RESEARCH ARTICLE

Spatio-temporal prediction of land surface temperature using semantic kriging

Shrutilipi Bhattacharjee[1] (iD) | **Jia Chen[1]** (iD) | **Soumya K. Ghosh[2]** (iD)

[1]Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany

[2]Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur, India

Correspondence

Jia Chen and Shrutilipi Bhattacharjee, Department of Electrical and Computer Engineering, Technical University of Munich, Munich 80333, Germany.
Email: shrutilipi.2007@gmail.com (S. B.) and jia.chen@tum.de (J. C.)

Abstract

Spatio-temporal prediction and forecasting of land surface temperature (LST) are relevant. However, several factors limit their usage, such as missing pixels, line drops, and cloud cover in satellite images. Being measured close to the Earth's surface, LST is mainly influenced by the land use/land cover (LULC) distribution of the terrain. This article presents a spatio-temporal interpolation method which semantically models LULC information for the analysis of LST. The proposed spatio-temporal semantic kriging (ST-SemK) approach is presented in two variants: non-separable ST-SemK (ST-SemK$_{NSep}$) and separable ST-SemK (ST-SemK$_{Sep}$). Empirical studies have been carried out with derived *Landsat 7* ETM+ satellite images of LST for two spatial regions: Kolkata, India and Dallas, Texas, U.S. It has been observed that semantically enhanced spatio-temporal modeling by ST-SemK yields more accurate prediction results than spatio-temporal ordinary kriging and other existing methods.

1 | INTRODUCTION

The increasing availability of spatio-temporal land surface temperature (LST) satellite data has invoked the need to analyze historical information to achieve better accuracy for LST prediction and forecasting (Deng et al., 2018; Lussana, Tveito, & Uboldi, 2018; Romaguera et al., 2018). LST is an influential factor for climate change, global warming, urban heat islands, drought, and other environmental threats (Trigo et al., 2011). One of the major challenges reported in the LST literature is the problem of handling missing pixels, line gaps, cloud cover, etc. in

satellite raster images (Chen, Huang, Chen, & Xu, 2017; Gerber, de Jong, Schaepman, Schaepman-Strub, & Furrer, 2018; Weiss et al., 2014). This work uses *Landsat 7* ETM+ satellite images; Figure 1a shows a raw band image (http://landsat.usgs.gov/) with line gaps. A magnified view is shown in Figure 1b, where small white patches of cloud cover are also visible.

To generate a comprehensive image of LST for further analysis, there is a need to analyze its past trends. Geo-statistical spatial interpolation (Dixon & Uddameri, 2016) is a popular technique in the remote sensing of LST which predicts missing pixels from the surrounding measured locations. The first law of geography (Tobler, 1970) is the building block to model the dependency among the sampled locations in the interpolation process for assigning optimal weights. In traditional spatio-temporal interpolation, the weights assigned to the sample points depend on the Euclidean distances and the temporal measurement lag with respect to the prediction point. A trade-off analysis is needed between spatial and temporal scales to model optimal weights by the interpolation process. Variety exists among different techniques based on how pragmatically this spatio-temporal autocorrelation is modeled.

Several studies have been reported in the field of spatio-temporal interpolation of LST. The performance of the methods is mainly data-specific and application-dependent (Foster & Evans, 2008). The following paragraph has identified some recent works on the existing spatial interpolation methods (such as inverse distance weighting (IDW), nearest neighbors (NN), splines (SP) and different variants of kriging) for LST-based applications. Spadavecchia and Williams (2009) compared three kriging methods: simple kriging, ordinary kriging (OK) and kriging with an external drift (KED), with a baseline inverse distance weighting (IDW) algorithm (Shiode & Shiode, 2011). They found KED to perform better in the estimation of maximum and minimum temperatures and precipitation in terms of mean absolute error (MAE) and root mean square error (RMSE). Chen, Li, Chen, Rao, and Yamaguchi (2014) applied thin-plate splines (TPS) to sharpen thermal images, which is useful for any LST application, such as urban heat island assessment and drought monitoring. Carrera-Hernández and Gaskin (2007) analyzed the spatio-temporal variations of temperature and rainfall in the Mexico Basin and their correlation with elevation. They considered five kriging methods: OK, KED, block kriging with external drift, OK in a local neighborhood, and KED in a local neighborhood. Estimation accuracy was observed to be improved by taking elevation as the auxiliary variable. Colombi, De Michele, Pepe, Rampini, and Michele (2007) proposed a numerical method for calculating spatially distributed daily mean temperature and compared it with the IDW interpolation method. They obtained a satisfactory result, reporting an RMSE of approximately 2°C. Metz, Andreo, and Neteler (2017) combined temporal and spatial interpolation methods to reconstruct the MODIS LST images for central Europe and reported an RMSE of 0.5 K. They considered emissivity and elevation as the covariates for
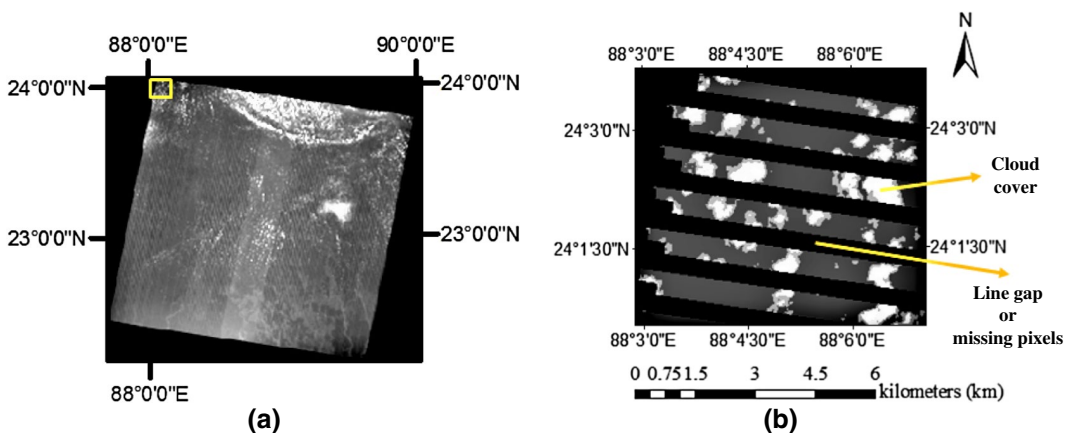

FIGURE 1 *Landsat 7* ETM+ satellite image (single band) with line gaps (black lines) and cloud cover (white patches): (a) Raw satellite image with line gaps; and (b) Magnified view of the selected area (yellow box)

the spatial interpolation. Yang, Wang, and August (2004) presented a method to estimate LST by calibrating four spatial interpolation methods using satellite-derived surface emissivity: IDW, spline, kriging and cokriging. They recommended kriging for LST interpolation if surface emissivity data are not available. Interpolation uncertainty was reduced here from 10 to 1.56°C after calibration. Xu, Wang, Hu, and Li (2013) introduced a novel technique called the point estimation model of Biased Sentinel Hospitals-based Area Disease Estimation (P-BSHADE) to interpolate missing data in temperature data sets and compared it against three estimators: kriging, IDW and spatial regression test (SRT). They applied and validated the new method with an annual Chinese temperature data set and found better results than others. Hengl, Heuvelink, Tadić, and Pebesma (2012) presented an interesting idea of predicting LST in Croatia from MODIS LST images. They found that LST is a function of coordinate location, LULC, orography, precipitation, and industrial activities.

As identified in previous work (Bhattacharjee, Mitra, & Ghosh, 2014; Hengl et al., 2012), LST is highly influenced by the LULC distribution on the Earth's surface (Hulley, Veraverbeke, & Hook, 2014; Tsendbazar, De Bruin, Fritz, & Herold, 2015). Similarly, the temporal change in LULC is also correlated with the temporal dynamics of LST. Hence, this auxiliary knowledge of LULC can be incorporated into the spatio-temporal LST interpolation process to achieve better accuracy. Here we have assumed that the amount of residual change in LST at a given location is introduced due to the temporal change in LULC. We present a new spatio-temporal interpolation method by extending the spatial semantic kriging (SemK) method (Bhattacharjee & Chen, 2019b; Bhattacharjee & Ghosh, 2017), which is intended for LST mapping (Bhattacharjee & Ghosh, 2015b). The SemK-based interpolation method has been applied for the incorporation of LULC information into the prediction process of LST. LULC is considered as the semantic knowledge of the terrain, modeled by the processing of its ontology hierarchy (Bhattacharjee & Ghosh, 2014; Bhattacharjee, Prasad, Dwivedi, Dasgupta, & Ghosh, 2012). However, this method can be applied for any meteorological variable that is influenced by LULC dynamics. The temporal behavioral change of LULC is modeled as the change in "semantics" of the sample points and is incorporated into the prediction process. The proposed method belongs to the kriging family (Hui, Hu, Yevenyo, & Yu, 2016) and is referred to as spatio-temporal semantic kriging (ST-SemK). The major advantage of ST-SemK over spatial SemK is that the ST-SemK is capable of modeling time-series LST data; hence, it can facilitate forecasting. It is also better than SemK when sample density is low in the prediction time instance, but higher in the past. However, the performance of ST-SemK is highly dependent on the study region, and spatial autocorrelation, LULC–LST correlation, and the diversity of the LULC distribution (entropy) are the major contributing factors. Thus, while choosing the region of interest (RoI) for the empirical study, the LULC-based entropy analysis (Bhattacharjee et al., 2014), LULC–LST correlation study, and the performance evaluation lemmas (Bhattacharjee & Ghosh, 2015a) can be performed beforehand to check the suitability of ST-SemK for that particular RoI. The objectives of the proposed work are as follows:

- to analyze the behavior of temporal semantic change of the LULC distribution and incorporate this behavior to model the residual LST change for each sample point;
- to extend the traditional spatio-temporal autocorrelation model with the semantic LULC information and carry out the spatio-temporal prediction and forecasting of LST; and
- to compare the proposed method in terms of accuracy with some popular spatio-temporal interpolation methods.

