



HAL
open science

MIA - Mathématiques et informatique appliquées

Rapport Hcéres

► **To cite this version:**

Rapport d'évaluation d'une entité de recherche. MIA - Mathématiques et informatique appliquées. 2009, AgroParisTech - Institut des sciences et industries du vivant et de l'environnement, Institut national de la recherche agronomique - INRA. hceres-02032799

HAL Id: hceres-02032799

<https://hal-hceres.archives-ouvertes.fr/hceres-02032799v1>

Submitted on 20 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport d'évaluation

Unité de recherche :

Mathématiques et Informatique Appliquées

UMR 518

d'AgroParisTech



mars 2009



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport d'évaluation

Unité de recherche

Mathématiques et Informatique Appliquées
d'AgroParisTech



Le Président
de l'AERES

Jean-François Dhainaut

Section des unités
de recherche

Le Directeur

Pierre Glorieux

mars 2009



Rapport d'évaluation



L'Unité de recherche :

Nom de l'unité : Mathématiques et Informatique Appliquées

Label demandé : UMR

N° si renouvellement : 518

Nom du directeur : M. Stéphane ROBIN

Université ou école principale :

AgroParisTech

Autres établissements et organismes de rattachement :

INRA

Date(s) de la visite :

10 février 2009



Membres du comité d'évaluation

Président :

M. Anestis ANTONIADIS, Université Joseph Fourier, Grenoble

Experts :

M. Eric MOULINES, Ecole Nationale Supérieure de Telecom, Paris

M. Jean CLOBERT, CNRS, Moulis Saint-Girons

Expert(s) représentant des comités d'évaluation des personnels (CNU, CoNRS, CSS INSERM, représentant INRA, INRIA, IRD...) :

Au titre de la CSS INRA, M. Alain DENISE

Au titre de la CNECA, M. Gilles CARAUX

Observateurs

Délégué scientifique de l'AERES :

M. Pascal AUSCHER

Représentant de l'université ou école, établissement principal :

Mme Mariane LEFORT, Directrice Scientifique d'AgrosParisTech

Représentant(s) des organismes tutelles de l'unité :

M. Bruno GOFFINET, Directeur Département MIA, INRA



Rapport d'évaluation

1 • Présentation succincte de l'unité

- L'UMR 518 est structurée en deux équipes comptabilisant au total un effectif de 30,2 (ETP), composé de 9 enseignants-chercheurs permanents dont 2 PR, 2 ingénieurs du GREF assimilés à des PR AgroParisTech et 5 MCF, 1,5 chercheurs INRA dont 1 DR et 0,5 CR, 2 chercheurs invités, 1,8 ingénieurs, 15 doctorants et 0,2 technicien et 0,7 administratif.
- Six chercheurs sont HDR ayant tous encadré ou encadrent actuellement des thèses.
- 1 HDR et 7 thèses ont été soutenues lors des 4 dernières années. La durée moyenne des thèses est de 3 ans. 2 autres thèses ont été soutenues après la rédaction du bilan de l'unité et il y a actuellement 12 thèses en cours.
- Tous les permanents sont publiants.

2 • Déroulement de l'évaluation

L'évaluation a commencé en présence de tout le personnel de l'unité par une présentation synthétique, faite par son directeur, des activités et réalisations des équipes du laboratoire MIA, appuyée par le dossier de demande de reconnaissance communiqué préalablement aux experts. Les membres du comité ont demandé des éclaircissements et fait les commentaires qu'ils ont jugés opportuns. Ont suivi cinq exposés scientifiques par des chercheurs du laboratoire et en guise de conclusion une présentation sur l'intégration de la composante apprentissage dans l'unité ainsi qu'une présentation du projet pour le prochain quadriennal. L'après-midi le comité rencontra les personnels administratifs et techniques, puis s'entretint avec les doctorants et post-docs en séjour au laboratoire. La fin de la visite fut consacrée à un entretien avec les tutelles. Les membres du comité se sont enfin réunis à huis clos pour un bilan de leurs impressions et pour s'entendre sur le contenu du reste du rapport.

