Spatial Interpolation to Predict Missing Attributes in GIS Using Semantic Kriging

Shrutilipi Bhattacharjee, *Student Member, IEEE*, Pabitra Mitra, *Member, IEEE*, and Soumya K. Ghosh, *Member, IEEE*

***Abstract*—Prediction of spatial attributes has attracted significant research interest in recent years. It is challenging especially when spatial data contain errors and missing values. Geostatistical estimators are used to predict the missing attribute values from the observed values of known surrounding data points, a general form of which is referred as *kriging* in the field of geographic information system and remote sensing. The proposed *semantic kriging* (*SemK*) tries to blend the semantics of spatial features (of surrounding data points) with *ordinary kriging* (OK) method for prediction of the attribute. Experimentation has been carried out with *land surface temperature* data of four major metropolitan cities in India. It shows that SemK outperforms the OK and most of the existing spatial interpolation methods.**

***Index Terms*—Data semantics, geographic information system (GIS), kriging, ontology, prediction.**

I. Introduction

PREDICTION of spatial attributes is a challenging task in the field of remote sensing (RS) and geographic information system (GIS). RS satellite imagery is processed further through several intermediate steps to create derived spatial attributes. These derived attributes are usually stored as the vector data in the form of thematic layers in the GIS. Prediction is required when these data contain missing values and errors. For example, if an RS image is captured by some defective sensors, raw image information may be missing for some regions, which, in turn, may introduce errors in the derived data. Similarly, due to the error in the intermediate processing, the derived data extraction process may be erroneous. It is necessary to replace the erroneous attribute value with some predicted value using interpolation. Moreover, the sensors related to land cover analysis are deployed over the region depending on the application. Measurements are taken for the sensor locations, and the values of the other points are estimated using interpolation.

RS literature reports several works for prediction of missing spatial attributes through image analysis, statistical interpolation methods, machine learning techniques, etc. In most GIS, it is the derived and processed attribute value with respect to each coordinate point [of region of interest (*RoI*)] that is stored in the database. Furthermore, spatial attributes are distinctive in nature. Most of them can be treated as random field parameters showing high spatial autocorrelation. In this circumstances, the interpolation from known data points is the most appropriate option which can handle spatial properties efficiently [1]. Spatial interpolation techniques incorporate geographic location information of the sample data points [2]. The methods based on regression exhibit better performances as spatial dependence within the RoI is incorporated in the regression process. The popular approaches for spatial interpolation include *inverse distance weighting* (IDW), *kriging*, *spline interpolation*, and *interpolating polynomials* [3]. Among them, *ordinary kriging* (OK) and IDW are the widely used, compared, and mostly recommended interpolation techniques [4].

There are several geoattributes stored in a GIS data repository. For example, geoattributes related to land cover analysis include *land surface temperature* (*LST*) and different indices like *normalized difference build up index*, *normalized difference vegetation index*, *moisture stress index* (*MSI*), *normalized difference water index* [5], etc. The information is stored in various formats, namely, *shapefile*, *GML encoded file*, *PostGIS*, etc., into the database. Attribute values are recorded against each of the coordinate point, represented as *latitude/longitude* convention. These attribute values may be missing or beyond the range of its possible values for the corresponding locations. This may lead to incorrect resolution of spatial query. Thus, there is a need for spatial interpolation from nearby (spatially related) location with known attribute values.

A. Contribution

Although regression-based interpolation methods show better performance in spatial prediction, there still exists scope of improvements through incorporation of local knowledge in the prediction process. Furthermore, if there is random sampling of the spatial data, error rate in prediction for these methods can be high (refer to Table III). For example, when GIS deals with weather attributes, different geofeatures show different behaviors and influence the spatial attributes in a varying manner. Geofeatures which are spatially related to the representative feature of the interpolation point influences the prediction attribute more than other features. The proposed work attempts to capture the semantics and spatial importance of the features which are near the prediction point and enhances the prediction in terms of accuracy and information content. The major contributions of this work can be summarized as follows:

1) capturing the semantics of the geofeatures within the RoI and their formal representation;

Manuscript received May 10, 2013; revised September 4, 2013; accepted September 4, 2013.

S. Bhattacharjee and S. K. Ghosh are with the School of Information Technology, Indian Institute of Technology Kharagpur, Kharagpur 721302, India (e-mail: shrutilipi.2007@gmail.com; skg@iitkgp.ac.in).

P. Mitra is with the Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India (e-mail: pabitra@gmail.com).

Digital Object Identifier 10.1109/TGRS.2013.2284489

2) analyzing the spatial importance and the semantic similarity between the geofeatures;
3) enhancing the prediction process by incorporating the knowledge regarding the geofeatures.

The rest of this paper is organized as follows. Section II describes the state of the art related to kriging and other interpolation methods. It also presents the overall objectives of this work. IDW and OK are presented in Sections III and IV. Section V presents the proposed interpolation method, namely, *semantic kriging* (*SemK*). The theoretical analysis of the proposed SemK is presented in Section VI. The experimental results of SemK and comparisons with other prediction techniques are shown in Section VII. Finally, the conclusion is drawn in Section VIII.

II. Review of Interpolation Methods and Kriging

Spatial interpolation methods have been applied in many disciplines. These methods are mainly data specific and dependent on the type of application. Many factors affect the estimations of the methods, which include sample size, distance from the prediction point, and others. Among the deterministic interpolation methods, kriging, named after D. G. Krige [6], is one of the most popular techniques based on linear regression. It is an optimal geostatistical interpolation [7], [8] and has invited significant research interests for the last few decades. It predicts the value of a spatial attribute at a particular location from known neighboring locations by taking spatial dependence and autocorrelation into account. There are several variants of kriging, namely, OK, *simple kriging*, *universal kriging*, etc. Among all of the existing interpolation methods, the OK and the IDW are the two mostly reviewed and applied methods [9]. This section focuses on the state of the art of two mostly recommended deterministic interpolation methods, namely, OK and IDW.

Heap *et al.* [4], [9] have reported OK and IDW to be the two most compared and reviewed spatial interpolation methods. According to Karydas *et al.* [10] and Wein *et al.* [11], the IDW method and its modifications are the ones most often applied in spatial interpolation. References [12] and [13] have mentioned that kriging and IDW are the most commonly used methods in GIS applications. Several works have compared these two methods. In some cases, kriging outperforms IDW [14], [15]. Chen *et al.* [14] evaluated the effect of spatial autocorrelation to analyze the performance of the grid soil sampling of different sampling densities with these two interpolation procedures. Kriging, with known *variogram* parameters, performed significantly better than the IDW for most of the studied applications [14], [16]. Again, in some other studies, IDW shows a better result than kriging [10]. Mueller *et al.* [17] observed IDW-based interpolation performed generally equaled or better than the accuracy of kriging for the optimal parameters [17], [18]. However, some mixed results are also observed by [19]–[21]. Schloeder *et al.* [19] have reported that OK and IDW show the same accuracy. Among recent works, Yasrebi *et al.* [22] compared OK and IDW to determine the degree of spatial variability of soil chemical properties. OK performed much better than the IDW procedures in this study. Foster *et al.* [23] witnessed that, although kriging produces accurate results in many cases, other interpolation techniques, like *natural neighbor*, show a better performance in the reconstruction of the total electron content of the ionosphere images.

