Gabriele Accarino, Ph.D.

Associate Research Scientist, Columbia University
New York, NY, USA | gabriele.accarino@gmail.com | +1 (646) 243-7951
GitHub | Google Scholar | ResearchGate | LinkedIn |

Research Profile

I work at the intersection of machine learning and Earth system science, developing data-driven tools to model, emulate and evaluate complex physical processes. My research spans generative models for the emulation of physical systems (multi-scale and hierarchical architectures built on Flow Matching, with generation performed in wavelet space), wavelet-based multi-scale metrics for comparing model output and observations, generative super-resolution for downscaling, and deep learning for the detection and tracking of extreme events such as tropical cyclones and wildfires. The question driving my current work is whether explicitly encoding multi-scale physical structure is the key to stable, accurate and realistic emulation. My work appears at NeurIPS (main track, first author) and in PNAS, which received the ADIA Lab Best Paper Award in Climate Data Sciences (1st place, $50,000). Before Columbia, I led a 5-person machine learning research unit at the Euro-Mediterranean Center on Climate Change (CMCC Foundation) and contributed to three EU-funded projects that brought over €1.3M to CMCC, including as Work Package Leader and Steering Committee member. Experience teaching undergraduate and graduate courses in High-Performance Computing, Machine Learning, and Information Processing Systems.

Appointments & Research Experience

Oct. 2026 – present
Associate Research Scientist
Columbia University, New York, NY, USA
Department of Earth and Environmental Engineering · LEAP NSF STC
PIs: Prof. Viviana Acquaviva (CUNY) and Prof. Pierre Gentine (Columbia University)

- Extending Wavelet Flow Matching (NeurIPS 2026) to weather and climate emulation on spherical geometries (using wavelets on the sphere).
- Reconstruct global surface-ocean ΔpCO_2 fields using Earth system model data and tabular diffusion models, extending work accepted at the NeurIPS 2026 Workshop on RPS.

Oct. 2024 – Sep. 2026
Postdoctoral Research Scientist
Columbia University, New York, NY, USA
Department of Earth and Environmental Engineering · LEAP NSF STC
Advisors: Prof. Viviana Acquaviva (CUNY), Prof. Sara Shamekh (NYU), Prof. Duncan Watson-Parris (UC San Diego)

- Developed Wavelet Flow Matching, a multi-scale generative framework that performs emulation in *wavelet space*, introduced as a new space for generative modeling. It better preserves spectral consistency and outperforms state-of-the-art neural operators on three complex physical systems; first-author paper accepted at NeurIPS 2026 (main track).
- Developed WaveSim, a wavelet-based multi-scale similarity metric for evaluating weather and climate fields; released as open source (github.com/gabrieleaccarino/wavesim), preprint under review at JAMES, presented at Climate Informatics 2025, the 2025 Gordon Research Conference on ML for Actionable Climate Science, and AGU 2024 and 2025.
- Developed a tabular diffusion model to estimate the conditional distribution of surface-ocean ΔpCO_2 from sparse SOCAT observations, using 13 environmental, spatial, and temporal predictors; accepted at the NeurIPS 2026 Workshop on Representations for the Physical Sciences (RPS).

Oct. 2024 – present
Scientific Research Affiliate
Euro-Mediterranean Center on Climate Change (CMCC Foundation), Lecce, Italy
Advanced Digital Innovation Center (ADIC)

- Contribute to research on tropical cyclone detection and tracking (ByteStorm, *Computers & Geosciences*, 2026) and on Bayesian optimization for model tuning and its integration into MLOps and provenance frameworks (OBSERVE, SC26 Workshops; MLOps+, under review at FGCS).
- Lecturer in Machine Learning at the CMCC Advanced Training and Education Center (see Teaching).

Feb. 2023 – Sep. 2024
Junior Scientist & Head of the Machine Learning Research Unit
Euro-Mediterranean Center on Climate Change (CMCC Foundation), Lecce, Italy
Advanced Scientific Computing Division, Institute for Earth System Prediction

- Led and mentored a team of 5 scientists: set priorities, coordinated tasks, and aligned individual research with project milestones and strategic goals.
- Built a collaborative learning culture through regular group meetings for idea exchange, expertise sharing and critical discussion of the ML literature.
- Coordinated the unit's scientific output and dissemination: peer-reviewed publications, talks and posters at EGU and Climate Informatics, and collaborations with CMCC divisions and institutes, as well as other international institutions.
- Conducted and contributed to ML projects on data-driven wildfire risk indices, extreme-weather event detection, generative super-resolution for climate downscaling, and data-driven ocean emulators.
- Scouted EU funding opportunities and contributed to proposals (drafting technical sections, coordinating inputs); managed projects through their full lifecycle, including periodic reporting, technical documentation, deliverables and reviews.

