutant, using observations and a dispersion model (from the near-field to the continental scale)? How to optimally predict the evolution of a plume? Hence, how to help people in charge of risk evaluation for the population?
- The definition of non-Gaussian approaches for data assimilation.
- The assimilation of satellite measurements of troposphere chemistry.

The activities of Clime in air quality are supported by the development of the Polyphemus air quality modeling system. This system has a modular design, which makes it easier to manage high level applications such as inverse modeling, data assimilation and ensemble forecast.

4.3. Oceanography

The capacity of performing a high quality forecast of the state of the ocean, from the regional to the global scales, is of major interest. Such a forecast can only be obtained by systematically coupling numerical models and observations (*in situ* and satellite data). In this context, being able to assimilate image structures becomes a key point. Examples of such image structures are:

- apparent motion linked to surface velocity;
- trajectories, obtained either from tracking of features or from integration of the velocity field;
- spatial objects, such as fronts, eddies or filaments.

Image Models for these structures are developed and take into account the underlying physical processes. Image data are assimilated in Image Models to derive pseudo-observations of state variables, which are further assimilated in numerical ocean forecast models.

4.4. Meteorology

Meteorological forecasting constitutes a major applicative challenge for Image Assimilation. Although satellite data are operationally assimilated within models, this is mainly done on an independent pixel basis: the observed radiance is linked to the state variables via a radiative transfer model, that plays the role of an observation operator. Indeed, because of their limited spatial and temporal resolutions, numerical weather forecast models fail to exploit image structures, such as precursors of high impact weather:

- cyclogenesis related to the intrusion of dry stratospheric air in the troposphere (a precursor of cyclones),
- convective systems (supercells) leading to heavy winter time storms,
- low-level temperature inversion leading to fog and ice formation, *etc*.

To date, there is no available method for assimilating such data, which are characterized by a strong coherence in space and time. Meteorologists have developed qualitative Conceptual Models (CMs), for describing the high impact weathers and their signature on images, and tools to detect CMs on image data. The result of this detection is used for correcting the numerical models, for instance by modifying the initialization. The aim is therefore to develop a methodological framework allowing to assimilate the detected CMs within numerical forecast models. This is a challenging issue given the considerable impact of the related meteorological events.

FLUMINANCE Project-Team

4. Application Domains

4.1. Introduction

By designing new approaches for the analysis of fluid-image sequences the FLUMINANCE group aims at contributing to several application domains of great interest for the community and in which the analysis of complex fluid flows plays a central role. The group focuses mainly on two broad application domains:

- Environmental sciences;
- Experimental fluid mechanics and industrial flows.

4.2. Environmental sciences

The first huge application domain concerns all the sciences that aim at observing the biosphere evolution such as meteorology, climatology or oceanography but also remote sensing study for the monitoring of meteorological events or human activities consequences. For all these domains image analysis is a practical and unique tool to *observe, detect, measure, characterize or analyze* the evolution of physical parameters over a large domain. The design of generic image processing technique for all these domains might offer practical software tools to measure precisely the evolution of fluid flows for weather forecasting or climatology studies. It might also offer possibilities of closed surveillance of human and natural activities in sensible areas such as forests, river edges, and valley in order to monitor pollution, floods or fire. The need in terms of local weather forecasting, risk prevention, or local climate change is becoming crucial for our tomorrow's life. At a more local scale, image sensors may also be of major utility to analyze precisely the effect of air curtains for safe packaging in agro-industrial.

4.3. Experimental fluid mechanics and industrial flows

In the domain of **experimental fluid mechanics**, the visualization of fluid flows plays a major role, especially for turbulence study since high frequency imaging has been made currently available. Together with analysis of turbulence at different scales, one of the major goals pursued at the moment by lot of scientists and engineers consists in studying the ability to manipulate a flow to induce a desired change. This is of huge technological importance to enhance or inhibit mixing in shear flows, improve energetic efficiency or control the physical effects of strain and stresses. This is for instance of particular interest for:

- military applications, for example to limit the infra-red signatures of fighter aircraft;
- aeronautics and transportation, to limit fuel consumption by controlling drag and lift effects of turbulence and boundary layer behavior;
- industrial applications, for example to monitor flowing, melting, mixing or swelling of processed materials, or preserve manufactured products from contamination by airborne pollutants, or in industrial chemistry to increase chemical reactions by acting on turbulence phenomena.

.

MAGIQUE-3D Project-Team

4. Application Domains

4.1. Seismic Imaging

The main objective of modern seismic processing is to find the best representation of the subsurface that can fit the data recorded during the seismic acquisition survey. In this context, the seismic wave equation is the most appropriate mathematical model. Numerous research programs and related publications have been devoted to this equation. An acoustic representation is suitable if the waves propagate in a fluid. But the subsurface does not contain fluids only and the acoustic representation is not sufficient in the general case. Indeed the acoustic wave equation does not take some waves into account, for instance shear waves, turning waves or the multiples that are generated after several reflections at the interfaces between the different layers of the geological model. It is then necessary to consider a mathematical model that is more complex and resolution techniques that can model such waves. The elastic or viscoelastic wave equations are then reference models, but they are much more difficult to solve, in particular in the 3D case. Hence, we need to develop new high-performance approximation methods.

Reflection seismics is an indirect measurement technique that consists in recording echoes produced by the propagation of a seismic wave in a geological model. This wave is created artificially during seismic acquisition surveys. These echoes (i.e., reflections) are generated by the heterogeneities of the model. For instance, if the seismic wave propagates from a clay layer to sand, one will observe a sharp reflected signal in the seismic data recorded in the field. One then talks about reflection seismics if the wave is reflected at the interface between the two media, or talks about seismic refraction if the wave is transmitted along the interface. The arrival time of the echo enables one to locate the position of this transition, and the amplitude of the echo gives information on some physical parameters of the two geological media that are in contact. The first petroleum exploration surveys were performed at the beginning of the 1920's and for instance, the Orchard Salt Dome in Texas (USA) was discovered in 1924 by the seismic-reflection method.

MOISE Project-Team

4. Application Domains

4.1. Introduction

The evolution of natural systems, in the short, mid, or long term, has extremely important consequences for both the global Earth system and humanity. Forecasting this evolution is thus a major challenge from the scientific, economic, and human viewpoints.

Humanity has to face the problem of **global warming**, brought on by the emission of greenhouse gases from human activities. This warming will probably cause huge changes at global and regional scales, in terms of climate, vegetation and biodiversity, with major consequences for local populations. Research has therefore been conducted over the past 15 to 20 years in an effort to model the Earth's climate and forecast its evolution in the 21st century in response to anthropic action.

With regard to short-term forecasts, the best and oldest example is of course **weather forecasting**. Meteorological services have been providing daily short-term forecasts for several decades which are of crucial importance for numerous human activities.

Numerous other problems can also be mentioned, like **seasonal weather forecasting** (to enable powerful phenomena like an El Niño event or a drought period to be anticipated a few months in advance), **operational oceanography** (short-term forecasts of the evolution of the ocean system to provide services for the fishing industry, ship routing, defense, or the fight against marine pollution), **air pollution** prediction systems, the prediction of **floods**, or the simulation of **mud flows** and **snow avalanches** for impact studies and regional planning.

As mentioned previously, mathematical and numerical tools are omnipresent and play a fundamental role in these areas of research. In this context, the vocation of MOISE is not to carry out numerical prediction, but to address mathematical issues raised by the development of prediction systems for these application fields, in close collaboration with geophysicists.

4.2. Oceanography and the Ocean-Atmosphere System

Participants: Eric Blayo, Pierre-Antoine Bouttier, Vincent Chabot, David Cherel, Laurent Debreu, Jérémie Demange, Marc Honnorat, Christine Kazantsev, Eugène Kazantsev, François-Xavier Le Dimet, Bénédicte Lemieux-Dudon, Xavier Meunier, Maëlle Nodet, Antoine Rousseau, Manel Tayachi, Arthur Vidard.

Multi-resolution, Coupling Methods, Data Assimilation, Ocean, Atmosphere

Understanding and forecasting the ocean circulation is currently the subject of an intensive research effort by the international scientific community. This effort was primarily motivated by the crucial role of the ocean in determining the Earth's climate, particularly from the perspective of global change. In addition, important recent research programs are aimed at developing operational oceanography, i.e. near real-time forecasting of ocean circulation, with applications for ship routing, fisheries, weather forecasting, etc. Another related field is coastal oceanography, dealing for example with pollution, littoral planning, or the ecosystems management. Local and regional agencies are currently very interested in numerical modelling systems for coastal areas.

Both ocean-alone models and coupled ocean-atmosphere models are being developed to address these issues. In this context, the MOISE project-team conducts efforts mainly on the following topics:

- *Multi-resolution approaches and coupling methods*: Many applications in coastal and operational oceanography require high resolution local models. These models can either be forced at their boundaries by some known data, or be dynamically coupled with a large-scale coarser resolution model. Such model interactions require specific mathematical studies on open boundary conditions, refinement methods (like mesh refinement or stochastic downscaling), and coupling algorithms. The latter have also to be studied in the context of ocean-atmosphere coupled systems.

- *Advanced numerical schemes*: Most ocean models use simple finite difference schemes on structured grids. We are seeking for better schemes allowing both accuracy and good conservation properties, and dealing with irregular boundaries and bottom topography.
- *Data assimilation methods for ocean modelling systems*: The main difficulties encountered when assimilating data in ocean or atmosphere models are the huge dimension of the model state vector (typically 10^6-10^8), the strongly nonlinear character of the dynamics, and our poor knowledge of model error statistics. In this context, we are developing reduced order sequential and variational data assimilation methods addressing the aforementioned difficulties. We are also working on the assimilation of lagrangian data, of sequences of images, and on the design of data assimilation methods for multi-resolution models and for coupled systems.

