

Pierre Pudlo — Curriculum vitæ

Professeur à l'Université d'Aix-Marseille, Faculté des Sciences
CNU 26 – Mathématiques appliquées et applications des mathématiques

I2M — Institut de Mathématiques de Marseille UMR
3 place Victor Hugo ; Case 39
13331 Marseille CEDEX 3, France
Tél. : 06 85 17 78 46

Email : pierre.pudlo@univ-amu.fr
URL : <https://pierre-pudlo.pedaweb.univ-amu.fr/>

Né le 20 septembre 1977 à Villers-Semeuse (08 – Ardennes)
Nationalité française.

Études et diplômes

2014 HABILITATION À DIRIGER DES RECHERCHES, université Montpellier 2 décernée après les rapports de Mark BEAUMONT (U. Bristol), Gérard BIAU (U. P. & M. Curie), Chris HOLMES (U. Oxford) et Adrian E. RAFTERY (U. Washington) et la soutenance de le jury présidé par Gilles CELEUX (INRIA), composé de Mark BEAUMONT (U. Bristol), Gérard BIAU (U. P. & M. Curie), Arnaud ESTOUP (INRA), Jean-Michel MARIN (U. Montpellier 2) et Didier PIAU (U. J. Fourier)

2001–2004 THÈSE au laboratoire de Probabilités, Combinatoire et Statistique, université Claude Bernard Lyon 1 sous la direction de Didier PIAU *Estimations précises de grandes déviations et applications à la statistique des séquences biologiques*

1998–2002 Élève à l'École Normale Supérieure de Lyon : MAGISTÈRE mathématiques et applications (MMA)
LICENCE, MAÎTRISE ET DEA à l'université Lyon 1 ; AGRÉGATION de mathématiques (reçu 52ème)

Postes occupés

2015– Professeur à l'université d'Aix-Marseille & Institut de Mathématiques de Marseille UMR CNRS 7373

2006–2015 Maître de conférences à l'université Montpellier 2 & Institut de Mathématiques et Modélisation de Montpellier UMR CNRS 5149

2005–2006 ATER, université de Franche-Comté, Laboratoire de Mathématiques de Besançon (UMR 6623)

2002–2005 Allocataire-moniteur, université Lyon 1

1998–2002 Fonctionnaire stagiaire, École Normale Supérieure de Lyon

Thèmes de recherche

Classification : Clustering spectral • Clustering basé sur la densité • Machine learning • Théorèmes asymptotiques • constante de Cheeger • Graphes aléatoires.

Probabilités numériques et statistique bayésienne : méthodes de Monte Carlo • Génétique des populations • ABC (Approximate Bayesian Computation) • échantillonnage préférentiel • Vraisemblance empirique.

Données complexes et applications : ajustement de modèles dynamiques • données temporelles (signal, son) • Machine learning (forêt aléatoire, réseaux de neurones) • données fonctionnelles

Étudiants en thèse

Mohammed SEDKI (2009-2012, actuellement MC à la faculté de Médecine de l'université Paris Sud), en co-encadrement avec J.-M. Marin (PR, Montpellier) : Échantillonnage préférentiel adaptatif et méthodes bayésiennes approchées appliquées à la génétique des populations.

Julien STOEHR (2012-2015, ancien élève de l'ENS Rennes, actuellement MC à Paris Dauphine), en co-encadrement avec J.-M. Marin (PR, Montpellier) : Statistical inference methods for Gibbs random fields.

Coralie MERLE (2013-2016, actuellement ATER à Paris Dauphine), en co-encadrement avec J.-M. Marin (PR, Montpellier) et F. Rousset (DR, Montpellier) : Nouvelles méthodes d'inférence de l'histoire démographique à partir de données génétiques.

Paul-Marie GROLLEMUND (2014-2017, MC à Clermont Auvergne, IUT site d'Aurillac), en co-encadrement avec C. Abraham (PR, SupAgro) et M. Baragatti (MC, SupAgro) : Régression linéaire bayésienne sur données fonctionnelles.

Grégoire AUFORT (2018-2022, actuellement en postdoc à l'IAP), en co-direction avec Denis Burgarella (Laboratoire d'Astrophysique de Marseille) : Contraindre l'évolution des galaxies à partir de données panchromatiques : repenser l'analyse bayésienne des distributions spectrales d'énergie.

