Understanding the impact of urbanisation on wetlands using random forest machine learning algorithm in Google Earth Engine: a case of East Kolkata Wetlands, India

Tanmay Mondal* and Preeti Onkar

Department of Architecture and Planning, Maulana Azad National Institute of Technology, Bhopal 462 003, India

Land cover maps help in understanding the spatial and temporal dynamics of a particular area in a given time. The present study was done in the Kolkata Metropolitan Area (KMA) and East Kolkata Wetland (EKW) in India using the machine learning (ML) algorithm of random forest classification and prediction in Google Earth Engine. The ML algorithm was used to classify the land use land cover (LULC) maps of KMA and EKW for 2002, 2013 and 2023. The ML prediction model further used these LULC maps to predict the future land cover map for the year 2033, both for KMA and EKW. A correlation analysis was done between the built-up of KMA and the built-up of EKW to assess the impact of the urban expansion of KMA on the wetlands of EKW. The present study explored the impact of the increase in built-up area on the wetland and it was seen that the built-up of EKW was increasing due to the increase in the urban expansion of KMA. As a result, the wetland is experiencing a decrease in biodiversity and quality of water and soil.

Keywords: Change detection, Google Earth Engine, machine learning, prediction, random forest.

THE natural

Remote sensing and geographic information systems (GIS) are used extensively to understand and analyse the effect of changing LULC on the environment[5]. They also help manage land and environment and long-term tracking of the spatio-temporal changes of the LULC[6]. Land prediction models are used to analyse the LULC transition and forecast future land cover changes[7]. Some of the LULC prediction tools are DYNAMICS EGO, IDRISI Selva, CA-MARKOV, QGIS MOLUSCE plugin, and CLUE-S/ Dyna-CLUE. In recent times, several prediction models, such as CA-MARKOV, QGIS MOLUSCE and CLUE have been used to predict the future LULC using a transition potential matrix. However, all these prediction models use primitive GIS techniques for LULC classification.

Various indices play an important role in detecting the land cover types. Normalised difference vegetation index (NDVI) is the most commonly used for classifying land cover types. But NDVI is outperformed by various other indices like the normalised difference soil index (NDSI), normalised difference water index (NDWI), and normalised difference built-up index (NDBI) where the spectral influence of soil is quite low and thus provides more accurate classification of land uses. The