Most of these studies are led in strong interaction with geophysicists, in particular from the Laboratoire des Ecoulements Géophysiques et Industriels (LEGI, Grenoble).

4.3. Glaciology

Participants: Eric Blayo, Bertrand Bonan, Bénédicte Lemieux-Dudon, Maëlle Nodet, Habib Toye Mahamadou Kele.

Inverse Methods, Data Assimilation, Glaciology, Ice Core Dating

The study of past climate is a means of understanding climatic mechanisms. Drillings in polar ice sheets provide a huge amount of information on paleoclimates: correlation between greenhouse gases and climate, fast climatic variability during the last ice age, etc. However, in order to improve the quantitative use of the data from this archive, numerous questions remain to be answered because of phenomena occurring during and after the deposition of snow. An important research aim is therefore to optimally model ice sheets in the vicinity of drilling sites in order to improve their interpretation: age scale for the ice and for the gas bubbles, mechanical thinning, initial surface temperature and accumulation when snow is deposited, spatial origin of ice from the drilling.

In another respect, ice streams represent an important feature of ice flows since they account for most of the ice leaving the ice sheet (in Antarctic, one estimates that ice streams evacuate more than 70% of the ice mass in less than 10% of the coast line). Furthermore, recent observations showed that some important ice streams are presently accelerating. Thus, we seek to improve models of ice sheets, by developing data assimilation approaches in order to calibrate them using available observations.

Another objective is the evaluation of the state of the polar ice caps in the past, and their interactions with the other components of the earth climate, in order to forecast their evolution in the forthcoming centuries. The joint use of models and data, through data assimilation techniques, to improve system description is relatively new for the glaciological community. Therefore inverse methods have to be developed or adapted for this particular purpose.

By gaining and loosing mass, glaciers and ice-sheets are playing a key role in the sea level evolution. This is obvious when regarding past as, for example, collapse of the large northern hemisphere ice-sheets after the Last Glacial Maximum has contributed to an increase of 120 m of sea level. This is particularly worrying when the future is considered. Indeed, recent observations clearly indicate that important changes in the velocity structure of both Antarctic and Greenland ice-sheets are occurring, suggesting that large and irreversible changes may have been initiated. This has been clearly emphasized in the last report published by the Intergovernmental Panel on Climate Change (IPCC). IPCC has further insisted on the poor current knowledge of the key processes at the root of the observed accelerations and finally concluded that reliable projections of sea-level rise are currently unavailable. In this context, our general aim is to develop data assimilation methods related to ice flow modelling purpose, in order to provide accurate and reliable estimation of the future contribution of ice-sheets to Sea Level Rise.

Development of ice flow adjoint models is by itself a scientific challenge This new step forward is clearly motivated by the amount of data now available at both the local and the large scales.

4.4. River Hydraulics

Participants: Eric Blayo, Antoine Rousseau, Manel Tayachi.

Shallow Water (SW) models are widely used for the numerical modeling of river flows. Depending on the geometry of the domain, of the flow regime, and of the level of accuracy which is required, either 1D or 2D SW models are implemented. It is thus necessary to couple 1D models with 2D models when both models are used to represent different portions of the same river. Moreover, when a river flows into the sea/ocean (e.g. the Rhône river in the Mediterranean), one may need to couple a 2D SW with a full 3D model (such as the Navier-Stokes equations) of the estuary. These issues have been widely addressed by the river-engineering community, but often with somehow crude approaches in terms of coupling algorithms. This may be improved thanks to more advanced boundary conditions, and with the use of Schwarz iterative methods for example. We want to tackle these issues, in particular in the framework of a partnership with the French electricity company EDF.

POMDAPI Project-Team (section vide)

SAGE Project-Team

4. Application Domains

4.1. Geophysics

The team has chosen a particular domain of application, which is geophysics. In this domain, many problems require solving large scale systems of equations, arising from the discretization of coupled models. Emphasis is put on hydrogeology, but the team also investigates geodesy, heat and mass transfer in soil, and granular materials. One of the objectives is to use high performance computing in order to tackle 3D large scale computational domains with complex physical models.

4.2. Hydrogeology

This is joint work with Geosciences Rennes, University of Poitiers and CDCSP at University of Lyon. It is also done in the context of the group Momas and Andra grant.

Many environmental studies rely on modelling geo-chemical and hydrodynamic processes. Some issues concern water resources, aquifer contamination, underground waste disposal, clean-up of former waste deposits, acid mine drainage remediation. Other issues, also related to energy, concern geothermy, unconventional gas, enhanced oil recovery, underground storage of CO2, underground storage of nuclear waste.

Simulation of contaminant transport in groundwater is a highly complex problem, governed by coupled linear or nonlinear PDAEs. Moreover, due to the lack of experimental data, stochastic models are used for dealing with heterogeneity. The main objective of the team is to design and to implement efficient and robust numerical models, including Uncertainty Quantification methods.

Recent research showed that rock solid masses are in general fractured and that fluids can percolate through networks of inter-connected fractures. Fractured media are by nature very heterogeneous and multi-scale, so that homogenisation approaches are not relevant. The team develops a numerical model for fluid flow and contaminant transport in three-dimensional porous fractured media.

An important output is the parallel scientific platform H2OLab, running on clusters, grids and machines available in supercomputing centers.

STEEP Exploratory Action

4. Application Domains

4.1. Urban economy and land use and transport modeling

Modern urban regions are highly complex entities. The understanding of the phenomena underlying urban sprawl and peri-urbanization is a key element to control the dynamics structuring urban space. Clearly, urban transport systems are intricately linked to urban structure and the distribution of activities, i.e., to land use. Urbanization generally implies an increase in travel demand. Cities have traditionally met this additional demand by expanding the transportation supply, through new highways and transit lines. In turn, an improvement of the accessibility of ever-farther land leads to an expansion of urban development, resulting in a significant feedback loop between transportation infrastructure and land use, one of the main causes of urban sprawl.

Several models have been developed in the field of urban economics to understand the complex relationship between transportation and land use and to facilitate the urban planning process. They enable the simulation of public policies and the quantification of indicators describing the evolution of urban structure. Key factors such as transport congestion, energy consumption, CO2 emissions etc., can be evaluated or estimated, and different urban development scenarios can be tested in a quantitative manner.

Yet, very few local authorities in charge of planning issues make use of these strategic models, mostly because they are difficult to calibrate and validate, two critical steps where systematic improvement would increase the level of confidence in the obtained results. These limitations prevent dissemination in local agencies. One goal of STEEP is therefore to meet the need of better calibration and validation strategies and algorithms. This research is the core of our projects CITiES (ANR Modèles Numériques) and TRACER (Ecos Nord Venezuela).

4.2. Ecological accounting and material flow analysis

One of the major issues in the assessment of the long-term sustainability of urban areas is related to the concept of "imported sustainability". Indeed, any city brings from the outside most of its material and energy resources, and rejects to the outside the waste produced by its activity. The modern era has seen a dramatic increase in both volume and variety of these material flows and consumption as well as in distance of origin and destination of these flows, usually accompanied by a spectacular increase in the associated environmental impacts. A realistic assessment of the sustainability of urban areas requires to quantify both local and distant environmental impacts; greenhouse gas emissions are only one aspect of this question.

In order to produce such an assessment for a given territory or urban area, one must first establish different types of ecological accounting: one must identify and quantify the different types of material and energy uses on the one hand, and the different types of impact associated to these uses. The first task is the object of Material Flow Analysis (MFA), while the second is more directly related to the logic of Life Cycle Analysis (LCA). One of the major challenges here is to obtain reliable MFA data at the region and *département* scales, either directly, or through appropriate disaggregation techniques.

The STEEP group has started a research program on this theme with three major aims in mind: 1) Creating a comprehensive database enabling such analyses; 2) Developing methodology and models resolving scaling issues, and developing algorithms allowing to rigorously and automatically obtain the adequate assessments; 3) Providing a synthetic analysis of environmental impacts associated to the major material flows, at various geographic levels (employment catchment area, *département* and *région*, for France). The database elaboration is already well underway, and apparently does not yet have any equivalent elsewhere in France. The impact evaluation and decision help strategy (i.e., evaluating alternative policy options in term of environmental impact) will be developed in collaboration with a private company ARTELIA, through a CIFRE PhD thesis that is scheduled to start early 2013, and will be based on existing LCA databases. The PhD student is co-directed with Denis Dupré from CERAG (Centre d'Etudes et de Recherches Appliquées à la Gestion).

4.3. Eco-system services

Long-term sustainability is closely related to the underlying ecosystems, on various fronts: production of renewable resources (either energy or biomass), waste and pollutant resorption, local and global climate regulations etc. These various functions constitute the "ecosystem services" provided to society by our natural environment.

The reduction of the adverse impacts of urban areas on the environment is linked not only to limiting urban sprawl and making more efficient use of the available resources, but also to developing a better grasp of the interrelations between urban/peri-urban areas and their agricultural and semi-natural surroundings. In particular, reducing distant impacts while making a better use of local resources is a major challenge for the coming decades.

In this context, the STEEP team is involved in a project bearing on the characterization of local Ecosystem Services NETworks (ESNET), piloted by LECA (*Laboratoire d'Ecologie Alpine*), and in collaboration with a number of other research laboratories (most notably, IRSTEA Grenoble) and a panel of local stakeholders; the scale of interest is typically a landscape (in the ecologic/geographic sense, i.e., a zone a few kilometers to a few tens of kilometers wide). The project aims at developing a generic modelling framework of ecosystem services, and studying their behavior under various scenarios of coupled urban/environment evolution. The contribution of the STEEP team is centered on the Land Use/Land Cover Change (LUCC) model that will be one of the major building blocks of the whole model, with the help of specifically hired personal.

