Facundo Fabián Sapienza

Doctoral Candidate
Department of Statistics, UC Berkeley

Contact: fsapienza@berkeley.edu
Website: facusapienza.org
Citizenship: Argentinian, Italian
Languages: Spanish (native), English (fluent)

I am a motivated young scientist pursuing my doctoral studies in statistics with a background in mathematics, physics, and data science. My research focuses in physics-informed machine learning methods applied for ice flow modeling. I am broadly interested in ways of how artificial intelligence will shape our future understanding of the natural world, scientific software development and open science.

1 Education

Ph.D. in Statistics, 2019-2024
University of California, Berkeley
Dissertation: *Physics-informed machine learning for glacier modelling based on differentiable programming*
Advisors: *Fernando Pérez, Jonathan Taylor*

BSc/MSc in Mathematics[1]*, 2018*
Universidad de Buenos Aires, Argentina
Thesis: *Fermat distance and geodesics in Euclidean percolation: theory and applications to machine learning.*
Advisor: *Pablo Groisman, Matthew Jonckheere*

BSc/MSc in Physics[2]*, 2017*
Universidad de Buenos Aires, Argentina
Thesis: *Resource theory of quantum thermodynamics: correlation in the processes of formation of resources.*
Advisor: *Augusto Roncaglia*

2 Professional experience

Senior data scientist, Aristas SRL, Argentina
March 2015 - August 2018

Visiting scholar, Courant Insitute of Mathematical Sciences, USA
October 2018

Research intern, Borealis AI, Canada
November 2018 - March 2019

Private instructor, Argentina
March 2012 - December 2014

[1] Licenciatura en Ciencias Matemáticas (5.5 years for completion)
[2] Licenciatura en Ciencias Físicas (6 years for completion)

3 Teaching experience

This is a list of academic courses I instructed as teaching assistant.

- STAT 201A: Introduction to Probability at an Advanced Level, UC Berkeley, Fall 2023
- STAT 159/259: Collaborative and Reproducible Data Science, UC Berkeley, Spring 2023
- STAT 159/259: Collaborative and Reproducible Data Science, UC Berkeley, Spring 2022
- Topics of Physics for Math majors, Universidad of Buenos Aires, Fall 2017

Developing STAT159-Collaborative and Reproducible Data Science

Together with Fernando Pérez (creator of IPython and co-founder of the Jupyter Project) we designed the curriculum and instructed STAT159/259-Collaborative and Reproducible Data Science at UC Berkeley The course is attended by for both undergraduate and graduate students, and introduces good computational skills and practices in the context of scientific computing and data science, necessary skills for reproducible research in modern science. Examples of this include how to package, test and document code in Python; fundamental and practical aspects of working collaboratively in Git and GitHub, including automation and deployment of pages; and tools from the Jupyter ecosystem (Notebook, Lab, Binder, Book). These are motivated with examples in Earth and Climate sciences, working around more advanced libraries such as Xarray. These tools constitute the building blocks of modern open, reproducible, and collaborative science, and are in alignment with the guidelines of the NASA Transform to Open Science (TOPS) program, and the 2023 Year of Open Science as announced along with the White House and other federal agencies.

4 Awards and fellowships

H2H8 Research support fellowship, USA, 2022-2023.

Berkeley Fellowship, USA, 2019-2021.

International Collegiate Programming Contest, South American Regionals finalist, 2014.

National Olympiads in Mathematics, Argentina, Annual Honorable Mentions.

5 Publication highlights

These are a few publications that reflect key topics in my research career. Section 7 contains a full list of papers and conference proceedings.

- BOLIBAR, J., SAPIENZA, F., MAUSSION, F., LGUENSAT, R., WOUTERS, B., and PÉREZ, F. "Universal Differential Equations for glacier ice flow modelling". In: *Geoscientific Model Development* 16.22 (2023), pp. 6671–6687. DOI: 10.5194/gmd-16-6671-2023

We introduce ODINN.jl, an open-source model for global glacier modelling making use of tools from the Scientific Machine Learning Julia ecosystem. ODINN uses Universal Differential Equations (UDEs) to learn subparts of a differential equation governing glacier ice flow. This work illustrates the first steps of a new global glacier modelling framework in Julia that allows the estimation of global empirical laws for the physical parameters. This paper was highlighted in Geoscientific Model Development.