Guillem BONAFOS (2019-2023, actuellement en postdoc à Saint-Etienne), en co-direction avec Arnaud Rey (LPC) et Jean-Marc Freyermuth (I2M) : Bourse CIFRE à Résurgences R&D sur l'analyse des vocalisations des bébés humains de 0 à 12 mois.

Ali SHAKIL (2021-2024), en co-encadrement avec A. Khalighi (Fresnel) : développement méthodologiques pour l'analyse de données massives IoT de capteurs pour comprendre la dynamique d'un réseau d'assainissement (thèse CIFRE avec la SERAMM - SUEZ).

Meriam EZZIATI (2021-...), en co-encadrement avec J.-G. Cuby (LAM) et F.-X. Dupé (LIS) : développement de méthodes bayésiennes pour la recherche de quasars à haut redshift dans les données Euclid. Soutenance prévue à l'automne 2024.

Marie-jose CHAAYA (2023-...), en co-encadrement avec F. Hubert (I2M), S. Chauvet (IBDM). Thèse CenTuri sur la modélisation de l'innervation des tumeurs, et l'impact du système immunitaire.

Tâches administratives et scientifiques

Responsabilités administratives et scientifiques

2021– Chargé de mission Projet TIGER et Institut d'Établissement pour la Faculté des Sciences

2021– Responsable de la mention « Mathématiques appliquées, statistique » de Master, portée par la Faculté des Sciences de l'Université d'Aix-Marseille. Évaluation HCERES 2023 et portage du projet pour le contrat 2024–2029.

2021– Responsable de l'équipe Statistique du pôle Mathématiques de l'aléatoire de l'I2M (12 enseignants-chercheurs permanents)

2016–2023 Responsable du M2 « Mathématiques et Informatique : Statistique, Signal et Santé » (MI3S), devenu le M2 « Data Science » de l'UFR Sciences.

2018 Co-responsable de la chaire Jean-Morlet de Kerrie Mengersen au CIRM (Marseille)

2017 Comité scientifique des Journées MAS 2018 (Dijon)

2016–2018 Comité scientifique de la conférence *Statistical Methods for Post Genomic Data*

2016– Membre élu du Conseil du département Mathématiques de la Faculté des Sciences

2016– Membre élu du Conseil de l'UFR Sciences, et de la commission enseignement

2015–2018 Membre des commissions informatique et scientifique de l'Institut de Mathématiques de Marseille (I2M)

2013–2015 Responsable de l'axe « *Algorithmes et calculs* » du Labex NUMEV

2010–2015 Membre élu du Conseil de l'UMR

2008–2014 Responsable du séminaire de probabilités et statistique de Montpellier commun à l'UM2 et à l'UMR MIS-TEA (INRA, SupAgro)

2012, 2013, 2014 Membre des comités d'organisation des Écoles-Ateliers « *Mathematical and Computational evolutionary biology* », juin 2012, 2013 et 2014.

2010 Membre du comité d'organisation des Journées de Statistiques du Sud à Mèze (juin 2010): « *Modelling and Statistics in System Biology* ».

2004–2005 Administrateur du serveur informatique du LaPCS (Laboratoire de Probabilités, Combinatoire et Statistique, Lyon 1)

Comités de sélection et d'évaluation

2022 Membre du comité de sélection à l'université de Rennes 2 pour le recrutement d'un maître de conférences en statistique (CNU 26)

2022 Membre du comité d'évaluation science (CE40) de l'ANR

2021 Membre du comité de sélection à l'université de Rennes 2 pour le recrutement d'un maître de conférences en statistique (CNU 26)

2021 Président du comité de sélection MC CNU 26 « statistique » à l'université d'Aix-Marseille & UFR Sciences

2021 Membre du comité d'évaluation science (CE40) de l'ANR

- 2021** Membre de la commission ad hoc groupe 5 de l'université d'Aix-Marseille pour l'avancement au grade des enseignants-chercheurs
- 2020** Membre du comité de sélection à l'université Paul Sabatier – Toulouse III pour recrutement d'un maître de conférences en statistique (CNU 26)
- 2019, 2020** Membre de la commission ad hoc groupe 5 de l'université d'Aix-Marseille pour l'avancement au grade des enseignants-chercheurs
- 2018** Membre du comité d'évaluation HCERES de l'UMR LAMME à l'université d'Evry Val d'Essonne
- 2017** Président du comité de sélection MC CNU 26 « statistique » à l'université d'Aix-Marseille & UFR Sciences
- 2016** Membre du comité de sélection à l'université Pierre et Marie Curie pour le concours de recrutement de maître de conférences en statistique (CNU 26)
- 2012** Membre du comité de sélection à l'université Lyon 1 pour recrutement d'un maître de conférences en statistique (CNU 26).
- 2009–2010** Membre de comités de sélection à Montpellier 2.
- 2008** Membre élu de la commission de spécialistes CNU 26 à Montpellier 2.