The project has been supported by FRB (*Fondation pour la Recherche en Biodiversité*) and will be funded by ONEMA (*Office National de l'Eau et des Milieux Aquatiques*) for the three years of its duration.

BANG Project-Team

4. Application Domains

4.1. Biology and medicine

The team is mostly involved in applications to biology and medicine. More precisely it aims at understanding biophysical mechanisms that sustain cell proliferation of malfunction. The main examples are biopolymers size repartition, cell self-organisation, tissu growth and cancer development or treatment.

4.2. Geophysical flows and environment

The team will split and give rise to another team ANGE specialised in complex geophysical flows in interaction with environment. Free surface flows as tsunamis, flows in river and costal areas and their ecological consequences are typical examples of applications developed in the team based on algorithms for the free-surface Navier-Stokes equations.

BIGS Project-Team

4. Application Domains

4.1. Data analysis and local regression

Our expertise in data analysis and advanced statistics methods has given raise no a wide number of interdisciplinary collaborations. Among those, here are the most challenging at a scientific level:

(i) Health inequalities: We have recently developed a statistical procedure in order to create a neighborhood socioeconomic index and investigate its influence on health inequalities. The study setting is composed with 3 major French metropolitan areas (Lille, Lyon and Marseille), and we collaborate for this project with a medical team at EHESP (Ecole des Hautes Etudes en Santé Publique) lead by D. Zmirou (see [19] for further details).

(ii) Fetal pathology: An ongoing work concerning local regression techniques is related to Fetal Biometry, an investigation line suggested by a collaboration between our team and the *Centre de Placentologie et Foetopathologie de la Maternité Régionale de Nancy*, under the direction of Professor Bernard Foliguet. The methods involved in Fetal Biometry are usually based on the comparison of some measured values with the predicted values derived from reference charts or equations in a normal population. However, it happens that maternal and pregnancy characteristics have a significant influence on in-utero Fetal Biometry. We will thus produce some models allowing to construct customized fetal biometric size charts. In order to evaluate them, classical and polynomial regression can be used, but they are not the most appropriate to the kind data we have to handle. Hence, we plan to use local regression estimation in order to perform such an evaluation.

(iii) Cohorts analysis: Some medical teams in Nancy are faced with an overwhelming amount of data, for which a serious statistical assessment is needed. Among those let us mention the INSERM team of Pr. Jean-Louis Guéant and the Inria team Orpailleur (particularly with Marie-Dominique Desvignes and Malika Smail). The goal of this collaboration is to extract biological markers for different diseases (cognitive decline; inflammatory intestinal diseases; liver cancer). To this aim, the INSERM team provides us with several data cohorts with a high number of variables and subjects. As in many instances in Biostatistics, one is then faced with a very high dimensional data, from which we hope to extract a reduced number of significant variables allowing to predict the cardiovascular risk accurately. Moreover, these characters should be meaningful to practitioners. The objective for us is thus to design an appropriate variable selection, plus a classification procedure in this demanding context. Let us highlight an original feature of this collaboration: it combines our own data analysis techniques with those developed by the Orpailleur team, based on symbolic tools. We hope that this experience will enrich both points of view and give raise to new methods of data analysis.

4.2. Estimation for complex and biological systems

Our main application for this line of investigation is the photodynamic therapy developed by T. Bastogne. We shall also focus on bacteriophage therapies and subdiffusion within molecules.

(i) Photodynamic therapy. One of the main application we have in mind for our identification problems is to model photodynamic therapy. This promising cancer treatment involves selective uptake and retention of a photosensitive drug in a tumor, followed by irradiation with light at an appropriate wavelength. Photosensitizers are photoactive compounds such as for instance porphyrins and chlorins. The activated photosensitizer is thought to produce singlet oxygen at high doses and thereby to initiate apoptotic and necrotic death of tumor. Due to the lack of response reproducibility, the complexity of interactions between physical, chemical and biological aspects and the high cost of experiments, there is a real demand in good mathematical and physical models which might help to better control and understand PDT responses. We are particularly concerned with modeling the drug uptake into cancer cells, the photoreactions induced by light exposition and tumor growth kinetics.

(ii) Bacteriophage systems. A collaboration between our team, the Mathematics and the Genetics and Microbiology Departments at the *Universitat Autònoma de Barcelona* (UAB) is being set up, focusing on probabilistic aspects of bacteriophage therapies for animal diseases like hemorrhagic septicemia in cattle or atrophic rhinitis in swine. This kind of therapy consists in inoculating a (benign) virus to animals in order to kill the bacteria known to be responsible of the disease. It was in use in the Soviet Union until the 80s, and is now re-emerging, still at an experimental level, due to the progressive slowdown in antibiotic efficiency.

Within this context, our analysis of a noisy predator-prey competition modeling the treatment helps to calibrate and to understand better the behavior of the system in terms of fluctuations around an equilibrium. Note that our preliminary contacts with the Genetics and Microbiology Departments at UAB also open the way to a particle model in order to represent the couple bacteria/virus living on a surface.

(iii) Subdiffusion into molecules. Our purpose here is a better understanding of the phenomena observed in nanoscale Biophysics, as explained in the series of papers [52]. The technological advances in nanoscale technologies allow the observation of single molecules, and thus the description of newly observed phenomenon. A typical example of this new kind of observation is given by the fluctuations in the folding of a protein-enzyme compound called *Fre*, which is involved in the DNA synthesis of the (canonical) bacterium *E. Coli*.

More specifically, the paper [52] advocates for modeling this folding fluctuations by means of a Volterra type equation driven by a fractional Brownian motion. This convincing model is based on some experimental and physical evidences, and have also been observed in a wide number of recent biological experiments. However, the model exhibited in [52] also raises some unsolved questions: some stochastic equations appearing in the models are not properly defined and their long time behavior is still mysterious. The lack of a method in order to simulate and estimate coefficients of these equations on a solid mathematical ground should also be mentioned. This is the kind of topic we wish to address, for which a preliminary contact with S. Kou and N. Pillai (Princeton University, USA) has been established.

(iv) Osteoporosis. During the year 2011-2012, C. Lacaux has been visiting the MAP 5 (Paris Descartes University) laboratory and joined the ANR Project MATAIM (Modèles Anisotropes de Textures. Applications à l'Imagerie Médicale). This project, which involves both mathematicians and practitioners, is in particular interested in the osteoporosis diagnostic. The paper [34] is a first step in the direction of modeling trabecular bone x-ray images by some operator scaling fields. Actually the estimation of the matrix, which characterizes the anisotropy of the model, is crucial for practical purposes. Hermine Biermé (Paris Descartes Univesity) and Céline Lacaux are working on this problem using quadratic variations. Once the problem of estimation is solved, they plan a comparison of the theoretical model with real data provided by our Biologist colleagues of the MATAIM project. If the model corresponds to real data (as suggested in [34]), this approach may help for the diagnostic of osteoporosis: a numerical study has to be performed in order to find the parameter value which characterizes osteoporosis.

BIOCORE Project-Team

4. Application Domains

4.1. Bioenergy

Finding sources of renewable energy is a key challenge for our society. We contribute to this topic through two main domains for which a strong and acknowledged expertise has been acquired over the years. First, we consider anaerobic digesters, the field of expertise of the LBE members of the team, for the production of methane and/or biohydrogen from organic wastes. The main difficulty is to make these processes more reliable and exploit more efficiently the produced biogas by regulating both its quality and quantity despite high variability in the influent wastes. One of the specific applications that needs to be tackled is the production of biogas in a plant when the incoming organic waste results from the mixing of a finite number of substrates. The development of control laws that optimize the input mix of the substrates as a function of the actual state of the system is a key challenge for the viability of this industry.

The second topic consists in growing microalgae, the field of expertise of the LOV members of the team, to produce biofuel. These microorganisms can synthesize lipids with a much higher productivity than terrestrial oleaginous species. The difficulty is to better understand the involved processes, which are mainly transient, to stimulate and optimize them on the basis of modeling and control strategies. Predicting and optimizing the productivity reached by these promising systems in conditions where light received by each cell is strongly related to hydrodynamics, is a crucial challenge.

Finally, for the energy balance of the process, it is important to couple microalgae and anaerobic digestion to optimize the solar energy that can be recovered from microalgae, as was explored within the ANR Symbiose project (2009-2012) [72].

4.2. CO_2 fixation and fluxes

Phytoplanktonic species, which assimilate CO_2 during photosynthesis, have received a lot of attention in the last years. Microalgal based processes have been developed in order to mitigate industrial CO_2. As for biofuel productions, many problems arise when dealing with microalgae which are more complex than bacteria or yeasts. Several models have been developed within our team to predict the CO_2 uptake in conditions of variable light and nitrogen availability. The first modeling challenge in that context consists in taking temperature effects and light gradient into account.

The second challenge consists in exploiting the microalgal bioreactors which have been developed in the framework of the quantification of carbon fluxes between ocean and atmospheres. The SEMPO platform (simulator of variable environment computer controlled), developed within the LOV team, has been designed to reproduce natural conditions that can take place in the sea and to accurately measure the cells behavior. This platform, for which our team has developed models and control methods over the years, is an original and unique tool to develop relevant models which stay valid in dynamic conditions. It is worth noting that a better knowledge of the photosynthetic mechanisms and improved photosynthesis models will benefit both thematics: CO_2 mitigation and carbon fluxes predictions in the sea.