2 | SPATIO-TEMPORAL SEMANTIC KRIGING

This section presents a detailed description of the proposed ST-SemK approach. The spatio-temporal interpolation methods estimate the variable value Z at the location x_0 in the prediction time instance t_0 from the known sample points (x_i, t_{a_i}), where ith year $a_i \in \{0, \cdots, -\infty\}$. In this study, Z is LST, but can be any other variables, such as carbon dioxide or solar-induced fluorescence. However, the difference in spatio-temporal prediction and forecasting lies

in the choice of the sampled data from different time instances. For prediction, the samples from the previous, present and following years can be chosen to gap-fill the raster images with missing pixels, line drops, cloud cover, etc. For forecasting, the sampled data should be chosen from the previous years only. Here we choose present and past data for prediction in 2015. Hence, $a_i \in \{0, \cdots, -10\}$. The theoretical interpolation estimator is modeled as follows:

$$\hat{Z}(x_0, t_0) = \sum_{i=1}^{N} w_i Z(x_i, t_{a_i}), \text{ such that } a_i \in \{0, \cdots, -\infty\} \tag{1}$$

where $\hat{Z}(x_0, t_0)$ is the predicted LST value at the prediction point x_0 at t_0, $Z(x_i, t_{a_i})$ is the measured LST value at the ith sampled location x_i at t_{a_i}, a_i is either a present or past time instance. w_i is the weight assigned to the sample point (x_i, t_{a_i}) which forms the weight vector **W** for N interpolation points. Therefore, the weight vector is the function of spatial lag h and temporal lag t. The existing approaches of spatio-temporal interpolation are broadly classified into two categories: (a) separable (purely spatial and purely temporal models); and (b) non-separable (inseparable space-time trade-off model) (Yang et al., 2013). The components of spatio-temporal analysis are depicted in Figure 2.

For the separable approach, the spatial and the temporal component are shown separately in the figure. However, for the latter approach, a non-separable component is shown as the "hypotenuse" of the spatio-temporal lag triangle. The proposed ST-SemK is implemented for both the approaches as separable ST-SemK (ST-SemK$_{Sep}$) and non-separable ST-SemK (ST-SemK$_{NSep}$). It follows the concept of time-forward kriging (Yang et al., 2013) along the temporal and spatial axes and extends the first law of geography (Tobler, 1970) in spatial, temporal, and semantic dimensions. It assumes that the change in LST is the function of coordinate location, time, and its semantics (LULC). The process flow diagram of ST-SemK is depicted in Figure 3. The input to this framework is the spatio-temporal LST data, classified (supervised) LULC data, and the ontology hierarchy of LULC classes depicted in Figure 4 (Bhattacharjee et al., 2014).

First, for each of the interpolation points in t_0, N sample points are collected from the present and past time instances. The traditional temporal, spatial semivariogram models (for the separable approach), or the spatio-temporal semivariogram model (for the non-separable approach) are built by following the concepts of the traditional OK method (Snepvangers, Heuvelink, & Huisman, 2003). In these traditional semivariogram models, the spatial semivariance ($\gamma(h)$), temporal semivariance ($\gamma(t)$), and spatio-temporal semivariance ($\gamma(h,t)$) of LST are plotted against the spatial lag (h), temporal lag (t), and the anisotropy ratio (α), respectively (Arslan, 2012). The first two models are implemented in the separable ST-SemK approach and the third in the non-separable approach. α is defined as (h/t), that is, spatial lag divided by temporal lag (for example, kilometers per year). This whole process of traditional semivariogram modeling is referred to as "Euclidean-distance-based spatio-temporal proximity" modeling in Figure 3.

To model the "spatio-temporal change in semantic proximity" in Figure 3, the semantic metrics of spatial SemK (Bhattacharjee et al., 2014) are extended into temporal or spatio-temporal dimensions (for the separable and
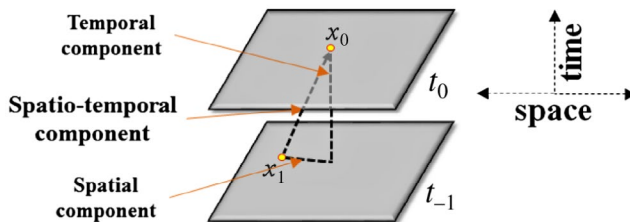
FIGURE 2 Simplified conceptual view of the components of spatio-temporal analysis, considering spatial, temporal and spatio-temporal lags between two LST sampled locations (x_0, t_0) and (x_1, t_{-1})
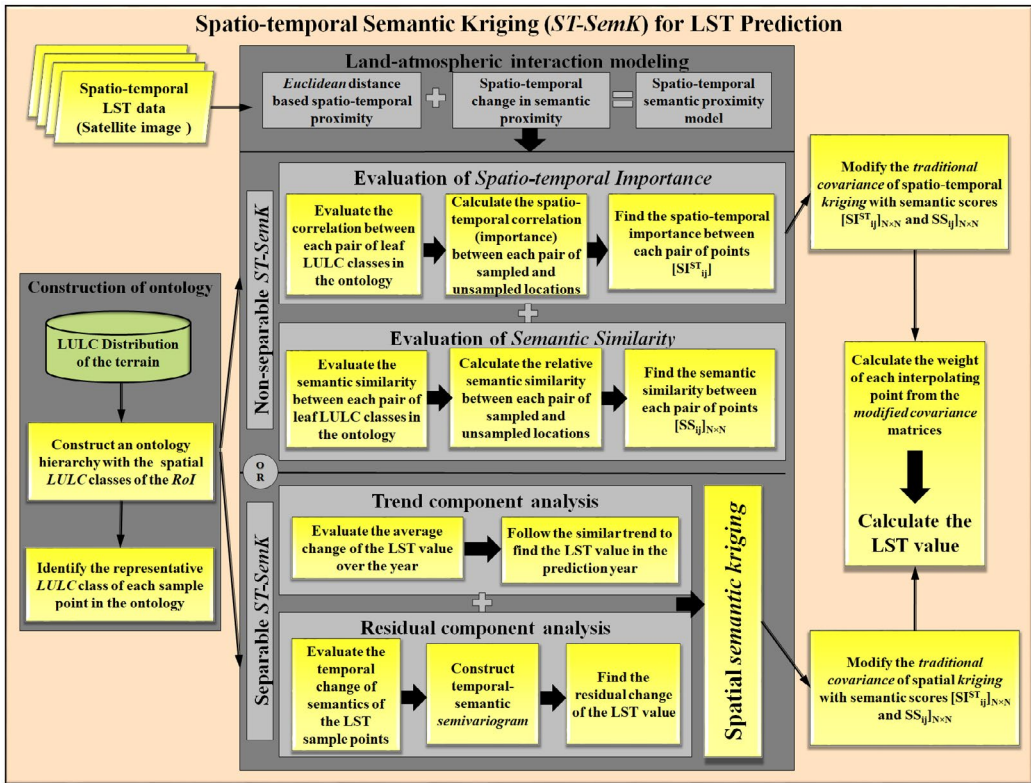
FIGURE 3 The ST-SemK framework considering semantic land–atmospheric interaction modeling as the core of the framework. The process flow of both separable and non-separable ST-SemK are presented in the block diagram
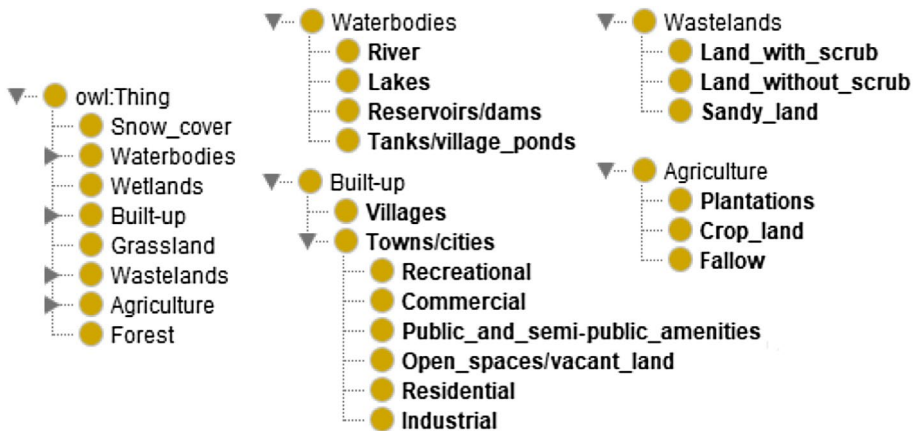

FIGURE 4 An example LULC ontology for LST interpolation. It is defined by the domain experts as LULC classification in Mendiratta et al. (2008)

non-separable approaches, respectively). Then the traditional semivariance and covariance scores are extended with these semantic metrics, which are used in the ST-SemK$_{NSep}$ and ST-SemK$_{Sep}$ approaches. These approaches are further elaborated in Sections 2.1 and 2.2, respectively.

2.1 | Non-separable ST-SemK

In ST-SemK$_{NSep}$, the space and time dimensions are made comparable to each other by defining an anisotropy ratio (α) to deal with different units and weighted distances along spatial and temporal dimensions. As reported in different literature, the separable model may suffer from some impractical hypotheses, which can be overcome by proper empirical estimation of the α value. Heuvelink & Griffith (2010) have explained the notion of traditional "sum metric" model for the spatio-temporal semivariogram modeling, which is adopted by ST-SemK$_{NSep}$. The semivariance (γ) of two sample points (x_i,t_{a_i}) and (x_j,t_{a_j}), h distance and t temporal lag apart, consists of three components as follows:

$$\gamma(h,t) = \gamma_{(x_i,x_j)}(h) + \gamma_{(t_{a_i},t_{a_j})}(t) + \gamma_{[(x_i,x_j),(t_{a_i},t_{a_j})]}(h,t) \tag{2}$$

where $\gamma_{(x_i,x_j)}(h)$ is a purely spatial semivariance, $\gamma_{(t_{a_i},t_{a_j})}(t)$ is a purely temporal semivariance, and $\gamma_{[(x_i,x_j),(t_{a_i},t_{a_j})]}(h,t)$ is a spatio-temporal semivariance where the two dimensions are made comparable with respect to α (see Figure 2). Evaluation of α between two dimensions is complex. This ratio can be empirically evaluated from the available sampled locations of the RoI. To make the proposed ST-SemK$_{NSep}$ process more pragmatic, the local anisotropy ratios can be calculated. This is because the α may vary (different semivariogram sills in different directions) with the group of sampled locations considered (Hengl et al., 2012; Snepvangers et al., 2003).

