CURRICULUM VITAE

Anne Sabourin
Married, two children

Professional address
MAP5, Université Paris Cité,
45 rue des Saint-Pères,
75270 Paris Cedex 06
`anne.sabourin@u-paris.fr`
`https://helios2.mi.parisdescartes.fr/~asabouri/`

POSITIONS HELD

- **2022- . . .** : Full Professor at Université Paris Cité (MAP5, UMR CNRS 8145)
- **2013 - 2022** : Maître de Conférences ($\simeq$ Associate professor) at Télécom Paris (LTCI, Institut polytechnique de Paris), research group S2A (Signal processing and Statistical learning)

RESEARCH

RESEARCH FIELDS

- Extreme value theory, multivariate extremes.
- Statistical learning, high dimensional problems, dimensionality reduction.
- Applications : anomaly detection, environmental sciences.

WORKING PAPERS

- N. Huet, S. Clémençon, and A. Sabourin. "On Regression in Extreme Regions". arXiv preprint arXiv :2303.03084 (2023).

PUBLICATIONS

Aghbalou, A. ; Bertail, P. ; Portier, F. ; Sabourin, A., 2024. Cross-validation on extreme regions. *Extremes*, p. 1-51.

Aghbalou, A. ; Portier, F. ; Sabourin, A. ; Zhou, C., 2024. Tail Inverse Regression : dimension reduction for prediction of extremes. *Bernoulli*. T. 30, nᵒ 1, p. 503-533.

Aghbalou, A. ; Sabourin, A. ; Portier, F., 2024. Sharp error bounds for imbalanced classification : how many examples in the minority class ? In : *International Conference on Artificial Intelligence and Statistics*, p. 838-846.

Clémençon, S. ; Huet, N. ; Sabourin, A., 2024. Regular variation in Hilbert spaces and principal component analysis for functional extremes. *Stochastic Processes and their Applications*. T. 174, p. 104375.

Lamalle, F. ; Feuillard, V. ; Sabourin, A. ; Clémençon, S., 2024. Weibull mixture estimation based on censored data with applications to clustering in reliability engineering. *Quality and Reliability Engineering International*.

Aghbalou, A. ; Sabourin, A. ; Portier, F., 2023. On the bias of K-fold cross validation with stable learners. In : *International Conference on Artificial Intelligence and Statistics*, p. 3775-3794.

Clémençon, S. ; Jalalzai, H. ; Lhaut, S. ; Sabourin, A. ; Segers, J., 2023. Concentration bounds for the empirical angular measure with statistical learning applications. *Bernoulli*. T. 29, nᵒ 4, p. 2797-2827.

Lhaut, S.; Sabourin, A.; Segers, J., 2022. Uniform concentration bounds for frequencies of rare events. *Statistics & Probability Letters*. T. 189, p. 109610.

Drees, H.; Sabourin, A., 2021. Principal component analysis for multivariate extremes. *Electronic Journal of Statistics*. T. 15, nº 1, p. 908-943.

Chiapino, M.; Clémençon, S.; Feuillard, V.; Sabourin, A., 2020. A multivariate extreme value theory approach to anomaly clustering and visualization. *Computational Statistics*. T. 35, nº 2, p. 607-628.

Jalalzai, H.; Colombo, P.; Clavel, C.; Gaussier, E.; Varni, G.; Vignon, E.; Sabourin, A., 2020. Heavy-tailed representations, text polarity classification & data augmentation. *Advances in Neural Information Processing Systems*. T. 33, p. 4295-4307.

Chiapino, M.; Sabourin, A.; Segers, J., 2019. Identifying groups of variables with the potential of being large simultaneously. *Extremes*. T. 22, nº 2, p. 193-222.

Jalalzai, H.; Clémençon, S.; Sabourin, A., 2018. On binary classification in extreme regions. In : *Advances in Neural Information Processing Systems*, p. 3092-3100.

Achab, M.; Clémençon, S.; Garivier, A.; Sabourin, A.; Vernade, C., 2017. Max K-armed bandit : On the ExtremeHunter algorithm and beyond. In : *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, p. 389-404.

Goix, N.; Sabourin, A.; Clémençon, S., 2017. Sparse representation of multivariate extremes with applications to anomaly detection. *Journal of Multivariate Analysis*. T. 161, p. 12-31.

Sabourin, A.; Segers, J., 2017. Marginal standardization of upper semicontinuous processes. with application to max-stable processes. *Journal of Applied Probability*. T. 54, nº 3, p. 773-796.

Thomas, A.; Clemencon, S.; Gramfort, A.; Sabourin, A., 2017. Anomaly Detection in Extreme Regions via Empirical MV-sets on the Sphere. In : *Artificial Intelligence and Statistics*, p. 1011-1019.

Chiapino, M.; Sabourin, A., 2016. Feature clustering for extreme events analysis, with application to extreme stream-flow data. In : *International Workshop on New Frontiers in Mining Complex Patterns*, p. 132-147.