Mar. 2024 – Jun. 2024
Visiting Research Scientist
Columbia University, New York, NY, USA
Department of Earth and Environmental Engineering · LEAP NSF STC
Host: Prof. Pierre Gentine

- Investigated deep learning-based atmospheric and ocean emulators, focusing on their coupling and integrated modeling.

Feb. 2022 – Jan. 2023
Research Fellow in Computer Science
University of Salento, Lecce, Italy
Department of Engineering for Innovation

- Applied machine learning research for Earth system science applications.

Dec. 2021 – Feb. 2022
Contractor
Italian Presidency of the Council of Ministers, Italy

- Developed a prototype automatic speech recognition system for investigative and forensic applications.

May 2018 – Oct. 2018
Post-Degree Research Assistant
Euro-Mediterranean Center on Climate Change (CMCC Foundation), Lecce, Italy
Advanced Scientific Computing Division

- Designed a multi-model deep learning framework based on LSTM networks for multi-step sea-level forecasting from observations, benchmarked against the Mediterranean Forecasting System (an operational numerical model); published in *Future Generation Computer Systems* (2021).

Education

Nov. 2018 – May 2022
Ph.D. in Environmental Sciences
University of Salento, Lecce, Italy
Thesis: *Exploring Machine Learning for Applications to the Climate Science Domain* (interdisciplinary thesis in computer and environmental science)
Advisors: Prof. Giovanni Aloisio and Prof. Luigi De Bellis

2016 – 2018
M.Sc. in Computer Engineering, *summa cum laude*
University of Salento, Lecce, Italy
Thesis: *On the use of Deep Learning in the Climate Change domain*

2012 – 2015
B.Sc. in Information Engineering
University of Salento, Lecce, Italy
Thesis: *Control systems with Smith's Predictor: analysis and fundamental properties*

Jul. 2020
Professional Qualification as Licensed Engineer (ICT Area)
University of Salento, Lecce, Italy
State examination qualifying for practice as a licensed engineer in Italy

Honors & Awards

2025 **ADIA Lab Best Paper Award in Climate Data Sciences, 1st place ($50,000)**
ADIA Lab, 2024 Call for Papers, Abu Dhabi, UAE
Awarded for Immorlano, F., Eyring, V., le Monnier de Gouville, T., **Accarino, G.**, et al., *Transferring climate change physical knowledge*, PNAS (2025). adialab.ae/call-for-papers-2024

2018 **Leonardo Innovation Award, Ph.D. category**
Leonardo S.p.A., Italy
For ANIMA, an artificial intelligence tool for migration analysis and projections

Publications

Peer-Reviewed Conference Papers & Proceedings

2026 **Accarino**, **G.**, Nathaniel, J., Roesch, C., Gentine, P., Shamekh, S., Watson-Parris, D., Acquaviva, V. *Wavelet Flow Matching for Multi-Scale Physics Emulation*. Advances in Neural Information Processing Systems (NeurIPS 2026), Main Track. 2026. *To appear.*

Peer-Reviewed Journal Articles

2026 Donno, D., Elia, D., **Accarino**, **G.**, De Carlo, M., Scoccimarro, E., Gualdi, S. *ByteStorm: A multi-step data-driven approach for Tropical Cyclones detection and tracking*. Computers & Geosciences 214, 106195 (2026). DOI: 10.1016/j.cageo.2026.106195.

2026 Epicoco, I., Donno, D., **Accarino**, **G.**, Norberti, S., Grandi, A., Giurato, M., McAdam, R., Elia, D., et al. *MedFormer: a data-driven model for forecasting the Mediterranean Sea*. Earth Science Informatics 19.8, 129 (2026). DOI: 10.1007/s12145-026-02169-9.

2026 Manzi, A., Bardaji, R., Rodero, I., Moltó, G., Fiore, S., Campos, I., Elia, D., Sarandrea, F., Millar, A. P., Spiga, D., Bunino, M., **Accarino**, **G.**, et al. *interTwin: Advancing Scientific Digital Twins through AI, Federated Computing and Data*. Future Generation Computer Systems 179, 108312 (2026). DOI: 10.1016/j.future.2025.108312.