Methods like OK [24] are used to model a *experimental semivariogram* which considers the interpolating points to be spatially related. It calculates the spatial relationships between interpolating points in terms of *covariance* between all pairs of points. Covariance is calculated from the *semivariogram* model which is built against the known data points. The covariance and the semivariogram do not consider the neighboring local properties of the sample points. They are the functions of distance (Euclidean distance in 2-D space) and independent of the nearby influencing spatial features. However, in some applications, like climatology, the nearby spatial features have a very significant effect on attribute value prediction. Spatial features exhibit different behaviors and influence different weather attributes in a varying manner. For example, to predict LST and MSI at a certain location, different geofeatures exhibit different behavior. Reference [25] has reported that "on a hot, sunny summer day, the sun can heat dry, exposed urban surfaces, like roofs and pavement, to temperatures 50 to 90 °F (27 to 50 °C) hotter than the air, while shaded or moist surfaces—often in more rural surroundings remain close to air temperatures." In other words, the properties of the geofeatures like *building*, *road surface*, *water body*, etc., influence weather attributes significantly. For example, a *multistoreyed building* absorbs and emits more heat than a water body, so a nearby multistoreyed building will have a more significant effect than a water body for LST value of a certain location. However, in case of predicting MSI, an opposite behavior can be observed for these two features. Therefore, spatial interpolation methods, for applications like weather attribute prediction, should incorporate the behavior and the semantics of the spatial features of the sample points.

A. Objectives

This work proposes a new scheme of spatial interpolation, namely, SemK, which is based on mean-square error (MSE) minimization. It modifies the existing OK method by incorporating the semantics of the spatial features into the interpolation method. The differentiation between spatial features is carried out using ontology [26], [27]. It is used to capture the knowledge of the spatial features [28] and to organize them into a hierarchy. Semantic similarity between features can be captured from ontology. Spatial importance between each pair of features is also measured through *a priori* correlation study. These two metrics (semantic similarity and spatial importance) are used to incorporate the spatiosemantic relations between the spatial features in the interpolation method. The overall objectives of the work are as follows:

1) building a spatial feature ontology based on the spatial attribute to be predicted and *a priori* correlation study (spatial importance) between each pair of leaf features in the ontology;
2) modifying the covariance as well as the weights assigned by OK considering spatial importance and the semantic similarity between spatial features;
3) mathematical formalization of the modified weight matrix and other related parameters;

4) performance evaluation and comparison of SemK with some of the existing methods with real LST data.

Before describing SemK, some fundamentals related to IDW and OK are presented in Sections III and IV, respectively.

III. IDW

The IDW [10], [11] is one of the mostly applied and compared interpolation techniques in the field of environmental science. In IDW, estimates are made based on nearby known locations which are weighted only by *distance* as it assumes that the value of a point is influenced by the nearby points. The weights assigned to the interpolating points are the inverse of its distance from the interpolation point. Therefore, the nearby points are supposed to have more weights (so, more impact) than distant points and vice versa. The known points are assumed to be independent from each other. The basic method of IDW interpolation is also known as *Shepard method* [29]. The estimated attribute value $\hat{Z}(x_0)$ at the prediction point (x_0) is given as

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \frac{w_i x_i}{\sum_{j=1}^{N} w_j} \tag{1}$$

where N is the total number of interpolating points, x_i is the attribute values at the ith interpolating point, and w_i is the weight assigned to each of the interpolating point. The weight function is given by

$$w_i = \frac{1}{d(x_0, x_i)^p} \tag{2}$$

where $d(x_0, x_i)$ is the distance from the interpolating point x_i to the prediction point x_0 (usually taken as the Euclidian distance), N is the total number of known points used for interpolation, and p is the *power parameter* and is defined as the rate of the reduction of weight with increasing distance [30]. The value of p depends on the dimension of the space where the prediction is carried out. For 2-D space, $p \leq 2$.

IV. OK

Kriging [7], [31] represents the full family of generalized least-square regression algorithms with the aim of minimizing MSE in prediction. It advances upon the existing interpolation techniques through the use of the underlying spatial relationships among the interpolating points. Sample points are not treated as independent; spatial autocorrelation influences their behavior. The kriging estimators use (3) with minor modifications of it

$$\hat{Z}(x_0) - \mu = \sum_{i=1}^{N} w_i \left[Z(x_i) - \mu(x_0)\right] \tag{3}$$

where $\hat{Z}(x_0)$ is the estimated random field (prediction attribute) value at point x_0 and μ is the stationary mean treated as the constant over the whole RoI [32].

OK assumes the stationarity of the first moment of all random variables, i.e., $E\{Z(x_i)\} = E\{Z(x_0)\} = \mu = \mu(x_0)$, where μ is unknown. The parameter w_i is the weight assigned to the ith interpolating point, calculated from the semivariogram, and N represents the number of interpolating points used for the estimation, which depends on the size of the search window. *Semivariance* $(\gamma(h))$ provides the knowledge about the underlying relationships and the amount of autocorrelation with respect to the distance. It is half the *variance* of the differences between the random field values of all of the sample points, separated by lag distance h. Spatial covariance is the function of Euclidian distance between sample points in 2-D space, which is calculated from the semivariogram. A plot of semivariance versus distance between known sample points is termed as semivariogram. Semivariance $(\gamma(h))$ of the random field Z between two data points (h distance apart) is defined as

$$\gamma(h) = \frac{\sum_{i=1}^{N} \left[Z(x_i) - Z(x_i + h)\right]^2}{2N} \tag{4}$$

where $\gamma(h)$ is the semivariance for the lag interval h, $Z(x_i)$ is the measured attribute value at a point x_i, $Z(x_i + h)$ is the measured attribute value at the point which is separated by lag distance h from x_i, and N is the total number of sample points within lag interval h. Trend analysis plot of $\gamma(h)$ against h gives the experimental semivariogram, which displays several important properties [33]. This experimental semivariogram model is used to measure the covariance of all of the interpolating points with respect to the prediction point.

Let us assume $\epsilon(x_0)$ be the amount of error in estimation of the random field value Z at x_0. If $Z(x_0)$ and $\hat{Z}(x_0)$ are the original and estimated random field values at x_0, then $\epsilon(x_0)$ is given as

$$\epsilon(x_0) = \hat{Z}(x_0) - Z(x_0) \tag{5}$$

$$= \sum_{i=1}^{N} w_i Z(x_i) - Z(x_0) \tag{6}$$

where w_i is the weight assigned to the ith interpolating point and $Z(x_i)$ is the random field value at x_i. For OK, as the random function is stationary, the expected value of the error is supposed to be zero

$$E\left(\epsilon(x_0)\right) = 0 \tag{7}$$

$$\sum_{i=1}^{N} w_i \times E\left(Z(x_i)\right) - E\left(Z(x_0)\right) = 0 \tag{8}$$

$$\mu \sum_{i=1}^{N} w_i - \mu = 0 \tag{9}$$

$$\sum_{i=1}^{N} w_i = 1 \tag{10}$$

$$\mathbf{1}^T \mathbf{W} = 1. \tag{11}$$

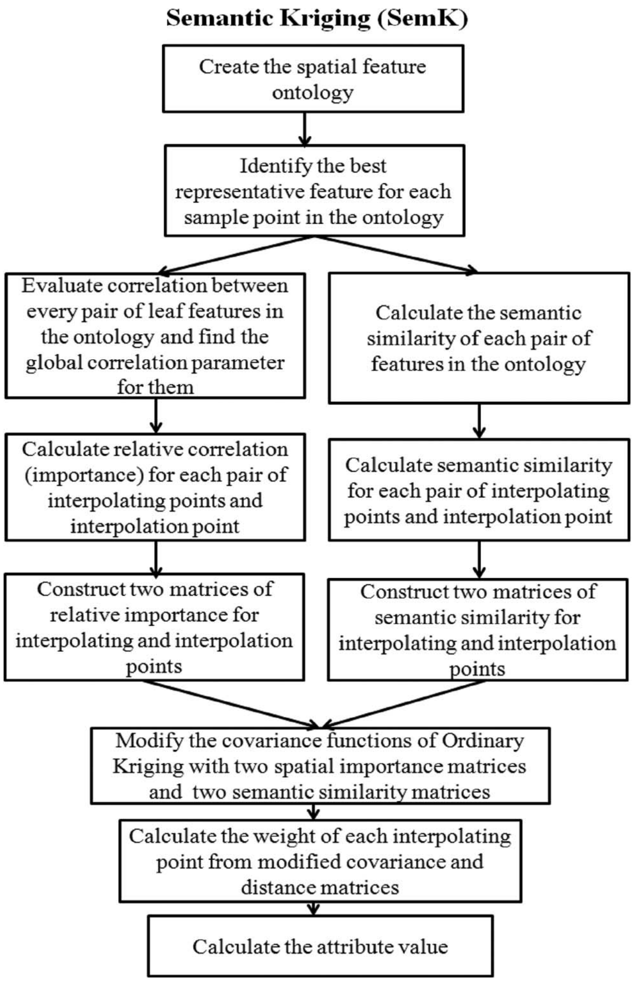

Fig. 1. Proposed SemK framework.