4.3. Biological control for plants and micro-plants production systems

This work concentrates on the protection of cultures of photosynthetic organisms against their pests or their competitors. The forms of cultures that we study are crop and micro-algae productions. In both cases, the devices are more or less open to the outside, depending on the application (greenhouse/field, photobioreactor/raceway) so that they may give access to invading species which can be harmful to the cultures; we opt for protecting the culture through the use of biocontrol agents which are, generically, natural enemies of these noxious populations [9].

In crop production, biological control is indeed a very promising alternative to pesticide usage; the use of predators, parasitoids or pathogens of crop pests in order to fight them has many advantages with respect to environmental protection, health of the consumers and the producers, the limited development of resistance (compared to chemicals),... It is however not widespread yet because it often lacks efficiency in real-life crop production systems (while its efficiency in the laboratory is much higher) and can fail to be economically competitive. Our objective is to propose models that would help to explain which factors are locks that prevent the smooth transition from the laboratory to the agricultural crop as well as develop new methods for the optimal deployment of the pests natural enemies.

Microalgae production is faced with exactly the same problems since predators of the produced microalgae (e.g. zooplankton) or simply other species of microalgae can invade the photobioreactors and outcompete or eradicate the one that we wish to produce. Methods need therefore to be proposed for fighting the invading species; this could be done by introducing predators of the pest and so keeping it under control, or by controling the conditions of culture in order to reduce the possibility of invasion; the design of such methods could greatly take advantage of our knowledge developed in crop protection since the problems and models are related.

4.4. Biological depollution

These works will be carried out with the LBE , mainly on anaerobic treatment plants. This process, despite its strong advantages (methane production and reduced sludge production) can have several locally stable equilibria. In this sense, proposing reliable strategies to stabilize and optimise this process is a key issue. Because of the recent (re)development of anaerobic digestion, it is crucial to propose validated supervision algorithms for this technology. A problem of growing importance is to take benefit of various waste sources in order to adapt the substrate quality to the bacterial biomass activity and finally optimize the process. This generates new research topics for designing strategies to manage the fluxes of the various substrate sources meeting at the same time the depollution norms and providing a biogas of constant quality. In the past years, we have developed models of increasing complexity. However there is a key step that must be considered in the future: how to integrate the knowledge of the metabolisms in such models which represent the evolution of several hundreds bacterial species? How to improve the models integrating this two dimensional levels of complexity? With this perspective, we wish to better represent the competition between the bacterial species, and drive this competition in order to maintain, in the process, the species with the highest depollution capability. This approach, initiated in [103] must be extended from a theoretical point of view and validated experimentally.

4.5. Experimental Platforms

To test and validate our approach, we use experimental platforms developed by our partner teams; these are highly instrumented for accurately monitoring the state of biological species:

- A photobioreactor (SEMPO) for Lagrangian simulation of the dynamical environment of marine microalgae (LOV) with computer controlled automata for high frequency measurement and on-line control. This photobioreactor is managed by Amélie Talec.
- Several pilot anaerobic digesters that are highly instrumented and computerized and the algotron, that is the coupling of a digester and a photobioreactor for microalgae production (LBE). Eric Latrille is our main contact for this platform at LBE.
- Experimental greenhouses of various sizes (from laboratory to semi-industrial size) and small scale devices for insect behavior testing at ISA.

Moreover, we may use the data given by several experimental devices at EPI IBIS/ Hans Geiselmann Laboratory (University J. Fourier, Grenoble) for microbial genomics.

CARMEN Team

4. Application Domains

4.1. Scientific context: the Liryc

Our fields of application are naturally: electrophysiology and cardiac physiopathology at the tissue scale on one side; medical and clinical cardiology on the other side.

The team's research project is part of the IHU Liryc project, initiated by Pr. M. Haissaguerre. It is concerned by the major issues of modern electrocardiology: atrial arrhythmias, sudden death due to ventricular fibrillation and heart failure related to ventricular dyssynchrony.

We aim at bringing applied mathematics and scientific computing closer to biomedical research applied to cardiac rhythmology and clinical cardiology. It aims at enhancing our fundamental knowledge of the normal and abnormal cardiac electrical activity, of the patterns of the electrocardiogram; and we will develop new simulation tools for training, biological and clinical applications.

4.2. Basic experimental electrophysiology

Our modeling is carried out in coordination with the experimental teams from the Liryc. It will help to write new concepts concerning the multiscale organisation of the cardiac action potentials and will serve our understanding in many electrical pathologies:

- At the atrial level, we apply our models to understand the mechanisms of complex arrythmias and the relation with the heterogeneities at the insertion of the pulmonary vein.
- At the ventricula level, we focus on (1) modeling the complex coupling between the Purkinje network and the ventricles and (2) modeling the structural heterogeneities at the cellular scale, taking into account the complex organisation and disorganisation of the myocytes and fibroblasts. Point (1) is supposed to play a major role in sudden cardiac death and point (2) is important in the study of infarct scars for instance.

4.3. Cardiac electrical signals

The Liryc use, on a daily basis and in the clinical context, complex electrical imaging systems, like intracardiac catheters and the CardioInsight vest with 252 body syrface electrodes.

The numerical models can guide the analysis of these signals and conversely, the models can be guided by the signals.

Other applied questions can be addressed by modeling, like the nature of the various electrical signals measured by catheters, that heavily depends on the nature and spatial localisation of the electrodes.

DRACULA Project-Team

4. Application Domains

4.1. Normal hematopoiesis

4.1.1. Introduction

Modelling normal hematopoiesis will allow us to explore the dynamical appearance of the various cell types, originating from the stem cell compartment, through the bone marrow development up to the blood stream. The differentiated cell types will both fulfill physiological functions, and play a key role on the feedback control on homeostasis (balance of the system) in their own lineages. We will describe the hematopoiesis from three different points of view:

- The initial cell type, the hematopoietic stem cell (HSC);
- The lineage choice question;
- Three differentiated lineages that are responsible for specific function, namely oxygen transport, immune response and coagulation.

The basic mechanisms of our modelling approach are as follows:

- Any cell type can have two possibilities at each time step: to divide or to die.
- At any division step, the cell can either give rise to two daughter cells which are identical to the mother cell (self-renewal) or that are more advanced in their differentiation.

All these processes will be first modelled at the cellular level. In parallel, we will develop models of intra-cellular molecular networks (as some proteins controlling the cell cycle) influencing this decision making process, so as to be able to describe both micro-to-macro effects (molecules influencing the global cell behaviour) as well as macro-to-micro effects (like the global state of the cell population influencing the molecular behaviour).

4.1.2. Hematopoietic stem cells (HSC)

Although widely studied by biologists, HSC are still poorly understood and many questions remain open: How fast and how frequently do they divide? How many of them are in the bone marrow and where? How is their behaviour modified under stress conditions such as blood loss or transfusion?

Our modelling approach will be based on two methods: deterministic and stochastic differential equations with delays (discrete and distributed), on one hand, and the DPD method using the individual based modelling on the other hand. The differential equation models based on the work initiated by Mackey [38] will describe the HSC compartment in normal conditions and the behaviour of these cells under some stress. The DPD method, as a complementary approach, will emphasize the spatial regulation of stem cell behaviour, and we will focus our attention to give a possible answer regarding their location in the bone marrow and the roles of the niche, their number in the system, their possible role under stress (that is their reaction under the different feedback controls).

4.1.3. Lineage choice

Positive and negative feedbacks in intra-cellular regulatory networks create a bistable or multistable situation where different cell populations can co-exist. This allows the production of different blood cells beginning from stem cells. It is an important property of hematopoietic cell populations, which is not yet completely understood. We will focus on the erythroid/myelomonocytic choice, which is governed by a balance of lineage-affiliated transcription factors, such as GATA1 and PU.1. How the ratios of lineage-determining transcription factors stabilize progenitor cells and resolve their indeterminacy to commit them to discrete, mutually exclusive fates remains unexplained.

We will analyze the dynamics of a binary fate decision governed by a gene-circuit containing auto-stimulation and cross-inhibition, as embodied by the GATA1-PU.1 paradigm. We will use mathematical models based on ordinary and partial differential equations and individually based modelling to study fundamental properties of hematopoiesis and its quantitative characteristics. We will also explore the fate decision process from a stochastic point of view.

4.1.4. Blood cell functions

(i) O2 transport: red lineage

O_2 transport is provided by red blood cells (RBC) also called erythrocytes. Many different stages of maturity (including progenitors, precursors, reticulocytes and erythrocytes) are necessary to achieve the complete formation of RBC. These latter are then released in the blood stream where they transport oxygen. The whole process is tightly dependent on a robust well-balanced equilibrium called homeostasis.

It has been shown in the 1990's that apoptosis is regulated by EPO, a growth factor released by the kidneys under hypoxia. But also, under severe stress (like an important blood loss) some other molecules known as glucocorticoids can be released leading to an increase of the self-renewing rate for each generation. This led to the formulation of a first model, demonstrating the role of self-renewal.

The study of the red blood cell lineage will involve different scale levels, from the molecular one, with the effects of the hormones on the surface and internal parts of the cell, the cell contacts in each stage of RBC formation, and the red branch population in its whole with all the interactions taken into account (see Figure 2) in normal and stress conditions.

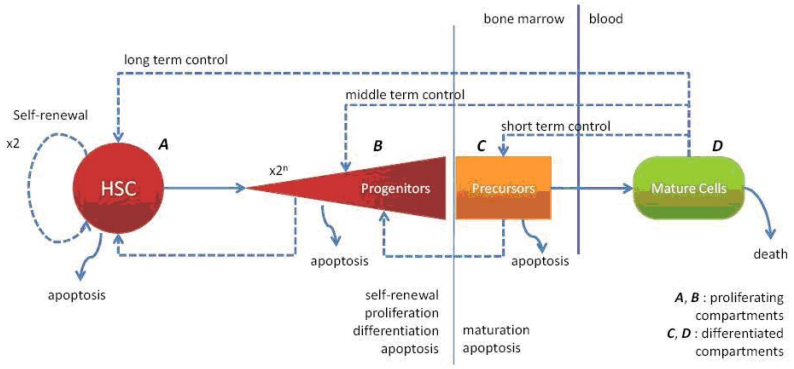

Figure 2. Scheme of Erythropoiesis Modelling. Without considering explicitly growth factor mediated regulation, all controls (proliferation, self-renewal, differentiation, apoptosis) are mediated by cell populations (dashed arrows). Mature cells can either regulate immature (HSC, progenitors) or almost mature (precursors) cells, precursors may act on progenitor dynamics, etc..

