

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- ...** : 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

- Aghbalou, A., Bertail, P., Portier, F., & Sabourin, A. (2022). Cross validation for rare events. arXiv preprint arXiv :2202.00488.
- N. Huet, S. Cléménçon, and A. Sabourin. "On Regression in Extreme Regions". arXiv preprint arXiv :2303.03084 (2023).

PUBLICATIONS

- Aghbalou, A., Sabourin, A., & Portier, F. (2023). On the bias of K-fold cross validation with stable learners. In *AISTATS proceedings*, PMLR.
- Aghbalou, A., Portier, F., Sabourin, A., Zhou, C. (2023). Tail inverse regression for dimension reduction with extreme response. To appear in *Bernoulli*. arXiv preprint arXiv :2108.01432.
- Cléménçon, S., Jalalzai, H., Sabourin, A., & Segers, J. (2023). Concentration bounds for the empirical angular measure with statistical learning applications. To appear in *Bernoulli*. arXiv preprint arXiv :2104.03966.
- Lhaut, S., Sabourin, A., & Segers, J. (2022). Uniform concentration bounds for frequencies of rare events. *Statistics & Probability Letters*, 189, 109610. arXiv preprint arXiv :2110.05826.
- Drees, H. & Sabourin, A. (2021) "Principal component analysis for multivariate extremes.", *Electronic Journal of Statistics*, 15(1), 908-943
- Jalalzai, H., Colombo, P., Clavel, C., Gaussier, E. Varni, G., Vignon, E. & Sabourin, A. (2020) "Heavy-tailed Representations, Text Polarity Classification and Data Augmentation", *NeurIPS*.
- Chiapino, M., Sabourin, A., & Segers, J. (2019). Identifying groups of variables with the potential of being large simultaneously. *Extremes*, 22(2), pp. 193-222.
- Chiapino, M., Cléménçon, S., Feuillard, V., Sabourin, A. (2019). A multivariate extreme value theory approach to anomaly clustering and visualization. *Computational Statistics*, 1-22.
- Jalalzai, H., Cléménçon, S., Sabourin, A. (2018). On binary classification in extreme regions. *NeurIPS* (pp. 3092-3100).
- Achab, M., Cléménçon, S., Garivier, A., Sabourin, A., Vernade, C. (2017). Max K-armed bandit : On the ExtremeHunter algorithm and beyond. *ECML- PKDD* (pp. 389-404)
- Thomas, A., Cléménçon, S., Gramfort, A., Sabourin, A. (2017). Anomaly Detection in Extreme Regions via Empirical MV-sets on the Sphere. *AISTATS* (Vol. 54, pp. 1011-1019).

- Sabourin, A., Segers, J. (2017). Marginal standardization of upper semicontinuous processes. with application to max-stable processes. *Journal of Applied Probability*, 54(3), 773-796.
- Goix, N., Sabourin, A., Cléménçon, S. (2017). Sparse representation of multivariate extremes with applications to anomaly detection. *Journal of Multivariate Analysis*, 161, 12-31.
- 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* (pp. 132-147). Springer, Cham.
- Goix, N., Sabourin, A., Cléménçon, S. (2016). Sparse Representation of Multivariate Extremes with Applications to Anomaly Ranking. *AISTATS* (pp. 75-83).
- Sabourin, A., and Renard, B. (2015). Combining regional estimation and historical floods : a multivariate semiparametric peaks-over-threshold model with censored data *Water resources research*, 51(12), 9646-9664.
- Goix, N., Sabourin, A., Cléménçon, S. (2015). On anomaly ranking and excess-mass curves. In *AISTATS* (pp. 287-295).
- Sabourin, A. (2015). Semi-parametric modeling of excesses above high multivariate thresholds with censored data. *Journal of Multivariate Analysis*, 136, 126-146.
- Sabourin, A., Naveau, P. (2014). Bayesian Dirichlet mixture model for multivariate extremes : a re-parametrization. *Computational Statistics and Data Analysis*, 71, 542-567..
- Sabourin, A. , Naveau, P., Fougères, A-L., (2013) Bayesian Model averaging for Multivariate extremes, *Extremes*, 1–26.

INVITED TALKS SINCE 2013

- International conferences : , 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 : Journée scientifique IpParis (2021); Research Training Group ‘High-dimensional Phenomena in Probability’, Ruhr University Bochum (2021); workshop ‘Valpred’ of ANR T-REX (2021); 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

- member of ANR T-REX (2021-2024, lead by Clément Dombry) : axis coordinator.
- member of ANR MELODY (2019-2023, lead 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, lead 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 (January 2023 - ...) (CIFRE) co-advised with Vincent Feuillard (Renault) and Stephan Cléménçon

- Nathan Huet (October 2021 - ...) co-advised with Stephan Clémentçon.
- Anass Aghbalou (December 2020 - ...) 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 Clémentçon.
- Mastane Achab (2016- 2020) : co-advised with Stephan Clémentçon.
- Maël Chiapino (2014-2018) : co-advised with François Roueff.
- Nicolas Goix (2013-2016) : co-advised with Stephan Clémentçon.

MASTER'S STUDENTS

Florian Lamalle (2021), Aya Slimen (2020) co-advised with Emmanuel Gobet and BNPP.

TEACHING

PRESENT COURSES (2022-23) :

- Non Parametric Statistics, Master's program (2nd year) Mathematics and Applications, Université Paris Cité
- Survival Analysis, Master's program (2nd year) Mathematics and Applications, Université Paris Cité
- Statistics, Master's program (1st year) Mathematics and Applications, Université Paris Cité
- Data Analysis, Master's program (1st year) Mathematics and Applications, Université Paris Cité
- Tail events analysis : Master's program 'data science' from École Polytechnique, Paris.

MAIN PAST COURSES : Between 180 and 240 HET per year (standard unit hours, 'heures équivalent TD' in France).

- 1st and 2nd years at Télécom Paris (graduate program) : Probability, Statistics, Numerical methods, Bayesian Filtering.
- Master's course 'Bayesian Statistics' (Ecole Polytechnique)

PEDAGOGICAL LEADERSHIP

- 2022- : Co-head of the second year of the Master 'Ingénierie Mathématique et Biostatistique', Université Paris Cité.
- 2021- : 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).
- Reviewer of PhD manuscripts : Kévin Ducharlet (2023, scheduled)
- Member of recruitment committees for MCF position (Section 26) : at 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)