Thus, the general estimation equation of OK can be given as follows:

$$\hat{Z}(x_0) = \sum_{i=1}^{N} w_i Z(x_i) \tag{12}$$

constrained by $\mathbf{1}^T\mathbf{W} = 1$, where $\mathbf{W}$ is a vector of size N and is given as $[w_1 w_2 \cdots w_N]^T$.

V. SemK

The proposed SemK extends the concept of OK by combining the semantics and the correlation between spatial features into the interpolation method to provide better estimation of the prediction point. It maps the traditional covariance to higher dimension to blend the semantic of the nearby features for more informative estimation. Differentiation between spatial features is done such that the correlated and semantically similar features (representing the sample points) will have more importance than the distant one. The overall flow of the SemK process is depicted in Fig. 1.

Based on the spatial RoI, the spatial feature ontology is constructed with all possible features of that region. The features are represented as the concepts in the ontology. They are organized into a hierarchy based on some standard semantic relations, such as *hyponym*, *meronym* [34], etc. According to the hierarchical ontology property, semantically similar features will be closer in the hierarchy than the dissimilar one. For example, in this work, four metropolitan cities in India are considered for the case study to predict the LST. The spatial region, Kolkata [a metropolitan city of India, central coordinate (22.567° N 88.367° E)], has several spatial features, namely, *built-up*, *agriculture*, *forest*, *wastelands*, *waterbodies*, *wetlands*, etc. They are further organized into a ontology hierarchy that is depicted in Fig. 2. It is constructed using *is-a* (hyponym) relation.

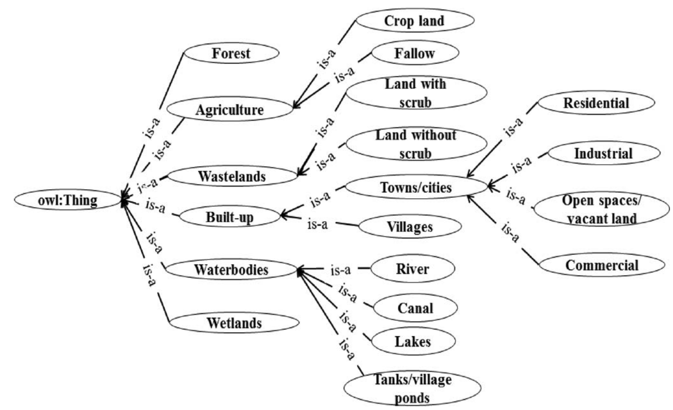

Fig. 2. Spatial feature ontology (for LST prediction).

This ontology is domain and region specific, i.e., it depends on the spatial RoI and the domain of prediction attribute. For a different domain, the construction of the ontology varies with the type of concepts that it is representing, hierarchy structure, etc. It may also vary with the number of concepts, relations, etc., with respect to the spatial RoI. This ontology is adaptive in nature. It is evident that each of the sample points can be represented by any of the leaf feature in the ontology. Next, the interpolation point and each of the interpolating points are mapped to the most appropriate representative leaf feature in the ontology hierarchy. This mapping is necessary for SemK to capture the representative feature information of each of the sample points. Once the best representative feature of each sample point is identified, the association between spatial features is evaluated. It is done in two ways, with evaluation of relative importance between a pair of leaf features and the semantic similarity between them in the ontology. These processes are named as *relative importance calculation* and *semantic similarity calculation* (described in Section V-A1 and A2, respectively). The relative importance measurement between each pair of leaf features is necessary to capture their correlation with respect to the prediction attribute in the RoI. The hop distance between a pair of features in the ontology can be taken as the heuristic to determine semantic similarity. These two parameters modify and map the standard covariance measure into higher dimension. Previously, in OK, the assigned weights are the function of distance only. The newly assigned weights in SemK are the function of distance as well as the semantic similarity and the spatial correlation of the representative features of the surrounding data points. As the covariance gets modified, the weight assigned by OK to each interpolating point also gets modified. Weights are normalized further and used to predict the attribute value. As the prediction process of

SemK has more number of decision parameters than OK, the former can be considered to be more informative than the latter. The information content in SemK and comparison with OK are presented in Section VI-A.

1) Relative Importance Calculation: Relative importance between each pair of leaf features in the ontology can be calculated by the correlation analysis between them, with respect to the prediction attribute. For this measurement, RoI is divided into some nonoverlapping zones. If k is the predefined number of sample points taken for the correlation study, the spatial region is divided into k random zones (R_k) such that $\bigcup_{i=1}^{k} R_k = RoI$. For pairwise correlation study between the spatial features, k pairs of sample points are chosen from each of the zones. Each of the pairs is chosen such that they are within a certain distance. For example, to study the correlation between *industrial* and *river* (refer to Fig. 2) with respect to LST in *Kolkata* (RoI), 50 $(k = 50)$ random sample points are selected for the feature industrial from the whole study region, and their LSTs are measured. Next, k sample points, representing the feature river, are identified against each of the k industrial, and their LST values are also measured. The global correlation value is measured for these 50 pairs of sample points over the whole study area. For this example, the correlation between industrial and river is found to be **0.81**. The correlation between any pair of leaf features in the ontology can be evaluated by this method. The correlation values, ranging between [−1, 1], are normalized to a positive range (e.g., [1, 3]) to avoid the negative mapping of the covariance. This correlation study between each pair of features assumes the following properties.

1) The correlation study is dependent on the value of the attribute to be predicted. For example, the correlations between industrial and river are **0.81** and **0.22** for prediction of LST and MSI, respectively, for study region *Kolkata*.
2) It is *a priori* correlation study, i.e., the correlation between a pair of features is determined without considering the impacts of other features.
3) The correlation parameters are taken as global for the whole study region.

Once representative features of interpolation and all of the interpolating points are identified in the ontology, each interpolating point is assigned a weight as per the correlation between the representative feature of the interpolation point and itself. This weight is also termed as *importance*. Let the representative feature of interpolation point x_0 be f_0 and the ith interpolating point x_i be f_i. Let the importance for x_i be se_i, and it is given as

$$se_i = Corr_{\text{prediction\_attribute}}(x_0, x_i) \tag{13}$$

$$= Corr_{\text{prediction\_attribute}}(f_0, f_i) \tag{14}$$

$$= \frac{\sum_{j=1}^{k} \left(Z(f_{0_j}) - Z(\bar{f}_0)\right)\left(Z(f_{i_j}) - Z(\bar{f}_i)\right)}{\sqrt{\sum_{j=1}^{k} \left(Z(f_{0_j}) - Z(\bar{f}_0)\right)^2 \sum_{j=1}^{k} \left(Z(f_{i_j}) - Z(\bar{f}_i)\right)^2}} \tag{15}$$

where $Z(f_{p_q})$ represents the random field value of the qth sample point, representing the feature f_p, and $Z(\bar{f}_p)$ represents the average of the random field value of the feature f_p over k sample points. For all of the interpolating points, it forms an $[N \times 1]$ vector, given as $SI^T = [se_1 se_2 \cdots se_N]$.