In order to couple the cellular behaviour to explicit molecular events, we will describe the events through a molecular network that is based upon the work of [42]. A first version of this model is shown in Figure 3 .

(ii) Immune response

We will focus on the production of T-cells during an immune response. This represents an important activity of the lymphoid branch, part of leucopoiesis (white blood cell production). Several models of the myeloid branch of leucopoiesis have been investigated in the frame of specific diseases (for instance cyclical neutropenia ([37], [34]), chronic myelogenous leukemia [39]).

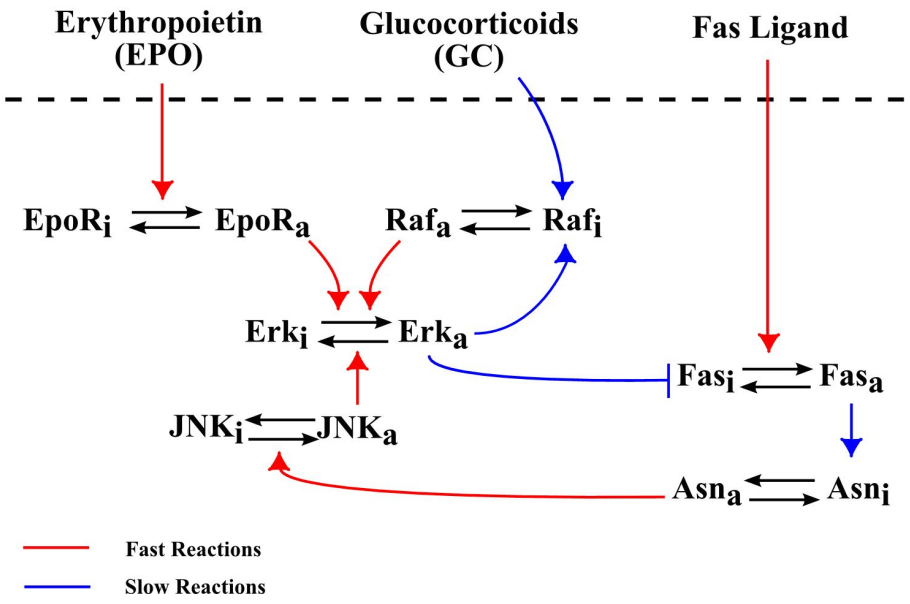

Figure 3. A schematic description of the intra-cellular molecular events that are relevant for decision making in an erythroid progenitor. The non active form of the protein is labeled i, the active form a. Blue lines indicate transcriptional regulation, red lines indicate biochemical regulation.

Time evolution of T-cell counts during an infection is well known: following the antigen presentation, the number of cells quickly increases (expansion), then decreases more slowly (contraction) and stabilizes around a value higher than the initial value. Memory cells have been produced, and will allow a faster response when encountering the antigen for a second time. Mechanisms that regulate this behaviour are however not well known.

A recent collaboration just started with immunologists (J. Marvel, Ch. Arpin) from the INSERM U851 in Lyon, who provide experimental data that are essential to assess the significance of models, based on strongly nonlinear ordinary differential equations, that can be proposed for T-cell production (Figure 4). By considering molecular events leading to cell activation when encountering a virus, we will propose a multi-scale model of the immune response.

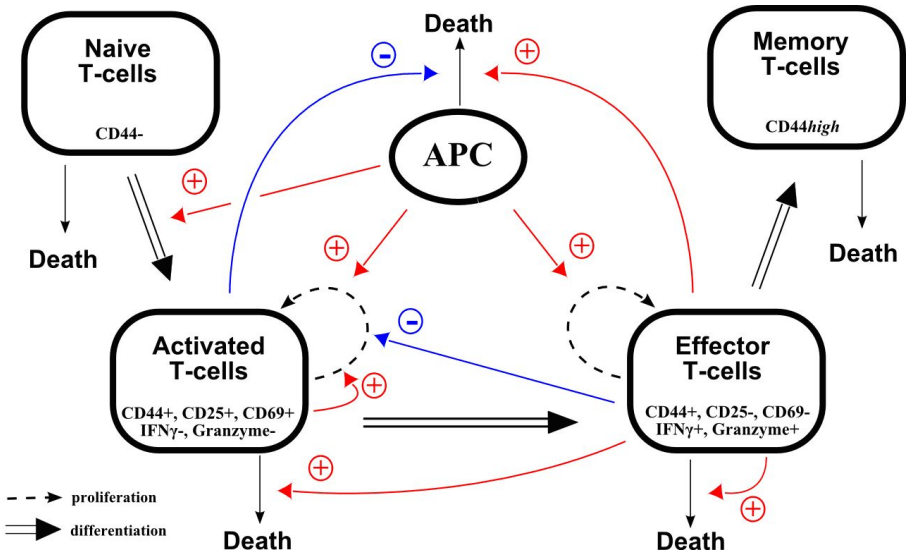

Figure 4. Model of the immune response resulting in the generation of CD8 memory T cells. The response starts with a viral infection resulting in the presentation of viral antigens through antigen presenting cells (APC) to naive T-cells. These latter, once activated, differentiate into activated cells which, under specific feedback loops will either die, differentiate into effector cells or self-renew. Differentiation of effector cells (killer cells) will result in the production of memory cells.

(iii) Coagulation: platelet lineage

Thrombopoiesis, the process of production and regulation of platelets, is similar to erythropoiesis although important differences are observed. These two processes have an immature progenitor (MEP) in common. Platelets are involved in blood coagulation, and can be the source of blood diseases (thrombopenia, thrombocytosis). Their production is mainly regulated by thrombopoietin (TPO), a growth factor similar to EPO.

It is important to mention that very few experimental data exist in the literature, and mathematical modelling of thrombopoiesis did not attract so much attention in the past 20 years. However, collaboration with some leading hematologists in this domain will allow us to get updated and new data regarding this process.

Deterministic models, in the form of structured transport partial differential equations, will be proposed to describe platelet dynamics, through the description of HSC, megakaryocytic progenitor and megacaryocyte (platelet precursor) compartments. Circulating TPO, regulated by platelets, will induce feedback loops in

thrombopoiesis, and we will investigate the dynamics of platelet production and emergence of platelet-related diseases.

4.2. Pathological hematopoiesis

The knowledge of hematopoiesis and related diseases has evolved to become a great deal in the past years, and Mackey's previous models (ref. [32]) do not allow us to correctly answer current questions that are clearly oriented toward the investigation of cell signalling pathways. These models nevertheless bring relevant ideas about the essential features of such modelling. It is also noteworthy that even though models of hematopoiesis have existed for quite a long time, their application to questions of explanation and prediction of hematopoiesis dynamics that are encountered in the clinic is still not sufficiently frequent, even though much progress has been achieved in the cooperation between hematologists and mathematicians [40]. This is in the optic of testable experimental predictions that the multi-scale model for pathological hematopoiesis will be developed. For instance, we will concentrate on myeloid leukemias (CML and AML) and their treatment.

4.2.1. Leukemia Modelling

(i) Chronic Myeloid Leukemia

The strong tyrosine kinase activity of the BCR-ABL protein is the basis for the main cell effects that are observed in CML: significant proliferation, anti-apoptotic effect, disruption of stroma adhesion properties, genomic instability. This explains the presence in CML blood of a very important number of cells belonging to the myeloid lineage, at all stages of maturation.

We will consider models based on ordinary differential equations for the action of the main intra- and extra-cellular proteins involved in CML (as BCR-ABL protein), and of transport equations (with or without delay, physiologically structured or not to represent healthy and leukemic cell populations, take into account many interactions between proteins (especially BCR-ABL), cells (anti-apoptotic effect, etc.), and their environment (disruption of stroma adhesion properties, for example). Transport pertains thus to cells from one compartment (or a group of compartments) to another compartment, with a determined speed of aging or maturation. These compartments may be detailed or not: the less mature are stem cells, then progenitor cells, etc.

(ii) Acute Myeloid Leukemia

The natural history of CML leads to its transformation ("blast crisis") in acute myeloid leukemia (AML), following supplementary genetic alterations that produce a maturation arrest (myeloid in $3/4$ of cases, lymphoid in $1/4$ of cases, confirming the insult to pluripotent stem cells), leading to an accumulation of immature cells in the bone marrow and in the general circulation, resulting in deep medullary impairment and fast fatal outcome, in spite of chemotherapy. This phenomenon is the same as the one observed in de novo AML, i.e., AML without a previous chronic phase.

The different modelling methods of AML will be similar to the ones described for CML, with some exceptions: the appearance of BCR-ABL mutations, which are not relevant in the case of AML, the appearance of a gene (spi-1) involved in the differentiation arrest, and constitutive activation of EPO receptor or Kit activating mutations promote proliferation and survival. This explains the accumulation of immature cells in the bone marrow and in the blood stream.