If Cov$^{ST}(Z_i,Z_j)$ is the spatio-temporal covariance between the sample points (x_i,t_{a_i}) and (x_j,t_{a_j}) with distance lag h and temporal lag t, and the space-time anisotropy ratio α, then Cov$^{ST}(Z_i,Z_j)$ is modeled as follows:

$$\begin{aligned} \text{Cov}^{ST}(Z_i,Z_j) &= \text{Cov}(h,t) \\ &= \text{Cov}(h) + \text{Cov}(t) + \text{Cov}\left(\sqrt{h^2+(\alpha t^2)}\right) \\ &= \leq(\gamma_{(x_i,x_j)}(h) + \gamma_{(t_{a_i},t_{a_j})}(t) + \gamma_{[(x_i,x_j),(t_{a_i},t_{a_j})]}(h,t)) \end{aligned} \tag{3}$$

Equation (3) indicates that the covariance is a function of semivariance which can be evaluated with the experimental semivariogram models. For ST-SemK$_{NSep}$, the traditional spatio-temporal covariance matrix ($\mathbb{C}^{ST}_{ij\ [N\times N]}$) and distance matrix ($\mathbb{D}^{ST}_{0i\ [N\times 1]}$) are given below, where Var$^{ST}(Z_i)$ denotes the covariance of (x_i,t_{a_i}) with itself (i.e., variance):

$$\mathbb{C}^{ST}_{ij\ [N\times N]} = \begin{bmatrix} \text{Var}^{ST}(Z_1) & \text{Cov}^{ST}(Z_1,Z_2) & \cdots & \text{Cov}^{ST}(Z_1,Z_N) \\ \text{Cov}^{ST}(Z_2,Z_1) & \text{Var}^{ST}(Z_2) & \cdots & \text{Cov}^{ST}(Z_2,Z_N) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}^{ST}(Z_N,Z_1) & \text{Cov}^{ST}(Z_N,Z_2) & \cdots & \text{Var}^{ST}(Z_N) \end{bmatrix}, \mathbb{D}^{ST}_{0i\ [N\times 1]} = \begin{bmatrix} \text{Cov}^{ST}(Z_0,Z_1) \\ \text{Cov}^{ST}(Z_0,Z_2) \\ \vdots \\ \text{Cov}^{ST}(Z_0,Z_N) \end{bmatrix} \tag{4}$$

2.1.1 | Evaluation of spatio-temporal semantic dependence

To evaluate semantic dependence in the non-separable ST-SemK approach, the two spatial semantic metrics of SemK (Bhattacharjee et al., 2014), semantic similarity and spatial importance, are extended in the spatio-temporal domain (Bhattacharjee & Ghosh, 2017). As the semantic relations between the LULC classes in the ontology hierarchy remain unchanged over time, the structure of the ontology remind constant over time. The ontology hierarchy consists of all possible LULC classes of every time instance and is exhaustive. The semantic similarity metric is dependent on the structure of the ontology. As the temporal lags among the sampled locations do not affect this structure, this metric does not need modifications for its spatio-temporal change analysis. Therefore,

the process of evaluating the spatio-temporal semantic similarity will be the same as that of the spatial semantic similarity evaluation process in Bhattacharjee et al. (2014).

However, as the spatio-temporal importance evaluation process deals with correlation analysis with sampled LST data, the temporal lag is significant here. This is because the same locations on the Earth's surface may have different LST measures at two distinct time instances. To evaluate the spatio-temporal importance, the whole RoI is divided into k random zones. Given two sample points (x_i, t_{a_i}) and (x_j, t_{a_j}) with their representative LULCs f_i and f_j, k pairs of sample points are chosen for both f_i and f_j, but from different time instances, t_{a_i} and t_{a_j}, respectively. Therefore, the extended spatial importance metric is termed spatio-temporal importance (SI^{ST}), and for (x_i, t_{a_i}) and (x_j, t_{a_j}) it is modeled as follows:

$$SI_{ij}^{ST} = \mathrm{Corr}_{LST}(f_i^{t_{a_i}}, f_j^{t_{a_j}}) = \frac{\sum_{m=1}^{k} (Z^{t_{a_i}}(f_{i_m}) - \overline{Z^{t_{a_i}}(f_i)})(Z^{t_{a_j}}(f_{j_m}) - \overline{Z^{t_{a_j}}(f_j)})}{\sqrt{\sum_{m=1}^{k} (Z^{t_{a_i}}(f_{i_m}) - \overline{Z^{t_{a_i}}(f_i)})^2 \sum_{m=1}^{k} (Z^{t_{a_j}}(f_{j_m}) - \overline{Z^{t_{a_j}}(f_j)})^2}} \tag{5}$$

where $\overline{Z^{t_{a_p}}(f_{p_q})}$ is the LST value (Z) of the qth sample point representing the LULC class f_p at the time instance t_{a_p}. $\overline{Z^{t_{a_p}}(f_p)}$ is the average of the LST values of the LULC class f_p over the k sample points at the time instance t_{a_p}. Being a correlation coefficient score, it lies in $[-1,1]$. It is further normalized to a positive range to avoid the negative mapping of the covariances. Now, the total spatio-temporal semantic score for the pair (f_i and f_j) is given as SIS_{ij}^{ST}, which is equal to $SI_{ij}^{ST} \cdot SS_{ij}$.

In spatio-temporal analysis, the semivariance scores are plotted against the anisotropy ratio α to build the experimental semivariogram model in the spatio-temporal domain. An example of an experimental semivariogram model is presented in Figure 5 to show that the semivariance scores vary with respect to the semantic scores, and is simulated using some randomly chosen locations of the spatial region of Kolkata. In this figure, the Y-axis represents the $\gamma_{ij}^{ST-SemK}$ scores against α on the X-axis, and the corresponding fitted semivariograms are shown. Three different semivariograms are shown in solid lines for three semantic scores (0.25, 0.5, and 0.75). The violet dotted line represents the semivariograms for the baseline ST-OK method. The experimental details of these models are
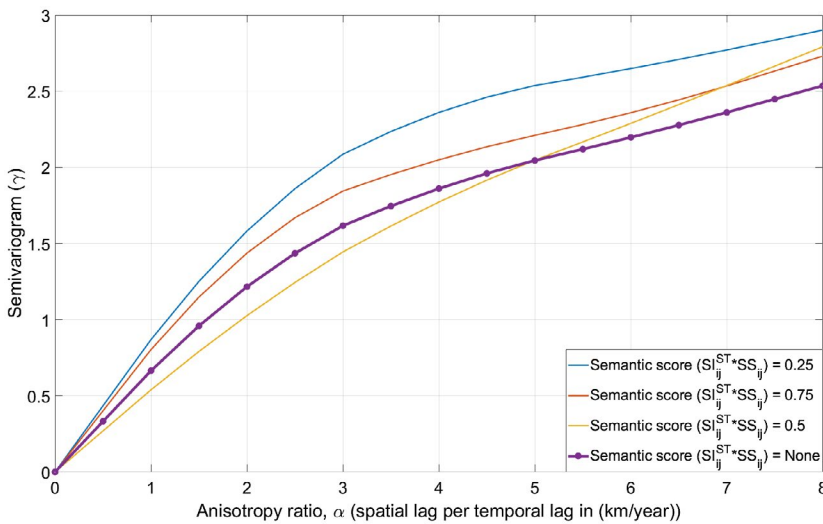
FIGURE 5 Example of experimental spatio-temporal semivariograms of LST for different semantic scores ($SI_{ij}^{ST} \cdot SS_{ij}$) (RoI: Kolkata). Here, the three example scores are: 0.25 (blue line), 0.5 (yellow line), 0.75 (red line). Different semivariogram behavior can be observed for different scores. The semivariogram with semantic score 1, which is similar to "None" (violet dotted line), stands for the ST-OK method where all the LULCs are considered to be identical

discussed further in Section 3.3. Here our goal is to update the traditional spatio-temporal covariance scores with the semantic scores. In ST-SemK, the traditional semivariance score ($\gamma(\alpha)$) is revised using the semantic score SIS_{ij}^{ST}, and the revised semivariance score $\gamma_{ij}^{ST-SemK}$ is given as follows:

$$\gamma_{ij}^{ST-SemK} = \gamma(\alpha)^{ST-SemK} = \frac{\gamma(\alpha)}{SIS_{ij}^{ST}} \tag{6}$$

The covariance scores are obtained from the semivariogram model as follows: $Cov_{ij} = Sill - \gamma_{ij}$ (Esri, 2019; Wagner, 2003). Another alternative approach is to update the covariance scores using semantic scores as both are the similarity metrics. The semantic covariance matrix $\mathbb{C}_{ij}^{SemK}$ and the semantic distance matrix $\mathbb{D}_{0i}^{SemK}$ are given as follows:

$$\mathbb{C}_{ij}^{ST-SemK} = [Cov_{ij}^{ST-SemK}]_{N \times N}, \ \mathbb{D}_{0i}^{ST-SemK} = [Cov_{0i}^{ST-SemK}]_{N \times 1} \tag{7}$$

Considering ST-OK as the base method, the weight vector of ST-SemK ($\mathbb{W}^{ST-SemK}$) is modeled as follows: $\mathbb{W}^{ST-SemK} = [\mathbb{C}_{ij}^{ST-SemK}]^{-1}[[\mathbb{D}_{0i}^{ST-SemK}] - \lambda^{ST-SemK}\mathbb{1}]$, where $\lambda^{ST-SemK}$ is the Lagrange multiplier of ST-SemK.