Similarly, due to spatial *autocorrelation*, relative importance exists between each pair of interpolating points. It is calculated as the correlation between their representative features. Therefore, the relative importance between ith and jth interpolating points (se_{ij}) is given as

$$se_{ij} = Corr_{\text{prediction\_attribute}}(x_i, x_j) \tag{16}$$

$$= Corr_{\text{prediction\_attribute}}(f_i, f_j) \tag{17}$$

$$= \frac{\sum_{m=1}^{k} \left(Z(f_{i_m}) - Z(\bar{f}_i)\right)\left(Z(f_{j_m}) - Z(\bar{f}_j)\right)}{\sqrt{\sum_{m=1}^{k} \left(Z(f_{i_m}) - Z(\bar{f}_i)\right)^2 \sum_{m=1}^{k} \left(Z(f_{j_m}) - Z(\bar{f}_j)\right)^2}}. \tag{18}$$

For N interpolating points, it forms an $[N \times N]$ symmetric matrix, named as *relative importance matrix* and denoted as W_2.

2) Semantic Similarity Calculation: The semantic similarity between two sample points or their representative features in the ontology is calculated using modified *context resemblance* method, described in [35]. A higher distant interpolating point in the ontology will be less similar and vice versa. This semantic similarity is proportional to the weight assigned to any interpolating point for prediction. The semantic similarity of the ith interpolating point from the interpolation point $[i \in 1 \ldots N]$ is referred to as sd_i and can be given as

$$sd_i = \frac{\frac{m_i}{|f_i|} + \frac{m_0}{|f_0|}}{2} \tag{19}$$

where $|f_i|$ and $|f_0|$ are the total number of features in the ith interpolating feature path and interpolation feature path in the ontology, respectively. m_i and m_0 represent the number of features matching in the ith interpolating feature path and interpolation feature path, respectively. With reference to the interpolation point, it forms an $[N \times 1]$ vector for all of the interpolating points, and it is given as $SD^T = [sd_1 sd_2 \cdots sd_N]$.

Since the features are spatially related, semantic dependence exists among all of the interpolating sample points and their representative features. Therefore, the relative semantic similarities are to be calculated between each pair of interpolating points too. The relative semantic similarity between the ith and jth interpolating points, $i, j \in 1 \ldots N$, is referred to as sd_{ij} and is calculated from the following

$$sd_{ij} = \frac{\frac{m_i}{|f_i|} + \frac{m_j}{|f_j|}}{2} \tag{20}$$

where $|f_i|$ and $|f_j|$ are the total number of features in the ith and jth interpolating feature paths and m_i and m_j represent the number of features matching in the ith and jth interpolating feature paths, respectively. For all of the interpolating points, it forms an $[N \times N]$ symmetric matrix, named as *semantic similarity matrix* and denoted as W_3.

These four matrices ($\mathbf{W_2}$ $[N \times N]$, $\mathbf{W_3}$ $[N \times N]$, $\mathbf{SI}$ $[N \times 1]$, and $\mathbf{SD}$ $[N \times 1]$) modify the *semivariance matrix* $(\mathbf{C})$ and the *distance matrix* $(\mathbf{D})$ of OK. The mathematical formulation of SemK and its analysis are given in Section VI.

VI. Theoretical Error Analysis of SemK

For theoretical error analysis of SemK, different parameters and constrains are formalized in this section. Let the random field value at point x_0 be $Z(x_0)$. The prediction is done based on the known interpolating points $Z^T = \{Z(x_1), \ldots, Z(x_N)\}$, where N is the number of interpolating points and $Z(x_i)$ is the random field attribute value at the point x_i. For simplicity, $Z(x_i)$ is denoted as Z_i. Let us assume $\hat{Z}_0$ to be the estimated value of the attribute at the point x_0. For OK, $\hat{Z}_0 = \sum_{i=1}^{N} w_i Z_i = \mathbf{W^T Z}$ such that $\mathbf{W^T 1} = 1$, or $\sum_{i=1}^{n} w_i = 1$.

OK is based on the notion of MSE minimization. The aim of OK is to choose the weight vector $\mathbf{W}$ such that estimation variance $\sigma_E^2 = E([Z_0 - \hat{Z}_0]^2)$ is minimized. Let the two matrices semivariance matrix $(\mathbf{C})$ and distance matrix $(\mathbf{D})$ be defined as

$$\mathbf{C} = \begin{bmatrix} Var(Z_1) & Cov(Z_1, Z_2) & \cdots & Cov(Z_1, Z_N) \\ Cov(Z_2, Z_1) & Var(Z_2) & \cdots & Cov(Z_2, Z_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(Z_N, Z_1) & Cov(Z_N, Z_2) & \cdots & Var(Z_N) \end{bmatrix} \tag{21}$$

$$\mathbf{D} = \begin{bmatrix} Cov\{Z_1, Z_0(r)\} \\ Cov\{Z_2, Z_0(r)\} \\ \vdots \\ Cov\{Z_N, Z_0(r)\} \end{bmatrix} \tag{22}$$

where $Cov(Z_i, Z_j)$ denotes the covariance between $Z(x_i)$ and $Z(x_j)$ and $Var(Z_i, Z_i)$ denotes the covariance of $Z(x_i)$ with itself, i.e., self-covariance. $\mathbf{C}$ can also be termed as $\mathbf{W_1}$ as it is the existing weight component in SemK, extended from OK. In OK, the weight vector $\mathbf{W}$ can be calculated from the matrix of semivariance between interpolating points, i.e., semivariance matrix $\mathbf{C}$ and the matrix of estimated semivariance between the interpolating points and the point at which variable Z is to be predicted, i.e., distance matrix $\mathbf{D}$.

In the proposed SemK, the covariance is influenced by the uncertainty of estimation produced by the local variables such as the impact of the surrounding spatial features. This is how the covariance between data points can be mapped to higher dimension by incorporating local properties. In this paper, the semantics and the spatial correlation between the spatial features are captured by four components $\mathbf{W_2}$, $\mathbf{W_3}$, $\mathbf{SI}$, and $\mathbf{SD}$. As the covariance captures the *variance of the difference between field values*, both semantic similarity and relative importance are inversely proportional to the covariance between random field values of any two sample points. Therefore, the modified covariance between the ith and jth interpolation points can be represented as $(C_{ij}/(se_{ij} \cdot sd_{ij}))$. The physical significance of this mapping is that, having the same distance from the interpolation point, the covariance between two interpolating points can be different depending on their representative features' relative importance and semantic similarity. The covariance value (of the same distance h) increases if semantic similarity and relative importance are less and vice versa. Therefore, the modified semivariance matrix $\mathbf{C}'$ and the modified distance matrix $\mathbf{D}'$ are given as

$$\mathbf{C}' = \frac{\mathbf{C}}{(\mathbf{W_2} \circ \mathbf{W_3})} \tag{23}$$

$$\mathbf{D}' = \frac{\mathbf{D}}{(\mathbf{SI} \circ \mathbf{SD})} \tag{24}$$

where "$\circ$" and "$-.-.-$" denote the *Hadamard product* and the *Hadamard division* between matrices, respectively. The newly modified weight matrix for SemK is termed as $\mathbf{W}'$ with dimension $[N \times 1]$. The MSE at x_0 is given as σ_{SemK}^2. As it is a variant of OK and also assumes the *mean* to be constant over the whole region, $E(\sigma_S^2 emK) = 0 \rightarrow \mathbf{1}^T\mathbf{W}' = 1$. Therefore, the MSE expression for SemK can be given as

