Coupling of Hydrological And Machine Learning Models For Reservoir Storage And Inflow Forecasting

By Team AQRITY
November 3, 2021

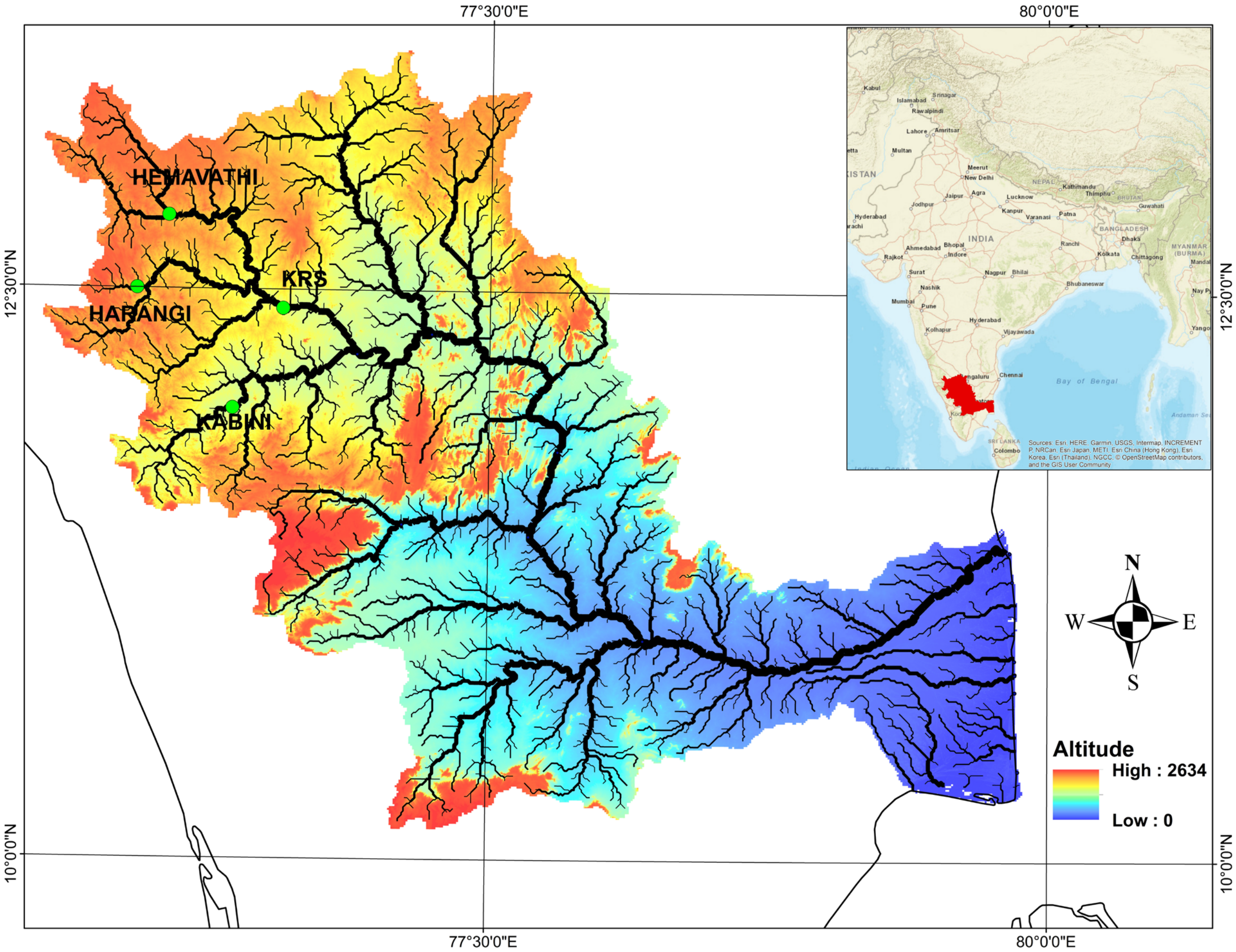
INTRODUCTION

Climate change is intensifying the water cycle. This brings more intense rainfall and associated flooding, as well as more intense droughts in many regions. This poses an undeniable threat to the water resources of the world. As Auden puts it, "Thousands have lived without love, not one without water." ***Water stress is critical in the Global South, and more so in India with just 4% of the world's freshwater, vis-a-vis 16% of the global population. Urban areas in India, like Bengaluru, are hotspots for urban water crises.*** The city has gone from the Garden City to the Silicon Valley of India. However, there is no denying that it has, sadly, also gone from the city of lakes to the city of concrete and become highly dependent on reservoirs in the Cauvery basin to maintain its urban dwellers.

