

📍 Bâtiment Sophie Germain, office 5032
8 place Aurélie Nemours, 75013 Paris, France

🌐 <https://hassothea.github.io/>
✉ sothea.has@lpsm.paris

EDUCATION

- France
2022 - Present **LPSM - Université Paris Cité and LMD - École Polytechnique**
Postdoctoral research under supervision of Riwal Plougonven (LMD - École Polytechnique) and Aurélie Fischer (LPSM - Université Paris Cité).
Research task *Improving parametrizations in climate modeling using statistical and machine learning methods (tree-based ensemble algorithms and deep learning). Modeling observed Gravity Wave Momentum Fluxes (a very important quantity in climate modeling) and extracting important features (from simulations and observations). Interpret and provide information of the stochastic component of GWMFs.*
- France
2018 - 2022 **Sorbonne University Pierre and Marie Curie - Paris 6**
Ph.D. in Applied Mathematics
Research topic *Consensual aggregation and distance measurements for statistical learning. Theoretical contributions and applications to the field of energy.*
🔗 KFC procedure: data modeling using clustering, local modeling and consensual aggregation methods.
🔗 Gradient COBRA: an aggregation technique for combining several predicted features for regression tasks.
🔗 Aggregation on randomly projected features for regression: an extension of the second work to high-dimensional predicted feature spaces.
- France
2017 - 2018 **University Paris Diderot - Paris 7**
Master 2 Random Modelling and Data Science (M2MO)
Project Data Science for Company, Massive Data Processing (R-programming).
Exam Statistical Learning, Statistical Modeling, Diffusion Statistics, Stochastic Calculus.
Both Machine Learning (Python), Monte Carlo Method (C++).
- France
2016 - 2017 **École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise - ENSIE**
Master 1 Applied Mathematics
Project Time Series, Simulation Methods, Research Project in Finance, Machine Learning.
Exam Stochastic Process, Operation Research, Stochastic Calculus in Finance.
Both Data Analysis, Numerical Methods for PDE, C++.
- Cambodia
2009 - 2015 **Royal University of Phnom Penh - RUPP**
Bachelor's Degree and Master 1 of pure mathematics

PUBLICATIONS

- 2023 **Estimating balloon-observed gravity wave momentum fluxes using ML and input from reanalysis.**
Status *In progress, with R. Plougonven, A. Fischer, R. Rani, F. Lott, A. Hertzog, A. Podglajen, M. Corcos.*
- 2023 **Gradient COBRA: A kernel-based consensual aggregation for regression.**
Status *Accepted for publication at Journal of data Science Statistics and Visualisation.*
- 2022 **A consensual aggregation on randomly projected high-dimensional features of predictions for regression.**
Status *Published in HAL.*
- 2022 **Machine learning methods applied to the global modeling of event-driven pitch angle diffusion coefficients during high-speed streams.**
Research topic *Coupled Feedback Mechanisms in the Magnetosphere-Ionosphere System,*
Status *Published in Frontiers Physics, with G. Kluth, J.F. Ripoll, A. Fischer, M. Mougeot, and E. Camporeale.*
- April 2021 **KFC: A clusterwise supervised learning procedure based on aggregation of distances.**
Status *Published in Journal of Statistical Computation and Simulation, with Aurélie Fischer and Mathilde Mougeot.*

EXPERIENCES

- 2022-Present **CNRS, LPSM Université Paris Cité and LMD École Polytechnique**
Position *Postdoctoral research in atmospheric science using statistical and applied machine learning.*
- 2018 - 2022 **LPSM (UMR 8001) - Sorbonne Université**
Position *Ph.D. research in applied mathematics and theoretical machine learning.*
- 2018 - Present **UFR Mathematics Université de Paris**
Position *Teaching assistant and temporary contractor of teaching and research (ATER)*
🔗 Practical class of Data Analysis with **R** and **Rstudio**, Master1 ISIFAR.
🔗 Practical class of Data Mining with **R** and **Rstudio**, Master2 ISIFAR.
🔗 Practical class of Exploratory Data Analysis with **R** and **Rstudio**, Master1 EDA.
🔗 Practical class of Algorithm and Programming with **Python**, Licence2 MIASHS.
🔗 Practical class of Big Data Technologies with **Python** and **Spark**, Master1 MATINF.
🔗 Tutorial class of Statistical Inference and Data Modeling, Master2 Random and Data Modeling (M2MO).
- 2018 April - Sep **LPSM (UMR 8001) Université de Paris**
M2 internship: predictive models based on clustering with Bregman divergences and local predictions
Analysis of sensitivity of K-means clustering with Bregman divergences on several types of datasets. Different clustering structures of inputs are estimated, then local predictive models are trained on separate clusters. Such a two-step modeling is applied in many domains, especially in energy modeling.
- 2017 June - Sep **Laboratory of TELECOM SudParis**
M1 internship: optimization problem with simulated annealing algorithm
Understanding the convergence property of simulated annealing algorithm, which is a probabilistic method for estimating the global optimizer of a given function (deterministic or non-deterministic).

SCHOLARSHIP & AWARDS

- 2018 - 2022 **LPSM Scholarship**
Ph.D. and research funds.
- 2017 - 2018 **ENSIIE Scholarship**
Second year Master's degree of M2MO at Université Paris Diderot (Paris 7).
- 2016 - 2017 **Erasmus+ Scholarship**
First year Master's degree of Applied Mathematics at ENSIIE, France.
- 2014 - 2016 **International Mathematics Union (IMU)**
2-year Master's degree of Pure Mathematics at Royal University of Phnom Penh.
- 2009 - 2013 **Ministry of Education of Cambodia Scholarship**
4-year Bachelor's degree of Mathematics at Royal University of Phnom penh.

LANGUAGES & PROGRAMMING

- Languages Khmer (Mother tongue), English (fluent), French (conversational)
- Programming **R**: tidyverse, dplyr, ggplot, plotly, ...
Python: TensorFlow, pandas, scikit-learn, PySpark, ...
Others: C++, Matlab, Scilab, \LaTeX .

PERSONAL INTEREST

- Reading Behavioral science, meditation and AI researches.
- Sports Volleyball, basketball and football.
- Other interests Music, guitar, a little bit of drum and drawing.