4.2.2. Treatment

As far as treatment of pathological hematopoiesis is concerned, two main strategies currently exist that aim at slowing down or eliminating damaged cell proliferation. The first of these strategies consists in launching the apoptotic process during the cell division cycle. This process is activated, for example when the cell is unable to repair damages, e.g., after exposure to cytostatic drugs. A typical example is apoptosis induced by chemotherapy-induced DNA damage: The damage is recognised by the cell, which then activates the sentinel protein p53 ("guardian of the genome") that arrests the cell cycle to allow, if possible, damage repair. If the latter is unrecoverable, then p53 activates the endogenous apoptotic processes.

The second strategy aims at pushing damaged cells toward the differentiation that has been stopped in the course of their genetic mutation. Since a few years back, a new approach has been developed around the strategy of differentiation therapy. This therapy relies on molecules (growth factors and specific cytokines) that are able to re-initialise the cell differentiation programs that have been modified during malignant transformation. The cancer that is most concerned by the development of this differentiation therapy is AML whose malignant cells present highly undifferentiated features and the ones that present a translocation responsible for the differentiation (PML/RAR of the promyelocytic form, AML1/ETO and CBFbeta/MyH11, involving Core Binding Factors alpha and beta).

Mathematical models based on ordinary differential equations will be developed to describe the action of drugs (in the two cases mentioned above). They will take into account interactions between drugs and their environment. Our goal will be the optimization of possible synergies between drugs acting on distinct cellular targets, and the control of resistances to these treatments as well as their toxicities.

Curative and palliative strategies must take into account the dynamics of healthy and leukemic hematopoietic cells at multiple scales. In time, from optimal scheduling of combination therapy (hours) to avoiding the development of resistances and relapse (months to years). In space, from the stem cell niche to circulating blood. In organization, from gene and signalling networks (JAK/STAT, BCR-ABL) to cell populations and cytokine regulation (EPO, CSFs). Several recent qualitative models have provided insight in the complex dynamics of the disease and the response to treatments. Many of these models focus on the control or regulation processes that promote homeostasis or oscillatory behavior in cell number. However, as A. Morley points out, "once the control-systems features of hematopoiesis are accepted, the ability to construct a model that shows oscillatory behavior, even if the model incorporates the latest advances in hematopoietic cell biology, really adds little new knowledge. Rather, the challenge to modellers would seem to be to provide detailed predictions for the input-output characteristics of the different parts of the various control systems so that these predictions can be tested by experimental hematologists and a truly quantitative description of hematopoiesis can emerge".

We propose for instance, to use models in the form of structured transport partial differential equations (with or without delay, physiologically structured or not) to represent the competition between target, resistant and healthy cell populations. The resulting models to describe the dynamic of these cell populations under the action of drugs are multi-scale systems of the form (Hyperbolic PDE)-ODE or DDE-ODE. For instance, we will develop mathematical models of chronotherapy and pharmacotherapy for CML and AML.

MACS Project-Team

4. Application Domains

4.1. Application domains

Our researches have natural applications in all sectors of the mechanical industry: car and naval industries; aeronautics and space; civil engineering; tires; MEMs and nanotechnologies...

We also actively seek new applications in biotechnologies, although of course the economy and structuring of this sector is not as developed yet.

MASAIE Project-Team

4. Application Domains

4.1. Metapopulation models

Heterogeneity plays an important role in many infectious disease processes. For instance, spatial heterogeneity is a strong determinant of host-parasite relationships. In modeling spatial or geographic effects on the spread of a disease, a distinction is usually made between diffusion and dispersal models. In diffusion models, spread is to immediately adjacent zones, hence the phenomenon of traveling waves can appear. These models traditionally use partial differential equations. However, there are some important situations that cannot be modeled by PDE. This is the case when the space considered is discrete. For example, when we have to consider sparsely populated regions, the human population is located in patches. The organization of human-hosts into well-defined social units such as families, villages or cities, are good examples of patches. Another example arises in the study of the human African Trypanosomiasis. The vector is the tse-tse fly, and it is known that flies take fewer blood meals in villages than in coffee plantations where the villagers work during the day. For such situations where human or vectors can travel a long distance in a short period of time, dispersal models are more appropriate. These models consider migration of individuals between patches. The infection does not take place during the migration process. The situation is that of a directed graph, where the vertices represent the patches and the arcs represent the links between patches. Recently, there has been increased interest in these deterministic metapopulation disease models. We have generalized to n patches the Ross-Macdonald model which describes the dynamics of malaria. We incorporate in our model the fact that some patches can be vector free. We assume that the hosts can migrate between patches, but not the vectors. The susceptible and infectious individuals have the same dispersal rate. We compute the basic reproduction ratio $\mathcal{R}_0$. We prove that if $\mathcal{R}_0 \leq 1$, then the disease-free equilibrium is globally asymptotically stable. When $\mathcal{R}_0 > 1$, we prove that there exists a unique endemic equilibrium, which is globally asymptotically stable on the biological domain minus the disease-free equilibrium.

MASAIE is developing, in the framework of the CAPES-COFECUB project (see international program), a metapopulation model for dengue. This model is for the state of Rio and is using the data of foundation FIOCRUZ.

4.2. Estimating total parasite load in falciparum malaria patients

We give a brief review of the biological features of malaria. Malaria in a human begins with an inoculum of *Plasmodium* parasites (sporozoites) from a female *Anopheles* mosquito. The sporozoites enter the liver within minutes. After a period of asexual reproduction in the liver, the parasites (merozoites) are released in the bloodstream where the asexual erythrocyte cycle begins. The merozoites enter red blood cells (RBC), grow and reproduce over a period of approximately 48 hours after which the erythrocyte ruptures releasing daughter parasites that quickly invade a fresh erythrocyte to renew the cycle. This blood cycle can be repeated many times, in the course of which some of the merozoites instead develop in the sexual form of the parasites : gametocytes. Gametocytes are benign for the host and are waiting for the mosquitoes. An important characteristic of *Plasmodium falciparum*, the most virulent malaria parasite, is sequestration. At the half-way point of parasite development, the infected erythrocyte leaves the circulating peripheral blood and binds to the endothelium in the microvasculature of various organs where the cycle is completed. A measurement of *Plasmodium falciparum* parasitaemia taken from a blood smear therefore samples young parasites only. Physician treating malaria use the number of parasites in peripheral blood smears as a measure of infection, this does not give the total parasite burden of the patient. Moreover antimalarial drugs are known to act preferentially on different stages of parasite development. Our work consists in developing tools for estimating the sequestered parasites and hence the total parasite burden of the patient.

MODEMIC Project-Team

4. Application Domains

4.1. Preservation of water resources

The biological decontamination of wastewater is our main application domain, in the continuation of the long collaboration with the INRA research laboratory LBE. We target applications from the decontamination industry, held by large groups as well as small companies specialized in specific pollutants (for instance in fish farming). We aim also to study connected application domains for

- the aquatic ecology where microorganisms play an important role in the quality of natural water resources,
- the re-use of water in arid climates for countries of North of Africa, within the euro-Mediterranean project TREASURE.

4.2. Microbial ecology of soil

This application domain is more recent for the team members. We target

- the research questions raised by agronomists, about the better understanding of the interactions and the biodiversity of microbial communities in soils, with the help of models and numerical simulations,
- the role of spatial structures on the functions or *ecological services* of microbial ecosystems, notably the soil fertility and the carbon sequestration.

4.3. Control of fermentation processes

Very closely to our studies about wastewater bioreactors and chemostat models, we target applications in fermentation processes:

- either for agro-food products. A typical application is the control of cascade fermenters in the study of wine fermentation with UMR SPO (Sciences Pour l'Œnologie, Montpellier), within the European project CAFE.
- either for the green chemistry. A typical application is the consideration of spatialization in enzymatic models of production of agro-polymers with UMR IATE (Ingénierie des Agro-polymères et Technologies Émergentes, Montpellier).

4.4. Animal digestive ecosystem

Ruminants absorb plant cells, mainly constituted by cellulose, from which the microbial population of their digestive system extracts carbon and energy to provide proteins and energetic molecules. This bio-conversion produces also important quantities of methane, a gas responsible of part of the greenhouse effect (the billion of cows on earth reject 20% of the methane linked to human activities). INRA researchers have shown that this methane production could be reduced by 30% by changing the proportion of fat acids in the their food, that also implies that the composition of their microbial ecosystem is modified.

This application domain of the microbial ecology is at an early stage. URH team (Unité de Recherche sur les Herbivores, Clermont) has developed an artificial rumen that is close to a chemostat, for testing different kind of nutrition diets. Preliminaries contacts have been taken, and a modelling demand has been clearly formulated and will be taken up my MODEMIC. This theme falls into the research priorities for the environment preservation.

NUMED Project-Team

4. Application Domains

4.1. Stroke

Stroke is a very complex pathology, involving many different time scales and phenomena. Numed is currently developping various models to describe some important aspects of stroke.

4.1.1. Inflammation modelling

MA Dronne has designed a first model of inflammation at cellular level, based on ordinary differential equations. To take into account spatial phenomena, a first partial differential equation based model is under study. Together with Taissia Lelekov Boissard (post doc of the ANR contrat "AVC in silico"), they tried to find biological data to parametrize these models, and to build a basis of qualitative facts that must be reproduced by the model.

MA Dronne has also developped a collaboration with the Mario Negri institute (Milano) through the team "inflammation and nervous system diseases" (MG de Simoni). This teams currently runs in vivo experiments in rodent that should provide new data to investigate the temporal evolution of various variables of the model.

The study and validation of these two models of inflammation will continue with the study of in silico experiments which will simulate the action of various anti inflammatory drugs, acting at various levels of the inflammatory reaction, work in common with biologists (INSERM 842, neurooncologie et neuro inflammation, Lyon), with clinicians (Creatis, Umr 5515, Inserm U 630 Lyon).