- SAPIENZA, F., GALLO, L. C., ZHANG, Y., VAES, B., DOMEIER, M., and SWANSON-HYSELL, N. L. "Quantitative Analysis of Paleomagnetic Sampling Strategies". In: *Journal of Geophysical Research: Solid Earth* 128.11 (2023), e2023JB027211. DOI: https://doi.org/10.1029/2023JB027211

Since 2020, I engaged in a collaboration in Paleomagnetism aiming to introduce novel statistical tools to understand the past motion of tectonic plates. This involves the modeling of directional data corresponding to the time-varying Earth magnetic field. This paper constitutes a major advance in understanding how different field work sampling strategies impact the uncertainty on the final magnetic pole estimation, which has a major impact in the later uncertainty of continental drift paths. This paper was highlighted in the Eos article Should I Stay or Should I Go…To Another Paleomagnetic Site?.

- SAPIENZA, F., CERISOLA, F., and RONCAGLIA, A. J. "Correlations as a resource in quantum thermodynamics". In: *Nature communications* 10.1 (2019), p. 2492

The role of correlation between the parts of a quantum system are introduced and quantified in their relation with the energy and work required to form these states. This work represents a major advance in the conciliation of classical thermodynamics with quantum mechanics vie the theory of resource formalism. This paper is the result of my Master thesis in Physics at the Universidad de Buenos Aires, Argentina. (46 citations).

- SAPIENZA, F. et al. "An Analytical Model of Magnetic Field Draping in Induced Magnetospheres". In: (2024). In preparation.

As part of a project founded by the SMD Artificial Intelligence NASA initiative to accelerate the use of AI in planetary sciences, we applied Bayesian methods to model the effect of the solar wind in Mars' space environment using data from the MAVEN mission.

6 Professional Service and Synergistic Activities

Academic service

At UC Berkeley, I had serve as co-president of the Statistic Graduate Student Association (SGSA) in the period 2021-2022. I was also part of the diversity and seminar committee. I further serve as part of the graduate admission committee.

I was invited in two opportunities (2018, 2023) to the Heidelberg Laureate Forum, Germany, an event that gathers Fields, Abel, Turing, and Nevarlina laureates with a selected group of young researchers. In 2023, I was invited as one of the 20 alumni to comeback an serve as guides. I was a panelist in the panel *Career Paths for Mathematicians and Computer Scientist in Academia* and instructed the workshop *Physics-informed machine learning*.

I serve as co-convener in the following conference sessions about machine learning in the cryosphere.

1. American Geophysical Union Fall Meeting 2022, Chicago, USA; *A Data-Driven Cryosphere: Insights from Machine Learning and Other Statistical Methods*. Together with Whyjay Zheng, Tasha Snow, Jordi Bolibar, and Yara Mohajerani.
2. The 28th International Union of Geodesy and Geophysics (IUGG) General Assembly 2023, Berlin, Germany. *Data Driven Cryospheric Sciences: Machine Learning, Data Assimilation and Inverse Methods for the Cryosphere*. Together with Fabien Maussion, Celia Baumhoer, Hui Tang, and Jordi Bolibar.

Diversity, Equality and Inclusion

At UC Berkeley I had served in the diversity committee, committed to lower the access barrier to science. As a Latinx international scientist, I am aware of the difficulties that students have to go through in order to succeed in their academic path. During the last years, I have mentored undergraduate students from the Latinx community. It is my belief that the open science practices I promote will further help the bring members from underrepresented minorities to the scientific community.

Machine learning in glaciology: organizer and instructor

I organized and instructed the Machine Learning in Glaciology Workshop in Finse, Norway, during its two realization in September 2022 and April 2023. The workshop consist in 5 days during which the instructors give general lectures on data science and specific application to the cryosphere, and the participants (in their majority graduate, postdoctoral and researches working in glaciology with interest in machine learning) work in different programming projects. Together with Jordi Bolibar, I lead the machine learning for glacier mass balance modelling project. The workshop has been co-organized by Ellianna Abrahams (UC Berkeley), Jordi Bolibar (Utrecht University, TU Delft), Regine Hock (University of Oslo), Konstantin Maslov (University of Twente), Fernando Pérez (UC Berkeley), Benjamin Robson (University of Bergen), and Thomas Schellenberger (University of Oslo).