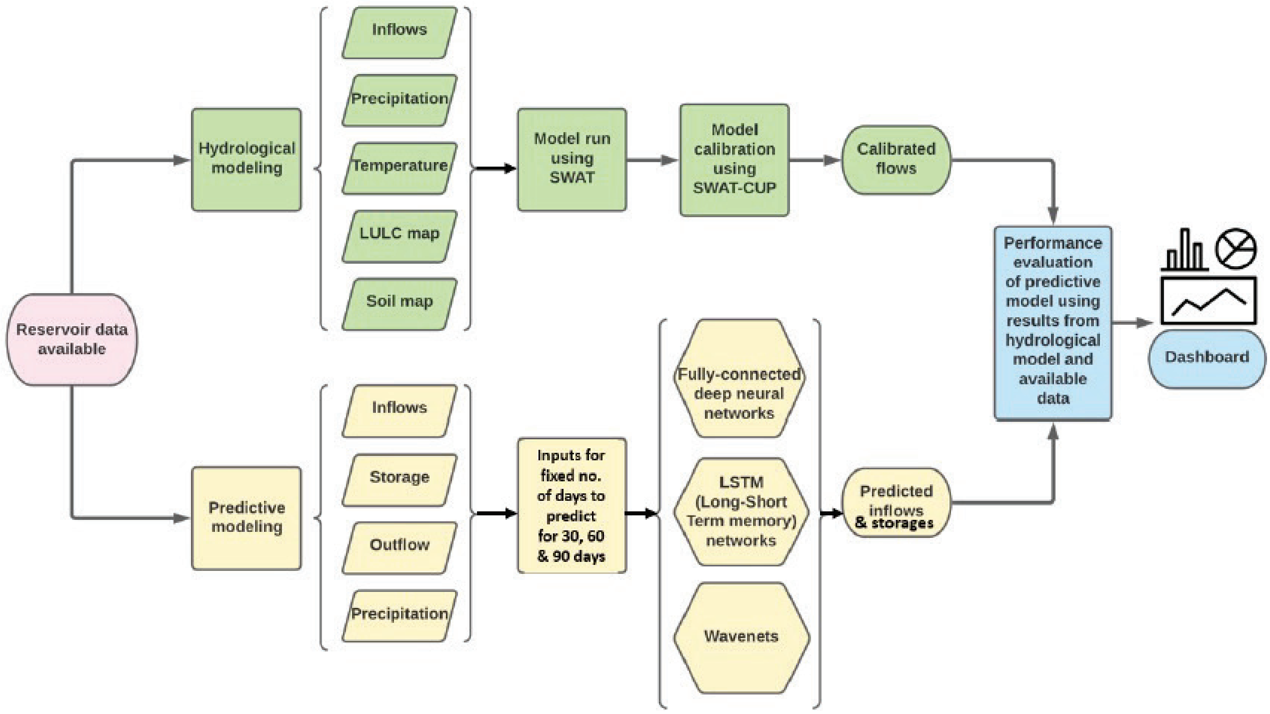
Short-term forecasting of water availability in these reservoirs can help policymakers, water supply planners, and risk managers mitigate water insecurity. ***Effective communication of this information is critical in water supply risk management. An interdisciplinary approach encompassing artificial intelligence, deep learning, water physics, hydrological science, and geoinformatics is important for such forecasts.*** Such predictive modelling exercise yields meaningful results when it is complemented with human understanding of the physical processes in the river basin. This can be achieved by combining machine learning and hydrological modelling, which is what team Aqrity attempted as part of the Wave2Web Hackathon.

