Arpit Kapoor

PhD Candidate (Machine Learning & Hydrology)
Phone: +61 412 982 005 | Email: kapoor.arpit97@gmail.com
www.arpit-kapoor.com | [GitHub (arpit-kapoor)](GitHub) | Google Scholar | LinkedIn

---

Research Summary

PhD candidate in Mathematics & Statistics at UNSW, specialising in data-driven hydrology and water resources modelling. My research develops hybrid and Bayesian deep learning approaches that integrate physical models with machine learning to advance rainfall–runoff modelling, flood forecasting, and groundwater flow emulation, with publications in leading journals. I have expertise in Bayesian inference, machine learning, uncertainty quantification, and high-performance computing, alongside experience with large-scale environmental and remote sensing datasets. Beyond academia, I worked as a Climate Science Support Officer at the Bureau of Meteorology and in data science roles, developing scalable ML pipelines, climate model bias correction, and reproducible workflows. I aim to create innovative, interpretable models that improve prediction and decision-making in water resources while advancing scientific machine learning.

Education

Doctor of Philosophy (PhD), Mathematics and Statistics   Aug 2022 – Feb 2026 (Expected)
University of New South Wales, Sydney
Grants: *UNSW University International Postgraduate Award (UIPA)*

Bachelor of Technology (B.Tech)   Jul 2015 – May 2019
SRM Institute of Science and Technology, India
Major: *Computer Science & Engineering* **(GPA: 9.01/10.0)**
Final Project: *Deep reinforcement learning methods for locomotion of humanoid robot*

PhD Projects

Enhancing environmental process modelling with scientific deep learning

- **Cyclone trajectory and intensity prediction with uncertainty quantification –** Applying variational recurrent neural networks to forecast cyclone behaviour with probabilistic confidence estimates.
- **DeepGR4J: Hybrid rainfall–runoff modelling –** Designing a deep learning hybridization of the GR4J conceptual hydrological model to advance streamflow prediction.
- **Quantile-based ensemble deep learning for flood forecasting –** Building ensemble frameworks for streamflow and flood prediction in Australian catchments, with emphasis on uncertainty quantification.
- **QDeepGR4J: Extreme value–aware rainfall–runoff modelling –** Extending hybrid models with quantile regression to better capture extremes and propagate uncertainty.
- **Neural operators for groundwater modelling –** Developing geometry-informed neural operator (GINO) approaches to emulate 3D groundwater flow with transient conditions and variable density.

Professional Experience

Bureau of Meteorology, Australia   Feb 2023 – Apr 2025
Research Support Scientist

- Implemented and evaluated a **multivariate bias correction method (MRNBC)** for climate projections to support the **Australian Climate Service**, improving reliability of projections for downstream tasks of hydrological modelling.
- Developed a modern **Python interface for legacy FORTRAN software**, enabling integration with current scientific workflows and facilitating reproducibility
- Applied distributed computing with **Dask on the National Computing Infrastructure (NCI), reducing model runtimes by 80%**.
- Co-authored an **abstract for EGU 2025** on bias adjustment techniques for climate data and contributed to a **forthcoming Bureau research report** on MRNBC implementation

Quince, India   Mar 2022 – Aug 2022
Data Scientist

- Applied **self-supervised learning** techniques to model and predict customer churn, which informed strategies leading to a **15% lift in user engagement**.
- Developed a **gradient boosting–based forecasting model** (logistics), highlighting skills in model design, optimisation, and validation for large-scale applied datasets.
- Worked in an interdisciplinary setting, integrating machine learning outputs into decision systems — experience directly transferable to applied environmental modelling pipelines.

3Qi Labs, India   Nov 2019 – Nov 2021
Data Scientist

- Designed and deployed an LSTM AutoEncoder for **anomaly detection** in high-volume data pipelines, demonstrating **expertise in sequence modelling and evaluation of unsupervised methods**.
- This approach led to a **40% uplift in the detection accuracy** compared to the existing approach
- Built **automated machine learning** workflows using distributed computate platforms (PySpark & DataBricks), developing reproducible pipelines **suited to large-scale scientific data**.
- Introduced and deployed **MLOps infrastructure** (MLflow & Docker), improving model lifecycle management.

BoMoTix, India   Jan 2019 – Nov 2019
Machine Learning Engineer

- Translate state-of-the-art research in computer vision to applications in sports analysis, bridging academic methods with real-world applications.
- Designed **deep learning** based solutions for **computer vision** problems like player **detection and tracking** (based on DeepSORT), 3D player **pose estimation** (based on DeepHAR)
- Improved model efficiency using **mixed-precision and GPU optimisations**, leading to a 15% reduction in model training costs
- Cleaned, processed, and curated video data for model training and validation
- Led the effort on methodological documentation and authored technical white papers.