7 Publications

First author publications

1. BOLIBAR, J., SAPIENZA, F., MAUSSION, F., LGUENSAT, R., WOUTERS, B., and PÉREZ, F. "Universal Differential Equations for glacier ice flow modelling". In: *Geoscientific Model Development* 16.22 (2023), pp. 6671–6687. DOI: 10.5194/gmd-16-6671-2023.
2. SAPIENZA, F., GALLO, L. C., ZHANG, Y., VAES, B., DOMEIER, M., and SWANSON-HYSELL, N. L. "Quantitative Analysis of Paleomagnetic Sampling Strategies". In: *Journal of Geophysical Research: Solid Earth* 128.11 (2023), e2023JB027211. DOI: https://doi.org/10.1029/2023JB027211.
3. SAPIENZA, F., CERISOLA, F., and RONCAGLIA, A. J. "Correlations as a resource in quantum thermodynamics". In: *Nature communications* 10.1 (2019), p. 2492.
4. GROISMAN, P., JONCKHEERE, M., and SAPIENZA, F. "Nonhomogeneous Euclidean first-passage percolation and distance learning". In: *Bernoulli* 28.1 (2022), pp. 255–276.
5. SAPIENZA, F., GROISMAN, P., and JONCKHEERE, M. "Weighted Geodesic Distance Following Fermat's Principle". In: (2018).

First-author publication in preparation are in the *In Preparation* section.

Collaborative publications

6. AZARI, A. R., ABRAHAMS, E., SAPIENZA, F., MITCHELL, D. L., BIERSTEKER, J., XU, S., BOWERS, C., PÉREZ, F., DIBRACCIO, G. A., DONG, Y., and CURRY, S. "Magnetic Field Draping in Induced Magnetospheres: Evidence From the MAVEN Mission to Mars". In: *Journal of Geophysical Research: Space Physics* 128.11 (2023), e2023JA031546. DOI: https://doi.org/10.1029/2023JA031546.
7. CHAZAL, F., FERRARIS, L., GROISMAN, P., JONCKHEERE, M., PASCAL, F., and SAPIENZA, F. "Choosing the parameter of the Fermat distance: navigating geometry and noise". In: (2023). arXiv: 2311.18663 [stat.ML].

8. GALLO, L. C., DOMEIER, M., SAPIENZA, F., SWANSON-HYSELL, N. L., VAES, B., ZHANG, Y., ARNOULD, M., EYSTER, A., GÜRER, D., KIRÁLY, Á., ROBERT, B., ROLF, T., SHEPHARD, G., and BOON, A. VAN DER. "Embracing Uncertainty to Resolve Polar Wander: A Case Study of Cenozoic North America". In: *Geophysical Research Letters* 50.11 (2023), e2023GL103436. DOI: https://doi.org/10.1029/2023GL103436.
9. CERISOLA, F., SAPIENZA, F., and RONCAGLIA, A. J. "Heat engines with single-shot deterministic work extraction". In: *Physical Review E* 106.3 (2022), p. 034135.
10. GALLO, L. C., SAPIENZA, F., and DOMEIER, M. "An optimization method for paleomagnetic Euler pole analysis". In: *Computers & Geosciences* 166 (2022), p. 105150.
11. SMUCLER, E., SAPIENZA, F., and ROTNITZKY, A. "Efficient adjustment sets in causal graphical models with hidden variables". In: *Biometrika* 109.1 (2022), pp. 49–65.
12. VASQUEZ, C. A., SAPIENZA, F., SOMACAL, A., and FAZZITO, S. Y. "Anhysteretic remanent magnetization: model of grain size distribution of spherical magnetite grains". In: *Studia Geophysica et Geodaetica* 62 (2018), pp. 339–351.

In preparation

13. SAPIENZA, F. et al. "A review of sensitivity methods for differential equations". In: (2024).
14. SAPIENZA, F. et al. "Fitting curves in the sphere using universal differential equations". In: (2024).
15. SAPIENZA, F. et al. "An Analytical Model of Magnetic Field Draping in Induced Magnetospheres". In: (2024).
16. AZARI, A. et al. "A Virtual Upstream Solar Wind Monitor for Mars with Uncertainty Quantification from Gaussian Processes". In: *Geophysical Research Letters,* (2024).
17. GALLO, L. C. et al. "On the feasibility of paleomagnetic Euler pole analysis". In: (2024).

Software

18. BOLIBAR, J. and SAPIENZA, F. *ODINN-SciML/ODINN.jl: v0.2.0*. Version v0.2.0. June 2023. DOI: 10.5281/zenodo.8033313. URL: https://doi.org/10.5281/zenodo.8033313.
19. SAPIENZA, F., GALLO, L. C., ZHANG, Y., VAES, B., DOMEIER, M., and SWANSON-HYSELL, N. *PolarWandering/PaleoSampling (Version 1.0.0)*. Comp. software. Version 1.0.0. 2023. DOI: https://doi.org/10.5281/zenodo.8347149.
20. SAPIENZA, F. and OTHER. *fermat*. Version v0.2.7. Mar. 2023. URL: https://pypi.org/project/fermat/.