2025 Azmi, E., Alibabaei, K., Kozlov, V., Krijger, T., **Accarino**, **G.**, et al. *Best practices for AI-based image analysis applications in aquatic sciences: The iMagine case study*. Ecological Informatics 91, 103306 (2025). DOI: 10.1016/j.ecoinf.2025.103306.

2025 **Accarino**, **G.**, De Carlo, M. M., Ruiz Atake, I., Elia, D., Dissanayake, A. L., Sepp Neves, A. A., Peña Ibañez, J., Epicoco, I., Nassisi, P., Fiore, S., Coppini, G. *Improving oil slick trajectory simulations with Bayesian optimization*. Ecological Informatics 91, 103368 (2025). DOI: 10.1016/j.ecoinf.2025.103368.

2025 Immorlano, F., Eyring, V., le Monnier de Gouville, T., **Accarino**, **G.**, Elia, D., Mandt, S., Aloisio, G., Gentine, P. *Transferring climate change physical knowledge*. Proceedings of the National Academy of Sciences (PNAS) 122.15, e2413503122 (2025). DOI: 10.1073/pnas.2413503122. *ADIA Lab Best Paper Award in Climate Data Sciences, 1st place.*

2023 **Accarino**, **G.**, Donno, D., Immorlano, F., Elia, D., Aloisio, G. *An ensemble machine learning approach for tropical cyclone localization and tracking from ERA5 reanalysis data*. Earth and Space Science 10.11, e2023EA003106 (2023). DOI: 10.1029/2023EA003106.

2022 Aloisi, V., Gatto, A., **Accarino**, **G.**, Donato, F., Aloisio, G. *The effect of known and unknown confounders on the relationship between air pollution and Covid-19 mortality in Italy: A sensitivity analysis of an ecological study based on the E-value*. Environmental Research 207, 112131 (2022). DOI: 10.1016/j.envres.2021.112131.

2022 Gatto, A., Aloisi, V., **Accarino**, **G.**, Immorlano, F., Chiarelli, M., Aloisio, G. *An Artificial Neural Network-Based Approach for Predicting the COVID-19 Daily Effective Reproduction Number Rt in Italy*. AI 3.1 (2022), pp. 146–163. DOI: 10.3390/ai3010009.

2021 **Accarino**, **G.**, Chiarelli, M., Fiore, S., Federico, I., Causio, S., Coppini, G., Aloisio, G. *A multi-model architecture based on Long Short-Term Memory neural networks for multi-step sea level forecasting*. Future Generation Computer Systems 124 (2021), pp. 1–9. DOI: 10.1016/j.future.2021.05.008.

2021 **Accarino**, **G.**, Chiarelli, M., Immorlano, F., Aloisi, V., Gatto, A., Aloisio, G. *MSG-GAN-SD: A Multi-Scale Gradients GAN for Statistical Downscaling of 2-Meter Temperature over the EURO-CORDEX Domain*. AI 2.4 (2021), pp. 600–620. DOI: 10.3390/ai2040036.

2021 Gatto, A., **Accarino**, **G.**, Aloisi, V., Immorlano, F., Donato, F., Aloisio, G. *Limits of Compartmental Models and New Opportunities for Machine Learning: A Case Study to Forecast the Second Wave of COVID-19 Hospitalizations in Lombardy, Italy*. Informatics 8.3, 57 (2021). DOI: 10.3390/informatics8030057.

2021 **Accarino**, **G.**, Lorenzetti, S., Aloisio, G. *Assessing correlations between short-term exposure to atmospheric pollutants and COVID-19 spread in all Italian territorial areas*. Environmental Pollution 268, 115714 (2021). DOI: 10.1016/j.envpol.2020.115714.

Peer-Reviewed Workshop Papers

2026 **Accarino**, **G.**, Amirian, M. M., Heimdal, T. H., Fay, A. R., McKinley, G. A., Acquaviva, V. *Learning Calibrated and Samplable Predictive Distributions of Surface-Ocean ΔpCO_2 Using Tabular Diffusion Models*. NeurIPS 2026 Workshop on Representations for the Physical Sciences (RPS). 2026. *To appear*.