The University of Sydney   Jan 2018 – Sep 2018
Visiting Student Researcher

- Contribute to research on **Bayesian methods** for uncertainty quantification in neural networks
- Co-developed a parallel **Markov chain Monte-Carlo (MCMC)** based approach for **Bayesian transfer learning in neural networks**
- Provided programming support for the development of a Bayesian inversion approach in geoscientific models using parallel-tempering MCMC methods
- Co-authored publications on Bayesian machine learning at top-tier ML journals

Grants & Awards

- PhD Scholarship, Australian Research Council Training Centre in Data Analytics for Resources and Environment (DARE) Centre
- PhD Scholarship, UNSW University International Postgraduate Award (UIPA)
- Finalist at the IEEE/RSJ IROS 2017 Humanoid Application Challenge (Canada)
- Secured 1 Gold, 2 Silver, 1 Bronze medal at Robogames 2017 (USA)

Academic Activities

- Co-organiser, MODSIM 2025, Session *"C7. Advancing Machine Learning in Hydrology: Ensuring Reliable, Reproducible, and Validated Applications"*
- Presented a 10-minute thesis at the DARE Symposium hosted by the ARC training centre for Data Analytics for Resources and Environment (DARE Centre) (2025) (video link: https://www.youtube.com/watch?v=3zhcXW1qfXQ)
- Selected Participant, Information Resilience PhD School, University of Queensland (2024)
- Challenge Facilitator, Data Study Group, Alan Turing Institute (UK, May 2024)
- Invited for talk on DeepGR4J project at the HydroCafe, The Bureau of Meteorology (2024)
- Talk on scientific machine learning at the Postgraduate Conference, School of Mathematics and Statistics, University of New South Wales, Sydney (2024)
- Poster presentation on hybrid hydrological modelling at the Postgraduate Conference, School of Mathematics and Statistics, University of New South Wales, Sydney (2023)
- Presentor at UNSW Transitional AI research group seminar series (video link: https://www.youtube.com/watch?v=F9Gi8xVlpQU, https://www.youtube.com/watch?v=QR2xoyzNKyo)
- Co-organised and hosted the UNSW Transitional AI research group seminar series (2023-2024)
- Tutored for postgraduate machine learning courses at UNSW (MATH3856, MATH5836 and COMP9417)
- Affiliate student researcher at the UNSW AI Institute (formerly UNSW Data Science Hub)
- Team Leader - SRM Team Humanoid, student-led humanoid robotics research (2018-2019)

Publications

2023-2025

Kapoor, A., Chandra, R., QDeepGR4J: A quantile-based ensemble of hybrid deep learning rainfall-runoff models catering to extreme values and uncertainties, Journal of Hydrology **(under review)**

Takbash, A., Irving, D., Peter, J., Dao, T. L., **Kapoor, A.,** Gammon, A., Dowdy, A., Black, M., Bende-Michl, U., Jakob, D., and Grose, M.: A Comprehensive Assessment of Climate Data Bias-Adjustment Techniques Over Australia, EGU General Assembly 2025, Vienna, Austria, 27 Apr–2 May 2025, EGU25-252

Chandra, R., **Kapoor, A.**, Khedkar, S., Ng, J., & Vervoort, R. W. (2024). Ensemble quantile-based deep learning framework for streamflow and flood prediction in Australian catchments. *arXiv preprint arXiv:2407.15882*.

Kapoor, A., Pathiraja, S., Marshall, L., & Chandra, R. (2023). DeepGR4J: A deep learning hybridization approach for conceptual rainfall-runoff modelling. *Environmental Modelling & Software*, 169, 105831.

Kapoor, A., Negi, A., Marshall, L., & Chandra, R. (2023). Cyclone trajectory and intensity prediction with uncertainty quantification using variational recurrent neural networks. *Environmental Modelling & Software*, 162, 105654.

2020 - 2022

Kapoor, A., Nukala, E., & Chandra, R. (2022). Bayesian neuroevolution using distributed swarm optimization and tempered MCMC. *Applied Soft Computing*, 129, 109528.

Chandra, R., Azam, D., **Kapoor, A.**, & Müller, R. D. (2020). Surrogate-assisted Bayesian inversion for landscape and basin evolution models. *Geoscientific Model Development*, 13(7), 2959-2979.

Chandra, R., Jain, K., **Kapoor, A.**, & Aman, A. (2020). Surrogate-assisted parallel tempering for Bayesian neural learning. *Engineering Applications of Artificial Intelligence*, 94, 103700.

Chandra, R., & **Kapoor, A.** (2020). Bayesian neural multi-source transfer learning. *Neurocomputing*, 378, 54-64.

2017 - 2019

Sripada, A., Asokan, H., Warrier, A., **Kapoor, A.**, Gaur, H., Patel, R., & Sridhar, R. (2018). Teleoperation of a humanoid robot with motion imitation and legged locomotion. In *2018 3rd International Conference on Advanced Robotics and Mechatronics (ICARM)* (pp. 375-379). IEEE.

Sripada, A., Warrier, A., **Kapoor, A.**, Gaur, H., & Hemalatha, B. (2017). Dynamic lateral balance of humanoid robots on unstable surfaces. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 1-6). IEEE.