Theses

21. SAPIENZA, F. *Distancia de Fermat y geodésicas en percolación euclidea: teoría y aplicaciones en Machine Learning*. Master's thesis. Buenos Aires, Argentina, Aug. 2018.
22. SAPIENZA, F. *Teoría de recursos de la termodinámica cuántica : correlaciones en los procesos de formación de recursos*. Master's thesis. Buenos Aires, Argentina, Dec. 2017.

8 Conference Presentations

Selected Talks

Invited speaker at the 2023 AGU Fall Meeting, *ODINN.jl: Leveraging the Julia and Python Open Science Ecosystems to Enhance Glacier Modeling with Scientific Machine Learning* (2023).

Speaker at the 2021 AGU Fall Meeting, *From MAVEN to Bayes: A Data-Driven Approach to Mars' Space Environment* (2021).

Special seminar at Courant Institute of Mathematical Sciences, *The Fermat distance, percolation theory and applications in machine learning* (2018).

Workshops and Symposiums

AI and Physical Sciences for Climate Innovation Symphosium, Berkeley USA (2024)

Berkeley Atmospheric Sciences Center Symposium, Berkeley USA (2024)

Machine Learning in Glaciology Workshop, Norway (2022 and 2023)

Write a Compiler by David Beazley, Online (2022)

Open Global Glacier Model workshop, Norway (2022)

Karthaus Summerschool on Ice Sheets and Glaciers in the Climate System, Italy (2021)

ICESat-2 Hackweek, online (2020)

CIMPA School - X Escuala Santaló and ICM Rio Satellite Geometry and scaling of random structures, Argentina (2018)

American Geophysical Fall (AGU) Meeting

Large part of my reseach is represented every year in the American Geophysical Union fall meeting. This is a list of my different presentations per year.

2023 AGU Fall Meeting (December 12-16, San Francisco)

Sapienza, F., Bolibar, J., Gimenes, L. , Pérez, F. (2023) Leveraging the Julia and Python Open Science Ecosystems to Enhance Glacier Modeling with Scientific Machine Learning.

Sapienza, F., Gallo, L., Zhang, Y., Vaes, B., Domeier, M., Swanson-Hysell, N. (2023), Quantitative analysis of Paleomagnetic Sampling Strategies.

Gallo, L., Domeier, M., Sapienza, F., Swanson-Hysell, N., Vaes, B., Zhang, Y., Arnould, M., Gurer, D., Kiraly, A., Robert, B., Rolf, T., Shephard, G., van der Boon, A. (2023). Embracing uncertanty to resolve polar wander.

Azari, A., Helekas, J., Abrahams, E., Biersteker, J., Sapienza, F., Mitchell, D., Pérez, F., Marquette, M. (2023) Virtual Solar Wind Monitor for Mars: Enabling an Updated Understanding of Induced Magnetospheres.

Snow, T., Millstein, J., Sauthoff, W., Scheick, J., Colliander, J., Leong, W., Munroe, J., Pérez, F., Felikson, D., Sutterley, T., Fisher, M., Sapienza, F., Abrahams, E., Zheng, W., and Siegfried, M. (2023) Accelerating scientific discovery for NASA Cryosphere communities with the CryoCloud JupyterHub.

Snow, T., Sauthoff, W., Scheick, J., Leong, W., Colliander, J, Munroe, J, Pérez, F., Felikson, D., Sutterley, T., Fisher, M., Sapienza, F., Abrahams, E., Zheng, W., and Siegfried, M.(2023) CryoCloud JupyterHub for NASA Cryosphere communities: Open science in the cloud as a process, not a product.

2022 AGU Fall Meeting (December 12-16, Chicago)

Sapienza, F., Bolibar, J., Maussion, F., Lguensat, R., Pérez, F, Wouters, B. (2022). Functional Inversion of Glacier Rheology from Ice Velocities using ODINN.jl.

Pérez, F., Sapienza, F., Sundell, E., Panda, Y. (2022). Design and Teaching of Collaborative and Reproducible Data Science with Environmental Applications.

Abrahams, E., Snow, T., Lee, E., Zheng, W., Field, M., Savidge, E., Sapienza, F., Grigsby, S., Taylor, J., Siegfried, M., Pérez, F. (2022). Automated Detection of West Antarctic Persistent Polynas using Physics-Featurized Neural Networks.

Zheng, W., Perez, F., Holdgraf, C., Sundell, E., Siegfried, M., Snow, T., Shane, G., Sapienza, F., Taylor, J. and Executable Books Community (2022). Jupyter Book-based Supplemental Material: a FAIR Practice to Connect Research Articles with Scientific Data.