$$\begin{aligned} \sigma_{\mathrm{SemK}}^2 &= Var\left(\left[\mathbf{W}'^T - 1\right] \times [Z(x_1) \cdots Z(x_N) Z(x_0)]^T\right) && (25) \\ &= \left[\mathbf{W}'^T - 1\right] \times \left[Var\left([Z(x_1) \cdots Z(x_N) Z(x_0)]^T\right)\right] \\ &\quad \times \left[\mathbf{W}'^T - 1\right]^T && (26) \\ &= C'_{00} + \mathbf{W}'^T\mathbf{C}'\mathbf{W}' - 2\mathbf{W}'^T\mathbf{D}' && (27) \end{aligned}$$

where $C'_{00} = C_{00}/(se_{00} \cdot sd_{00})$, C_{00} is $Cov\{Z_0(r), Z_0(r)\}$, and se_{00} and sd_{00} are the relative importance and relative semantic similarity between (f_0, f_0), respectively. Being a least-square regression algorithm, SemK tries to minimize the MSE σ_{SemK}^2 by minimizing the following equation:

$$C'_{00} + \mathbf{W}'^T\mathbf{C}'\mathbf{W}' - 2\mathbf{W}'^T\mathbf{D}'; \ni \mathbf{W}'^T\mathbf{1} = 1. \tag{28}$$

To solve it without constraints, a *Lagrange multiplier* can be introduced to the error expression. It is a technique for converting a constrained minimization problem into an unconstrained one. Let us take Lagrange multiplier -2λ and make (28) as unconstrained. Let K be the unconstrained error expression for SemK, and it is given as

$$K = C'_{00} + \mathbf{W}'^T\mathbf{C}'\mathbf{W}' - 2\mathbf{W}'^T\mathbf{D}' + 2\lambda\left(\mathbf{W}'^T\mathbf{1} - 1\right). \tag{29}$$

In order to get the minimum variance of error, the partial first-order derivative of (29) with respect to each unknown, $\mathbf{W}'$ and λ, are calculated as follows:

$$\frac{\delta K}{\delta \mathbf{W}'} = \left(\mathbf{C}' + \mathbf{C}'^T\right)\mathbf{W}' - 2\mathbf{D}' + 2\lambda\mathbf{1} \tag{30}$$

$$\frac{\delta K}{\delta \mathbf{W}'} = 2\mathbf{C}'\mathbf{W}' - 2\mathbf{D}' + 2\lambda\mathbf{1}\left(\because \mathbf{C}'^T = \mathbf{C}'\right) \tag{31}$$

$$\frac{\delta K}{\delta \lambda} = 2\mathbf{W}'^T\mathbf{1} - 2. \tag{32}$$

As SemK can be mapped to a minimization problem, the partial first-order derivative (of the error expression with respect to $\mathbf{W}'$

and λ) can be set to zero to find the value at minima. Setting $(\delta K)/(\delta \mathbf{W}') = 0$ and $(\delta K)/(\delta \lambda) = 0$, we get

$$2\mathbf{C}'\mathbf{W}' - 2\mathbf{D}' + 2\lambda\mathbf{1} = 0 \quad (33)$$

$$2{\mathbf{W}'}^T\mathbf{1} - 2 = 0. \quad (34)$$

From (33) and (34)

$$\mathbf{C}'\mathbf{W}' + \lambda\mathbf{1} = \mathbf{D}' \quad (35)$$

$${\mathbf{W}'}^T\mathbf{1} = 1. \quad (36)$$

From (35), we can write

$$\mathbf{C}'\mathbf{W}' + \lambda\mathbf{1} = \mathbf{D}' \quad (37)$$

$$\Rightarrow \mathbf{W}' = {\mathbf{C}'}^{-1}[\mathbf{D}' - \lambda\mathbf{1}]. \quad (38)$$

Substituting $\mathbf{C}' = \begin{bmatrix} \mathbf{C} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{W_2} \circ \mathbf{W_3}) \end{bmatrix}$ and $\mathbf{D}' = \begin{bmatrix} \mathbf{D} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{SI} \circ \mathbf{SD}) \end{bmatrix}$

$$\mathbf{W}' = \begin{bmatrix} \mathbf{C} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{W_2} \circ \mathbf{W_3}) \end{bmatrix}^{-1} \left[\begin{bmatrix} \mathbf{D} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{SI} \circ \mathbf{SD}) \end{bmatrix} - \lambda\mathbf{1} \right]. \quad (39)$$

From (37), multiplying both sides by ${\mathbf{C}'}^{-1}$ and $\mathbf{1}^T$, respectively

$$\mathbf{C}'\mathbf{W}' + \lambda\mathbf{1} = \mathbf{D}'$$

$$\mathbf{W}' + \lambda{\mathbf{C}'}^{-1}\mathbf{1} = {\mathbf{C}'}^{-1}\mathbf{D}'$$

$$\mathbf{1}^T\mathbf{W}' + \lambda\mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{1} = \mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{D}'$$

$$1 + \lambda\mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{1} = \mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{D}'$$

[As, $\mathbf{1}^T\mathbf{W}' = 1$]

$$\lambda = \frac{\mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{D}' - 1}{\mathbf{1}^T{\mathbf{C}'}^{-1}\mathbf{1}}. \quad (40)$$

Again substituting $\mathbf{C}' = \begin{bmatrix} \mathbf{C} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{W_2} \circ \mathbf{W_3}) \end{bmatrix}$ and $\mathbf{D}' = \begin{bmatrix} \mathbf{D} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{SI} \circ \mathbf{SD}) \end{bmatrix}$, we get

$$\Rightarrow \lambda = \frac{\mathbf{1}^T \begin{bmatrix} \mathbf{C} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{W_2} \circ \mathbf{W_3}) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{D} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{SI} \circ \mathbf{SD}) \end{bmatrix} - 1}{\mathbf{1}^T \begin{bmatrix} \mathbf{C} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{W_2} \circ \mathbf{W_3}) \end{bmatrix}^{-1} \mathbf{1}}. \quad (41)$$

$\mathbf{W}'$ and λ can be calculated from (39) and (41), where $\mathbf{W}'$ represents the final weight matrix of dimension $[N \times 1]$. $\mathbf{W}'$ is a vector which is given as $\mathbf{W'^T} = [w'_1 w'_2 \cdots w'_N]$. It can be normalized further to satisfy the constraint ${\mathbf{W}'}^T\mathbf{1} = 1$. Initially, λ is evaluated as it is independent of $\mathbf{W}'$. The desired attribute value can be obtained by $\hat{Z}(x_0) = \sum_{i=1}^{N} w'_i Z(x_i)$, where w'_i is the weight assigned to the ith interpolating point by SemK.

The minimum variance of error in SemK is given by

$$\sigma^2_{\text{SemK}} = C'_{00} + {\mathbf{W}'}^T\mathbf{D}' - \lambda \quad (42)$$

$$= \frac{C_{00}}{(se_{00} \cdot sd_{00})} + {\mathbf{W}'}^T \begin{bmatrix} \mathbf{D} \\ \text{---}\,.\,\text{---}.\,\text{---} \\ (\mathbf{SI} \circ \mathbf{SD}) \end{bmatrix} - \lambda. \quad (43)$$

As σ^2_{SemK} represents the variance of error, it should be as small as possible, preferably closer to zero. To prove the improved performance of SemK, the relation $\sigma^2_E > \sigma^2_{\text{SemK}}$ must be satisfied for a given surface for any interpolation methods. In terms of information content, SemK is better than OK and other interpolation methods. Sections VI-A and VII give the comparison of SemK with other methods in terms of information content and accuracy, respectively.

