

SAMUEL JOHNSTON

johnston.samuelj@gmail.com | linkedin.com | github.com

EXPERIENCE

Data Scientist – Lake ice cover applications

H2O Geomatics

May 2024– Present
and Jan 2023 – Sept 2023

- Optimized a petabyte-scale MODIS processing pipeline, achieving a 12× speedup and 16× reduction in memory usage on HPC systems through algorithmic improvements, memory-efficient data structures, and elimination of redundant operations.
- Developed and deployed 3D CNN models for cloud removal from lake ice products, achieving 97% classification accuracy.
- Adapted a spatiotemporal transformer for multi-step forecasting of lake ice cover, achieving mean absolute errors under 7 days and outperforming the FLake model. *see presentations*.
- Used HPC resources to scale data processing and training across large geospatial datasets.

Research Assistant – Agricultural land use analyst

University of Waterloo

May 2024 – Aug 2024
part-time

- Used Google Earth Engine and Python to analyze terabyte-scale geospatial datasets (Annual Crop Inventory), identifying long-term crop trends across the Canadian prairies.

Visiting Scientist – Incorporation of snow over ice in freshwater lake model

European Centre for Medium Range Weather Forecasting (ECMWF)

Oct 2023 – Dec 2023

- Integrated snow accumulation and insulation processes into the FLake parameterization in ECMWF's Integrated Forecasting System, improving lake ice thickness and break-up date forecasts.
- Conducted model development and testing on ECMWF's high-performance computing infrastructure in collaboration with the Coupled Processes team.

Geospatial Data Scientist – Land use classification

Lakes Environmental Software

Sept 2022 – Dec 2022
part-time

- Built a machine learning-ready dataset linking Sentinel-1 SAR observations to Dynamic World V1 land use classifications using geospatial libraries (Rasterio).

Machine Learning Researcher – Modelling of fluid dynamics

Lakes Environmental Software

Sept 2021 – Apr 2022
see publications

- Designed deep learning models (e.g. FNO) for spatiotemporal fluid dynamics simulations, matching the accuracy of OpenFOAM solvers.
- Evaluated generalization of FNO on classic CFD benchmarks with mass conservation aware training routines.

Research Assistant – Machine learning for nutrient load prediction

University of Waterloo

Jan 2021 – Aug 2021
see publications

- Built Random Forest models to predict nutrient concentrations (N, S, P) from land use with >80% accuracy.
- Used model interpretation (PDPs) to extract meaningful ecological insights and validate against hydrology literature.

EDUCATION

University of Waterloo

Bachelor of Computer Science, Honours

Diploma in Sustainability

Sept 2019 – Apr 2024
Major Average: 90.85

PUBLICATIONS

- Costa Rocha, Paulo A., **Samuel J. Johnston**, Victor Oliveira Santos, Amir A. Aliabadi, Jesse V.G. Thé, and Bahram Gharabaghi. 2023. “Deep Neural Network Modeling for CFD Simulations: Benchmarking the Fourier Neural Operator on the Lid-Driven Cavity Case.” *Applied Sciences*, 13, (5): 3165-65. <https://doi.org/10.3390/app13053165>
- Basu, Nandita B., J. Dony, K.J. Van Meter, **Samuel J. Johnston**, and Anita T. Layton. 2023. “A Random Forest in the Great Lakes: Stream Nutrient Concentrations across the Transboundary Lake Basin.” *Earth's Future*, 11, (4). <https://doi.org/10.1029/2021EF002571>

ORAL PRESENTATIONS

- Johnston, Samuel J.** “A Deep Learning Approach for Lake Ice Cover Forecasting.” Conference Presentation, Hydrospace, Lisbon, Portugal, November 30, 2023 [20 minutes]. https://youtu.be/f2Zv0HwVC8Q?si=XaDJjAw0T9_DWz83&t=14662

SKILLS

Remote Sensing & Geospatial Technologies

- Remote Sensing:** MODIS, Landsat, Sentinel-1/2/3
- GIS Tools:** Google Earth Engine, QGIS
- Geospatial Libraries:** Xarray, Rasterio, Pyproj, Geopandas

Data Science & Machine Learning

- Models:** Transformers, CNNs, RNNs, LSTMs, GANs, MLPs
- Tools:** PyTorch, Scikit-Learn, Keras, TensorFlow
- Statistical Analysis:** NumPy, Pandas, SciPy

Data Engineering & Computing

- High Performance Computing (HPC):** Digital Research Alliance of Canada, Atos (ECMWF)
- Big Data Processing:** Dask, Apache Spark

Programming & Development

- Languages:** Python, C, C++, R, JavaScript, Julia, Fortran, MATLAB
- Version Control:** Git, Github
- Operating Systems:** Linux, Unix, Windows

Natural Languages

- English** – Native Speaker
- French** – Professional working proficiency (DELF B2)
- Spanish** – Basics

REFERENCES

Claude Duguay – H2O Geomatics
Founder & Scientific Advisor

Relation: Employer and Supervisor
Contact: claudio.duguay@h2ogeomatics.com

Margarita Choulga – ECMWF
Research Scientist: Earth System Modelling, Coupled Processes

Relation: Research Advisor
Contact: margarita.choulga@ecmwf.int

Jesse Van Griensven – Lakes Environmental Software
Founder & CEO

Relation: Employer and Supervisor
Contact: jesse.the@weblakes.com