3 • Analyse globale de l'unité, de son évolution et de son positionnement local, régional et européen

Les paragraphes qui suivent proposent une synthèse des avis exprimés par les membres du comité d'évaluation. Chacun des évaluateurs a fourni un rapport succinct contenant ses impressions sur les activités de l'unité de recherche dans son ensemble ainsi qu'un avis plus détaillé sur le projet de développements futurs. La synthèse qui suit a été effectuée par le Président du comité et «amendée» par tous les évaluateurs.

Les évaluateurs tiennent en premier lieu à souligner la clarté et l'aspect synthétique du dossier qui leur a été présenté, décrivant de façon organisée les activités de recherche de l'unité. La seule remarque éventuelle à formuler est le manque d'indicateurs bibliographiques qui auraient simplifié le travail du lecteur. Cependant ces éléments ont été fournis lors de l'exposé oral.



Le laboratoire de mathématiques et informatique appliquées (MIA), est une petite unité associée à l'INRA, et résulte de la fusion au sein d'AgroParisTech de deux équipes aux thématiques distinctes, l'une «Statistique et Génome» centrée autour des applications de la statistique à la génomique et l'autre «MORSE» aux applications de la statistique aux sciences de l'environnement. Un regroupement physique, dans des locaux communs, s'est opéré durant la période examinée lors de cette évaluation. Les choix thématiques, et la nature des travaux de cette UMR, sont tout à fait pertinents au sein d'un établissement formant des ingénieurs généralistes du vivant. Cette structure est adaptée à de jeunes enseignants chercheurs et leur offre un cadre propice à la recherche, notamment en statistique ou en informatique, tout en ayant une finalité très bien connectée aux autres disciplines et aux besoins des étudiants. Les enseignants chercheurs sont très investis dans les enseignements de statistique et d'informatique de l'Ecole. Ils participent également fortement à l'organisation de formations spécialisées, à destination doctorale ou de chercheurs, notamment dans l'analyse des données transcriptomiques.

La production scientifique est d'excellente qualité (160 publications listées sur la période et plusieurs développements de logiciels) mais avec une certaine disparité entre les deux équipes, disparité qui se retrouve également dans les comptes-rendus des activités et des projets d'avenir. Tous les chercheurs de l'unité sont « publiants », ce qui est un signe de la vitalité scientifique du laboratoire. Le laboratoire occupe une place nationale centrale d'un point de vue thématique qui le conduit à interagir fortement avec plusieurs autres laboratoires de la région parisienne, notamment le laboratoire MIG (Mathématiques et Informatique pour la Génomique), unité propre de l'INRA, le laboratoire Statistique et Génome de l'Université d'Evry, l'équipe SELECT de l'INRIA et l'Institut Curie entre autres. Il anime, ou participe à, plusieurs séminaires ou groupes de travail parisiens. Globalement, le laboratoire a une très bonne visibilité dans sa communauté de référence et jouit d'une reconnaissance nationale et internationale, tant auprès des statisticiens que des biologistes et des bioinformaticiens.

4 • Analyse équipe par équipe et par projet

L'équipe « Statistique et Génome » est remarquable dans son positionnement et dans son activité. Les approches sont originales et innovantes, traitées avec rigueur et pertinence. Ce groupe est très actif, bien inséré dans la communauté internationale, et a un rayonnement important. Il est animé par deux chercheurs seniors très productifs, qui jouissent d'une grande reconnaissance nationale et internationale. Les thèmes de recherche de l'équipe sont tous motivés par l'application de méthodes statistiques inférentielles pour l'analyse des données de la génomique et de la post-génomique, notamment la recherche de motifs exceptionnels dans les séquences biologiques, l'analyse des données de microarrays, l'étude des réseaux d'interactions biologiques. Des avancées très significatives ont été obtenues dans ces trois domaines. Le thème « réseaux d'interactions biologiques » est plus récent au sein de l'unité. L'analyse statistique et probabiliste des réseaux est un sujet en plein essor, qui se développe aux frontières entre les communautés d'informatique théorique, de probabilité appliquée et de statistique. L'équipe « statistique et génome » joue un rôle pionnier en France dans l'application de ces techniques aux réseaux métaboliques, de régulation de gènes et aux interactions entre protéines. La composante informatique de cette équipe est récente. Elle remonte au recrutement d'un professeur travaillant dans le domaine de l'apprentissage informatique. Les approches dans ce domaine sont étroitement liées aux statistiques, laissant présager des collaborations fructueuses à condition que des objectifs communs soient définis dans les projets futurs de l'équipe.