2.2 | Separable ST-SemK

ST-SemK$_{Sep}$ facilitates spatio-temporal prediction and forecasting of LST. It attempts to estimate the trend component and the residual component over the temporal axis. These two components approximate each past LST interpolating point (x_i, t_{a_i}) to the prediction time instance t_0 as (x_i, t_0). This is a purely temporal prediction process. Then the spatial semantic kriging is carried out with all the approximated (x_i, t_0)s to estimate the LST value at the prediction point (x_0, t_0), which is a purely spatial process. Thus, the properties of ST-SemK$_{Sep}$ can be summarized as follows.

- The trend component captures the average behavioral change in LST over the temporal dimension.
- The residual component deals with the change in LST that is due to the temporal changes of the semantics or LULC distribution of the terrain. This change is captured by building a three-dimensional (3D) temporal semantic semivariogram model of LST.
- With these two components, once all the interpolating points are approximated to the prediction time instance t_0, the spatial semantic kriging is carried out by building a 3D spatio-semantic semivariogram of LST at t_0 using all the approximated (x_i, t_0)s.

The general spatio-temporal estimation equation of ST-SemK$_{Sep}$ is given by:

$$\hat{Z}(x,t) = m(x,t) + \epsilon(x,t) \tag{8}$$

where $\hat{Z}(x,t)$ is the predicted LST value at each distinct (x,t), $m(x,t)$ is the trend component (i.e., the deterministic part of the random variable), and $\epsilon(x,t)$ is the corresponding residual part. These two components are described further.

2.2.1 | Trend component analysis

Many variants of the kriging method have been reported in the literature, depending on how $m(x,t)$ is modeled. The default approach is to assume that the trend is an unknown constant mean over space and time. As both OK and spatial SemK adopt this notion in the purely spatial domain, the proposed ST-SemK can be extended further in the temporal domain by adopting the same notion to model $m(x,t)$. However, the assumption of constant mean over the temporal axis is not a very realistic approach as it assumes that the same LST trend is followed at every

time instance. Furthermore, the temporal uncertainties associated with the de-trending process are not taken into account in some approaches, such as ST-SK (Snepvangers et al., 2003). The existing methods do not consider the temporal change in semantics, which indirectly affects the change in LST beyond its average trend. The proposed ST-SemK method models the trend component, which is a function of location and time (x,t). In addition, the residual $\varepsilon(x,t)$ models the additional change in LST due to the semantic change in the terrain.

In this study the $m(x,t)$ analysis is carried out for the individual LST sample points. A predefined radius is considered against each past sample point (x_i,t_{a_i}). The N_{trend} additional sample points are considered within that radius and their LST value is measured at every time instance, starting from t_{a_i} to t_{-1} in the forward direction. A trend line then approximates this temporal behavior in the prediction year or beyond, which may also facilitate forecasting (see Figure 6). Thus, the trend component $m(x_i,t_0)$ of the interpolating point (x_i,t_{a_i}), approximated in the prediction time instance t_0 is given by:

$$m(x_i,t_0) = \frac{\sum_{q=a_i}^{-1} \frac{\sum_{p=1}^{N_{trend}} Z(x_p,t_q)}{N_{trend}}}{|a_i|} \tag{9}$$

where $a_i \in \{0, \cdots, -\infty\}$.

Figure 6 shows an example of average LST behavioral change at a particular location in the Kolkata region over the years 2005 to 2014. The Y-axis shows the average LST for the corresponding year on the X-axis. The graph in Figure 6a shows the LST trend over the years, including the average measured LST in 2015. Figure 6b shows the approximated average LST in the year 2015 by following the trend line of the past data, estimated by Equation (9).

2.2.2 | Residual component analysis

The spatio-temporal 2D semivariogram between two interpolating points (x_i,t_{a_i}) and (x_j,t_{a_j}) is modeled by $\gamma((x_i,t_{a_i}),(x_j,t_{a_j}))$. It captures the additional behavioral change as the variance between random fields between the pair of locations. Considering the respective stationarity assumptions in both space and time dimensions, the traditional semivariance γ for spatio-temporal analysis is a function of spatial lag h and the temporal lag t between the pair of locations (x_i,t_{a_i}) and (x_j,t_{a_j}). It is given by:

$$\gamma(h,t) = \frac{1}{2M} \sum_{i=1}^{M} [\epsilon(x_i,t_{a_i}) - \epsilon(x_i+h,t_{a_i}+t)]^2 \tag{10}$$
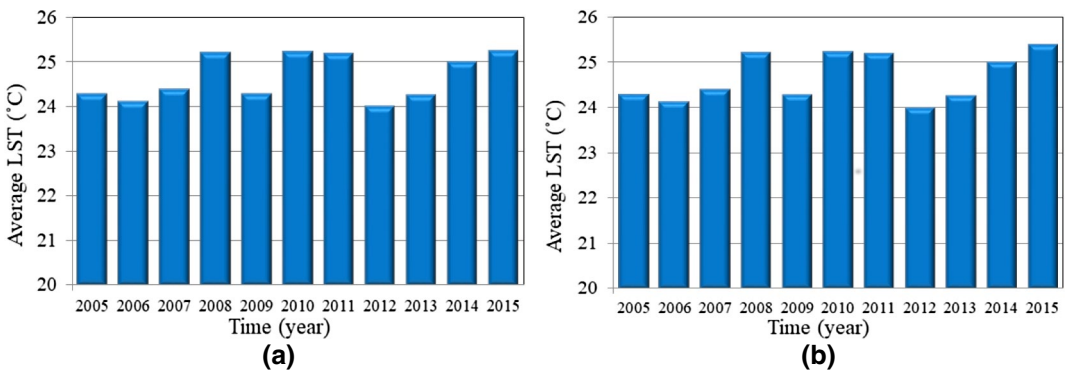
FIGURE 6 Average temporal change of LST at a particular location (RoI: Kolkata): (a) Temporal change of LST from 2005 to 2015; and (b) Predicted LST in 2015 considering the linear LST trend from 2005 to 2014

where M is the pair of sampled locations within spatial lag h and temporal lag t.

Since ST-SemK first attempts to approximate each interpolating point (x_i, t_{a_i}) in the prediction time instance t_0 as (x_i, t_0), the residual $\varepsilon(x,t)$ is modeled by a purely temporal semivariogram with spatial lag $h=0$. Considering 0 and t lags in the spatial and the temporal axes respectively in Equation (10), the purely temporal semivariogram model of ST-SemK is given by:

$$\gamma(0,t) = \frac{1}{2M} \sum_{i=1}^{M} [\varepsilon(x_i, t_{a_i}) - \varepsilon(x_i, t_{a_i} + t)]^2 \tag{11}$$

As identified by ST-SemK, the representative LULC of a particular location may change over time. The residual $\varepsilon(x,t)$ of LST should also capture this change in semantics of the terrain through temporal semivariance modeling. Therefore, the temporal semantic semivariance, $\gamma(0,t,\Delta sem_{ij})$, is given by:

$$\gamma(0,t,\Delta sem_{ij}) = \frac{1}{2M} \sum_{i=1}^{M} [\varepsilon(x_i, t_{a_i}, f_i) - \varepsilon(x_i, t_{a_i} + t, f_j)]^2 \tag{12}$$

where Δsem_{ij} is the semantic change of a sample point x_i, due to its representative LULC change from f_i to f_j over the temporal lag t, and M is the pair of sampled locations that experienced this LULC change.

This temporal change Δsem_{ij} is evaluated by analyzing the ontology in Figure 4 with two metrics: temporal semantic similarity and temporal importance. The evaluation processes of both the metrics are different in ST-SemK$_{Sep}$, compared to the ST-SemK$_{NSep}$ approach. As the structure of the ontology remains static over time, the change with respect to the temporal semantic similarity metric can be evaluated from the ontology itself. For this metric, the temporal change in semantics (ΔSS_{ij}^t) of a particular sample point x_i, represented by the LULC classes f_i and f_j at two different time instances t_{a_i} and t_{a_j}, is modeled by:

$$\Delta SS_{ij}^t = 1 - \frac{m_i / |f_i^{t_{a_i}}| + m_j / |f_j^{t_{a_j}}|}{2} \tag{13}$$

Here, $|f_i^{t_{a_i}}|$ and $|f_j^{t_{a_j}}|$ are the total number of concepts in the paths of f_i and f_j in the ontology, staring from the root LULC. The m_i and m_j represent the number of matching concepts in the paths of f_i and f_j in the ontology ($m_i = m_j$).