TABLE I
SAMPLE SCENARIO TO ANALYZE THE INFORMATION CONTENT OF OK AND SEMK

Scenario			
A_h	A_{SS}	A_{RI}	Assigned Weight
h_1	SS_1	RI_1	w'_1
h_1	SS_1	RI_2	w'_2
h_2	SS_2	RI_3	w'_3
h_3	SS_3	RI_4	w'_4
h_3	SS_3	RI_4	w'_5
h_4	SS_4	RI_5	w'_6

A. Information Content in SemK

This section gives an analysis of SemK in terms of information content. In this regard, the information content of OK is also analyzed to be compared with SemK. Let us consider a sample scenario with each tuple describing the specification of each interpolating point in terms of their supporting attributes and class label attributes. Let us assume six $(N = 6)$ interpolating points with supporting attributes, namely, *distance from the interpolation point* (A_h), *semantic similarity with respect to the interpolation point* (A_{SS}), and *relative importance with respect to the interpolation point* (A_{RI}). The scenario is given in Table I.

In this context, the number of tuples in the table resembles to the number of interpolating points. With respect to six interpolating points, the class label attribute has six values. In case of OK, the assigned weight to any interpolating point depends upon the distance only, so for the given scenario, $w'_1 = w'_2$, and $w'_4 = w'_5$. The reason is, for both cases, the supporting attribute A_h (distance) corresponds to the same value, h_1 and h_3, respectively. Let us assume that the number of distinct class label for OK is m_{OK}, and each of them is named as C_i $(i = 1, \ldots, m_{\text{OK}})$. Each class label C_i associates to a distinct weight assigned to the interpolating points. For the sample scenario in Table I, $m_{\text{OK}} = 4$, $C_1 = w'_1$ or w'_2, $C_2 = w'_3$, $C_3 = w'_4$ or w'_5, and $C_4 = w'_6$. Let C_i^{OK} be the set of tuples of class C_i. Therefore, $\sum_{i=1}^{m_{\text{OK}}} C_i^{\text{OK}} = N$ (number of interpolating points). If p_i^{OK} is the probability that the assigned weight to an interpolating point is assigned to class C_i, then

TABLE II
STANDARD METRICS TO MEASURE ERROR IN PREDICTION

Error Metrics	
Metric	**Definition**
Mean square reduced error (MSRE)	$\frac{1}{N}\sum_{i=1}^{n}(p_i-a_i)^2/s^2$
Mean absolute error (MAE)	$\frac{1}{N}\sum_{i=1}^{n}\lvert p_i-a_i\rvert$
Mean square error (MSE)	$\frac{1}{N}\sum_{i=1}^{n}(p_i-a_i)^2$
Root mean square error (RMSE)	$[\frac{1}{N}\sum_{i=1}^{n}(p_i-a_i)^2]^{\frac{1}{2}}$

$p_i^{\mathrm{OK}} = C_i^{\mathrm{OK}}/N$. The expected information content for OK (I_{OK}) is given as

$$I_{\mathrm{OK}} = -\sum_{i=1}^{m_{\mathrm{OK}}} p_i^{\mathrm{OK}} \log_2\left(p_i^{\mathrm{OK}}\right). \tag{44}$$

For example, the information content in OK for the given scenario in Table I is given as

$$I_{\mathrm{OK}} = -\frac{2}{6}\log_2\left(\frac{2}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) \tag{45}$$

$$= 1.92 \tag{46}$$

whereas for SemK, although $w'_4 = w'_5$ (as they match in all of the supporting attributes), $w'_1 \neq w'_2$. The reason is that, although the first and second tuples have the same distance as h_1, they vary in the supporting attribute A_{RI} (RI_1 and RI_2, respectively). This fact is captured exclusively by SemK. As the number of supporting attributes is high, the number of distinct class label of SemK (m_{SemK}) is always greater or equal to that of OK, i.e., $m_{\mathrm{SemK}} \geq m_{\mathrm{OK}}$. For example, from Table I, the distinct number of class labels for SemK is 5. For the same reason, $m_{\mathrm{SemK}} \geq m_{\mathrm{OK}} \rightarrow p_i^{\mathrm{OK}} \geq p_i^{\mathrm{SemK}}$. The expected information content for SemK (I_{SemK}) is given as

$$I_{\mathrm{SemK}} = -\sum_{i=1}^{m_{\mathrm{SemK}}} p_i^{\mathrm{SemK}} \log_2\left(p_i^{\mathrm{SemK}}\right). \tag{47}$$

For example, the information content in SemK for the given scenario in Table I is given as

$$I_{\mathrm{SemK}} = -\frac{1}{6}\log_2\left(\frac{1}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) \tag{48}$$

$$= 2.25. \tag{49}$$

Therefore, $I_{\mathrm{SemK}} \geq I_{\mathrm{OK}}$, i.e., the information content of SemK is always greater or equal to OK. Similarly, it can be proved that for, any scenario, the information content of SemK is always greater than or equal to most of the other existing interpolation techniques.

VII. EXPERIMENTAL RESULTS AND COMPARISON

Experimentation has been carried out with LST data obtained from satellite images. The primary source of these data is Landsat ETM+ satellite imagery, offered by the United States Geological Survey (USGS).[1] The spatial resolutions of these data are 30 m (for bands 1–5 and 7) and 60 m (for band 6). Satellite image processing tool has been used in generating the LST data from the thermal band (band 6.1) of the Landsat ETM+ satellite imagery. Each pixel of the satellite image has been converted to the corresponding kinetic temperature or the LST value. Thus, the data (LST values) are derived from the original satellite imagery. The LSTs of four metropolitan cities in India, namely, *Kolkata*, *Chennai*, *Delhi*, and *Mumbai*, are considered for the case study. Actual LST is measured from the image itself. It is assumed that, for some locations, the LST values are missing. Prediction is carried out on these locations. The proposed SemK is compared with other prediction and spatial interpolation methods. Five types of prediction techniques are used for comparisons, namely, *simple spatial averaging* (*Average*) of the known interpolating points, *multilayer-perceptron*-based prediction considering the position (latitude/longitude) as the input parameter and the attribute value as the output parameter, *nearest* neighbor, IDW, and OK. All of the sample points are specified by its corresponding latitude/longitude. Predicted value is compared with the actual value, and different error metrics are measured for the performance analysis of each of the methodologies. The experimentation has the following specifications.

1) A 35-km radius is taken against each interpolation point for selection of interpolating points.
2) Twenty interpolating points are considered randomly (uniformly random) against each of the interpolation points. All of these points are taken within the predefined radius.