Goix, N.; Sabourin, A.; Clémençon, S., 2016. Sparse representation of multivariate extremes with applications to anomaly ranking. In : *Artificial Intelligence and Statistics*, p. 75-83.

Goix, N.; Sabourin, A.; Clémençon, S., 2015a. On anomaly ranking and excess-mass curves. In : *Artificial intelligence and statistics*, p. 287-295.

Goix, N.; Sabourin, A.; Clémençon, S., 2015b. Learning the dependence structure of rare events : a non-asymptotic study. In : *Conference on Learning Theory*, p. 843-860.

Sabourin, A., 2015. Semi-parametric modeling of excesses above high multivariate thresholds with censored data. *Journal of Multivariate Analysis*. T. 136, p. 126-146.

Sabourin, A.; Renard, B., 2015. Combining regional estimation and historical floods : A multivariate semiparametric peaks-over-threshold model with censored data. *Water Resources Research*. T. 51, nº 12, p. 9646-9664.

Sabourin, A.; Naveau, P., 2014. Bayesian Dirichlet mixture model for multivariate extremes : A re-parametrization. *Comput. Stat. Data Anal.* T. 71, p. 542-567.

Sabourin, A.; Naveau, P.; Fougères, A.-L., 2013. Bayesian model averaging for multivariate extremes. *Extremes*. T. 16, p. 325-350.

INVITED TALKS SINCE 2013

- International conferences : Bernoulli-IMS world congress (2024), Extreme Value Analysis (2015, 2019, 2023), IMS annual meeting (2022), CMstatistics (2017, 2018, 2019,2022) ISNPS (2016), IMS China (2015), JSM (2014)

- Workshops and seminars : I2M, Marseille (2024); Workshop 'Eurandom 25th anniversary', Eindhoven (2024); Journée scientifique IpParis (2021); Research Training Group 'High-dimensional Phenomena in Probability', Ruhr University Bochum (2021); workshop 'Valpred' of ANR T-REX (2021,22); MIA seminar, AgroParisTech (2020, 2021); chair Stress Test's seminar - Ecole Polytechnique/BNPP (2020); Workshop AI4OAC, ANR MELODY, Brest

(2020) ; 'séminaire Parisien de Statistique' (2020) ; CMAP's seminar, Ecole Polytechnique (2019) ; 'Rencontres Statistiques Lyonnaises', Lyon I (2019) ; Working-Group-On-Risk, Crear-Essec (2018) ; workshop 'Extreme Values in applications', EPFL (2017) ; workshop 'Statistics / Learning Paris-Saclay', IHES (2017) ; workshop 'STOR-i extremes', Université de Lancaster (2016) ; workshop 'Extreme value modeling and Water resources' - Aussois (2016) ; team seminar at University of Hamburg (2016) ; MAP5 seminar, Paris-Descartes (2016) ; workshop 'Extremes, Copula and Actuarial Sciences', Luminy (2016) ; workshop 'Dependence, stability and Extremes', Toronto (2016) ; AppliBUGS (2015) ; MODAL'X seminar, Nanterre (2015) ; workshop 'Statistics of Extreme Events', KAUST (2014) ; Rencontres statistiques de Rochebrune (2014).

RESEARCH PROJECTS, GRANTS AND PRIZES
- Principal Investigator of ANR EXSTA (2024–2029).
- member of ANR T-REX (2021-2024, led by Clément Dombry) : axis coordinator.
- member of ANR MELODY (2019-2023, led by Ronan Fablet)
- PEPS JCJC INS2I 2015 'Agreed' : Apprentissage statistique pour l'analyse des événements rares, dans des problèmes de grande dimension.
- member of ANR Ameriska (2014-2016, led by Olivier Wintenberger)
- 2011 : Award for best contribution at the workshop 'Environmental Risk and Extreme Events', Ascona.

SHORT VISITS
- University of Hamburg (Holger Drees), 2016
- Ulm University (Evgeny Spodarev), 2015
- Université de Louvain-La-Neuve (Johan Segers), 2013

PHD STUDENTS
- Florian Lamalle (2023 - . . .) (CIFRE) co-advised with Vincent Feuillard (Renault) and Stephan Clémençon.
- Nathan Huet (2021 - . . .) co-advised with Stephan Clémençon.
- Anass Aghbalou (2020 - 2024) co-advised with François Portier, Patrice Bertail.
- Hamid Jalalzai (2017 - 2020) : co-advised with Eric Gaussier and Chloé Clavel.
- Robin Vogel (2017- 2020) : co-advised with Stephan Clémençon.
- Mastane Achab (2016- 2020) : co-advised with Stephan Clémençon.
- Maël Chiapino (2014-2018) : co-advised with François Roueff.
- Nicolas Goix (2013-2016) : co-advised with Stephan Clémençon.

MASTER'S STUDENTS
Alexandre Mansire (2024), Florian Lamalle (2021), Aya Slimen (2020) with the chair X-BNPP.