L'équipe « Modélisation et risque en statistique environnementale (MORSE) » a une assise beaucoup plus large et diversifiée. Les domaines d'applications visés par les travaux de cette équipe sont beaucoup plus vastes que ceux de l'équipe « Statistique et Génome » au point que l'on peut se demander si cette dispersion n'est pas un facteur d'éparpillement nuisible à la visibilité et la cohérence de l'équipe, qui ne jouit d'ailleurs pas, dans la communauté nationale et internationale, de la même reconnaissance. Cette équipe composée de 4 permanents développe trois thèmes de recherche. Le premier axe concerne la modélisation par raisonnement conditionnel bayésien en sciences de l'environnement. Alors que l'utilisation des méthodes bayésiennes pour l'inférence des méthodes hiérarchiques est un sujet de recherche aujourd'hui bien balisé, le groupe devrait renforcer ses liens avec des groupes utilisateurs de ces méthodes, pour produire un recherche «réellement» exploitable en écologie.



Le risque est grand, dans ce type de recherche, de produire des modèles dont la pertinence (valeur diagnostique ou prédictive) soit faible (problème soulevé pendant la visite du comité). Le niveau des publications (choix des revues) est moyen et devrait être l'objet d'un effort significatif lors du prochain quadriennal (viser des revues de meilleur niveau). Le deuxième axe est la modélisation spatiale pour l'environnement. Ce thème est porté par deux MdC dont un en détachement ; les résultats sont satisfaisants mais le positionnement du groupe par rapport au reste de l'équipe et l'originalité de ses thématiques par rapport à la recherche en statistique spatiale en France et dans le monde n'apparaissent pas clairement dans le rapport. Le troisième thème est l'étude théorique de la dépendance pour des processus temporels. Ce thème de recherche est porté par un enseignant chercheur. Ce thème a disparu de la prospective du groupe et n'a pas été présenté lors du comité. Il a pourtant donné lieu à des publications dans des revues internationales de bon niveau (ESAIM P&S, Markov Proc. And Rel. Fields, JASA) et à un livre (écrit en collaboration et publié à Springer Series in Statistics). Si les recherches dans cette thématique sont intéressantes et solides sur le plan scientifique, il est un peu délicat de comprendre leurs relations avec l'ensemble des travaux du laboratoire (de toute évidence, de telles relations pourraient se développer, à l'intérieur de l'équipe MORSE - en particulier avec l'axe de recherche sur les statistiques spatiales, ou avec l'équipe « Statistique et Génome », notamment pour la modélisation de séquences biologiques).

5 • Analyse de la vie de l'unité

La visite au cours de l'évaluation a permis aux membres du comité de se rendre compte que les contacts humains y sont très chaleureux et qu'il y règne une excellente entente. La petite taille du laboratoire favorise les contacts et sont une force en termes de réactivité scientifique et d'organisation. Les locaux sont exigus et éclatés. Il serait tout à fait souhaitable de réduire cet éclatement. La direction est collégiale et semble fonctionner à la satisfaction de tous. Le soutien aux chercheurs est remarquable, tant au niveau du nombre que de la qualité des personnes et de leur enthousiasme vis-à-vis des activités et de la mission du laboratoire. Le comité a remarqué un faible investissement de l'INRA dans les moyens humains de l'UMR même si son directeur est DR - INRA.

6 • Conclusions

— Points forts :

- la qualité des recherches de l'équipe Statistique et Génome
- le dynamisme de l'unité
- projets de développement en cours de grande qualité

Le comité d'évaluation a particulièrement apprécié la qualité des recherches effectuées au Laboratoire de Mathématiques et Informatique Appliquée (nombreuses publications dans des journaux de premier plan et plusieurs développements de logiciels). Il a tout aussi apprécié le dynamisme de l'unité : sans délaisser les domaines qui ont contribué à son rayonnement, elle a su développer avec succès de nouvelles thématiques et dispose d'enseignants chercheurs de premier plan international tant en mathématiques appliquées qu'en informatique et bioinformatique. Les projets de développement en cours (Modélisations poissonniennes plutôt que gaussiennes pour l'analyse des puces à très haute-densité ; apprentissage avec données dépendantes; analyse de réseaux ; développements d'algorithmes stochastiques pour l'inférence) sont de grande qualité et le comité les appuie avec la plus grande vigueur.