TABLE III
COMPARISON STUDY OF THE PROPOSED SEMK WITH EXISTING PREDICTION TOOLS

Comparison of Error Metrics with *SemK* and other Methods					
City	**Prediction Method**	**Error in Prediction**			
		MSRE	**MAE**	**MSE**	**RMSE**
Kolkata	Average	0.80	1.35	2.70	1.64
	MLP	0.90	1.45	2.78	1.67
	NN	0.70	1.39	2.96	1.72
	IDW	0.83	1.28	2.35	1.53
	OK	0.77	1.42	3.15	1.77
	SemK	**0.79**	**0.62**	**0.61**	**0.78**
Chennai	Average	1.30	1.82	4.11	2.03
	MLP	1.06	2.56	9.12	3.02
	NN	0.74	0.78	1.05	1.03
	IDW	0.75	0.61	0.62	0.79
	OK	0.32	0.89	2.27	1.51
	SemK	**0.79**	**0.65**	**0.68**	**0.83**
Delhi	Average	1.77	6.35	44.34	6.66
	MLP	1.20	2.09	5.68	2.38
	NN	0.31	0.92	1.35	1.16
	IDW	0.41	0.90	1.72	1.31
	OK	0.37	1.70	7.48	2.73
	SemK	**0.69**	**0.81**	**1.01**	**1.01**
Mumbai	Average	0.54	1.28	3.44	1.85
	MLP	0.84	1.61	4.40	2.10
	NN	0.25	0.58	0.79	0.88
	IDW	0.34	0.40	0.45	0.67
	OK	0.24	2.25	39.41	6.28
	SemK	**0.34**	**0.45**	**0.64**	**0.80**

[1] http://www.usgs.gov/

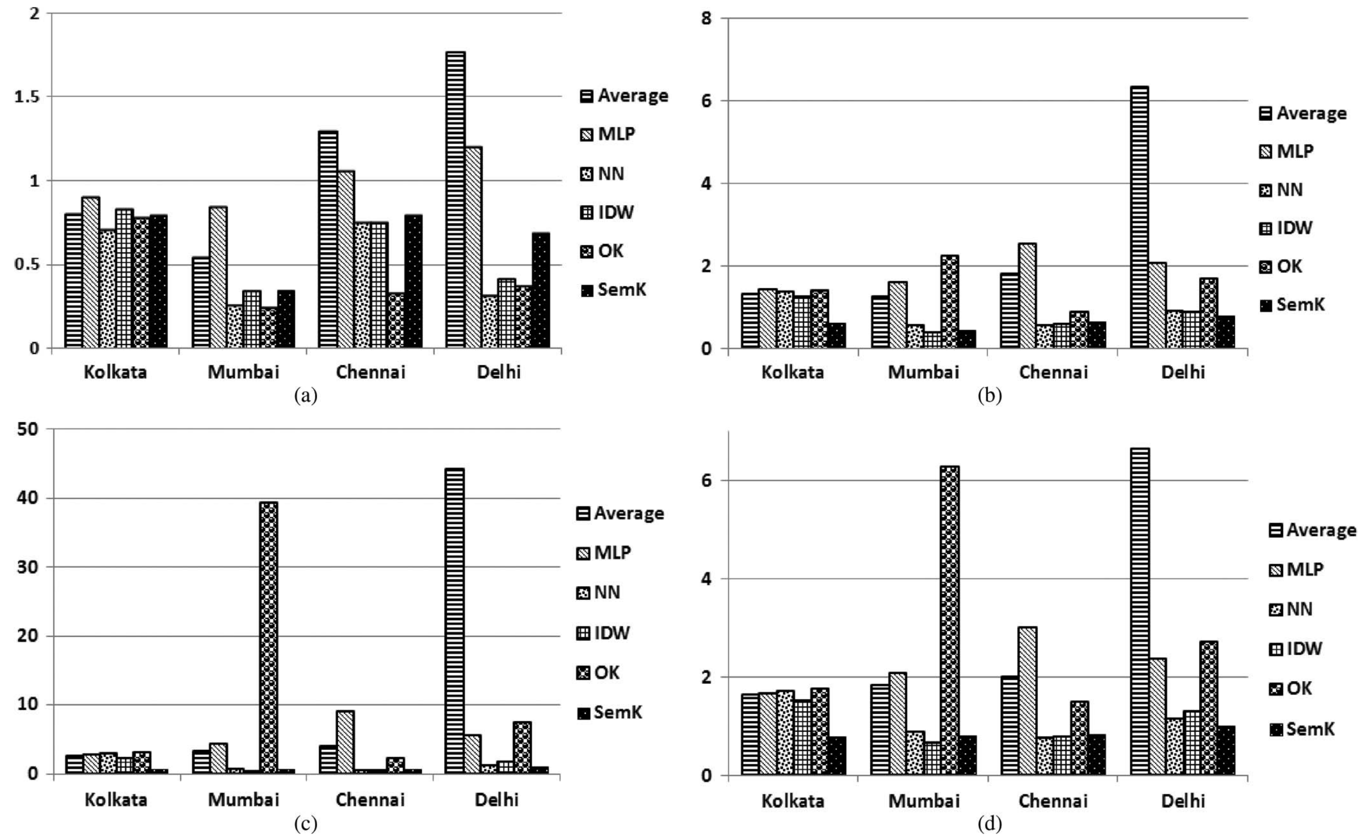

Fig. 3. Comparison study of error metrics (MSRE, MAE, MSE, and RMSE).

3) The experimental semivariogram is modeled by taking 50 random (uniformly random) sample points with lag distance $h = 5$ km.

These specifications may vary depending upon the prediction algorithm, its efficiency, accuracy in measurement, etc. The performance of each of the methodology is specified by some standard error metrics for prediction. The metrics are *mean-square reduced error* (*MSRE*), *mean absolute error* (*MAE*), MSE, and *root MSE* (*RMSE*). Their mathematical experiences are given in Table II.

Here, N is the number of interpolating points, p_i is the predicted or estimated value, a_i is the actual or observed value at the ith interpolating point, and s is the standard deviation of the estimated error. Li *et al.* [9] have discussed the evaluation criteria and physical significance of each of the metric. A model can be considered as better than others if the corresponding MSRE approaches 1, the *ME* and MSE are closer to 0, and the RMSE is smaller than others [9]. If RMSE > 1, the method underestimates the primary variable, or else, it overestimates the primary variable. According to these criteria, it has been found that the proposed SemK outperforms most of the existing methodologies, mainly the OK. The performance analysis with the comparison study is given in Table III.

A. Discussion

Table III shows the comparative study of the proposed SemK with some well-known prediction and interpolation methods. All of the methods are compared on the basis of four well-known and standard error metrics for prediction. Based on the evaluation criteria, it has been found that the proposed SemK, in general, outperforms OK and other prediction methodologies. In Fig. 3, graphical representations of error metrics are shown. In case of MSRE, as it should approach to 1, SemK and, in some cases, IDW give better performances. In case of other metrics, i.e., MAE, MSE, and RMSE, the performance of SemK is usually better than others. It is evident from Table III and Fig. 3 that the proposed SemK performs better compared to OK and most of the other prediction techniques. This may be due to the incorporation of spatiosemantic relation of the local spatial features in SemK.

VIII. CONCLUSION

Prediction of spatial attributes is one of the challenging problems in the field of GIS. The prediction of spatial attributes through interpolation is an obvious choice for GISs. The proposed SemK extends a popular regression-based interpolation technique, namely, OK, by blending the local knowledge of the surrounding spatial features. It considers the local properties of the spatial features, and the information content of SemK is higher than most of the existing interpolation techniques. The spatial autocorrelation is modeled not only in terms of distance but also considers semantic similarity and the spatial correlation of the spatial features. This autocorrelation model changes dynamically depending on the prediction attribute. Experiment has shown that the proposed method yields better performance than most of the popular interpolation techniques, mainly, OK. The prediction and forecasting of the time series data using SemK can be considered as the future extension of the work.

References

[1] F. Maselli and M. Chiesi, "Evaluation of statistical methods to estimate forest volume in a Mediterranean region," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2239–2250, Aug. 2006.

[2] N. Cressie and C. K. Wikle, *Statistics for Spatio-Temporal Data*. Hoboken, NJ, USA: Wiley, 2011.

[3] D. Zimmerman, C. Pavlik, A. Ruggles, and M. P. Armstrong, "An experimental comparison of ordinary and universal kriging and inverse distance weighting," *Math. Geol.*, vol. 31, no. 4, pp. 375–390, May 1999.

[4] J. Li and A. D. Heap, "A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors," *Ecol. Inf.*, vol. 6, no. 3/4, pp. 228–241, Jul. 2011.

[5] X. L. Chen, H. M. Zhao, P. X. Li, and Z. Y. Yin, "Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes," *Remote Sens. Environ.*, vol. 104, no. 2, pp. 133–146, Sep. 2006.