2021 AGU Fall Meeting (December 13-17, New Orleans and online)

Sapienza, F., Grigsby, S., Zheng, W., Taylor, J., Snow, T., Pérez, F., & Siegfried, M. (2021). Spectral Unmixing of Antarctic Snow Grain Size Distribution: A Data-Driven Perspective.

Snow, T., Sapienza, F., Grigsby, S., Taylor, J., Savidge, E., Zheng, W., Alley, K. E., Pérez, F., & Siegfried, M. (2021). Basal channel outflow inferred from persistent polynya variability at the Eastern Thwaites Ice Shelf.

Grigsby, S., Sapienza, F., Zheng, W., Taylor, J., Snow, T., Savidge, E., Pérez, F., & Siegfried, M. (2021). Mission in a minute: Complex Spatial Query and Data Access in the Cloud for the ICESat-2 Mission.

Sapienza, F., Abrahams, E., Azari, A., Biersteker, J., Pérez, F., & Mitchell, D. (2021). From MAVEN to Bayes: A Data-Driven Approach to Mars' Space Environment.

Azari, A., Mitchell, D., DiBraccio, G. A., Sapienza, F., Abrahams, E., Biersteker, J., & Pérez, F. (2021). Mars' Global Magnetic Field Response To Localized Crustal Fields and Interplanetary Magnetic Field Direction.

Bolibar, J., Sapienza, F., Lguensat, R., Wouters, B., & Pérez, F. (2021). Optimizing and Discovering Models of Glacier Processes with Neural Networks Embedded in Differential Equations.

2020 AGU Fall Meeting (December 1-17, Online)

Grigsby, S., Sapienza, F., Snow, T., Cima, A., Heagy, L. J., Siegfried, M., Pérez, F., & Taylor, J. (2020). *Spatio-Temporal Interpolation of Cloud Data*. In AGU Fall Meeting 2020.

Pérez, F., Banihirwe, A., Holdgraf, C., Sundell, E., Sapienza F., Paul, K., & Heagy L. (2020). *Jupyter meets the Earth: connecting Jupyter development with geoscience research*. In JupyterCon 2020

Pérez, F., Hamman, J., Larsen, L., Paul, K., Heagy, L. J., Moges, E., Banihirwe, A., Cima, A., Sapienza F., Sundell, E., & Holdgraf C. (2020) *Jupyter meets the Earth: advancing an open ecosystem that supports science*. In AGU Fall Meeting 2020.

Sapienza, F., Snow, T., Cima, A., Grigsby, S., Heagy, L. J., Pérez, F., Siegfried, M., & Taylor, J. (2020). *Multi-modal Dataset Integration for Cloud Masking of ICESat-2*. In AGU Fall Meeting 2020.

2023 JuliaCon (Massachusetts, USA). Facundo Sapienza, Jordi Bolibar, Redouane Lguensat, Bert Wouters and Fernando Pérez. Modeling Glacier Ice Flow with Universal Differential Equations.

2023 IUGG (Berlin, Germany).

2022 Open Global Glacier Model (OGGM) workshop (Finse, Norway) Facundo Sapienza and Jordi Bolibar. ODINN.

2022 EarthCube Annual Meeting (June 14-16, San Diego (La Jolla)). Zheng, W., Holdgraf, C., Pérez, F., Sundell, E., Siegfried, M. R., Snow, T., Grigsby, S., Sapienza, F., Taylor, J., & Executable Books Community. (2022). Let supplemental material be FAIR: using narrative and reusable Jupyter Book to complement your journal paper.

2021 Young CEED (The Center for Earth Evolution and Dynamics) (Drovak, Norway) workshop.

2018 6th International Conference on Learning Representations (Vancouver, Canada). Sapienza F, Jonckheere M, Groisman P; Weighted Geodesic Distance Following Fermat Principle.

2018 VIII Conference on Quantum Foundations (Buenos Aires, Argentina) F Sapienza, F Cerisola, A. Roncaglia; Correlations as a resource in quantum thermodynamic processes of state formation.

CIMPA School - X Escuala Santaló and ICM Rio Satellite Geometry and scaling of random structures (Buenos Aires, Argentina). Sapienza F, Groisman P, Jonckheere M; Geodesics in First-Passage Percolation and Distance Learning.

1st Quantum Mechanics School 2019=8 (Cordoba, Argentina). Sapienza F. et al.; Correlations as a Resource in Quantum Thermodynamics

Annual Conference of the Argentinian Physics Society 2016 (Tucuman, Argentina). Sapienza F, Somacal A; Study of the grain size distribution of SD magnetite from ARM magnetization measurements