— **Points à améliorer :**

- la petite taille du laboratoire qui peut devenir une faiblesse
- un déficit sur les collaborations internationales et la participation à des projets internationaux
- l'absence de fiches individuelles n'a pas facilité l'évaluation globale
- locaux éclatés
- faible investissement de l'INRA dans les moyens humains

La petite taille du laboratoire a des avantages au niveau du fonctionnement et des communications internes mais, peut aussi devenir une faiblesse, car reposant sur un petit nombre de leaders scientifiques. La politique de recrutements futurs est donc très importante. On peut aussi remarquer un déficit sur les collaborations internationales et la participation à des projets internationaux et des efforts seraient souhaitables pour assurer la présence et la visibilité internationale du laboratoire. Un effort significatif de l'équipe MORSE devrait être fait pour définir une stratégie et un positionnement plus clair quant aux thématiques de recherche qu'elle souhaite développer lors du prochain quadriennal. À propos des documents fournis, l'absence de fiches individuelles n'a pas facilité l'évaluation globale.

— **Recommandations :**

- l'équipe MORSE devrait travailler à avoir une identité scientifique plus marquée
- le niveau des publications de cette équipe (choix des revues) devrait faire l'objet d'un effort significatif lors du prochain quadriennal
- il ne faudrait pas entretenir une présentation séparée en deux sous-équipes mais plutôt rechercher une production scientifique commune
- améliorer l'installation physique du laboratoire

Les travaux de l'équipe MORSE sont à finalité très appliquée et semblent manquer un peu d'envergure et d'ambition scientifique. A la lecture des documents on ressent une activité à vocation trop finalisée, répondant à des sollicitations d'opérateurs, ayant des besoins opérationnels. Tout en conservant une activité de recherche évidente en statistique, l'équipe reste dépendante de ses commanditaires, notamment de la qualité et de la pertinence des choix scientifiques qu'ils opèrent. L'équipe devrait travailler à avoir une identité scientifique plus marquée et une personnalité plus emblématique de son image, dans le paysage des statisticiens. Une réflexion de nature méthodologique sur les méthodes d'échantillonnage, l'estimation des erreurs, etc. est effectivement souhaitable (le fait de n'être qu'utilisateur de méthodes d'inférence semble un peu dangereux). Le niveau des publications de cette équipe (choix des revues) devrait faire l'objet d'un effort significatif lors du prochain quadriennal (viser des revues de meilleur niveau).

A la suite du renforcement de la composante informatique, des inflexions se sont opérées en conduisant des travaux d'apprentissage informatique sur des applications à la bioinformatique. Il semble cependant, à la lecture du projet de l'unité, que la richesse pouvant ressortir de ces deux approches statistique/informatique n'ait pas encore été totalement exploitée et qu'il serait pertinent d'être plus volontariste dans ce domaine. Le bilan d'activités de l'unité révèle une présentation séparée qui semble traduire deux sous-équipes qu'il ne faudrait pas durablement entretenir. Une production scientifique commune devrait être recherchée.

A l'avenir, le déménagement d'AgroParisTech sur le plateau de Saclay dans le cadre du plan campus devra faire l'objet d'une réflexion précoce et intégrative. L'identification d'alliances scientifiques, et la recherche d'une organisation plus fédérative est à conseiller. Les ouvertures potentielles inévitablement associées à ce déménagement seront une véritable mutation et une révolution culturelle pour cette unité. D'une équipe pouvant apparaître comme une unité propre d'AgroParisTech, même si l'INRA y est associée, il faudra éventuellement envisager une intégration dans une entité plus vaste. Cette évolution ne devra pas mettre en péril le positionnement et la dynamique scientifique actuelle qui reste efficace et pertinente.