However, the temporal importance between any pair of LULC classes evolves over time. Thus, the change with respect to this metric is the change in correlation of a particular point x_i, or its representative LULC classes ($f_i^{t_{a_i}}$ and $f_j^{t_{a_j}}$) over the time interval t. If the representative LULC class of the point x_i changes from f_i to f_j over the temporal lag t, the temporal correlation score, SI_{ij}^t, is modeled as $SI_{ij}^t = SI_{ij}^{ST}$, where SI_{ij}^{ST} is defined in Equation (5). The semantic correlation score of the location x_i is maximal when it is represented by the same LULC class over time interval t. This score is referred to as SI_{ii}^t and is modeled by Equation (5). Hence, the change in the temporal importance metric for a given sample point x_i over the temporal lag t is modeled as:

$$\Delta SI_{ij}^t = |SI_{ii}^t - SI_{ij}^t| \tag{14}$$

The total semantic change (Δsem) is then evaluated over time interval t if both the scores are non-zero (Bhattacharjee & Ghosh, 2015a). Once the temporal change in semantics is evaluated through the temporal semantic similarity and the temporal importance metrics, a 3D temporal semantic semivariogram is modeled further. The temporal semantic semivariogram captures the residual change in LST (ΔZ) over time and the change in semantics. An example temporal semantic semivariogram for the Kolkata study region is shown in Figure 7, which is simulated using a few selected locations. The X-axis shows the inter-annual temporal lag (t) and the Y-axis shows the semantic change (Δsem). The Z-axis shows the corresponding temporal semantic semivariance scores
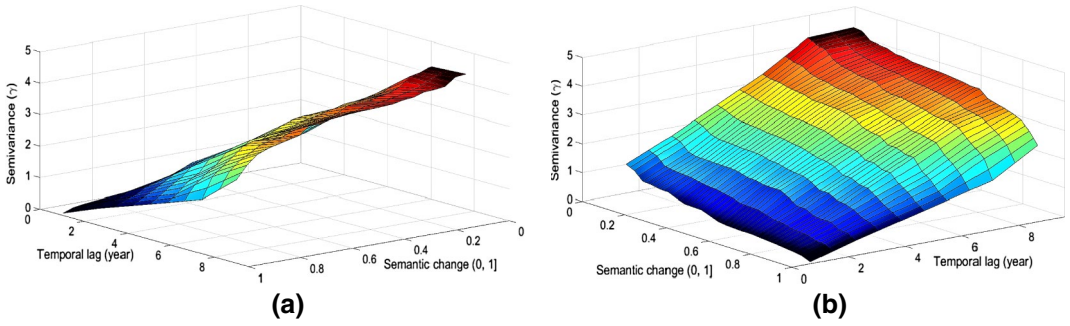
FIGURE 7 Example of 3D temporal semantic semivariogram of LST considering the temporal lag on the X-axis, the semantic change on the Y-axis, and the semivariance (γ) on the Z-axis (RoI: Kolkata). Two views of the same semivariogram show how the semivariance score is changing with both time and semantic (LULC) metrics. The semivariance score increases with the increment of temporal lag and the decrement of semantic score

($\gamma(0,t,\Delta sem_{ij})$), varying with respect to both time and semantics of the terrain. The details of experimental specifications are discussed in Section 3.3.

Following the characteristic of the temporal semantic semivariogram, the residual $\varepsilon(x,t)$ of LST can be approximated over the temporal axis for future time instances to facilitate forecasting. Finally, the LST value of each interpolating point $Z(x_i,t_{a_i})$ is approximated at the prediction time instance as $\hat{Z}(x_i,t_0) = m(x_i,t_0) + \epsilon(x_i,t_0)$ (see Equation 8). Once all the interpolatied points are approximated to the same prediction time instance t_0, the spatial SemK is carried out to predict $\hat{Z}(x_0,t_0)$ with the approximated $\hat{Z}(x_i,t_0)$s. This models 3D spatio-semantic semivariograms to capture the change in LST with respect to spatial lag and semantics. The spatial semivariance, $\gamma(h)$, and the spatio-semantic semivariance, $\gamma(h)^{SemK}$, are given by:

$$\gamma(h) = \frac{1}{2M} \sum_{i=1}^{M} [Z(x_i,t_0) - Z(x_i+h,t_0)]^2, \quad \gamma(h)^{SemK} = \frac{\gamma(h)}{SIS_{ij}} \tag{15}$$

where M is the pair of sampled locations, (x_i,t_0) and (x_i+h,t_0), that are h spatial lag and 0 temporal lag apart, Z is the LST value of the corresponding point, and SIS_{ij} is the semantic score for the spatial analysis (Bhattacharjee et al., 2014). The covariance scores can also be modified directly by the semantic scores as both are similarity metrics. This model is further used to determine the experimental spatio-semantic semivariogram, the weights vector, and the LST value at the prediction process. The detailed mathematical modeling of spatial SemK is discussed in Bhattacharjee et al. (2014).

The experimental spatio-semantic semivariograms for two study regions, Kolkata, India and Dallas, TX, are simulated using a few randomly selected locations and depicted together in Figure 8. Here, the X-axis shows the spatial lag (h) between sample points at the prediction time instance t_0 and the Y-axis is the same as that of the temporal semivariogram model (see Figure 7), representing the semantic change (Δsem). The semantic change is modeled with respect to the semantic similarity and spatial importance metrics which lie in (0,1] (Bhattacharjee et al., 2014). The Z-axis shows the spatial semivariance score ($\gamma(h)$), varying with respect to both space and semantics of the terrain. The experimental details are specified in Section 3.3.

3 | EMPIRICAL EXPERIMENT

The details of the empirical experiment are given in this section. The study regions chosen, their terrestrial and climatic details, the specifications of the data sets considered, and empirical details of the interpolation processes are discussed.
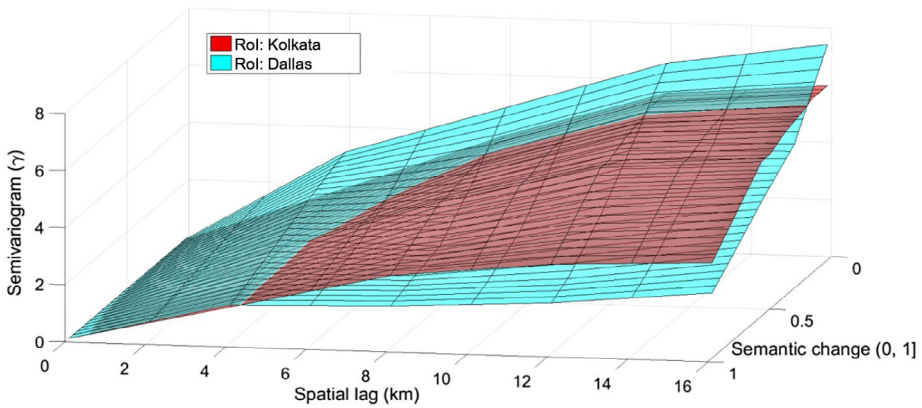
FIGURE 8 Example of 3D spatio-semantic semivariogram of LST considering the spatial lag on the *X*-axis, the semantic change on the *Y*-axis, and the semivariance (γ) on the *Z*-axis for both the RoIs: Kolkata and Dallas. The semivariance score increases with the increment of spatial lag and the decrement of semantic score

3.1 | Data specifications

An empirical experiment was carried out with the derived satellite images of LST to check the performance of the proposed method and compare it with existing popular interpolation methods. *Landsat 7* ETM+ satellite images (http://landsat.usgs.gov/) are considered for two RoIs, Kolkata and Dallas. The satellite images consist of seven spectral bands with spatial resolutions 30 m (for bands 1–5, 7) and 60 m (for band 6). The LST is derived from this band information in degrees Celsius (°C). We considered data for the years of 2005–2015, with prediction for the year 2015 being carried out by considering the time series input to the interpolation framework.

3.2 | Study regions

The Kolkata study region is located in the eastern part of India (central coordinate 22°34′N 88°22′E) (see Figure 9). The annual mean temperature of Kolkata is 26.8°C and the monthly mean temperature varies in the range 19–30°C. The Dallas study region is the largest urban center and the fourth most populous metropolitan area in the U.S. (central coordinate 32°46′N 96°47′W). This metropolitan area is shown in Figure 10. Dallas has a humid and hot subtropical climate with a mean temperature of about 39°C during summer and heat-humidity indexes soaring as high as 47°C. Both of the RoIs have very rich and diverse types of LULC distributions. Some common LULC classes for both the RoIs are: residential areas, commercial areas, waterbodies, vegetation, croplands, and industrial areas.

3.3 | Experimental specifications

To model the semantic properties of the terrain, the raw satellite images underwent classification separately. As the properties of every LULC class may change over time, different signature sets should be considered for different time instances. Domain expert knowledge was considered in order to identify the optimal classification properties. For each of the RoIs, five different zones were identified for further analysis. These zones are shown in Figures 9 and 10. The empirical study has the following specifications.

- The total number of pixels in each zone of Kolkata is approximately 205 × 241. A minimum of 26 × 27 sampled LST pixels are randomly (uniformly random) chosen from each zone from 10 years of time series data, which is approximately one seventieth of the total number of pixels.

Fast color composite map of Kolkata (November 2015)
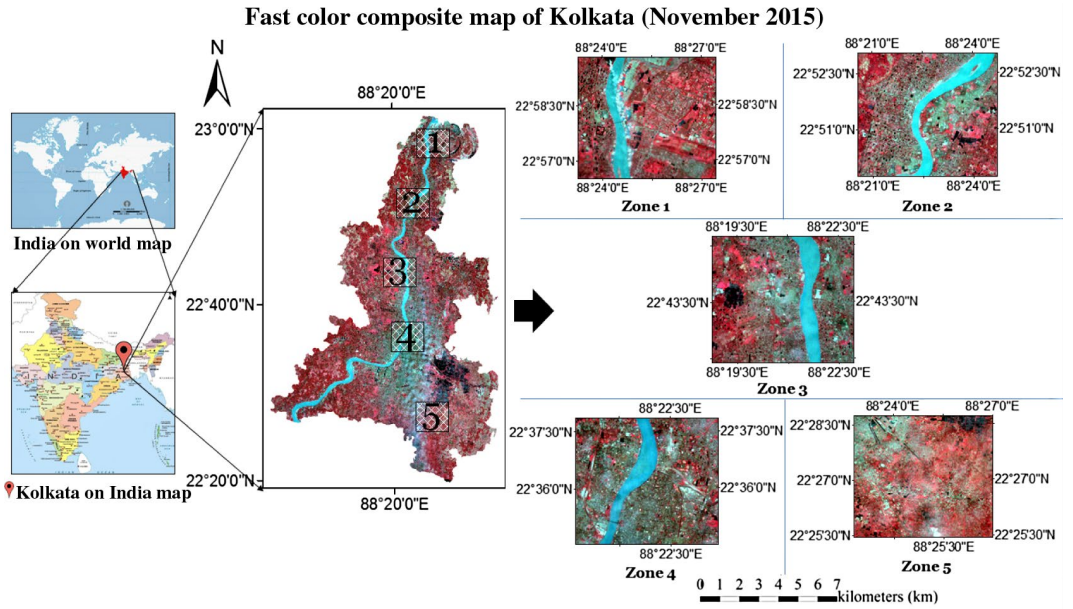
FIGURE 9 Spatial zones of Kolkata selected for analysis