AQRITY'S WORKING METHODOLOGY FOR RESERVOIR STORAGE AND INFLOW FORECASTING

Cauvery river originates in the Western Ghats of India at Coorg district in Karnataka, India. It traverses 800 km and drains an area of 82,000 square kilometers before it flows into the Bay of Bengal. The Upper Cauvery has four major reservoirs, namely - Hemavathy, Harangi, Kabini, and Krisharajasagara. The reservoirs mainly depend on the southwest monsoon rainfall with high inflows during the months of June-September while October - May is considered a lean period.



The transformation of rainfall to runoff is a highly complex process that can only be understood by adopting a hydrological model to simulate the behavior of the river basin.



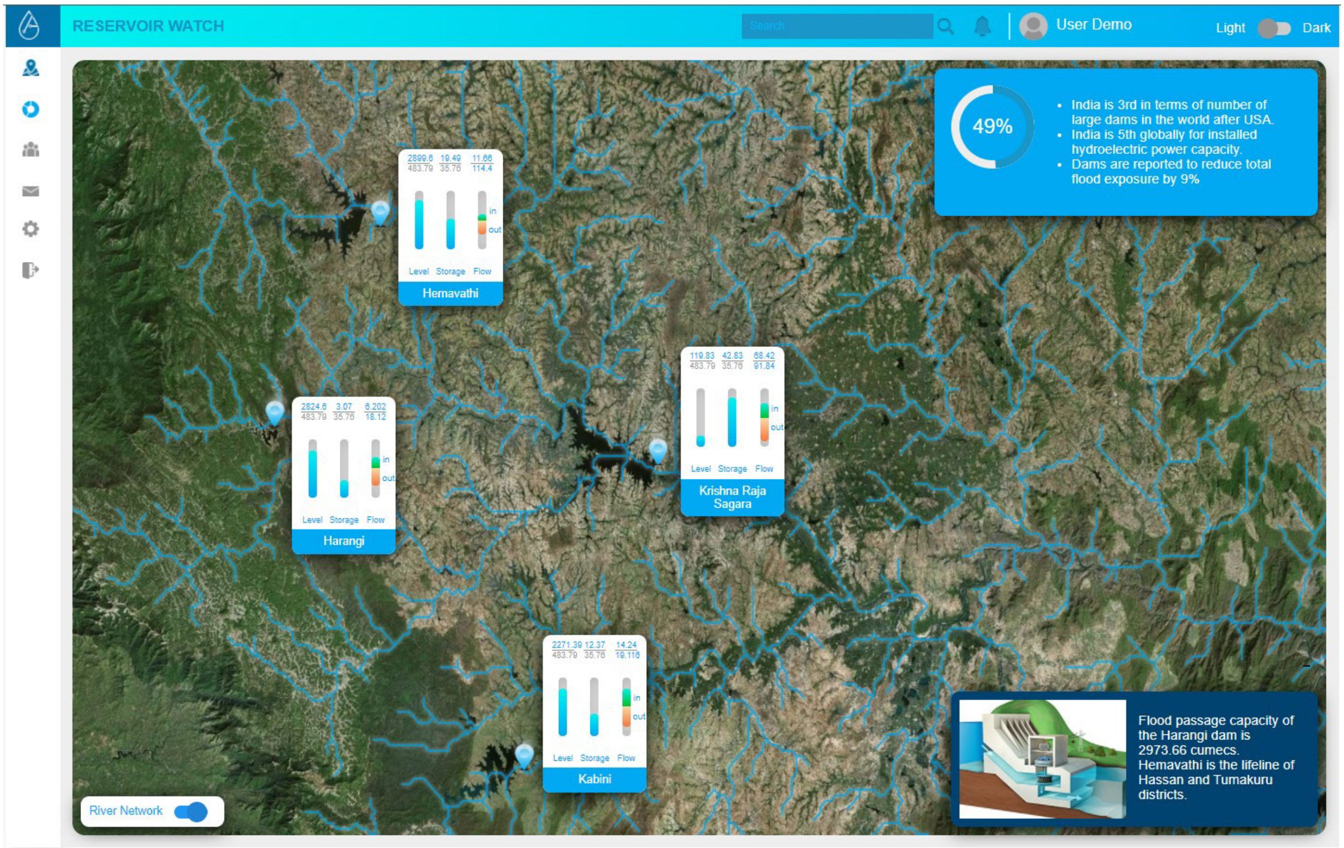
A semi-distributed hydrological model (SWAT) takes precipitation, temperature, wind speed, solar radiation, relative humidity, land use data, and soil data as inputs. ***This hydrological model outputs variables like inflow, outflow, and storage for four reservoirs in the Upper Cauvery River Basin.***
These outputs from the calibrated and the validated hydrological model, along with observed data, are fed into machine learning models like fully connected deep neural networks, long-short term memory (LSTM) networks, and WaveNets.

