Maria-Theodora Folina

LinkedIn: Maria-Theodora Folina | GitHub: m-fol | Google Scholar: Maria Folina | Email: maryfolina@gmail.com | Website: mfolina.com

CERN Technical Student, ESA Academy Alumna and incoming **SNSF-funded ETH Zurich Automation Fellow**. Scientific ML researcher on stochastic dynamical systems, neural operators, observability and **uncertainty quantification**, with research experience in **European Commission projects** and **LIGO**. **Valedictorian** with **20+ peer-reviewed publications** and **50+ citations**.

RESEARCH IN DYNAMICAL SYSTEMS & UNCERTAINTY

Visiting Researcher - Aalborg University, Automation & Control Group — 08/2026

Smart Water Infrastructures Laboratory (SWIL) | Supervisor: Prof. Rafal Wisniewski

- Established **unobservability due to graph symmetry** in partially observed water networks, deriving **ε^2 scaling** under pipe perturbations that break symmetry, with exact arithmetic certificates.
- Applied an orbit-breaking placement rule needing at most **$\log_2|G|$ sensors** and compared **Bayesian posterior covariance** criteria targeting weakly observable directions on linear **Gaussian** models and EPANET simulations.
- Benchmarked **Kalman/RTS, IMM filters** and tuned recurrent, GNN, PointNet and PointNet++ observers under stochastic disturbances and model mismatch.
- Tested observers on a synthetic water network with six states **(5 training seeds, 3 test sets of 1,000 trajectories)**.

Research Collaborator - Functional Autoencoders for Operator Learning — 07/2026 - Present

Imperial College London (LOGML) | Supervisor: Dr. Hefin Lambley (Freie Universitat Berlin)

- Investigated **discretization bias in mesh invariant neural operators** under nonuniform sampling of point clouds, comparing Monte Carlo, quadrature with Voronoi weights and kernel quadrature on two **Navier-Stokes datasets**.
- Extended functional autoencoders with **Nyström approximation for attention pooling** and alternative mappings from functions to vectors (log-integral-exp, max), for reconstruction accuracy, numerical stability and computational cost.
- Compared reconstruction errors using **paired statistical tests with Holm correction**, finding improvements with low dropout and shorter training runs (**50,000 steps**).

Research Assistant - International Hellenic University — 07/2024 - 01/2026

- Reconstructed **ocean pCO_2 fields** and estimated air-sea CO_2 flux from sparse in situ observations, fusing Sentinel-3, wind and SOCAT data across the North Atlantic using knowledge graphs and SHAP attribution.
- Examined the **2023 marine heatwave**, identifying uncertainty in extreme event prediction as an open problem.
- Developed a **Riemannian trajectory learning framework** to forecast environmental degradation over 12 months across **49 sites**, representing satellite and climate data on symmetric positive definite (SPD) manifolds.
- Used **Log-Euclidean geodesic distances and second order trajectory regularization** to model environmental dynamics and cumulative stress.
- Achieved **$R^2 = 0.753$ on unseen sites**, with the largest errors during periods of high environmental stress.
- Led work packages on EU projects involving digital twins and sustainability (IndustriSphere, Seek & Go).

OTHER RESEARCH

Technical Student - CERN (European Organization for Nuclear Research) — 01/2026 - Present

Software Frameworks and Tools (EP-SFT)

- Brought a shared software build system for LHC experiments **from prototype to production**, automating tests across platforms and compilers, nightly builds and package distribution.
- Formulated a **Reproducibility Debt Index** combining Gower distances for mixed variables, pairwise behavioral **divergence** and **entropy** measures to quantify variation across software environments.

- Estimated uncertainty using **bootstrap resampling at the build level and Wilson confidence intervals**, examining dependence between pairwise observations.
- Evaluated build fragility and **sensitivity to declared and resolved configurations** using real CERN CI data.
- Developed **Python/C++ tools** for automated scientific data transfers, stress testing and parallelisation using CERN's cloud infrastructure.

NCCR Automation Fellow - ETH Zurich (SNSF) 03/2027 - 06/2027 (incoming)

Automatic Control Laboratory (Prof. Dr. John Lygeros's Group) | Supervisor: Dr. Efe Balta

- Study software deployment and reproducibility as a control and optimization problem with measurable guarantees.
- Apply optimization methods to make **scientific computing** pipelines verifiable and **reliable.**

Research Collaborator - LIGO Scientific Collaboration 07/2025 - 01/2026

Missouri University of Science & Technology | Supervisor: Prof. Marco Cavaglia

- Developed **ML pipelines** (LSTM-pointcloud classifier, transformer-based localizer) to classify and localize transient noise in **500 GB+** of gravitational wave detector data.

Junior Research Lead - Software and Data Engineering Laboratory, UoM 06/2024 - 01/2026

- **Supervised 4 Machine Learning** and Data Engineering **theses as an undergraduate.**
- Peer reviewer: **Cambridge University Press** (*Data & Policy)* and **Springer Nature** *(Discover Food, Discover Sustainability)*

Selected Projects

- **VICTORIOUS (EDF-2023)**: Contributed to **XAI-driven** observation data pipelines for **space defence simulation tools.**
- **AgriVision (CORTEX2) & AgriBot (HORIZON)**: **ML geospatial data** analysis, **90%** complex environment accuracy.
- **RAISESuite (HORIZON/EOSC)**: FAIR-by-design automated research workflows; **ML metadata modeling**.

