Steven **Golovkine**

PhD in Applied Mathematics, Statistics

(+33) 06 72 28 22 71 | steven_golovkine@icloud.com | stevengolovkine.netlify.app | StevenGolovkine | Steven Golovkine

Current Position

Postdoctoral Researcher on the FAST project

Limerick, Ireland

University of Limerick

Feb. 2022 - Present

- Develop novel, computationally efficient statistical models and algorithms for the modelling of multivariate sensor data.
- Demonstrate increased computational efficiency, so that the methodologies are applicable to modern, large-scale datasets.
- Supervisor: Norma Bargary (UL) and Andrew Simpkin (NUI Galway)

Experience

Data Scientist on the EMPOWER project

Paris, France

IRMES (Insep)

Jun. 2021 - Jan. 2022

- Maximize the performance of elite female athletes by optimizing their training responses with adapted workloads in synergy with their physiology and menstrual cycle.
- Develop statistical models to determine athletes hormonal profiles and analyze responses to training and competition loads.
- Link: https://labos-recherche.insep.fr/fr/empower

Research Engineer

Guyancourt, France

Technocentre, Renault

Jan. 2018 - Mar. 2021

- Develop clustering methods for the analysis of autonomous vehicle Advanced Driver-Assistance Systems data.
- Create a Python package for Functional Data Analysis: FDApy.
- Use of Google Cloud Platform (Compute Engine and BigQuery) for the analysis of vehicle data.
- Software: Python and **R**.

Data Scientist (Intern)

Guyancourt, France

Technocentre, Renault

Apr. 2017 - Oct. 2017

- Implement a massive data analysis methodology for the validation of driving assistance systems.
- Software: Matlab.

Research Assistant

Fort Collins, USA

Colorado State University

Jun. 2016 - Aug. 2016

- Compare survey estimators for the *National Survey of College Graduation*.
- Software: **R**.

Education

PhD. in Applied Mathematics, Statistics

Guyancourt, France

Technocentre, Renault and CREST (Ensai)

Jan. 2018 - Jun. 2021

- **Title:** Statistical methods for multivariate functional data
- **Supervisors:** Valentin Patilea (Ensai, CREST), Nicolas Klutchnikoff (Univ Rennes, IRMAR)
- **Funding:** Partnership with Groupe Renault through a CIFRE convention.
- **Abstract:** The topic of this thesis is related to functional data analysis and is motivated by modern data from automobile industry. The standard functional data methods rely on the assumption that the curves are continuously observed, without error. However, in general, the real data is neither continuously nor exactly observed. Therefore, a crucial step is to recover the trajectories from noisy measurements at discrete random points. For that, we propose an original point of view: the local regularity of the process generating the curves. Thus, combining information both within and across trajectories, we propose a simple estimator for this local regularity. Given this estimate, we build a nearly optimal local polynomial smoother of the curves from a sample of noisy trajectories. Nonparametric estimators for the mean and the covariance functions of functional data, using the local regularity of the process, are derived. Moreover, we propose a model-based clustering algorithm for a general class of functional data for which the components could be curves or images. Results of both simulated and real data show the good performances of this method. A Python package, implementing the methods and publicly available, has been developed.
- **Keywords:** Adaptive optimal smoothing, functional data analysis, gaussian mixtures, Hölder exponent, local polynomials, model-based clustering, multivariate functional principal component analysis, traffic flow.
- **Defense**: June 18, 2021 in front of the jury composed of André Mas (Professor, Montpellier University, President), Sophie Dabo-Niang (Professor, Lille University, Reviewer), Alois Kneip (Professor, Bonn University, Reviewer), Vincent Feuillard (Statistical Expert, Renault, Examiner) and Claire Gormley (Professor, University College Dublin, Examiner).
- **Manuscript**: tel-03540827

MSc in Big Data — *Rennes, France*

Ensai (National School for Statistics and Data Analysis) — *Sept. 2016 - Oct. 2017*

- Dual degree program with Ensai engineering degree.
- Main topics: Statistics, Applied Mathematics, Computer Science.
- Training topics: assess, treat, and analyze massive amounts of heterogeneous data.
- Program taught entirely in English.

M.S. in Statistics (*Diplôme d'ingénieur*) — *Rennes, France*

Ensai (National School for Statistics and Data Analysis) — *Sept. 2014 - Oct. 2017*

- Training topics: Statistics, Econometrics and Computer Science.

Statistics of random processes — *Aarhus, Danmark*

Aarhus University — *Jan. 2016 - Jun. 2016*

- ERASMUS exchange.

CPGE MPSI/MP — *Reims, France*

Lycée Clemenceau — *Sep. 2011 - Jun. 2014*

Skills

Programming	Python, **R**, Bash
Tools	Google Cloud Platform, Git, Continuous Integration
Languages	French (mother tongue), English (C1, fluent)

Teaching Experience

R for Statistical Data Science — *University of Limerick*

MS6071 — *Sep. 2023 - Dec. 2023*

- Link: r-programming-course.netlify.app

Engineering maths 5 — *University of Limerick*

MA4006 — *Feb. 2023*

- Vector fields and integration.

Duration models — *Ensai*

M.S in Statistics — *Apr. 2021 - Jun 2021*

- Theoretical and practical (with **R**) sessions.

Linear regression — *Ensai*

M.S in Statistics — *Sep. 2019 - Nov. 2021*

- Theoretical and practical (with **R**) sessions.