2026 De Carlo, M. M., Padovani, G., **Accarino**, **G.**, Anantharaj, V., Nassisi, P., Fiore, S. *OBSERVE: An Observable, Sustainable, and Fault-Tolerant Runtime for Hyperparameter Optimization on Production HPC Systems*. Proceedings of the SC '26 Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis. Chicago, IL, USA, 2026. *To appear*.

2023 Elia, D., Scardigno, S., Ejarque, J., D'Anca, A., **Accarino**, **G.**, Scoccimarro, E., et al. *End-to-End Workflows for Climate Science: Integrating HPC Simulations, Big Data Processing, and Machine Learning*. Proceedings of the SC '23 Workshops of the International Conference on High Performance Computing, Network, Storage, and Analysis. Denver, CO, USA, 2023, pp. 2042–2052. DOI: 10.1145/3624062.3624283.

Preprints & Papers Under Review

2026 Donno, D., Epicoco, I., Cafaro, M., **Accarino**, **G.**, Amirian, M. M., Acquaviva, V., Nassisi, P., Iovino, D., Bracco, A., Masina, S., Gentine, P. *Neptune: An AI model for Global Ocean Subseasonal Prediction*. 2026. arXiv: 2609.08606.

2026 Alvi, S., Fedele, G., **Accarino**, **G.**, Epicoco, I., Manco, I., Schiano, P. *OpFML: Pipeline for ML-based Operational Inference*. 2026. arXiv: 2601.11046.

2026 De Carlo, M. M., Padovani, G., **Accarino**, **G.**, Nassisi, P., Fiore, S. *MLOps+: Enhancing the MLOps Lifecycle with Traceable, Explainable, and Energy-Aware Bayesian Optimization*. 2026. *Under review at* Future Generation Computer Systems.

2025 **Accarino**, **G.**, Acquaviva, V., Shamekh, S., Watson-Parris, D., Lawrence, D. *WaveSim: A Wavelet-based Multi-scale Similarity Metric for Weather and Climate Fields*. 2025. arXiv: 2512.14656. *Under review at* Journal of Advances in Modeling Earth Systems (JAMES).

Conference Talks, Posters & Abstracts

2025 **Accarino**, **G.**, Acquaviva, V., Shamekh, S., Watson-Parris, D., Lawrence, D. *WaveSim: A Multi-Scale Wavelet-Based Similarity Metric for Climate Field Comparison*. AGU Fall Meeting 2025. New Orleans, LA, USA, 2025. *Poster*.

2025 **Accarino**, **G.**, Acquaviva, V., Shamekh, S., Watson-Parris, D. *A Benchmark for Evaluating Outputs of Earth System Models based on Perceptual Similarity Metrics*. Gordon Research Conference on Machine Learning for Actionable Climate Science. Bryant University, Smithfield, RI, USA, 2025. *Poster*.

2025 **Accarino**, **G.**, Acquaviva, V., Shamekh, S., Watson-Parris, D. *A Benchmark for Evaluating Outputs of Earth System Models based on Perceptual Similarity Metrics*. Climate Informatics 2025. Rio de Janeiro, Brazil, 2025. *Talk*.

2024 **Accarino**, **G.**, Acquaviva, V., Shamekh, S., Watson-Parris, D. *Improved Evaluation Methods for Earth System Models*. AGU Fall Meeting 2024. GC51X-0297. Washington, DC, USA, 2024.

2024 Elia, D., Antonio, F., Fiore, S., Donno, E., **Accarino**, **G.**, Nassisi, P., Aloisio, G. *A Data Space environment for big data and ML-based climate applications in the European Open Science Cloud*. EGU General Assembly 2024. EGU24-17472. Vienna, Austria, 2024.

2024 D'Anca, A., Scardigno, S., Ejarque, J., **Accarino**, **G.**, Peano, D., Immorlano, F., et al. *An end-to-end workflow for climate data management and analysis integrating HPC, Big Data and Machine Learning*. EGU General Assembly 2024. EGU24-11774. Vienna, Austria, 2024.

2024 Atake, I., **Accarino**, **G.**, De Carlo, M., Elia, D., Coppini, G., Aloisio, G. *Optimizing Medslik-II: Parametrization through a Bayesian search algorithm applied at the Baniyas oil spill incident (Syria, 2021)*. EGU General Assembly 2024. EGU24-7256. Vienna, Austria, 2024.

2023 **Accarino**, **G.**, Elia, D., Donno, D., Immorlano, F., Aloisio, G. *A machine learning-powered Digital Twin for extreme weather events analysis*. EGU General Assembly 2023. EGU23-6060. Vienna, Austria, 2023.

