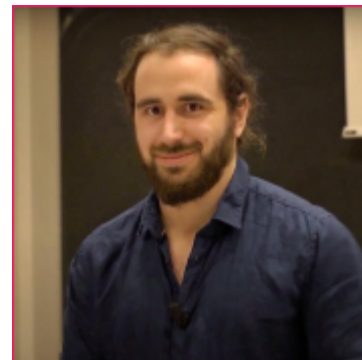


Date et lieu de naissance: 6 septembre 1991, Alençon, France
Nationalité: Française
Email: arthur.leroy.pro@gmail.com
Page web: <https://arthur-leroy.netlify.app/>
Téléphone: +33 770345284
Affiliation actuelle: The University of Manchester



Arthur Leroy

FORMATION ET EXPERIENCE DE RECHERCHE

- 2024 - : **Chargé de Recherche**, [INRAE](#), [GABI](#) & [MIA Paris-Saclay](#), France
CR en statistique et intelligence artificielle pour la génétique animale.
- 2021 - 2024: **Post-Doctorat (Research Associate)**, [The University of Manchester](#), Royaume-Uni
Projet de recherche: Modélisation et analyse de données moléculaires de grande dimension pour l'étude longitudinale des populations.
Encadrant: [Mauricio Alvarez](#)
- 2021 - 2022: **Post-Doctorat (Research Associate)**, [The University of Sheffield](#), Royaume-Uni
Projet identique (voir ci-dessus)
- 2020 - 2021: **ATER**, [Université Paris Cité](#), [MAP5](#), France
- 2017 - 2020: **Doctorat de mathématiques appliquées**, [Université Paris Cité](#), [MAP5](#), France
Titre: Apprentissage de données fonctionnelles par modèles multitâches : application à la prédiction de performance sportive.
Directrice : [Servane Gey](#) - Co-encadrants : [Pierre Latouche](#), [Benjamin Guedj](#)
- 2017 - 2017: **Ingénieur de recherche**, [Insep](#), France
Projets de recherche et d'accompagnement de diverses fédérations sportives (natation, pentathlon moderne, triathlon, aviron, judo) pour la performance.
- 2014 - 2014: **Stage de recherche**, [TAO](#), [Inria Paris Saclay](#), France
Titre: Modèle graphiques probabilistes pour l'analyse de systèmes complexes et application à des données électroencéphalographiques.
Encadrant: [Cyril Furtlehner](#)
- 2012 - 2014: **Master recherche mathématiques/statistiques**, [Université de Rennes 1](#), France

Publications [[Google Scholar](#)][[ResearchGate](#)][[ORCID](#)]

Articles publiés

- Avr. 2025: **Longitudinal prediction of DNA methylation to forecast epigenetic outcomes;**
[Leroy A.](#), [Teh A. L.](#), et al.;

[\[arxiv\]](#)[\[code\]](#)

Fév. 2025: ***Prospective Prediction of Body Mass Index Trajectories using Multitask Gaussian Processes;***

Leroy A., Gupta V., et al.;

International Journal of Obesity, 2025, 49, 340-347

[\[url\]](#)[\[arxiv\]](#)[\[code\]](#)[\[NeurIPS poster\]](#)

Jan. 2023: ***Cluster-Specific Predictions with Multi-Task Gaussian Processes;***

Leroy A., Latouche P., Guedj B., Gey S.;

Journal of Machine Learning Research, 2023, 24(5):1-49

[\[url\]](#)[\[pdf\]](#)[\[code\]](#)[\[R package\]](#)[\[ICML poster\]](#)

Mai 2022: ***MAGMA: inference and prediction with multi-task Gaussian processes with common mean;***

Leroy A., Latouche P., Guedj B., Gey S.;

Machine Learning, 2022, DOI: 10.1007/s10994-022-06172-1

[\[url\]](#)[\[arxiv\]](#)[\[code\]](#)[\[R package\]](#)

Nov. 2021: ***International Swimming League: Do Successive Events Lead to Improve Swimming Performance?;***

Pla R., Leroy A., Raineteau Y., & Hellard P.;

International Journal of Sports Physiology and Performance, 2021,

DOI: 10.1123/ijsp.2021-0375

[\[url\]](#)

Dec. 2020: ***Body and boat: significance of morphology on elite rowing performance;***

De Larochelambert Q., Del Vecchio S., Leroy A., et al.;

Frontiers in Sports and Active Living, 2020, DOI: 10.3389/fspor.2020.597676

[\[url\]](#)

Oct. 2019: ***Bayesian approach to quantify morphological impact on performance in international elite freestyle swimming;***

Pla R., Leroy A., et al.;

BMJ Open SEM, 2019, DOI: 10.1136/bmjsem-2019-000543

[\[url\]](#)

Oct. 2019: ***Robust Exponential Decreasing Index (REDI): adaptive and robust method for computing cumulated workload;***

Moussa I., Leroy A., et al.;

BMJ Open SEM, 2019, DOI: 10.1136/bmjsem-2019-000573

[\[url\]](#)[\[R package\]](#)[\[Web app\]](#)

Sept. 2018: ***Functional Data Analysis in Sport Science: Example of Swimmers' Progression Curves***

Clustering;

Leroy A., Marc A., Dupas O., Rey J., Gey S.;

Applied Sciences, 2018, DOI: 10.3390/app8101766

[url]

Pré-publications

Mai 2025: ***A Bayesian framework for multivariate differential analysis accounting for missing data;***

Chion M., Leroy A.;

[arxiv][R package][Web app]