Jurys de thèse et d'habilitation à diriger des recherches

Ne figure pas dans cette liste les jurys pour des étudiantes et étudiants que j'ai co-encadré.

- 2023** Rapporteur de la thèse de Roméo Tayewo (U. Bretagne Sud) *Decision Making on Spatio-temporal stochastic processes with Environmental Applications* et membre du jury.
- 2022** Rapporteur de la thèse d'Alice Martin (Institut Polytechnique de Paris) *Deep Learning models and algorithms for sequential data problems* et membre du jury.
- 2022** Rapporteur de la thèse d'Antoine Baudiquez (U. Franche-Comté) *Metrology and statistics: From clocks to millisecond pulsars* et membre du jury.
- 2022** Rapporteur de la thèse d'Eunice Okome Obiang (U. Angers) *High-dimensional statistics applied to the modeling of Triple-Negative Breast Cancer*.
- 2021** Président du jury de thèse de Giulia Mezzadri (U. Côte d'Azur) *Inférence statistique pour les modèles de catégorisation et ordre de présentation*.
- 2021** Rapporteur de la thèse de
- 2021** Rapporteur de l'habilitation à diriger des recherches de Marie-Pierre Etienne (U. Rennes 1) *The hidden part of Markovian stochastic processes for Biology and Ecology* et membre du jury.
- 2020** Rapporteur de la thèse de Christophe Dumora (U. Bordeaux) *Estimation de paramètres clés liés à la gestion d'un réseau de distribution d'eau potable : Méthode d'inférence sur les nœuds d'un graphe* et membre du jury.
- 2020** Membre du jury de la thèse de Thi Khuyen Le (AMU) *Sparse precision matrix estimation in high dimension and application to medical imaging. Hypothesis testing on some particular graphical models of GLASSO*.
- 2019** Président du jury de thèse de Jean-Michel Galharret (U. Nantes) *Statistique bayésienne appliquée en Sciences Humaines et Sociales*.
- 2019** Rapporteur de la thèse de Sandrine Boulet (U. Paris Descartes) *Développement de méthodes bayésiennes de sélection de variables pour intégrer l'information experte* et membre du jury.
- 2019** Rapporteur de la thèse de Manon Ansart (Sorbonne U.) *Design of data driven decision support systems for the early detection of subjects at risk to develop Alzheimer's Disease* et membre du jury.
- 2018** Membre du jury de l'habilitation à diriger des recherches de Sylvain LeCorff (U. Paris-Saclay) *Estimation et simulation de modèles à données latentes. Contributions non paramétriques et en grande dimension*.
- 2018** Rapporteur de l'habilitation à diriger des recherches de Sabine Le Mercier (U. Paul Sabatier – Toulouse 3) *Distribution du score local pour la mise en évidence de segments atypiques au sein de séquences* et membre du jury.
- 2017** Rapporteur de la thèse de Paul Bastide (AgroParisTech) *Processus stochastiques avec sauts sur arbres : application à l'évolution adaptatives sur des phylogénies* et membre du jury.
- 2017** Rapporteur de la thèse de Panisa Treepong (U. Franche-Comté) *Bioinformatic analysis of the genomes of epidemic Pseudomonas Aeruginosa* et membre du jury.

2016 Rapporteur de la thèse de Clara Grazian (U. Paris Dauphine et Sapienza U. Rome) *Contribution to Bayesian Computing for Complex Models* et membre du jury.

