

MARIE COURBARIAUX

91400 Orsay, France
✉ marie.courbariaux@gmail.com

DOCTEUR EN STATISTIQUES APPLIQUÉES
INGÉNIEUR AGRONOME AGROPARISTECH
34 ans, nationalité française.

THÈMES DE RECHERCHE : Analyse de données de séquençage et de signaux longitudinaux, modèles à variables latentes (classification non supervisée, données censurées, HMM, algorithme EM et ses variantes,...), régularisation en grande dimension (LASSO, adaptive ridge,...), détection de valeurs aberrantes, analyse de survie, statistiques Bayésiennes, réseaux de neurones (classification, auto-encodeurs,...), co-clustering.

EXPÉRIENCES PROFESSIONNELLES

depuis 2019 (5 ans) <i>Activités principales</i>	INGÉNIEUR DE RECHERCHE EN APPRENTISSAGE STATISTIQUE SUMMIT, Sorbonne Université Gestion de projets de recherche et de prestation. Conception et implémentation d'algorithmes. Préparation et présentation des livrables.
<i>Projets (exemples)</i>	Obépine et EMERGEN : Algorithmes de déconvolution et de lissage adaptés au suivi du SARS-CoV-2 et de ses variants dans les eaux usées (données de quantitative PCR, de digital droplet PCR et de séquençage). Projets avec entreprises : Détection non supervisée d'anomalies sur des signaux longitudinaux (autoencodeurs, clustering); Prédiction du risque de défaut de crédit et du risque de dysfonctionnement (analyse de survie, segmentation sur graphes).
<i>Autres activités</i>	Conception, programmation et animation des Webinaires Maths-Innovation .
<i>Articles scientifiques</i>	[Cluzel et al., 2022 , Courbariaux et al., 2022a , Lewis-Borrell et al., 2023 , Boni et al., 2023]
<i>Encadrement</i>	Chefferie de projets impliquant d'autres ingénieurs; (depuis 2021) Stage puis doctorat de El Hacene Djaout (co-encadrement Grégory Nuel, Vincent Maréchal) : <i>Approche probabiliste pour le suivi des proportions de variants de SARS-CoV-2 dans les eaux usées.</i>
2017–2018 (17 mois) <i>Intitulé</i>	POST-DOCTORAT EN STATISTIQUES APPLIQUÉES UMR 8071 Laboratoire de Mathématiques et de Modélisation d'Évry Identification de sous-types de la maladie de Parkinson avec association génétique, projet MeMoDeep.
<i>Direction</i>	Christophe Ambroise (Université d'Évry Val d'Essonne)
<i>Partenariats</i>	INSERM, Institut du Cerveau et de la Moelle épinière
<i>Articles scientifiques</i>	[Courbariaux et al., 2022b]
2013 – 2016 (3 ans) <i>Intitulé</i>	DOCTORAT EN STATISTIQUES APPLIQUÉES UMR 518 AgroParisTech/INRA Mathématiques et Informatique Appliquées Contributions statistiques à l'amélioration des prévisions hydrologiques probabilistes par méthodes d'ensemble [Courbariaux, 2017].
<i>Direction</i>	Éric Parent (AgroParisTech/INRAE)
<i>Partenariats</i>	EDF-DTG et Hydro-Québec (IREQ); séjour de 5 mois au Québec.
<i>Articles scientifiques</i>	[Courbariaux et al., 2017, 2019]
<i>Enseignement</i>	TD d'initiation à R (2 ^e année du cursus ingénieur, AgroParisTech)
<i>Activités annexes</i>	Relectures critiques pour le Journal of Hydrometeorology

COMPÉTENCES ET ACTIVITÉS ANNEXES

LANGUES

Anglais (courant, TOEIC : 990/990, longs séjours anglophones), Espagnol (intermédiaire), Russe (élémentaire)

PROGRAMMATION

Python (niveau intermédiaire), SQL (bases), git
R (niveau avancé, R Stan, bases de Shiny et INLA)

PSC1 (Prévention et Secours Civiques)

PERMIS B

BAFA (Fonctions d'Animateur)

SPORTS : Escalade (bloc), course à pied

FORMATION

depuis 2017 <i>Ateliers SFdS</i>	FORMATION CONTINUE (2023) Apprentissage par réseaux de neurones : analyse de textes et de séries temporelles (par Jairo Cugliari et Vincent Guigue)
<i>GdT SUMMIT</i> <i>Journée frENBIS</i> <i>Autoformation</i>	(2023) Structures des réseaux de neurones (PyTorch) (2023) Agrégation d'experts (par Pierre Gaillard) Python, auto-encodeurs, R Stan, R INLA, R Shiny, SQL
2012 – 2013 <i>Enseignements suivis</i> <i>(exemples)</i> <i>Projet</i>	MASTER MATHÉMATIQUES POUR LES SCIENCES DU VIVANT UNIVERSITÉ PARIS SUD, Mention Bien Apprentissage statistique en grande dimension ; Optimisation et simulation numérique ; Modèles pharmaco-cinétiques et pharmaco-dynamiques Méthodes statistiques pour la détection précoce d'épidémies de gastro-entérite (direction : Avner Bar-Hen).
2009 – 2013 <i>Activités associatives</i>	FORMATION D'INGÉNIEUR AGROPARISTECH Chef de projet chez AgroParisTech Service Études (junior entreprise).