2023 Elia, D., Scardigno, S., D'Anca, A., **Accarino**, **G.**, Ejarque, J., Immorlano, F., et al. *Convergence of HPC, Big Data and Machine Learning for Earth System workflows*. EGU General Assembly 2023. EGU23-6857. Vienna, Austria, 2023.

Research Projects & Funding

Three EU-funded projects with a combined budget of over €1.3M for CMCC. Amounts are the CMCC net EU contribution (source: CORDIS).

Oct. 2021 – Mar. 2025 **SILVANUS**
H2020 Innovation Action, GA 101037247, CORDIS
- **Role:** Scientific Officer, Work Package 4 Leader, Steering Committee member.
- **Activity:** Designed and implemented an operational pipeline for wildfire detection and risk mapping that combines satellite Earth-observation data, numerical modeling and deep learning.
- **Budget:** €571,375 for CMCC (total project cost €24.2M).

Sep. 2022 – Aug. 2025 **interTwin**
Horizon Europe RIA, GA 101058386, CORDIS
- **Role:** Co-supervisor and contributor in Work Packages 4, 6 and 7.
- **Activity:** Designed and developed two digital-twin demonstrators for extreme-weather event detection (tropical cyclones and wildfires) built on the project's Digital Twin Engine.
- **Budget:** €581,562.50 for CMCC (total project cost €11.7M).

Sep. 2022 – Aug. 2025 **iMagine**
Horizon Europe RIA, GA 101058625, CORDIS
- **Role:** Co-supervisor, Work Package 3, Use Case 4.
- **Activity:** Designed and implemented an operational pipeline for calibrating and optimizing the MEDSLIK-II oil-spill model with Bayesian optimization.
- **Budget:** €182,500 for CMCC (total project cost €4.5M).

Teaching & Mentoring

Summer 2026 **Instructor, LEAP Summer Bootcamp: Machine Learning for Climate Modeling**
Columbia University, LEAP NSF STC, New York, NY, USA
- Led Week 2 of the bootcamp and taught its first two days: artificial neural networks and deep learning architectures for image and sequence data.

Oct. 2024 – present **Lecturer in Machine Learning**
CMCC Advanced Training and Education Center (ATEC), Lecce, Italy
- Course for 30 scientists from diverse backgrounds: ML theory, models and common tasks in environmental applications, with hands-on climate case studies.

Nov. 2022 – Jan. 2023 **Adjunct Professor, Information Processing Systems**
University of Salento, B.Sc. in Nursing, Lecce, Italy
- Introductory course (3 ECTS credits) for 96 students: digital information representation, computer architectures, internet protocols and computer networking.

Feb. 2022 – Jun. 2022 **Adjunct Professor, High Performance Computing**
University of Salento, M.Sc. in Computer Engineering, Lecce, Italy
- Course (9 ECTS credits) for 15 students: supervised learning (linear, logistic and polynomial regression; Ridge/Lasso regularization; bias–variance trade-off), artificial neural networks and backpropagation with Keras, and hands-on Python labs in Jupyter.

Feb. 2023 – Sep. 2024 **Team Lead and Mentor, Machine Learning Research Unit**
CMCC, Lecce, Italy

- Mentored 5 scientists on research design, ML methods and scientific dissemination.

Technical Skills

Machine Learning	Generative models (Flow Matching, diffusion models, GANs), recurrent networks (LSTM), ensemble methods, Bayesian optimization, uncertainty quantification, super-resolution, surrogate modeling of physical systems
Frameworks	PyTorch, PyTorch Lightning, TensorFlow/Keras, scikit-learn
Scientific data	NumPy, SciPy, pandas, xarray, NetCDF, Zarr, Matplotlib; ERA5 reanalysis, satellite Earth-observation and ocean observational data
HPC & MLOps	Distributed training on HPC systems, parallel computing, end-to-end data and ML pipelines, hyperparameter optimization in MLOps workflows
Tools	Python, Git; HPC workload managers: SLURM, LSF, PBS
Languages	Italian (native), English (professional working proficiency)

Professional Service

Reviewer	**Conferences:** NeurIPS. **Journals:** Journal of Advances in Modeling Earth Systems (JAMES); journals of the American Meteorological Society (AMS); IOPscience Machine Learning: Earth; Future Generation Computer Systems; Environmental Pollution; Environmental Informatics; MDPI journals.