Fast color composite map of Dallas (November 2015)
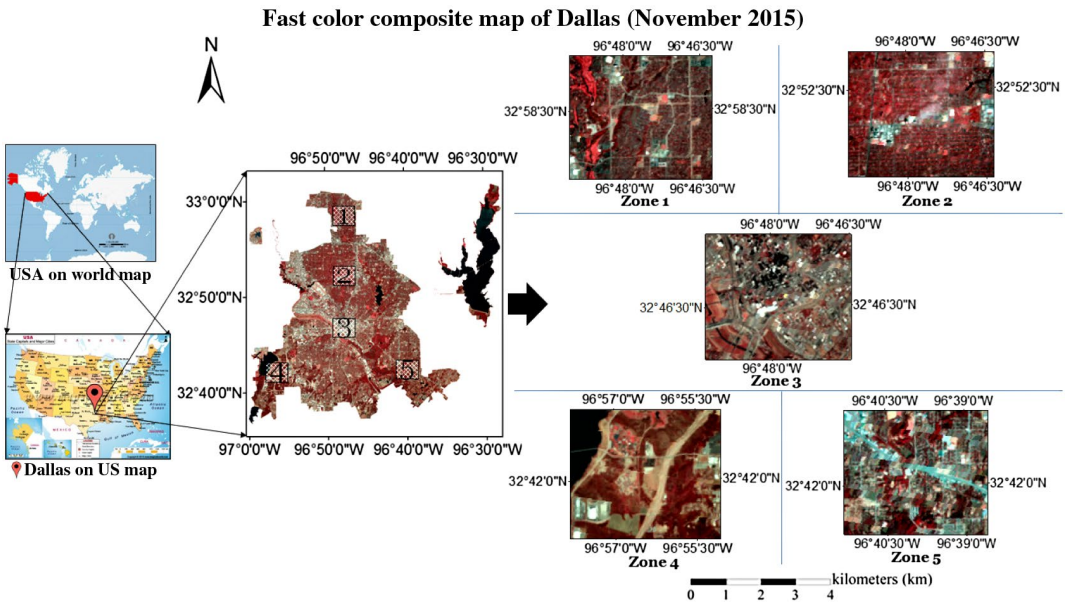
FIGURE 10 Spatial zones of Dallas selected for analysis

- A predefined radius is considered against each prediction point for the selection of interpolation points.
- For ST-SemK$_{NSep}$, the space-time anisotropy ratio α is modeled for each of the zones separately. An example spatio-temporal semivariogram model for the spatial region Kolkata was shown in Figure 5. Here, the semantic scores are chosen in the same interval of 0.25 within the range (0,1]. Unique behavior is observed for every semantic score, which also supports the fact that different change in LULC results in diverse temporal LST changes. All the semantic semivariograms are normalized to the original range of corresponding traditional

semivariograms. Another example experimental semivariogram plot is also reported for the ST-OK method (violet dotted line) in the same figure. This method does not differentiate between the sample points in terms of their representative LULC classes, hence the semantic score between any pair of sample points is always 1, which is the same as no semantics or "None."

- The trend component of ST-SemK$_{Sep}$ chooses N_{trend} LST sample points against each prediction point within a predefined radius. Figure 6 shows the average LST behavioral change at a particular location in the Kolkata region over the years 2005–2014. It is modeled by taking $N_{trend} = 30$ within a predefined radius.

- An example temporal semantic semivariogram for the Kolkata study region is presented in Figure 7, taking a few random sample points of this region . In this model, 11 years of data (2005–2015, inter-annual) are considered, taking the period from mid-October to mid-November to avoid seasonal effects. Hence, the inter-annual temporal lag (Δt) varies from 0 to 10 (on the X-axis). The semantic scores with respect to the change in semantic similarity and temporal importance metrics (Δsem) vary in (0,1] (on the Y-axis). The temporal semantic semivariance scores are plotted on the Z-axis. The characteristic of this semivariogram in Figure 7 is that the model varies with respect to both time and semantic scores. With more temporal lag, the semivariance score increases up to a certain point and then becomes stable. A similar trend is observed for the discrete semantic scores. However, the changing pattern is different for different scores.

- Example experimental spatio-semantic semivariograms of spatial SemK for both the study regions in 2015 are presented in Figure 8 together. This considers the satellite LST data from 2015. In this figure, both the RoIs can also be compared through their semivariograms. For example, at 8 km spatial lag and 0.5 semantic score, the semivariance scores for Kolkata and Dallas are 3.66 and 2.42, respectively. Similarly, for identical LULC classes (semantic score = 1) at 16 km lag, the semivariance scores are 3.36 and 2.30 for Kolkata and Dallas, respectively. These semivariograms are exponential and linear in nature for Kolkata and Dallas, respectively. It is observed for both the models that the semivariance score increases with the increase in spatial lag. These scores also vary with respect to the semantic scores. A similar trend is observed for every semantic score, but the change is different for different scores. The curve becomes stable beyond the range of spatial autocorrelation (Bhattacharjee et al., 2014).

4 | RESULTS AND DISCUSSION

This section compares the ST-SemK method with some existing interpolation methods. The error measurement metrics, their significance, and a discussion on the comparison study are presented.

4.1 | Comparison study

For the comparison study, the ST-SemK (both ST-SemK$_{NSep}$ and ST-SemK$_{Sep}$) method is compared with some popular spatio-temporal interpolation methods (as reported by Li & Heap, 2011), such as splines with spatio-temporal data (ST-SP), nearest neighbors with spatio-temporal data (ST-NN), IDW with spatio-temporal data (ST-IDW) and OK with with spatio-temporal data (ST-OK) (Li, 2008). These four methods do not include the contextual LULC information in the spatio-temporal interpolation process. Four metrics are calculated by comparing the predicted and measured LST: MAE, RMSE (Bhattacharjee et al., 2014; Li, 2008), mean bias (MB) which represents the mean of the difference between the predicted and measured LST, and standard deviation of error (SDE) where "error" represents the difference between the predicted and measured LST. For these metrics, the measured pixel values ($Z(x_i)$) are compared with the corresponding predicted pixels ($\hat{Z}(x_i)$) of the validation samples, and the average error is calculated (for MAE and MB). However, the RMSE and SDE both represent the deviation of the LST predictor and error. The evaluation criteria and the significance of each metric are discussed in Li (2008). The basic idea of

the error metrics is to check whether the predicted values are similar to the measured values, hence the difference is less. Therefore, the lower the MAE and RMSE values, the better the model (Li, 2008). For a better prediction model, the MB should be closer to zero and the SDE should be low. For pictorial comparison of the predicted LST images with respect to the measured LST images, the peak signal-to-noise ratio (PSNR; Bhattacharjee & Ghosh, 2017) is reported for each prediction method. The PSNR (in decibels) is a standard metric for predicted image analysis, which is a ratio between the maximum possible power of a signal and the power of corrupting noise. A higher PSNR indicates a better prediction model. The mathematical formulations of MAE, RMSE, MB, SDE and PSNR are defined as follows: in the Equations (16)–(20) respectively, where $Z(x_i)$ is the measured pixel value, $\hat{Z}(x_i)$ is the predicted pixel value by the respective model, N is the number of sampled locations, μ_u is the mean of the error surface and MAX_I is the maximum pixel value of the measured image:

$$\text{Mean absolute error (MAE)}_{LST} = \frac{\sum_{i=1}^{N} |Z(x_i) - \hat{Z}(x_i)|}{N} \, (°C) \tag{16}$$

$$\text{Root mean square error (RMSE)}_{LST} = \sqrt{\frac{\sum_{i=1}^{N} [Z(x_i) - \hat{Z}(x_i)]^2}{N}} \, (°C) \tag{17}$$

$$\text{Mean bias (MB)}_{LST} = \frac{\sum_{i=1}^{N} [Z(x_i) - \hat{Z}(x_i)]}{N} \, (°C) \tag{18}$$

$$\text{Standard deviation of error (SDE)}_{LST} = \sqrt{\frac{\sum_{i=1}^{N} [(Z(x_i) - \hat{Z}(x_i)) - \mu_u]^2}{N - 1}} \, (°C) \tag{19}$$

$$\text{Peak signal- to- noise ratio (PSNR)}_{LST} = 20 log_{10} \left(\frac{MAX_I}{RMSE} \right) (dB) \tag{20}$$

Here $Z(x_i)$ is the measured pixel value, $\hat{Z}(x_i)$ is the pixel value predicted by the respective model, N is the number of sampled locations, μ_u is the mean of the error surface, and MAX_I is the maximum pixel value of the measured image.

Tables 1 and 2 present the comparison of different spatio-temporal interpolation approaches with the ST-SemK for Kolkata and Dallas respectively, in terms of MAE, RMSE, MB, and SDE. The mapping images (measured and predicted) of LST for five zones of each of the RoIs are depicted in Tables 3 and 4. Each predicted image is compared with the measured image of the respective zone.

4.2 | Discussion

From Tables 1 and 2, it is evident that the proposed ST-SemK performs better than other methods for both the RoIs. Though ST-SP, ST-NN, ST-IDW, and ST-OK are popular interpolation methods in the existing literature (Li, 2008; Li & Heap, 2011), they do not consider the LULC information of the terrain explicitly for LST change modeling. It is also evident from the results that in most cases, these existing interpolation methods give similar and comparable results for LST prediction in terms of MAE and RMSE. The differences exist in these methods based on how they model the spatial autocorrelation/dependency among the sample points for LST analysis. Their performance also depends on the spatial region of interest and the diversity of the terrestrial distribution (Bhattacharjee & Ghosh, 2015a).

