26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Machine learning approach to irrigation demand forecasting in the Indus Basin using Google Earth Engine

Mohsin Hafeez[a], **Hafsa Aeman**[a] **and Umar Waqas Liaqat**[a]

[a]*International Water Management Institute (IWMI), Pakistan*
Email: h.aeman@cgiar.org

Abstract: The availability of cloud platforms like Google Earth Engine (GEE) has revolutionized large-scale geospatial analysis, enabling timely and spatially resolved agricultural irrigation demand forecasting. This is crucial for sustainable water use and agricultural resilience. GEE's cloud computing capabilities allow for parallel processing, reducing the need for local storage or high-performance computing infrastructure. Additionally, integrating remote sensing (RS) data archives with machine learning (ML) models and scalable application services creates a powerful environment for operational, data-driven irrigation applications. While many RS-based studies in GEE have focused on land use and crop classification, few have developed operational tools for irrigation forecasting using integrated climate and satellite information. This study focuses on Pakistan's Indus Basin region, a major agricultural region supported by canal-based irrigation. The terrain is largely flat with a gradual slope from south to west, and the soil is predominantly coarse-textured alluvium, with a mix of fine to medium silt and sand in the upper regions. The region receives most of its rainfall during the monsoon season, and evapotranspiration rates are generally high. This study addresses that gap by developing a cloud-based irrigation forecasting system within GEE, integrated with the Google Colab platform for training and testing geospatial remote sensing datasets to develop early irrigation forecasts.

The system integrates multi-sensor satellite information, including Landsat-8, Sentinel-2, and PlanetScope, with archival datasets from 2015 to 2025 at 30 m, 10 m, and 3 m spatial resolutions, respectively, combined with climate information. A Random Forest (RF) machine learning model was implemented and trained using an 80-20% split to estimate crop-specific water requirements, measured as actual evapotranspiration (ETact), across spatial and temporal scales. The framework automates data retrieval and irrigation forecast generation, demonstrating the potential of GEE as a cloud-native platform for real-time agricultural water management. The methodology involves the modular prediction of four key biophysical variables within GEE, including the Normalized Difference Vegetation Index (NDVI), Soil-Adjusted Vegetation Index (SAVI), Land Surface Temperature (LST), and Net Radiation (Rn).

Trained models demonstrated high validation accuracy for NDVI ($R^2 = 0.92$), LST ($R^2 = 0.86$), Rn ($R^2 = 0.84$), and SAVI ($R^2 = 0.87$). These predicted variables served as inputs to an RF model used to estimate ETact, the core metric for calculating irrigation demand. The forecasting information was applied to wheat cropping systems during the Rabi season. During late March, a critical grain-filling stage for wheat, the model predicted a peak irrigation demand of approximately 29.87 mm/day, consistent with seasonal agronomic expectations. Weekly ET maps generated by the system supported precise and timely irrigation scheduling. By integrating satellite-derived biophysical variables within the GEE cloud-based application system, it enables informed decision-making for farmers, extension agents, and water managers in advancing sustainable water resource planning under increasing climate variability.

Keywords: *Google Earth Engine, Evapotranspiration (ET), Irrigation Demand Forecasting, Remote Sensing, Cloud-Based Application System*