Note de l'unité	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A	A	A	A	A

A M. Pierre GLORIEUX,
Directeur de la section des unités de recherche de l'AERES

Monsieur,

Nous avons bien reçu le rapport du comité sur le bilan et le projet de recherche que nous avons présentés. Nous répondons ici sur certains points qui nous semblent devoir être rectifiés ou précisés.

UMR 518

Petite taille de l'unité et faible investissement de l'INRA dans les moyens humains.

Des moyens humains supplémentaires en termes de chercheurs, enseignants-chercheurs et ingénieurs de recherche permettraient de conforter la masse critique nécessaire à une recherche de pointe comme le suggère le rapport. Les prochains renouvellements d'enseignants-chercheurs aboutiront de fait à une augmentation de la capacité de recherche et permettront également de rééquilibrer les forces entre les deux équipes. Nous avons déposé une demande de poste d'IR (ou éventuellement de CR) auprès de l'INRA sur des compétences qui ne pourraient pas être exploitées en enseignement à AgroParisTech.

Déficit sur les collaborations internationales et la participation à des projets internationaux.

Les membres de l'UMR participent à diverses actions au niveau européen :

- Action COST en génétique génomique
- Erasmus Intensive program: Graduate course in Mathematics and Biology
- Réseau Pascal réseau d'excellence européen (Pattern Analysis, Statistical modelling and Computational Learning)

Notre rapport n'insiste sans doute pas assez sur ce point mais nous participons aux réseaux et actions européens quand cela est pertinent et n'implique pas un investissement administratif très lourd. Nous sommes ouverts à des collaborations internationales. Nous pouvons nous améliorer sur ce point ; une politique volontariste nécessiterait des moyens plus conséquents.

Absence de fiches individuelles n'a pas facilité l'évaluation globale.

Il s'agissait là d'une consigne conjointe d'AgroParisTech et de l'INRA : nous invitons nos tutelles et l'AERES à s'entendre sur ce point.

Projet « Saclay ».

L'UMR est consciente des enjeux : elle veillera à ne pas mettre en péril son positionnement en privilégiant, par exemple, une intégration « souple » (sous forme de séminaires ou de groupes de travail) dans une entité plus vaste dédiée aux mathématiques appliquées.

Equipe « Statistique et Génome »

Intégration de la thématique « Apprentissage » dans l'équipe « Statistique et génome ».

La jonction entre les deux « sous-équipes » est récente ; sa réussite constitue évidemment un enjeu du prochain quadriennal mais les sujets de recherche communs ont été bien identifiés.

Equipe « MOdélisation et Risque en Statistique Environnementale » (MORSE)

Notoriété de l'équipe.

La visibilité de l'équipe MORSE est sous-estimée par le comité. Ses membres sont extrêmement actifs dans l'animation des thématiques environnementales et spatiales : président fondateur, président en exercice et membres du bureau du groupe « Environnement et statistique » de la Société Française de Statistique (SFdS). L'équipe a organisé avec des membres de la Royal Statistical Society une journée "biodiversité et approches spatio-tempo-

relles" le 24 mai 2006 à Paris, et a noué des liens avec les membres italiens du GRASPA pour des actions communes. D'autre part des relations sont instaurées avec HydroQuébec (contrat en cours), l'université d'Ohio (séjour 2 mois) et l'université de Vancouver (délégation à partir de Sept 2009).

Production de l'équipe MORSE.

Il nous semble utile de diffuser des travaux dans des revues à l'usage de la communauté utilisatrice de nos méthodes. Le niveau statistique de certaines « revues applicatives » est d'ailleurs souvent tout à fait comparable aux revues de statistiques appliquées (par exemple, les idées originales de l'Approximate Bayesian Computing – ABC – ont été publiées dans *Genetics*). Ainsi, certaines publications de l'équipes dans *Genetics* ou *Mol. Ecol. Notes* sont citées plusieurs dizaines de fois. Enfin, des publications dans des revues françaises peuvent également être nécessaires pour diffuser des méthodes (ex. : publication dans *JASA* puis diffusion via la *Revue de Statistique Appliquée*).