It is also observed that for every zone, the most accurate interpolation method (excluding ST-SemK) is different. This is the reason for modeling the ST-SemK method as a generic framework to extend any interpolation method with semantic analysis. Again, both ST-SemK$_{NSep}$ and ST-SemK$_{Sep}$ yield comparable results in most cases

TABLE 1 Comparison of predicted LST with four error measurement metrics—MAE, RMSE, MB, and SDE (in °C)—for the Kolkata region. Lower MAE, RMSE, and MB closer to zero represent a better model and vice versa

RoI	Error measures						
Kolkata	Zone	ST-SP	ST-NN	ST-IDW	ST-OK	ST-SemK$_{NSep}$	ST-SemK$_{Sep}$
	Mean absolute error (MAE, °C)						
	Zone 1	0.480	0.475	0.504	0.476	0.290	0.315
	Zone 2	0.499	0.488	0.512	0.485	0.303	0.295
	Zone 3	0.530	0.517	0.557	0.606	0.483	0.470
	Zone 4	0.545	0.542	0.768	0.700	0.410	0.420
	Zone 5	0.367	0.350	0.351	0.408	0.330	0.322
	Root mean square error (RMSE, °C)						
	Zone 1	0.439	0.431	0.491	0.432	0.211	0.248
	Zone 2	0.504	0.504	0.580	0.493	0.279	0.249
	Zone 3	0.589	0.575	0.633	0.648	0.405	0.372
	Zone 4	0.539	0.541	0.638	0.571	0.275	0.308
	Zone 5	0.233	0.212	0.215	0.282	0.164	0.147
	Mean bias (MB, °C)						
	Zone 1	0.027	0.022	0.021	0.021	−0.002	−0.003
	Zone 2	−0.006	−0.004	−0.002	−0.004	0.001	0.001
	Zone 3	0.026	0.025	0.022	0.020	−0.007	0.003
	Zone 4	0.002	0.005	0.008	0.002	0.001	0.001
	Zone 5	0.009	0.008	0.007	0.007	0.001	0.001
	Standard deviation of error (SDE, °C)						
	Zone 1	0.457	0.453	0.487	0.453	0.425	0.421
	Zone 2	0.513	0.516	0.564	0.508	0.490	0.512
	Zone 3	0.555	0.556	0.589	0.729	0.507	0.555
	Zone 4	0.493	0.497	0.538	0.725	0.484	0.484
	Zone 5	0.314	0.299	0.303	0.340	0.284	0.274

in terms of MAE and RMSE. The corresponding PSNR values are also reported in Tables 3 and 4 for each of the predicted images by comparing them with the corresponding measured surface. To understand better, the PSNR values are compared graphically in Figure 11. The proposed ST-SemK reports the highest PSNR, ≈3–7 dB higher than other methods for both Kolkata and Dallas. The separable and non-separable approaches of ST-SemK yield comparable results. with ≈1 dB PSNR difference for both the RoIs.

For any spatial analysis with auxiliary information, it must be noted that the improvement of the estimation depends on the covariates and their correlation with the primary variable. If the auxiliary variable is strongly correlated with the primary variable, then only it can improve the analysis and would be useful to consider. Many works in the literature have already identified the correlation between LST and LULC and used one variable for the analysis of the other (Liu & Weng, 2009); Jiang & Tian, 2010); Tran et al., 2017); Fan, Rey, & Myint, 2017). The distinctive LST patterns are associated with the thermal characteristics of the LULC types (Xiong et al., 2012). Also, for the *Landsat* TM/ETM+ imagery, the land surface emissivity is often derived by the normalized difference vegetation index (NDVI) thresholds method, where the NDVI is an indirect representation of the LULC distribution

TABLE 2 Comparison of predicted LST with four error measurement metrics—MAE, RMSE, MB, and SDE (in °C)—for the Dallas region. Lower MAE, RMSE, and MB closer to zero represent a better model and vice versa

RoI	Error measures						
Dallas	Zone	ST-SP	ST-NN	ST-IDW	ST-OK	ST-SemK$_{NSep}$	ST-SemK$_{Sep}$
	Mean absolute error (MAE, °C)						
	Zone 1	0.543	0.541	0.575	0.481	0.360	0.385
	Zone 2	0.619	0.623	0.662	0.618	0.362	0.426
	Zone 3	0.694	0.693	0.738	0.673	0.519	0.494
	Zone 4	0.586	0.628	0.735	0.711	0.518	0.434
	Zone 5	0.706	0.700	0.765	0.757	0.528	0.437
	Root mean square error (RMSE, °C)						
	Zone 1	0.530	0.519	0.596	0.524	0.331	0.356
	Zone 2	0.716	0.716	0.829	0.703	0.362	0.367
	Zone 3	0.910	0.907	1.037	0.858	0.493	0.446
	Zone 4	0.688	0.756	0.794	0.640	0.400	0.370
	Zone 5	0.880	0.859	1.012	0.935	0.470	0.429
	Mean bias (MB, °C)						
	Zone 1	−0.015	−0.011	−0.012	−0.014	−0.008	−0.005
	Zone 2	0.002	−0.002	−0.004	−0.001	0.001	0.001
	Zone 3	−0.018	−0.011	−0.012	−0.012	0.001	0.001
	Zone 4	0.012	0.009	0.010	0.008	0.005	0.002
	Zone 5	−0.002	−0.003	−0.005	−0.005	−0.001	−0.001
	Standard deviation of error (SDE, °C)						
	Zone 1	0.485	0.475	0.515	0.660	0.458	0.423
	Zone 2	0.554	0.550	0.601	0.543	0.441	0.466
	Zone 3	0.654	0.653	0.702	0.695	0.573	0.550
	Zone 4	0.587	0.601	0.678	0.693	0.485	0.427
	Zone 5	0.629	0.607	0.653	0.665	0.537	0.587

(Oguz et al., 2013). Therefore, the LULC distribution of the terrain is generally correlated with the LST distribution if the considered LULC classes are more generic, for example, assuming all plantations, cropland, fallow land together as their generic class agriculture (refer to the ontology in Figure 4). Here, we show an empirical study in Table 5 to support it.

The LULC and LST distributions for Kolkata are analyzed to check whether the urban area is associated with high-LST zones and vice versa. The November 2015 data are chosen as this time instance belongs to the early winter season of the study region and does not exhibit any stress conditions. The same five zones as chosen in Figure 9 are investigated in terms of the mean LST within and across different LULC types. The five LULC classes that we have considered for this analysis are: (1) waterbodies, (2) agriculture, (3) grassland, (4) wetlands, and (5) built-up. The mean LSTs for each of these LULC types are evaluated and compared in Table 5. The second and third columns of Table 5 show the distribution of LULC and LST for the five zones in Kolkata. Pictorially, it is evident that, for every zone, the urban (built-up) area is associated with the high-LST pixels compared to the non-urban areas (such as waterbodies, agricultural land, vegetation or grassland). In the fourth column of the

TABLE 3 Comparison study for ST-SemK (RoI: Kolkata) with LST mapping images. The error is reported in gray scale, where black pixels represent the highest error and white pixels represent the lowest error. Each predicted LST image is compared with the corresponding measured LST surface, and the PSNR is reported
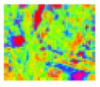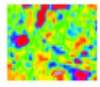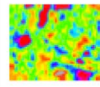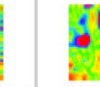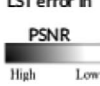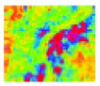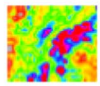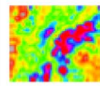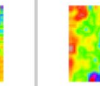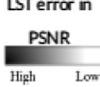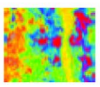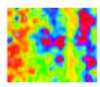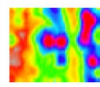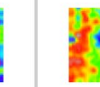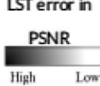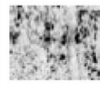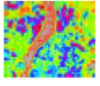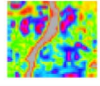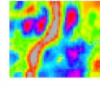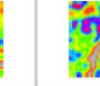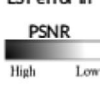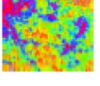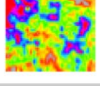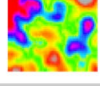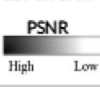
Zone	Measured LST	Predicted LST					
		ST-SP	ST-NN	ST-IDW	ST-OK	ST-SemK$_{NSep}$	ST-SemK$_{Sep}$
Zone 1	PSNR	38.58 dB	38.74 dB	37.60 dB	38.72 dB	44.93 dB	43.52 dB
Zone 2	PSNR	37.46 dB	37.45 dB	36.23 dB	37.65 dB	42.58 dB	43.57 dB
Zone 3	PSNR	36.25 dB	36.46 dB	35.62 dB	35.43 dB	39.51 dB	40.25 dB
Zone 4	PSNR	37.03 dB	37.00 dB	35.57 dB	36.53 dB	42.89 dB	41.89 dB
Zone 5	PSNR	42.67 dB	43.50 dB	43.36 dB	41.01 dB	45.75 dB	46.66 dB

TABLE 4 Comparison study for ST-SemK (RoI: Dallas) with LST mapping images. The error is reported in gray scale, where black pixels represent the highest error and white pixels represent the lowest error. Each predicted LST image is compared with the corresponding measured LST surface, and the PSNR is reported