Sept. 2023: ***Latent variable multi-output Gaussian processes for hierarchical datasets;***

Ma C., Leroy A., Alvarez M.;

[arxiv]

Contributions à des ouvrages collectifs

Jan. 2023: ***Statistics meets Sports: what can we learn from sports data;***

Editors: Dominicy Y., Ley C.;

[url]

COMMUNICATIONS ORALES (sélection choisie parmi 30+ répertoriées ici)

27-05-2025: Séminaire statistique - University College Dublin - Irlande [Slides]

10-09-2024: Gaussian Process Summer School 2024 - Manchester - Royaume-Uni [Vidéo] [Slides]

11-09-2023: Gaussian Process Summer School 2023 - Manchester - Royaume-Uni [Vidéo] [Slides]

25-07-2023: ICML 2023, Online [Poster] [Slides]

17-07-2023: Séminaire bio-statistique, University of North Carolina at Chapel Hill - USA [Slides]

02-12-2022: NeurIPS workshop on Gaussian Processes and Spatiotemporal Modelling [Poster] [Slides]

18-12-2021: CMStatistics 2021, Londres, Royaume-Uni [Slides]

09-12-2020: Soutenance de thèse, Paris, France, [Vidéo] [Slides]

ENSEIGNEMENT

2020 - 2021: **ATER**, IUT de Paris, Rives de Seine, 192 heures/an :

- | | | |
|---------------------------------|--------------------------|-----------------|
| - Analyse de données | - STID 2ème année & DU | - 70h - CM & TP |
| - Mathématiques avancées | - STID 2ème année | - 22h - CM & TD |
| - Recherche opérationnelle | - STID 2ème année | - 43h - CM & TD |
| - Statistiques avancées | - STID 2ème année | - 27h - TP |
| - Suivi de projets et de stages | - STID 1ère & 2ème année | - 30h |

2017 - 2020: **Doctorant chargé d'enseignement**, IUT de Paris, Rives de Seine, 64 heures/an :

- | | | |
|----------------|-------------------|------------|
| - Probabilités | - STID 1ère année | - 64h - TD |
|----------------|-------------------|------------|

REVUE PAR LES PAIRS

Journaux :

- Journal of Machine Learning Research - 4 articles
- Transactions on Machine Learning Research - 1 article
- Statistics and Computing - 1 article
- Journal of the Royal Statistical Society (Series C) - 1 article
- Journal of Sports Sciences - 1 article
- PeerJ - 1 article

Conférences :

- NeurIPS (2022, 2023, 2024, 2025) - 13 articles
- ICML (2023) - 3 articles
- AISTATS (2023, 2024, 2025) - 9 articles

ENCADREMENT

Stagiaires accompagnés (% d'encadrement) :

- Terence Vieillard (50%) - Master 2 - ENSAI & Rennes 1 - 2025
- Kamenan Ngadi (33%) - CDD 2ans - INRAE - 2024
- Paguidame Sambiani (50%) - Master 2 - Sorbonne Université - 2023
- Alexia Grenouillat (100%) - Master 1 - INSA Toulouse - 2022
- Hugo Lelièvre (100%) - Master 1 - INSA Toulouse - 2022
- Pierre Pathé (90%) - Master 2 - Université Paris 1 - 2021
- Léonard Gousset (70%) - STID 2A - IUT Paris Rives de Seine - 2020
- James Haziza (70%) - STID 2A - IUT Paris Rives de Seine - 2020
- Maxime Bellami (100%) - Master 1 - ENSAI - 2018
- Romain Massal (80%) - Master 1 - ENSAI - 2017

DIVERS

Sociétés savantes

- Trésorier (2022-2024), membre (2019-2024) du groupe Jeunes Statisticien.ne.s de la Société Française de Statistique (SFdS)
- Co-fondateur, vice-président (2018-) du groupe Statistique et Sport de la SFdS

Animation de la communauté

- Co-organisation de la Gaussian Process Summer School (2022, 2023, 2024)
- Co-organisation de séminaires à l'institut Henri Poincaré :
 - Journées Young Statisticians and Probabilists (2022, 2023, 2024)
 - Séminaire Statistique et Sport (2018, 2019, 2021, 2023)
- Co-organisation de sessions et événements aux Journées de Statistique :
 - Session Jeunes 2024 : Sexisme ordinaire, violences sexistes et sexuelles, biais de genre dans la recherche académique
 - Session Jeunes 2023 : Les enjeux d'une recherche éthique
 - Session Jeunes 2022 : La santé mentale des jeunes chercheurs
 - Déjeuners scientifiques (2021, 2022, 2023, 2024)

- Rencontre conférencier.e.s invité.e.s (2021, 2022, 2023, 2024)
- Animation de divers événements scientifiques ou de vulgarisation (data challenge, tables rondes, fête de la science, podcasts ...)

Bourses et financements obtenus

- Appel à projet ministériel de R&D appliqués au sport de haut niveau - 30 000€
- Bourse de contrat doctoral de l'ED 386
- Turing-Manchester Exchange Fellowship - £6000

Programmation [Github]

Auteur et maintenir des packages R (disponibles sur le CRAN) :

- MagmaClustR, Clustering and Prediction using Multi-Task Gaussian Processes
- ProteoBayes, Bayesian Statistical Tools for Quantitative Proteomics
- REDI, Robust Exponential Decreasing Index

Avancé: R, LaTeX, Markdown, HTML, CSS

Intermédiaire: Python, C++, Java, Matlab, Scilab

Langues

- Français : langue maternelle
- Anglais : courant
- Allemand, Tchèque, Italien : notions