TEACHING

PRESENT COURSES (2023-24) :
- Statistical Learning with Extreme Values, Master's program MVA, Université Paris-Saclay.
- High dimensional statistics, Master's program 'Mathematics and Applications', Université Paris Cité.

MAIN PAST COURSES : Between 180 and 240 HET per year (standard unit hours, 'heures équivalent TD' in France).
- Tail events analysis, Master's program 'Data Science' from Ecole Polytechnique, Institut polytechnique de Paris.
- Bayesian Statistics, Master's program 'Data Science', Ecole Polytechnique, Institut polytechnique de Paris.
- Extremes, Master's program (2nd year) 'Mathematics of Randomness', Orsay.

- Statistics, Master's program (1st year) 'Mathematics and Applications', Université paris Cité.
- Non Parametric Statistics, Master's program (2nd year) 'Mathematics and Applications', Université Paris Cité.
- Data Analysis, Master's program (1st year) 'Mathematics and Applications', Université Paris Cité.
- Probability, Statistics, Numerical methods, Bayesian Filtering : 1st and 2nd years at Télécom Paris (graduate program).

PEDAGOGICAL LEADERSHIP
- 2023- ; Head of the Master's program 'Mathématiques, Données, Apprentissage' (Mathematics and applications), Université Paris Cité
- 2022- : Co-head of the second year of the Master's track 'Ingénierie Mathématique et Biostatistique', Université Paris Cité.
- 2021-22 : Co-head of the Mastère Spécialisé 'Big Data' at Télécom Paris (One-year graduate professional training program).
- 2018-2021 : Co-head of the continuing education training program 'CES Data Scientist' at Télécom Paris.

MISCELLANEA
- 2016 : Invited course 'Théorie des valeurs extrêmes', thematic semester AMERISKA, Nanterre-UPMC.
- 2014 : Online Course 'Les fondamentaux du Big Data', Télécom Paris : 10 videos recorded (probability, statistics)

EDUCATION

2021: Habilitation (HDR), Institut polytechnique de Paris,
Extreme Value Theory and Machine Learning.
Jury : Stéphane Boucheron (Reviewer), Richard Davis (Reviewer), Matthieu Lerasle (Reviewer), Clément Dombry, Holger Drees, Johan Segers.

2010-2013: PhD in applied mathematics (Statistics)
Bayesian model mergings for multivariate extremes, Application to regional predetermination of floods with incomplete data
Jury : Anne-Laure Fougères (advisor) Philippe Naveau (co-advisor), Anthony Davison (reviewer), Johan Segers (reviewer), Clémentine Prieur, Stéphane Robin, Eric Sauquet.

2009-2010: Master 'Statistics and Information', at *Université Paris-Dauphine*, "Traitement Statistique de l'information", Paris, France. Joint diploma with ENSAE (National School of Statistics and Economical Administration).

2008-2009: Environmental engineering at *Ecole Nationale du Génie Rural, des Eaux et des Forêts*, Paris, France.

2005-2008: *Ecole Polytechnique*, Paris, France.

MISCELLANEA

- Maternity leaves : June to October 2020, April to September 2017.
- Organisation of events : Workshop Extremes, Statistics and Statistical Learning, 2021. workshop Rare Events, Extremes and Machine Learning, 2018.
- Organisation of conference sessions in international conferences 'Extreme value Analysis (2017, 2019, ISNPS (2016),
- 2019-2020 : Associate researcher in the industrial chair 'stress-test' from Ecole Polytechnique-BNP Paribas
- Since 2013 : active member of the industrial chair DSAIDIS (previously MLBD), scientific axis coordinator (Machine learning for trusted and robust decision).

- Reviewer for journals (*Bernoulli, Annals of Statistics, Journal of Applied probability, JASA, Biometrika, Journal of Multivariate Analysis, Extremes, Journal of Statistical Computation and Simulation, Water Resources Research, Journal of Statistical Planning and Inference, Computational Statistics and Data Analysis*) and for conferences (*NeurIPS, ICML, ACML, AISTATS, ECML-PKDD, UAI*). Meta-reviewer for AISTATS, 2018.
- Member of thesis committees (apart from co-advised students) : Albert Thomas (2017), Olivier Ho (2018), Alban Siffer (2019), Oihana Coustie (2021), Théobald de Riberolles (2021), Kamelia Daudel (2021), Samuel Stocksieker (2024)
- Reviewer of PhD manuscripts : Kévin Ducharlet (2023)
- Member of recruitment committees
    - for Mcf position (Section 26) : Sorbonne Université, LPSM (2018), Télécom Sud Paris (2021), Télécom Paris (2022), Université de Franche-Comté, Laboratoire de Mathématiques (2022,2023) ; Université Paris-Saclay, laboratoire de mathématiques d'Orsay (2023) ; IUT Paris Rives de Seine (2024), Université de Bourgogne, Institut Mathématiques de Bourgogne (2024).
    - for Full Professor positions (Section 26) : Université Paris-Cité, LPSM (2024)