Axe « Raisonnement conditionnel bayésien en sciences de l'environnement ».

Les méthodes bayésiennes sont aujourd'hui très populaires en environnement (cf. le numéro spécial de *Ecological Applications* sur les modèles bayésiens hiérarchiques en 2009) mais la mise en œuvre reste souvent un problème en soi. L'équipe est fréquemment sollicitée pour son expertise dans l'utilisation de ces méthodes. D'autre part, l'équipe a contribué au balisage du domaine (livre en 2007 chez Springer France sur le Raisonnement Bayésien, Modélisation et Inférence) et est en relation avec des personnalités françaises marquantes du domaine (ACI nouvelle interface des mathématiques en collaboration avec C. Robert et J.M. Marin).

Production « d'une recherche 'réellement' exploitable en écologie ».

En matière halieutique, l'équipe travaille avec des écologues de l'IRD et du CIRAD qui sont directement concernés par l'exploitation de modèle de dynamique de population de thons, avec les biologistes de l'INRA de St Pé et d'AgroCampus Rennes pour la gestion durable du saumon atlantique (membres du CIEM), avec les chercheurs d'Environnement Canada pour l'étude de la répartition spatiale des espèces et de la gestion de la biodiversité. Ces collaborateurs sont concernés au premier chef par l'« exploitabilité » des méthodes et des résultats.

Axe « Modélisation spatiale pour l'environnement »

Les membres de l'équipe ont commencé à travailler ensemble sur des problèmes de dispositifs spatiaux avec des techniques bayésiennes. Comme indiqué plus haut l'équipe organise à AgroParisTech le groupe spatial qui est très bien positionné par rapport à la recherche en statistique spatiale en France : la plupart des chercheurs en statistique spatiale (sauf en analyse d'image) en France connaissent le groupe et y ont exposé leurs travaux. Le thème des extrêmes en spatial est original et très peu traité en France (l'autre lieu est le LSCE / CEA-CNRS), et du point de vue des applications environnementales il est également très peu traité à l'étranger. Le projet d'ANR Modélisation Probabiliste pour l'Evaluation du Risque Avalanche, qui porte sur l'évaluation et la cartographie de l'aléa avalancheux (CNRS/LSCE, Cemagref, INSA Lyon), intègre une partie méthodologique sur ce thème.

Définition d'un thématique « emblématique » pour l'équipe.

Les thématiques abordées par l'équipe MORSE sont diverses du fait de l'encadrement d'un certain nombre de doctorants GREF fonctionnaires dont la thèse correspond à une commande précise du ministère de l'Agriculture. Les thèses correspondantes doivent être pluridisciplinaires et sont souvent à visées très appliquées. Les doctorants ainsi formés sont destinés à continuer une activité de recherche ou à prendre un poste d'expertise ou d'interface de l'administration avec les milieux scientifiques, en appui à la définition des politiques publiques. Ce travail d'encadrement doctoral fait partie des missions qui nous sont confiées par AgroParisTech.

De plus, cette équipe est encore jeune au sens où, sur 4 membres, un l'a rejoint il y a 2 ans et un autre est toujours localisé en dehors de l'UMR (mais la rejoindra d'ici 2 mois). Ses membres travaillent à la définition de sujets plus génériques permettant de fédérer plusieurs axes. Le projet PathTIS marque cette volonté. Tous les permanents de l'équipe sont impliqués dans ce projet d'ANR déposé dans le cadre de l'appel d'offre SYStème COmplexe et Modélisation Mathématique. L'objectif est de développer et d'appliquer de nouvelles méthodes pour l'analyse de trajectoires en milieu marin. Il comporte une partie portant sur l'étude théorique de processus stochastiques adaptés aux spécificités de ces trajectoires, une partie portant sur l'adaptation de techniques d'estimation (particules, ABC) et deux parties portant sur des applications écologiques. Outre l'occasion de rassembler toutes les compétences de l'équipe autour d'un projet, PathTIS permet la création d'une communauté de travail avec un ensemble de partenaires scientifiques en statistique pour l'écologie (INRA, IRD, AgroCampus, AgroParisTech).

S. Robin, le 8 avril 2009

