Thesis Abstract

Shrutilipi Bhattacharjee (11IT91p01)
Department of Computer Science and Engineering, IIT Kharagpur, India

Modeling climatological dynamics and weather patterns have been studied extensively using *remote sensing* (*RS*) and *geographic information system* (*GIS*). These analyses are crucial for various geospatial applications, such as, climatological trend analysis, prediction and forecasting, urban growth modeling, etc. The meteorological parameters, closely related to earth surface, play important roles for any climatological study. Prediction of these parameters is one of the crucial pre-processing steps involved in most of the analyses, specially when the datasets contain missing and erroneous values. The geostatistical spatial interpolation methods are reported to be the most efficient choice for predicting meteorological parameters that are derived from the satellite (raster) imagery. These methods facilitate improved modeling of spatial autocorrelation/proximity, hence producing minimal in estimation. From the state of the art, it can also be observed that the interdependencies between the meteorological and terrestrial dynamics play a critical role for the proximity estimation. The semantic modeling of these land-atmospheric interactions and analyzing the associations between different factors are obvious for better spatial prediction of the meteorological parameters.

This work focuses on the semantic land-atmospheric interaction modeling for the meteorological parameters that are influencing and correlated with terrestrial dynamics. Many lietratures have reported that the *land surface temperature* (*LST*) is one such parameter, which is highly influenced by the terrain information, mainly the *land-use/land-cover* (*LULC*) distribution. Hence, for better prediction/interpolation accuracy of this parameter, the terrestrial dynamics should be modeled and incorporated into the estimation process. However, for most of the existing interpolation methods, the proximity (autocorrelation) is modeled in terms *Euclidean* distance between selected sampled locations, and is independent of the representative *LULC* variation of these locations. Therefore, modeling and quantifying this inherent knowledge of the terrain, establishing its relationship with the existing knowledge and finally proposing a more pragmatic interpolation method for the meteorological parameters are needed to be studied further.

In this work, an extensive literature survey has been carried out of existing spatial interpolation methods and some contemporary groups are found based on their popularity in the literature. It is observed that though some literatures have identified the importance of *LULC* analysis for meteorological parameters', they failed to incorporate this knowledge into their prediction process. A new spatial interpolation method is proposed, namely *semantic kriging* (*SemK*), which is capable not only to model the *LULC* distribution of the terrain, but also incorporate this property into the exiting interpolation method to make the prediction process more pragmatic and yield better accuracy. It presents a novel approach to extend any spatial interpolation method (for meteorological parameters) with contextual/semantic *LULC* knowledge of the terrain. A hierarchical ontology based approach has been adopted to quantify the same. To blend this semantic knowledge into the interpolation process, the most popular interpolation method reported in literature, i.e., *ordinary kriging* (*OK*) has been extended further in semantic domain. Hence, like *OK*, the proposed approach, SemK can also be categorized as a geostatistical univariate spatial interpolation method, which aims to minimize the variance of estimation error.

This proposed univariate spatial interpolation method, semantic kriging has been extended further to come up with more accurate interpolation model by a-posterior correlation analysis between *LULC* classes. A fuzzy Bayesian analysis approach has been considered to evaluate the influence of other classes on a pair of *LULC* classes. It is more pragmatic solution than the basic *SemK* approach, with more accuracy. The *SemK* is further extended to spatio-temporal domain and past time-series data is utilized further to predict the parameter in future time instances. Next, an application of *LULC* distribution pattern analysis for urban landscape modeling is presented. Here, the basic spatio-temporal *SemK* is extended to a multivariate model in order to incorporate different meteorological parameters' information for forecasting *LULC* distribution of a terrain in future.

For the efficacy analysis of all the proposed approaches, an empirical experimentation have been carried out with *LST* data, derived from the satellite imagery. The proposed approaches are compared with relevant exiting spatial and spatio-temporal interpolation methods and found to produce more accurate results than others. For basic spatial *SemK* approach, theoretical analysis is also presented to prove that the extension of the basic interpolation framework with *LULC* distribution knowledge of the terrain is more reliable. With higher information content, it is accountable to produce better results than other existing methods.