Liste de publications

Articles publiés

- (A1) P. Pudlo (2009) Large deviations and full Edgeworth expansions for finite Markov chains with applications to the analysis of genomic sequences. *ESAIM: Probab. and Statis.* 14, pp. 435–455.
- (A2) B. Pelletier and P. Pudlo (2011) Operator norm convergence of spectral clustering on level sets. *Journal of Machine Learning Research*, 12, pp. 349–380.
- (A3) E. Arias-Castro, B. Pelletier and P. Pudlo (2012) The Normalized Graph Cut and Cheeger Constant: from Discrete to Continuous. *Advances in Applied Probability*, 44(4), dec 2012.
- (A4) B. Cadre, B. Pelletier and P. Pudlo (2013) Estimation of density level sets with a given probability content. *Journal of Nonparametric Statistics* 25(1), pp. 261–272.
- (A5) J.-M. Marin, P. Pudlo, C. P. Robert and R. Ryder (2012) Approximate Bayesian Computational methods. *Statistics and Computing* 22(6), pp. 1167–1180.
- (A6) A. Estoup, E. Lombaert, J.-M. Marin, T. Guillemaud, P. Pudlo, C. P. Robert and J.-M. Cornuet (2012) Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources* 12(5), pp. 846–855.
- (A7) Mengerson, K.L., Pudlo, P. and Robert, C. P. (2013) Bayesian computation via empirical likelihood. *Proc. Natl. Acad. Sci. USA* 110(4), pp. 1321–1326.
- (A8) Gautier, M., Foucaud, J., Gharbi, K., Cezard, T., Galan, M., Loiseau, A., Thomson, M., Pudlo, P., Kerdelhué, C., Estoup, A. (2013) Estimation of population allele frequencies from next-generation sequencing data: pooled versus individual genotyping. *Molecular Ecology* 22(14), pp. 3766–3779.
- (A9) Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., Cornuet, J.-M., Estoup, A. (2013) The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology* 22(11), pp. 3165–3178.
- (A10) Cornuet J.-M., Pudlo P., Veyssier J., Dehne-Garcia A., Gautier M., Leblois R., Marin J.-M., Estoup A. (2014) DIYABC v2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30(8), pp. 1187–1189. Voir <http://www1.montpellier.inra.fr/CBGP/diyabc/>
- (A11) Baragatti, M. and Pudlo, P. (2014) An overview on Approximate Bayesian computation. *ESAIM: Proceedings* 44, pp. 291–299.
- (A12) Leblois, R., Pudlo, P., Néron, J., Bertaux, F., Beeravolu, C. R., Vitalis, R. and Rousset, F. (2014) Maximum likelihood inference of population size contractions from microsatellite data. *Molecular Biology and Evolution*, 31(10), pp. 2805–2823.
- (A13) Stoehr, J., Pudlo, P. and Cucala, L. (2014) Adaptive ABC model choice and geometric summary statistics for hidden Gibbs random fields. *Statistics and Computing* 25 (1), pp. 129–141.
- (A14) Pudlo, P., Marin, J.-M., Estoup, A., Gautier, M., Cornuet, J.-M. and Robert, C. P. (2015) Reliable ABC model choice via random forests. *Bioinformatics*, 32 (6), 859–866.
- (A15) Stoehr, J., J.-M. Marin and P. Pudlo (2016) Hidden Gibbs random fields model selection using Block Likelihood Information Criterion. *Stat* 5 (1), pp. 158–172.
- (A16) Fraimout, A., *et al.* (2017) Deciphering the Routes of invasion of *Drosophila suzukii* by Means of ABC Random Forest. *Molecular Biology and Evolution* 34 (4), pp. 980–996.
- (A17) Merle, C., Leblois, R., Rousset, F. and Pudlo, P. (2017) Resampling: an improvement of Importance Sampling in varying population size models. *Theoretical Population Biology*, 114, pp. 70–87.
- (A18) Raynal, L., Marin, J.-M., Pudlo, P., Ribatet, M. and Robert, C. P. (2018) ABC random forests for Bayesian parameter inference. *Bioinformatic*, 35(10), pp. 1720–1728.
- (A19) Pudlo, P. and Sedki, M. (2018) Simulation of stochastic models of structured population in population genetics under neutrality. *Journal de la Société Française de Statistique*, 159(3), pp. 126–141.
- (A20) Grollemund, P.-M., Abraham, C., Baragatti, M. and Pudlo, P. (2019) Bayesian functional linear regression with sparse step functions. *Bayesian Analysis*, 14(1), pp. 111–135.