GRANTS & SELECTED PROGRAMMES

CERN 48th School of Computing 2026 — **Final Examination Top Performer (1st of 64) | CERN Grant | Lightning Talk Speaker**
AIDA AI/ML Summer School and Symposium — **Full Scholarship | Ranked 1st applicant**
BEiNG-WISE Training School on Cybersecurity — **INRIA & COST ITC Grant | Selected 1 of 2 in Greece**
ESA Innovation in Earth Observation Training Course (DICEO) — **ESA Grant | Selected among 30 globally**
NATO STO Summer School on 5G and Defence — **COST STSM Grant (InterCoML) | Selected among 40 globally**
13th Heidelberg Laureate Forum 2026 — **HLFF Travel Grant | One of 200 selected Young Researchers worldwide**
ACM-W WomENcourage 2026 — **ACM-W Conference Grant | Selected among 50 globally**

EDUCATION

MSc in Systems Engineering and Technology – Democritus University of Thrace (2025-2027)
Thesis (at CERN): Systems Engineering Frameworks for Reproducible Software Stacks in LHC Computing Ecosystems

BSc in Computer Science – University of Macedonia (2021-2025) — **GPA: 9.2/10 (Valedictorian - Excellence Award)**
Thesis: SmartESPD - Intelligent ML Extraction of Supplier Information — ***Featured at the EU Joint Research Centre***

HONOURS & AWARDS

IKY Excellence Award | Hellenic State Scholarships Foundation — **Valedictorian, Top Graduate (Computer Science)**
1st Place | HUAWEI Seeds for the Future — **National Innovation Winner (Featured in national news)**
Global Winning Team | MRC Robotics Olympiad — **Best Business Idea**

Best Student Poster Award | European Space Agency (ESA) **Model-Based Systems Engineering (MBSE 2025)**

ESA MBSE Conference Student Bursary | European Space Agency (ESA) **Selected as one of 4 students worldwide**

RELEVANT PUBLICATIONS & PRESENTATIONS

20+ peer-reviewed publications | 55 citations | h-index: 3

Aalborg University | When Symmetry Hides the State: Observability and Sensor Limits in a Partly Instrumented Water Network
*Authors: **Folina, M.-T**., et al. (in preparation)*

Imperial College London | New Point-Cloud Architectures for Mesh-Invariant Operator Learning
*Authors: Lambley, H., **Folina, M.-T**., et al. (in preparation)*

NATO ICMCIS 2027 | Not Observed Is Not Calm: Where Open Sensors Go Blind and How to Verify Alerts for C2
*Author: **Folina, M.-T**. (in preparation)*

ESA Ocean Carbon from Space 2025 | Towards an Explainable AI Framework for Quantifying Air–Sea CO_2 Fluxes: Multi-Sensor Satellite Data Fusion with Knowledge Graph Representation
*Authors: **Folina, M.-T**., Folinas, D., Kostavelis, I.*

LOGML 2026 | Geometric AI for Climate-Resilient Heritage: Riemannian Trajectory Learning for Predictive Monitoring
*Author: **Folina, M.-T**.*

CERN, AI for Gravitational Waves 2026 | GravCloud: A Stacked LSTM-Pointcloud Framework for LIGO Glitch Classification
*Authors: **Folina, M.-T.**, Parmar, S., Cavaglia, M.*

American Physical Society Global Physics Summit 2026 | Omniformer: A Real-Time Transient Noise Source Localizer
*Authors: Parmar, S., Cavaglia, M., Staats, K., **Folina, M.-T.***

CERN School of Computing 2026 | It Works on My Machine: Reproducibility Debt in HEP Software Stacks
*Author: **Folina, M.-T.***

ESA MBSE 2025 (Best Student Poster Award) | AeroMLM: Accelerating MBSE in Aerospace: From Unstructured Data to Intelligent Models
*Author: **Folina, M.-T.***

Elsevier, 2025 (ELKE Research Award, 2026) | Applications of Robotics and Extended Reality in Agriculture: A Review
*Authors: Anastasiou, E., et al. (incl. **Folina, M.-T.**)*

OPERATIONAL & LEADERSHIP EXPERIENCE

Experiments Guide, CERN (2026-) · Robotics Team Leader, HERON UoM (2024-25) · PR Manager, IEEE Student Branch (2023-24)

SKILLS

Applied Mathematics: Observability Analysis, Graph Symmetries, Stochastic Lyapunov Analysis, Numerical Quadrature, Riemannian Geometry, Numerical Optimization, Statistical Inference, Dynamical Systems.
Estimation & Uncertainty Quantification: Bayesian Posterior-Covariance Analysis, Optimal Sensor Placement, Observability Gramians, Kalman/RTS, IMM Filters, Bootstrap Resampling, Sensitivity Analysis, Bayesian Experimental Design.
Scientific & Geometric ML: Neural Operators, Functional Autoencoders, GNNs, Point-Cloud Learning, Manifold Learning, Equivariant/Spherical CNNs, Transformers, Recurrent Networks, Reinforcement Learning, Spatiotemporal Modeling.
Scientific Computing & Software Engineering: Python (Expert), C++ (Proficient), Java, SQL, MATLAB/Simulink, PyROOT, PyTorch, TensorFlow, NumPy, SciPy, Scikit-learn, Pandas.
Control & Verification: EPANET, Semidefinite Programming (SDP), SMT/Z3, MBSE (SysML), Dynamical System Simulation, Sensor Fusion, Real-Time Control.
Infrastructure & DevOps: Linux/Unix (Advanced), Git, Docker, CI/CD, GDB, System Architecture, Parallel Computing, Large-Scale Data Processing, Build & Debug Tooling, MySQL.
Languages: English (C2 - Proficient), Greek (Native), German (B1).