LIMITS TO MACHINE LEARNING MODELS

Machine learning models often suffer from overconfident predictions, uncertainty in observations, and uncertainty in the physical processes being modelled. Overconfident predictions mean that accuracy is likely to be lower than what is indicated by the predictive performance metrics. Observation uncertainty arises due to our limited ability to obtain precise data measurements. ***Uncertainty in the physical processes in the river basin arises due to the differential behavior of the river basin for the same weather inputs. Observational and process-based uncertainty estimation is done using the Monte-Carlo Dropout technique.*** Quality of predictions is found out by using two performance metrics: coefficient of determination and root mean square error. Our coefficient of determination values ranges from 0.65 to 0.98, which is reasonably good. Root mean square error ranges from 1.8 TMC in low flow periods to 110 TMC in high flow periods, which again suggests that our predictions are reliable.
The general philosophy of Occam's Razor states that there is no need to over-compliment a model, with many features and variables and algorithmic supplements when a simpler model would work. ***To reduce the number of input features, a correlation matrix of all variables is used. To simplify our model, the most important features are found and all other features are removed by using the Random Forests method.***

USE-CASES FOR AQRITY'S HYBRID FORECASTING APPROACH

This simple solution has relevance to multiple stakeholders. It can give near-real-time predictions to city water managers, farmers, and electricity generators. The dashboard also has some general trivia about dams and reservoirs which will spark interest in children and the public and motivate them to conserve water. The forecasts can decrease flood and drought risk by informing decision-makers on what crops to plant, when to restrict water use, or on how to prioritize emergency funds. These forecasts can also inform the energy sector and power plant managers to better understand power plant-specific water risks.



The results from this deep tech exercise are conveyed through an easy to grasp interface that is intuitive so that a first-time user can use the dashboard without prior training. The interactive dashboard helps stakeholders in making scientifically informed decisions. On the dashboard, visualization of geospatial and time-series data of a host of variables is readily accessible both from historical and future perspectives with management options. This eases the management of available water resources and improves risk preparedness. Predictions of inflows to the reservoir and storage capacity in the reservoir are done for 30, 60, and 90 days ahead, and these predictions are shown on the dashboard along with performance metrics.

Our combined model predicts inflows and storage. Water managers can use the predictions from inflows and storage to make risk-informed decisions on the outflows from the reservoir. The dashboard also shows historical information related to soil moisture and evapotranspiration in the basin upstream of the reservoir. This information is useful for decisions related to agriculture and irrigation.

Graphics are created by authors.