[6] H. Sheikhhasan, "A Comparison of Interpolation Techniques for Spatial Data Prediction," M.S. thesis, Dept. Comput. Sci., Univ. van Amsterdam, Amsterdam, The Netherlands, 2006.

[7] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*. New York, NY, USA: Springer-Verlag, 1999.

[8] A. Gelfand, P. J. Diggle, P. Guttorp, and M. Fuentes, *Handbook of Spatial Statistics*. Boca Raton, FL, USA: CRC Press, 2010.

[9] J. Li and A. D. Heap, *A Review of Spatial Interpolation Methods for Environmental Scientists*. Canberra, Australia: Geoscience Australia, 2008.

[10] I. A. Nalder and R. W. Wein, "Spatial interpolation of climatic normals: Test of a new method in the Canadian boreal forest," *Agr. Forest Meteorol.*, vol. 92, no. 4, pp. 211–225, Dec. 1998.

[11] C. G. Karydas, I. Z. Gitas, E. Koutsogiannaki, N. Lydakis-Simantiris, and G. Silleos, "Evaluation of spatial interpolation techniques for mapping agricultural topsoil properties in Crete," *EARSeL eProc.*, vol. 8, no. 1, pp. 26–39, Jan. 2009.

[12] D. W. Franzen and T. R. Peck, "Field soil sampling density for variable rate fertilization," *J. Prod. Agr.*, vol. 8, no. 4, pp. 568–574, Oct.–Dec. 1995.

[13] R. Weisz, S. Fleischer, and Z. Smilowitz, "Map generation in high-value horticultural integrated pest management: Appropriate interpolation methods for site-specific pest management of Colorado potato beetle (Coleoptera: Chrysomelidae)," *J. Econ. Entomol.*, vol. 88, no. 6, pp. 1650–1657, Dec. 1995.

[14] A. Kravchenko, "Influence of spatial structure on accuracy of interpolation methods," *Soil Sci. Soc. Amer.*, vol. 67, no. 5, pp. 1564–1571, Sep. 2003.

[15] F. Reinstorf, M. Binder, M. Schirmer, J. Grimm-Strele, and W. Walther, "Comparative assessment of regionalisation methods of monitored atmospheric deposition loads," *Atmos. Environ.*, vol. 39, no. 20, pp. 3661–3674, Jun. 2005.

[16] A. Kravchenko and D. G. Bullock, "A comparative study of interpolation methods for mapping soil properties," *Agronom. J.*, vol. 91, no. 3, pp. 393–400, May 1999.

[17] T. Mueller, N. Pusuluri, K. Mathias, P. Cornelius, R. Barnhisel, and S. Shearer, "Map quality for ordinary kriging and inverse distance weighted interpolation," *Soil Sci. Soc. Amer.*, vol. 68, no. 6, pp. 2042–2047, Nov. 2004.

[18] K. Krivoruchkoa and C. Gotwayb, "Using spatial statistics in GIS," in *Proc. Int. Congr. ModSim*, Townsville, Australia, Jul. 2003, pp. 713–736.

[19] C. Schloeder, N. Zimmerman, and M. Jacobs, "Comparison of methods for interpolating soil properties using limited data," *Soil Sci. Soc. Amer.*, vol. 65, no. 2, pp. 470–479, Mar. 2001.

[20] T. Mueller, F. Pierce, O. Schabenberger, and D. Warncke, "Map quality for site-specific fertility management," *Soil Sci. Soc. Amer.*, vol. 65, no. 5, pp. 1547–1558, Nov. 2001.

[21] D. R. Lapen and H. N. Hayhoe, "Spatial analysis of seasonal and annual temperature and precipitation normals in Southern Ontario, Canada," *J. Great Lakes Res.*, vol. 29, no. 4, pp. 529–544, 2003.

[22] J. Yasrebi, M. Saffari, H. Fathi, N. Karimian, M. Moazallahi, and R. Gazni, "Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some soil chemical parameters," *Res. J. Biol. Sci.*, vol. 4, no. 1, pp. 93–102, 2009.

[23] M. P. Foster and A. N. Evans, "An evaluation of interpolation techniques for reconstructing ionospheric TEC maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2153–2164, Jul. 2008.

[24] M. R. Inggs and R. T. Lord, "Interpolating satellite derived wind field data using ordinary kriging, with application to the nadir gap," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 250–256, Jan. 1996.

[25] Cool Roofs, Reducing Urban Heat Islands: Compendium of Strategies, EPA US Environ. Protection Agency, Washington, DC, USA, Tech. Rep.. [Online]. Available: http://www.epa.gov/hiri/resources/compendium.htm

[26] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing," *Int. J. Human Comput. Studies*, vol. 43, no. 5/6, pp. 907–928, Nov. 1995.

[27] S. Bhattacharjee, A. Dwivedi, R. R. Prasad, and S. K. Ghosh, "Ontology based spatial clustering framework for implicit knowledge discovery," in *Proc. Annu. IEEE INDICON*, 2012, pp. 561–566.

[28] S. Bhattacharjee, R. R. Prasad, A. Dwivedi, A. Dasgupta, and S. K. Ghosh, "Ontology based framework for semantic resolution of geospatial query," in *Proc. 12th Int. Conf. ISDA*, 2012, pp. 437–442.

[29] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in *Proc. 23rd ACM Nat. Conf.*, New York, NY, USA, 1968, pp. 517–524.

[30] M. Peralvo and D. Maidment, "Influence of DEM interpolation methods in drainage analysis," Univ. Texas Austin, Austin, TX, USA, 2003, vol. 4, *GIS Hydro*.

[31] M. H. Ribeiro Sales, C. M. Souza, and P. C. Kyriakidis, "Fusion of MODIS images using kriging with external drift," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2250–2259, Apr. 2013.

[32] H. Wackernagel, *Multivariate Geostatistics*. Berlin, Germany: Springer-Verlag, 2003.

[33] P. A. Burrough and R. McDonnell, *Principles of Geographical Information Systems*. London, U.K.: Oxford Univ. Press, 1998.

[34] Wordnet 3.0, A lexical database for english, Princeton Univ., Princeton, NJ, USA, Tech. Rep. 2006. [Online]. Available: http://wordnet.princeton.edu/

[35] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

Shrutilipi Bhattacharjee (S'13) received the B.Tech. degree in information technology from West Bengal University of Technology, West Bengal, India, and the M.Tech. degree in information technology from the National Institute of Technology, Durgapur, India. She is currently working toward the Ph.D. degree in information technology from the School of Information Technology, Indian Institute of Technology Kharagpur, Kharagpur, India.

Her research interests include spatial data mining, knowledge discovery, and spatial information retrieval.

Pabitra Mitra (M'12) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India, and the Ph.D. degree in computer science from the Indian Statistical Institute (ISI) Kolkata, Kolkata, India.

He is currently an Associate Professor with the Department of Computer Science and Engineering, IIT Kharagpur. Before joining IIT Kharagpur, he was a Scientist with the Centre for Artificial Intelligence and Robotics, Bangalore, India, and a Research Fellow with Machine Intelligence Unit, ISI. He has over 70 research publications in journals and conference proceedings. His research interests include machine learning, data mining, and information retrieval.

Soumya K. Ghosh (M'04) received the M.Tech. and Ph.D. degrees in computer science and engineering from the Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India.

He is currently an Associate Professor with the School of Information Technology, IIT Kharagpur. Before joining IIT Kharagpur, he was with the Indian Space Research Organization, working in the area of satellite remote sensing and geographic information system (GIS). He has over 100 research publications in journals and conference proceedings. His research interests include geoscience and spatial web services, mobile GIS, spatial information retrieval, and knowledge discovery.