- (A21) Marin, J.-M., Pudlo, P. and Sedki, M. (2019) Consistency of adaptive importance sampling and recycling schemes. *Bernoulli*, 25(3), pp. 1977–1998.
- (A22) Aufort, G., Ciesla, L., Pudlo, P. and Buat, V. (2020) Constraining the recent star formation history of galaxies: an approximate Bayesian computation approach. *Astronomy & Astrophysics* 635, A136.
- (A23) Guillot, J., Dominici, C., ..., Pudlo, P., ..., Chauvet, S. and Mann, F. (2022) Sympathetic axonal sprouting induces changes in macrophage populations and protects against pancreatic cancer. *Nature Communications* 13(1).
- (A24) Shakil, A., Khalighi, M.A., Pudlo, P., Leclerc, C., Laplace, D. and Hamon, F. (2023) Outlier detection in non-stationary time series applied to sewer network monitoring. *Internet of Things* 21, 100654.

Chapitres de livres, documents et programmes informatiques, et autres publications

- (C1) Cornuet J-M, Pudlo P., Veyssier J, Dehne-Garcia A, Estoup A (2013) DIYABC V2.0. a user-friendly package for inferring population history through Approximate Bayesian Computation using microsatellites, DNA sequence and SNP data. Programme disponible sur le site <http://www1.montpellier.inra.fr/CBGP/diyabc/>
- (C2) Cornuet J-M, Pudlo P., Veyssier J, Dehne-Garcia A, Estoup A (2013) DIYABC V2.0. a user-friendly package for inferring population history through Approximate Bayesian Computation using microsatellites, DNA sequence and SNP data. Notice détaillée d'utilisation de 91 pages disponible sur le site <http://www1.montpellier.inra.fr/CBGP/diyabc/>
- (C3) Stoehr J, Pudlo P., Friel N. GiRaF: A Toolbox for Gibbs Random Fields Analysis. Available on CRAN.
- (C4) Marin, J.-M., Pudlo, P., Estoup, A. and Robert, C. P. (2018) Likelihood-free model choice. In *Handbook of Approximate Bayesian Computation*, pp. 153–178.
- (C5) Estoup, A., Verdu, P, Marin, J.-M., Robert, C. P., Dehne-Garcia, A., Cornuet, J.-M. and Pudlo, P. (2018) Application of Approximate Bayesian Computation to infer the genetic history of Pygmy hunter-gatherers populations from Western Central Africa. In *Handbook of Approximate Bayesian Computation*, pp. 541–567.
- (C6) Stoehr, J., Pudlo, P., Friel, N. (2019) GiRaF: Gibbs Random Fields Analysis. *R package on CRAN.*
- (C7) Mengersen, K.L., Pudlo, P., Robert, C.P. (2020) *Case Studies in Applied Bayesian Data Analysis.* Springer.

Preprints et articles soumis

- (P1) Ratmann, O., Pudlo, P., Richardson, S. and Robert, C. P. (2011) *Monte Carlo algorithms for model assessment via conflicting summaries.* Voir [Arxiv:1106.5919](https://arxiv.org/abs/1106.5919)
- (P2) Sedki, M., Pudlo, P., J.-M. Marin, C. P. Robert and J.-M. Cornuet (2013) *Efficient learning in ABC algorithms.* Voir [Arxiv:1210.1388](https://arxiv.org/abs/1210.1388)
- (P3) Ebert, A., Pudlo, P., Mengersen, K.L., Wu, P. and Drovandi, C. *Combined parameter and state inference with automatically calibrated ABC.* Voir [Arxiv:1910.14227](https://arxiv.org/abs/1910.14227)
- (P4) Aufort, G., Pudlo, P. and Burgarella, D. (2022) *Tempered, Anti-truncated, Multiple Importance Sampling.* Voir [Arxiv:2205.01501](https://arxiv.org/abs/2205.01501)
- (P5) Bonafos, G., Pudlo, P., Freyermuth, J.-M., Legou, T., Fagot, J., Tronçon, S. Rey, A. (2023) *Detection and classification of vocal productions in large scale audio recordings.* Voir [HAL-04264454](https://hal.archives-ouvertes.fr/hal-04264454)
- (P6) Bonafos, G., Freyermuth, J.-M., Pudlo, P., Tronçon, S. Rey, A. (2023) *Topological data analysis of human vowels: Persistent homologies across representation spaces.* Voir [HAL-04264454](https://hal.archives-ouvertes.fr/hal-04264454)