Xingyan Li

Ph.D. Candidate in Information Systems

E-mail: xingyanli@umbc.edu | Google Scholar | Website: bdal.umbc.edu/people/xingyanli

Profile

- Ph.D. candidate in Information Systems at UMBC with five years of research experience in artificial intelligence (AI) for Earth science, including interdisciplinary collaborations with researchers at UMBC and NASA Goddard Space Flight Center. Expected graduation: Spring 2027.
- Develop accurate and computationally efficient AI methods to help scientists understand environmental conditions and anticipate hazardous weather. Research focuses on learning from environmental observations and adapting existing AI models to new scientific tasks.
- First author of two published conference papers and one accepted conference paper. Two dissertation studies are complete. The third has produced preliminary improvements over baseline methods; ongoing experiments focus on refining the approach and evaluating its performance.
- Experience building environmental datasets, training AI models on multiple GPUs, and supporting researchers through UMBC's research computing services. Developed practical tutorials and guided undergraduate researchers in applying AI methods to Earth science.

Education

University of Maryland, Baltimore County — Baltimore, U.S.
Ph.D. Candidate in Information Systems — *Aug. 2021 – Present*

- GPA: 3.97/4.0
- Advisor: Jianwu Wang (UMBC); co-advisor: Jie Gong (NASA GSFC)
- Dissertation proposal defended: September 2025

Beijing Jiaotong University — Beijing, China
Bachelor of Mechatronic Engineering — *Aug. 2017 – Jun. 2021*

- GPA: 3.89/4.0

Fellowships and Awards

- **Student Travel Award** — 2026
  ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.
- **Technuf Endowed Fellowship** — 2024
  Department of Information Systems, UMBC.
- **GESTAR II Graduate Fellowship** — 2022–2023
  Goddard Earth Sciences Technology and Research II, a UMBC–NASA research partnership.
- **Student Travel Award** — 2021
  IEEE International Conference on Big Data.
- **China National Scholarship** — 2019
  Ministry of Education of China (top 1%).

Publications

First-Authored Papers

[1] **X. Li**, J. A. Caraballo-Vega, J. Gong, M. L. Carroll, and J. Wang, "SPECTRA: Band-routed embedding and stage-wise LoRA for cross-sensor fine-tuning of geospatial foundation models," accepted for publication in *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2026.

[2] **X. Li**, A. M. Sayer, I. T. Carroll, X. Huang, and J. Wang, "MT-HCCAR: Multi-task deep learning with hierarchical classification and attention-based regression for cloud property retrieval," in *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Applied Data Science Track*, Lecture Notes in Computer Science, vol. 14950. Cham, Switzerland: Springer, 2024, pp. 3–18.

[3] **X. Li**, J. Li, Z. Williams, X. Huang, M. Carroll, and J. Wang, "Enhanced deep learning super-resolution for bathymetry data," in *Proceedings of the IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)*, 2022, pp. 49–57.

Co-Authored Papers

[4] D. Murphy, K. Zhang, C. Parten, A. Sterling, H. Zhang, **X. Li**, J. A. Caraballo-Vega, J. Gong, M. L. Carroll, and J. Wang, "Deep learning approaches for cloud property retrieval: Comparing foundation model fine-tuning with training from scratch," in *Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW)*, 2025, pp. 2837–2846.

[5] O. Faruque, **X. Li**, M. A. Khan, H. Alam, and J. Wang, "Comparative evaluation of causal discovery and inference approaches on Arctic sea ice time series data," in *Proceedings of the IEEE International Conference on Big Data*, International Workshop on Big Data Analytics for Sustainability (BDA4S), 2024, pp. 4588–4597.

[6] X. Wang, P. Guo, **X. Li**, A. Gangopadhyay, C. E. Busart, J. Freeman, and J. Wang, "Reproducible and portable big data analytics in the cloud," *IEEE Transactions on Cloud Computing*, vol. 11, no. 3, pp. 2966–2982, 2023.

[7] S. Ali, S. A. M. Mostafa, **X. Li**, S. Khanjani, J. Wang, J. Foulds, and V. Janeja, "Benchmarking probabilistic machine learning models for Arctic sea ice," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2022, pp. 4654–4657.

Preprint

[8] S. Ali, U. Hasan, **X. Li**, O. Faruque, A. Sampath, Y. Huang, M. O. Gani, and J. Wang, "Causality for Earth science – A review on time-series and spatiotemporal causality methods," 2024, arXiv:2404.05746.