4.1.2. Free radicals

A first model of free radical synthesis has been initiated by V. Lemesle (post doc of ANR AVC in silico). This model is under developpment with P. Vigneaux. A collaboration begins with Michel Plotkine (EA 2510 pharmacology of cerebral blood flow, Paris 5 university) to get experimental data on the temporal evolution of the various variables of the model.

This model will be used to manage in silico experiments in order to study the effects of various drugs.

4.1.3. Ionic motions

A mechanistic model of ionic motions has already been developped, studied and validated to study in silico the dual role of astrocytes during ischemia, and to study the effect of various ionic chanels blockers in man and roden.

This model is now used to study in silico the effect of the combination of several neuroprotectors acting on ionic channels, transporters or receptors. This work should help to understand antagonist or synergic effects of blockers.

4.1.4. Spreading depression

Spreading depressions are propagative waves which travel in brain during ischemia and which may have a major role in the extension of the ischemic core. Currently 3D computations in real geometry are run to study their speed and the role of brain anatomy in their propagation.

4.1.5. Apoptosis during stroke

A collaboration has begun with Christiane Charriaut Marlangue (INSERM U676, Hopital Robert Debré) to study the apoptotic cascade during stroke.

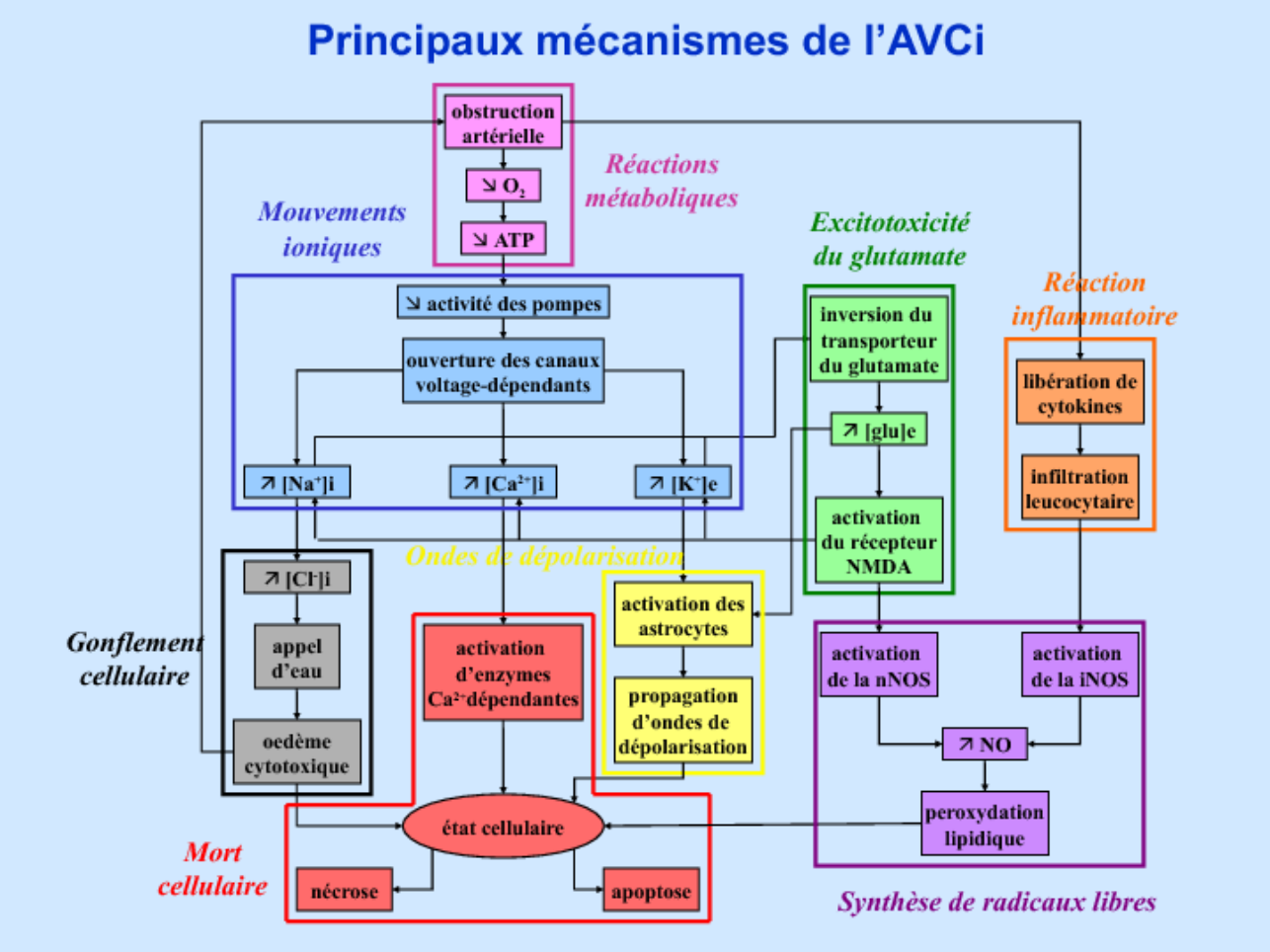

Figure 4. An example of a pdf map reconstructed by using geometrical methods in detecting landmarks

4.2. Oncology

4.2.1. Tumor growth in mice

Through a collaboration with University of Lyon and Lyon-Sud Hospital, we setup several mechanistic models to predict the evolution of tumor growth in mice including the complex biological process of angiogenesis. This work was presented at the eighteen PAGE (population approach group in Europe) meeting in Saint-Petersburg in June.

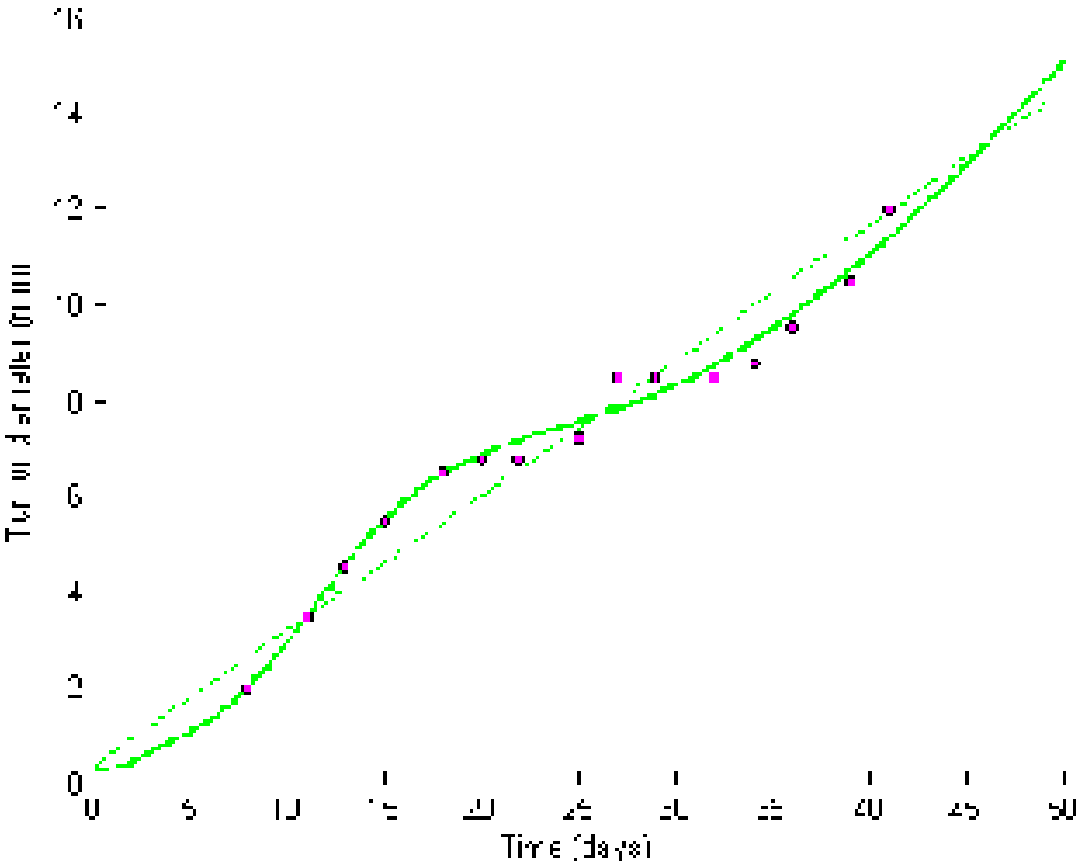

Figure 5. Modeling tumor growth in mice

4.2.2. Radioresistance

Within the framework of the project ETOILE, B. Bernard is part of the team that is building a predictive model of tumor responses to the conventional treatment and irradiation with carbon ions. During 2009, Branka has been working under supervision of Jean-Pierre Boissel and Benjamin Ribba. In collaboration with the group of Claire Rodriguez-Lafrasse (Radiobiology group, Hospital Lyon Sud), her research activities included the

analysis of microarray data from different head and neck cancer cell lines, irradiated with X-ray and carbon ions. They detected differences in the irradiation response of different cancer cell lines that underlie their different radiosensitivities. Within GRAAL project, a lot of radiobiological information will be acquired on a several glioma cell lines and cell lines representing healthy brain tissue. Therefore, our interest is to model the dynamics of the glioma tumor growth and its response to radiation therapy. At the moment, we are working on the estimation of parameters describing tumor growth and diffusion from the MRI images of glioblastoma patients (collaboration with Francois Ducray).

4.3. Virology

In collaboration with MERIAL SA and Edouard Heriot Hospital in Lyon, B. Ribba develops mathematical models to describe the dynamic of Feline immunodeficiency virus (FIV) in infected cats. A translational approach is developed in the context of parameter estimation for complex biologically-based model. He intensively uses mixed-effect modeling approaches and its SAEM algorithm implementation in MONOLIX (Inria Saclay).