COMMUNICATIONS ORALES

COMMUNICATIONS INVITÉES

Avril 2023	SÉMINAIRE DE MATHÉMATIQUES POUR LA BIOLOGIE, ROSCOFF Estimation des proportions de variants de SARS-CoV-2 à partir du séquençage des eaux usées
Juin 2022	SÉMINAIRE SUR LA SURVEILLANCE DE LA COVID-19 PAR WBE, PARIS Sequencing to monitor SARS-CoV-2 variants in wastewaters : a probabilistic approach
Mai 2022	ISDM-NUMEV Traitement des données Obépine pour le suivi du SARS-CoV-2

CONFÉRENCES INTERNATIONALES

Juin 2016	HEPEX WORKSHOP, QUEBEC An exchangeable construction for ensemble forecasts post-processing
Juin 2015	IAHS GENERAL ASSEMBLY, PRAGUE Bayesian modeling of rainfall-runoff uncertainty to improve probabilistic forecasts
Avr. 2015	EGU GENERAL ASSEMBLY, VIENNA Bayesian modeling of rainfall-runoff uncertainty to improve probabilistic forecasts
Sept. 2014	CWRA CONFERENCE, QUEBEC Taking into account rainfall-runoff models' uncertainty to forecast waterflows

CONFÉRENCES NATIONALES

Mars 2022	SÉMINAIRE DE STATISTIQUES DE ROCHEBRUNE Un lisseur flexible pour les données de suivi du SARS-CoV-2 dans les eaux usées
Juin 2018	JOURNÉES DE STATISTIQUE DE LA SFdS, PALAISEAU Disease subtyping with integrated genome association study
Juin 2016	JOURNÉES DE STATISTIQUE DE LA SFdS, MONTPELLIER Une construction statistique échangeable pour le post-traitement des ensembles de séries météorologiques
Mars 2016	RENCONTRES STATISTIQUES DE ROCHEBRUNE, ROCHEBRUNE Une construction statistique échangeable pour le post-traitement des ensembles de séries météorologiques
Juin 2015	JOURNÉES DE STATISTIQUE DE LA SFdS, LILLE Construction bayésienne de prévisions probabilistes à partir des sorties d'un modèle déterministe pluie-débit

- M Boni, S Wurtzer, JM Mouchel, Y Maday, SF Le Guyader, P Garry, I Bertrand, N Cluzel, M Courbariaux, S Wang, et al. Implementation of a national wastewater surveillance system in France as a tool to support public authorities during the COVID crisis : The Obepine project. Springer, 2023.
- Nicolas Cluzel, Marie Courbariaux, Siyun Wang, Laurent Moulin, Sébastien Wurtzer, Isabelle Bertrand, Karine Laurent, Patrick Monfort, Christophe Gantzer, Soizick Le Guyader, Mickaël Boni, Jean-Marie Mouchel, Vincent Maréchal, Grégory Nuel, and Yvon Maday. A nationwide indicator to smooth and normalize heterogeneous SARS-CoV-2 RNA data in wastewater. *Environment International*, 158 :106998, 2022. ISSN 0160-4120. doi : <https://doi.org/10.1016/j.envint.2021.106998>. URL <https://www.sciencedirect.com/science/article/pii/S0160412021006231>.
- Marie Courbariaux. *Statistical contributions to hydrometeorological forecasting from ensemble methods*. PhD thesis, Université Paris-Saclay, 2017. URL <https://tel.archives-ouvertes.fr/tel-01566195/>.
- Marie Courbariaux, Pierre Barbillon, and Éric Parent. Water flow probabilistic predictions based on a rainfall-runoff simulator : a two-regime model with variable selection. *Journal of Agricultural, Biological and Environmental Statistics*, 22(2) :194–219, 2017. URL <https://link.springer.com/article/10.1007/s13253-017-0278-5>.
- Marie Courbariaux, Pierre Barbillon, Luc Perreault, and Éric Parent. Post-processing multiensemble temperature and precipitation forecasts through an exchangeable normal-gamma model and its Tobit extension. *Journal of Agricultural, Biological and Environmental Statistics*, 24(2) :309–345, 2019.
- Marie Courbariaux, Nicolas Cluzel, Siyun Wang, Vincent Maréchal, Laurent Moulin, Sébastien Wurtzer, Obépine Consortium, Jean-Marie Mouchel, Yvon Maday, and Grégory Nuel. A flexible smoother adapted to censored data with outliers and its application to SARS-CoV-2 monitoring in wastewater. *Frontiers in Applied Mathematics and Statistics*, 8, 2022a. ISSN 2297-4687. doi : 10.3389/fams.2022.836349. URL <https://www.frontiersin.org/articles/10.3389/fams.2022.836349>.
- Marie Courbariaux, Kylliann De Santiago, Cyril Dalmaso, Fabrice Danjou, Samir Bekadar, Jean-Christophe Corvol, Maria Martinez, Marie Szafranski, and Christophe Ambroise. A sparse mixture-of-experts model with screening of genetic associations to guide disease subtyping. *Frontiers in Genetics*, 13, 2022b.
- Luke Lewis-Borrell, Jessica Irving, Chris J Lilley, Marie Courbariaux, Gregory Nuel, Leon Danon, Kathleen M O'Reilly, Jasmine MS Grimsley, Matthew J Wade, and Stefan Siegert. Robust smoothing of left-censored time series data with a dynamic linear model to infer SARS-CoV-2 RNA concentrations in wastewater. *AIMS Mathematics*, 8(7) :16790–16824, 2023.