Scientific production

Articles

A multivariate multilevel longitudinal functional model for repeatedly observed human movement data — *arXiv preprint*

Gunning E., Golovkine S., Simpkin A. J., Burke A., Dillon S., Gore S., Moran K., O'Connor S., Whyte E. & Bargary N. — *2024*

- Link: arXiv:2408.08481

Analysing kinematic data from recreational runners using functional data analysis — *arXiv preprint*

Gunning E., Golovkine S., Simpkin A. J., Burke A., Dillon S., Gore S., Moran K., O'Connor S., Whyte E. & Bargary N. — *2024*

- Link: arXiv:2408.08200

On the estimation of the number of components in multivariate functional principal component analysis *arXiv preprint*

Golovkine S., Gunning E., Simpkin A. J. & Bargary N. *2023*

- Link: arXiv:2311.04540

On the use of the Gram matrix for multivariate functional principal components analysis *arXiv preprint*

Golovkine S., Gunning E., Simpkin A. J. & Bargary N. *2023*

- Link: arXiv:2306.12949

Learning the smoothness of noisy curves with application to online curve estimation *EJS*

Golovkine S., Klutchnikoff N. & Patilea V. *2022*

- DOI: 10.1214/22-EJS1997

Clustering multivariate functional data using unsupervised binary trees *CSDA*

Golovkine S., Klutchnikoff N. & Patilea V. *2021*

- DOI: 10.1016/j.csda.2021.107376

Adaptive optimal estimation of irregular mean and covariance functions *arXiv preprint*

Golovkine S., Klutchnikoff N. & Patilea V. *2021*

- Link: arXiv:2108.06507

FDApy: a Python package for functional data *arXiv preprint*

Golovkine S. *2021*

- Link: arXiv:2101.11003

Conference proceedings

Functional multilevel modelling of the influence of the menstrual cycle on the performance of female cyclists *IWSM*

Golovkine S., Chassard T., Meignié A., Brunet E., Toussaint J.-F. & Antero J. *2023*

- Link: Proceedings of the 37th International Workshop on Statistical Modelling

Clustering multivariate functional data using unsupervised binary trees *EYSM*

Golovkine S., Klutchnikoff N. & Patilea V. *2021*

- Link: 22nd European Young Statisticians Meeting - Proceedings

Lissage de données fonctionnelles par estimation de leur régularité locale *JDS*

Golovkine S., Klutchnikoff N. & Patilea V. *2020*

- Link: Proceedings of the 52*èmes Journées de Statistiques de la Société Française de Statistique*

Softwares

FDApy (`https://github.com/StevenGolovkine/FDApy`)

Python package

denoisr (`https://github.com/StevenGolovkine/denoisr`)

R package

funestim (`https://github.com/StevenGolovkine/funestim`)

R package

Conferences

With talk

May 2024	**JDS,** 55rd Statistical Days	*Bordeaux, France*
Sep. 2023	**NSHPC,** National Sport and Human Performance Conference 2023	*Limerick, Ireland*
Jul. 2023	**IWSM,** 37th International Workshop on Statistical Modelling	*Dortmund, Germany*
Jul. 2023	**JDS,** 54rd Statistical Days	*Brussels, Belgium*
May 2023	**CASI,** 43rd Conference on Applied Statistics in Ireland	*Killarney, Ireland*
Aug. 2022	**SDS,** CSDA & EcoSta Workshop on Statistical Data Science (invited)	*Bologna, Italy*
Jun. 2022	**JDS,** 53rd Statistical Days	*Lyon, France*
Feb. 2022	**YSP,** 10th Young Statisticians and Probabilists day (invited)	*Virtual Conference*
Dec. 2021	**CMStatistics,** 14th conference on Computational and Methodological Statistics (invited)	*Virtual Conference*
Sep. 2021	**EYSM,** 22nd European Young Statisticians Meetings (invited)	*Virtual Conference*
Jun. 2021	**JDS,** 52nd Statistical Days	*Virtual Conference*
Mar. 2021	**Mathematics Seminars,** Hunter College (invited)	*Virtual Conference*
Dec. 2020	**CMStatistics,** 13th conference on Computational and Methodological Statistics (invited)	*Virtual Conference*
Nov. 2020	**StatMod2020,** Statistical Modeling with Applications (invited)	*Virtual Conference*
Jun. 2019	**JDS,** 51st Statistical Days	*Nancy, France*
Mar. 2019	**MASCOT-NUM,** Annual Conference	*Rueil-Malmaison, France*

Attended

Jun. 2019	**DS^3,** Data Science Summer School	*Saclay, France*

Student supervision

Postgraduate students

2024	**Abhilash Patade,** Bayesian analysis for football data.	*University of Limerick*
2023	**Ansari Z. H., Moore J. & Ndukwe R. N.,** Spatio-temporal models for disease mapping.	*University of Limerick*
2023	**McDermott H.,** The home advantage in rugby.	*University of Limerick*
2022	**Harrington T.,** Bayesian analysis for sport data.	*University of Limerick*
2022	**Scanlon S.,** Does grid position affect who wins the race in F1?	*University of Limerick*

Undergraduate students

2020	**Manaa A., Maissoro H. & Samaila Z. ,** Classification de courbes après recalage.	*ENSAI*
2020	**Gervillie R., Li J.-U. & Rousseaux F. ,** Détection de points d'impact dans des courbes.	*ENSAI*