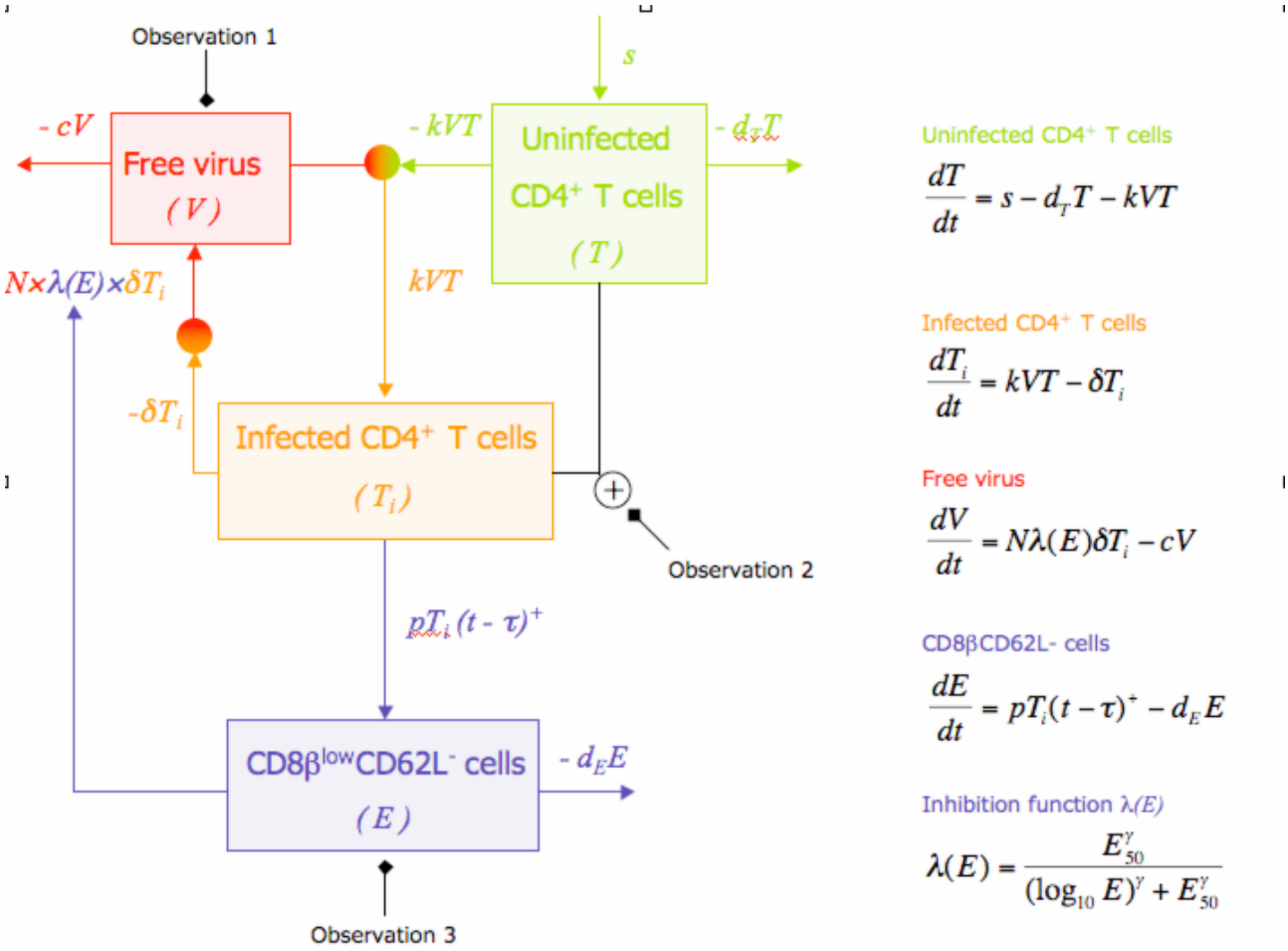

Figure 6. Model of FIV

4.4. Prion.

[In collaboration with M. Doumic (Inria Rocq.), P. Gabriel and B. Perthame (Paris 6) – ANR TOPPAZ]

We study mathematically and numerically the polymerization/fragmentation equation involved in prion aggregation. We have investigated first the case of a size-dependent polymerization rate motivated by recent experiments. We now focus on some issue in optimization of protocol. This is closely related to recent challenges in fitness optimization, and optimal control.

4.5. Atheroma

[In collaboration with N. Meunier (Paris 5)]

Following El Khatib et al. (2007) we have proposed a mathematical model for the inflammatory processes driving the growth of early atherosclerotic plaques. This model is coupled with blood flow, with particular emphasis on the influence of shear stress.

REO Project-Team

4. Application Domains

4.1. Blood flows

Cardiovascular diseases like atherosclerosis or aneurysms are a major cause of mortality. It is generally admitted that a better knowledge of local flow patterns could improve the treatment of these pathologies (although many other biophysical phenomena obviously take place in the development of such diseases). In particular, it has been known for years that the association of low wall shear stress and high oscillatory shear index give relevant indications to localize possible zones of atherosclerosis. It is also known that medical devices (graft or stent) perturb blood flows and may create local stresses favorable with atherogenesis. Numerical simulations of blood flows can give access to this local quantities and may therefore help to design new medical devices with less negative impacts. In the case of aneurysms, numerical simulations may help to predict possible zones of rupture and could therefore give a guide for treatment planning.

In clinical routine, many indices are used for diagnosis. For example, the size of a stenosis is estimated by a few measures of flow rate around the stenosis and by application of simple fluid mechanics rules. In some situations, for example in the case a sub-valvular stenosis, it is known that such indices often give false estimations. Numerical simulations may give indications to define new indices, simple enough to be used in clinical exams, but more precise than those currently used.

It is well-known that the arterial circulation and the heart (or more specifically the left ventricle) are strongly coupled. Modifications of arterial walls or blood flows may indeed affect the mechanical properties of the left ventricle. Numerical simulations of the arterial tree coupled to the heart model could shed light on this complex relationship.

One of the goals of the REO team is to provide various models and simulation tools of the cardiovascular system. The scaling of these models will be adapted to the application in mind: low resolution for modeling the global circulation, high resolution for modeling a small portion of vessel.

4.2. Respiratory tracts

Breathing, or “external” respiration (“internal” respiration corresponds to cellular respiration) involves gas transport though the respiratory tract with its visible ends, nose and mouth. Air streams then from the pharynx down to the trachea. Food and drink entry into the trachea is usually prevented by the larynx structure (epiglottis). The trachea extends from the neck into the thorax, where it divides into right and left main bronchi, which enter the corresponding lungs (the left being smaller to accommodate the heart). Inhaled air is then convected in the bronchus tree which ends in alveoli, where gaseous exchange occurs. Surfactant reduces the surface tension on the alveolus wall, allowing them to expand. Gaseous exchange relies on simple diffusion on a large surface area over a short path between the alveolus and the blood capillary under concentration gradients between alveolar air and blood. The lungs are divided into lobes (three on the right, two on the left) supplied by lobar bronchi. Each lobe of the lung is further divided into segments (ten segments of the right lung and eight of the left). Inhaled air contains dust and debris, which must be filtered, if possible, before they reach the alveoli. The tracheobronchial tree is lined by a layer of sticky mucus, secreted by the epithelium. Particles which hit the side wall of the tract are trapped in this mucus. Cilia on the epithelial cells move the mucous continually towards the nose and mouth.

Each lung is enclosed in a space bounded below by the diaphragm and laterally by the chest wall and the mediastinum. The air movement is achieved by alternately increasing and decreasing the chest pressure (and volume). When the airspace transmural pressure rises, air is sucked in. When it decreases, airspaces collapse and air is expelled. Each lung is surrounded by a pleural cavity, except at its hilum where the inner pleura give birth to the outer pleura. The pleural layers slide over each other. The tidal volume is nearly equal to 500 ml.

The lungs may fail to maintain an adequate supply of air. In premature infants surfactant is not yet active. Accidental inhalation of liquid or solid and airway infection may occur. Chronic obstructive lung diseases and lung cancers are frequent pathologies and among the three first death causes in France.

One of the goals of REO team in the ventilation field is to visualize the airways (virtual endoscopy) and simulate flow in image-based 3D models of the upper airways (nose, pharynx, larynx) and the first generations of the tracheobronchial tree (trachea is generation 0), whereas simple models of the small bronchi and alveoli are used (reduced-basis element method, fractal homogenization, multiphysics homogenization, lumped parameter models), in order to provide the flow distribution within the lung segments. This activity has been carried out in the framework of successive research programs: RNTS “R-MOD” until 2005, ACI “le-poumon-vous-dis-je” until 2007 and ANR M3RS until 2013.

4.3. Cardiac electrophysiology

The purpose is to simulate the propagation of the action potential in the heart. A lot of works has already been devoted to this topic in the literature (see *e.g.* [71], [75], [74] and the references therein), nevertheless there are only very few studies showing realistic electrocardiograms obtained from partial differential equations models. Our goal is to find a compromise between two opposite requirements: on the one hand, we want to use predictive models, and therefore models based on physiology, on the other hand, we want to use models simple enough to be parametrized (in view of patient-specific simulations). We are now working on using our ECG simulator to address the inverse problem of electrocardiology. In collaboration with the Macsproject-team, we are working on the electromechanical coupling in the myocardium. We are also interested in various clinical and industrial issues related to cardiac electrophysiology. In particular, we collaborated with ELA Medical company (pacemaker manufacturer, Sorin group).

SISYPHE Project-Team (section vide)

VIRTUAL PLANTS Project-Team (section vide)