Research Experience

Rainfall Estimation and Storm Forecasting with Foundation Models Jun. 2025 – Present

- Collaborate with NASA GSFC's Climate and Radiation Laboratory on AI methods for estimating rainfall and predicting how storms develop over the following hour.
- Connect pretrained Earth-observation and weather foundation models so detailed cloud observations can be interpreted alongside atmospheric temperature, moisture, and winds.
- Completed the literature review and baseline experiments; initial results show improvements over existing methods, with further experiments underway to refine and evaluate the proposed approach.
- Evaluate rainfall estimates and next-hour storm predictions using historical observations and atmospheric reanalysis, examining heavy rain, storm development, forecast lead times, and computational cost.

Efficient Adaptation of Geospatial Foundation Models Dec. 2024 – Jun. 2026

- Developed SPECTRA to adapt pretrained geospatial AI models to satellite sensors with different spectral bands, reusing learned representations while updating relatively few parameters.
- Designed band-routing and stage-wise low-rank adaptation methods to use all available sensor bands and focus trainable parameters on model stages most relevant to each task.
- Evaluated three pretrained models on four image-segmentation datasets, comparing predictive accuracy and trainable-parameter counts with full fine-tuning and standard low-rank adaptation.
- Completed the first-authored paper accepted at ACM SIGSPATIAL 2026; conducted model training and comparative experiments using multiple GPUs on UMBC's chip cluster.

Multi-Task Learning for Cloud-Property Retrieval Aug. 2022 – Oct. 2023

- Developed MT-HCCAR to jointly detect clouds, distinguish liquid and ice phases, and estimate cloud optical thickness, providing complementary information for atmospheric and environmental studies.
- Combined hierarchical classification with attention-based regression, allowing cloud detection and phase information to guide optical-thickness estimates within a shared model.
- Evaluated the model on simulated observations from three satellite instruments (OCI, VIIRS, and ABI), using four metrics to compare performance and guide model selection.
- Completed the research and first-authored paper published at ECML PKDD 2024, supported by the UMBC–NASA GESTAR II Graduate Fellowship.

Deep Learning for Bathymetry Super-Resolution Sept. 2021 – May 2022

- Collaborated with NASA GSFC's AI Center of Excellence (AI CoE) to develop AI methods that increase the spatial detail of bathymetric maps and improve the representation of underwater terrain.
- Implemented and compared SR-ResNet and SR-GAN models, using a VGG feature extractor to guide reconstruction of fine-scale structures from lower-resolution bathymetric data.
- Prepared, analyzed, and visualized geospatial raster data using Python, Rasterio, OSGeo, and QGIS.
- Completed the NASA-supported research and first-authored paper published at IEEE/ACM BDCAT 2022.

Research Computing and Mentoring

Research Assistant, UMBC High Performance Computing Facility Oct. 2025 – Present

- Support UMBC researchers using the chip cluster by troubleshooting Slurm, MPI, CUDA, and software environments, and improving job scheduling and computing-resource use.
- Served as a key author of Scale-Up: The HPC & AI Tutorial Series, a five-part program helping researchers progress from cluster access to reproducible workflows, parallel computing, and scalable AI.
- Developed instructional slides, hands-on examples, and in-person lab assignments for a flipped-classroom format combining independent preparation with practical activities.
- Reviewed students' assignments and provided feedback to help participants apply HPC and AI techniques in their research workflows.

Research Assistant, REU Summer Project Jun. – Jul. 2025

- Designed and delivered tutorials for undergraduate researchers on implementing machine learning models with PyTorch on UMBC's chip cluster.
- Guided undergraduate researchers in adapting geospatial foundation models for cloud-property retrieval.

Technical Skills

- **Programming and machine learning tools**: Python, C/C++, PyTorch, Git/GitHub.
- **AI and foundation-model adaptation**: Transfer learning; parameter-efficient fine-tuning with low-rank adaptation (LoRA); multi-task learning; fusion of geospatial and weather foundation models.
- **Computer vision and Earth-science applications**: Image segmentation, regression, and super-resolution; cloud-property retrieval; rainfall estimation and short-term storm forecasting.
- **Environmental data and geospatial tools**: Multispectral and spatiotemporal data processing; satellite-data collocation; alignment across sensors, resolutions, and observation times; training-dataset construction; Rasterio, OSGeo, and QGIS.
- **Model training and evaluation**: Distributed multi-GPU training with PyTorch; hyperparameter tuning; baseline comparisons and ablation studies; analysis of predictive accuracy, trainable parameters, and computational cost.
- **Research computing and reproducibility**: Slurm job scheduling on high-performance computing (HPC) clusters; MPI/CUDA and software-environment troubleshooting; reproducible research workflows; computing-resource allocation; experience with UMBC's chip cluster, AWS EC2, and TACC.