Role of machine learning techniques for spatial interpolation of environmental variables

Names and affiliations of the authors: Chorapalli Jayendra Praveen Kumar (1), (2)
Prabhakar Alok Verma (1), (2)
(1) Indian Institute of Remote Sensing,
(2) Indian Space Research Organization

Designated speaker: Chorapalli Jayendra Praveen Kumar

Machine Learning techniques are widely used in the Classification and Segmentation of Satellite Imagery, however, fewer works are carried out in the field of Spatial Interpolation. The research paper explains the concept of spatial interpolation and how to carry out spatial interpolation of a continuous variable using Geostatistical and Machine Learning techniques for a particular geographical area. Spatial Interpolation is the process of estimating the value of continuous target variable at unsampled locations based on the values at known sample locations whereas Supervised Machine Learning is a process of learning a function that maps an input to an output based on training input features and output labels. Air Temperature is an important climatic variable that plays a crucial role in global change studies thus accurate estimation of air temperature is important in several environmental studies. Usually, air temperature is measured at Automatic Weather Stations (AWS) at different heights and these provide very limited information about spatial patterns at regional or global scales because they are less in number. Hence, spatial interpolation of daily air temperature is carried out for Uttarakhand state in India to predict the values at unsampled locations and the results of geostatistical and machine learning techniques were compared. Daily temperature from 16 Automatic Weather Stations located in different parts of Uttarakhand state is considered as the primary dataset for Geostatistical techniques. MODIS Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), Digital Elevation Model (DEM), Distance from the nearest Water Body datasets are considered as input features and the daily temperature is considered as output label for the Machine Learning based approach. The concept of spatial interpolation was preserved during the training of a Machine Learning model by considering the distances among the Automatic Weather Stations and their respective daily temperature. The spatial distribution of daily temperature was estimated with the greatest accuracy of 96% by Machine Learning models when compared with Kriging based approaches. A continuous surface for daily temperature was created for all Geostatistical and Machine Learning Models for the Uttarakhand State. Keywords: Spatial Interpolation, Geostatistical techniques, Machine Learning techniques, Automatic Weather Stations (AWS).