Zone	Measured LST	ST-SP	ST-NN	ST-IDW	ST-OK	ST-SemK$_{NSep}$	ST-SemK$_{Sep}$
Zone 1	PSNR	36.28 dB	36.45 dB	35.26 dB	36.38 dB	40.37 dB	39.72 dB
Zone 2	PSNR	34.56 dB	34.56 dB	33.29 dB	34.72 dB	40.48 dB	40.36 dB
Zone 3	PSNR	33.37 dB	33.40 dB	32.23 dB	33.88 dB	38.70 dB	39.56 dB
Zone 4	PSNR	35.24 dB	34.42 dB	34.00 dB	35.87 dB	39.96 dB	40.62 dB
Zone 5	PSNR	32.60 dB	32.81 dB	31.39 dB	32.08 dB	38.05 dB	38.84 dB

The header spans "Predicted LST".
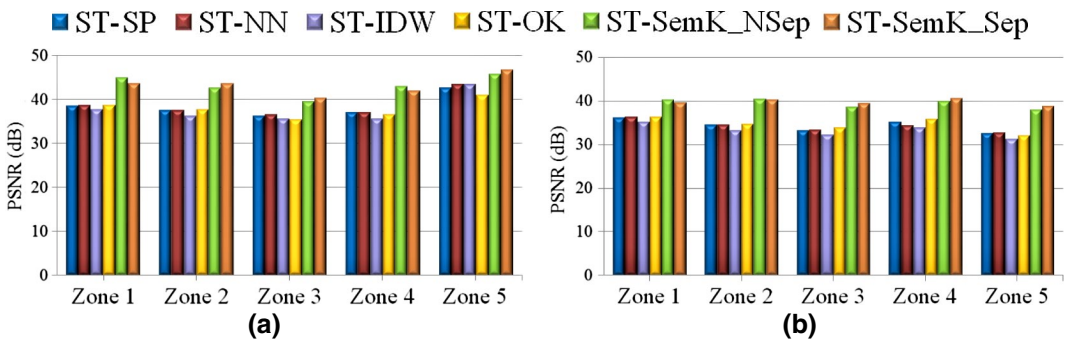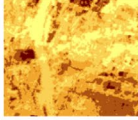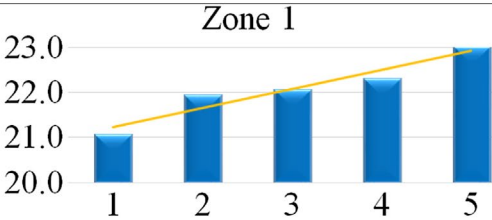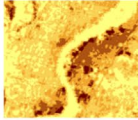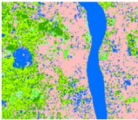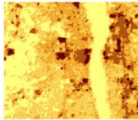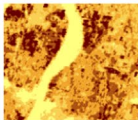
FIGURE 11 Comparison of LST prediction for two spatial regions, (a) Kolkata and (b) Dallas, with respect to the PSNR error measurement metric. The PSNR (in decibels) is reported on the *Y*-axis for the zones of the RoIs on the *X*-axis. A higher PSNR represents a better model, and vice versa

table, the correlation study has been carried out for every zone separately and the difference in mean LST across different LULC types is studied. It is observed that the relationship among the LULC classes in terms of LST is increasing from waterbodies to built-up areas. An exception is evident between the wetlands and built-up classes, mainly for zone 2, because most of the wetlands pixels are surrounded by the built-up pixels in this zone, such that it is difficult to distinguish these two classes. However, the increasing nature of the yellow linear trend line shows the general relationships among the LULC classes for every zone. Though the range of LST varies across the zones, the waterbodies are generally cooler than the plantation (agriculture and grassland) areas, followed by built-up areas.

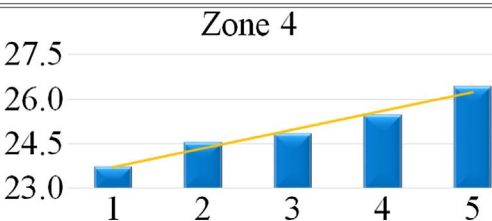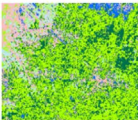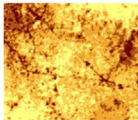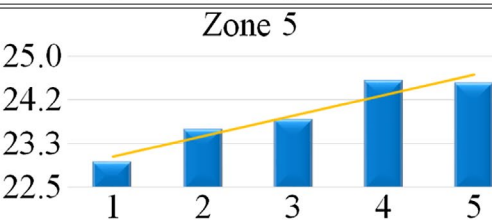The performance of the spatial interpolation method is highly dependent on the autocorrelation of the sampled location with respect to the prediction variable. Additionally, the performance of ST-SemK varies with the amount of entropy in the LULC distribution of the terrain. For example, if there is a single LULC class which covers the entire region (such as snow cover, desert), the performance of ST-SemK will be exactly equal to that of ST-OK. In contrast, if the LULC distribution is highly diverse, ST-SemK generally yields more accurate results than ST-OK and others (Bhattacharjee & Ghosh, 2015a). Therefore, an analysis of the terrain entropy can be carried out before the actual interpolation, by following the lemmas in Bhattacharjee and Ghosh (2015a), to check whether ST-SemK is suitable for the given study regions and the application. Further, it is also evident from the spatio-temporal analysis literature that experimental spatio-temporal semivariograms are difficult to construct. However, for remote sensing satellite data, most of the Earth's surface is measured, producing plenty of observations. Therefore, for the application of LST interpolation using *Landsat* ETM+ satellite images, the use of ST-SemK$_{NSep}$ is appropriate. The non-separable approach may not always be a suitable choice where the number of sampled locations is limited and may lead to over- or under-estimation of the prediction variable.

Finally, the basic SemK framework has been developed as a generic framework, which can be adopted for other variables and semantic knowledge. We have recently applied the semantic modeling of the SemK-based interpolation method for the mapping of XCO$_2$ (column-averaged dry-air mole fractions of atmospheric CO$_2$) measurements (Crisp et al., 2016) from NASA's *Orbiting Carbon Observatory 2* (OCO-2) satellite (Bhattacharjee & Chen, 2019a). The basic spatial SemK framework has been extended in a multivariate scenario to model these data sets together for predicting XCO$_2$. Incorporation of LULC and other CO$_2$ emission information into the interpolation process of XCO$_2$ enhanced the prediction accuracy. We are now improving the method by modeling atmospheric dynamics and its effect on CO$_2$ emission. This supports the wide application of the SemK framework and its variants.

TABLE 5 LULC–LST correlation study for five zones in Kolkata

5 | CONCLUSIONS

The ST-SemK method is motivated by the fact that the spatio-temporal LULC dynamics have significant impacts in deciding the patterns of land surface temperature. Both separable and non-separable variants of ST-SemK semantically analyze the space-time trade-off for assigning priority/weight to the LST sampled locations. From the empirical study, it is observed that ST-SemK yields more accurate LST prediction results than most of the popular

spatio-temporal interpolation techniques in terms of four error metrics: MAE, RMSE, MB and SDE. ST-SemK also reports ≈3–7 dB higher PSNR for both the RoIs when compared with the mapping images of LST.

The prediction accuracy reported by the ST-SemK$_{Sep}$ method is a little higher than ST-SemK$_{NSep}$ in some cases. However, the latter approach is less complex than the former. Therefore, depending on the required accuracy and the complexity specifications, either of the approaches can be chosen for the given spatio-temporal application. The ST-SemK is a generic approach that can be applied to any meteorological variables influenced by semantic knowledge of the terrain. Moreover, the same semantic modeling framework can be applied to extend any traditional interpolation method to its semantic extension. Though the method introduces additional overhead for incorporating auxiliary information into the interpolation process, the accuracy–complexity trade-off analysis proves the efficacy and the applicability of the ST-SemK method.

For future development of the ST-SemK framework, we want to combine OCO-2 data and ground-based remote sensing CO_2 data (Chen et al., 2016; Heinle & Chen, 2017) with the LULC to predict the missing pixels/footprints of OCO-2 data (Bhattacharjee & Chen, 2019) . ST-SemK can be utilized to predict other variables' missing footprints—for example, NDVI or solar-induced fluorescence (Bhattacharjee, Chen, & Rammig, 2019)—using relevant semantic knowledge, such as urban/population density.

ORCID

Shrutilipi Bhattacharjee 🔟 http://orcid.org/0000-0002-6288-0942
Jia Chen 🔟 http://orcid.org/0000-0002-6350-6610
Soumya K. Ghosh 🔟 http://orcid.org/0000-0001-8359-581X

REFERENCES

ArcGIS-Desktop:ArcMap. (2019). Semivariogram and covariance functions. Retrieved from http://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/semivariogram-and-covariance-functions.htm

Arslan, H. (2012). Spatial and temporal mapping of groundwater salinity using ordinary kriging and indicator kriging: The case of Bafra Plain, Turkey. *Agricultural Water Management*, *113*, 57–63.

Bhattacharjee, S., & Chen, J. (2019a). Global mapping of CO_2 concentration of OCO-2 by statistical modeling of anthropogenic emission dataset. In *Geophysical research abstracts* (vol. *21*, p 1.). Vienna, Austria.

Bhattacharjee, S., & Chen, J. (2019b). Semantic kriging (SemK) example code. Retrieved from https://www.esm.ei.tum.de/index.php?id=105

Bhattacharjee, S., Chen, J., & Rammig, A. (2019). Mapping of sif of OCO-2 from local to global scale: A multivariate statistical analysis approach. In *Geophysical research abstracts*, 17231–1. Vienna, Austria.

Bhattacharjee, S., & Ghosh, S. K. (2014). Automatic resolution of semantic heterogeneity in GIS: An ontology based approach. In *Advanced computing, networking and informatics* (Vol. *1*, pp. 585–591). Kolkata, India: Springer.

Bhattacharjee, S., & Ghosh, S. K. (2015a). Performance evaluation of semantic kriging: A Euclidean vector analysis approach. *IEEE Geoscience and Remote Sensing Letters*, *12*, 1185–1189.

Bhattacharjee, S., & Ghosh, S. K. (2015b). Time-series augmentation of semantic kriging for the prediction of meteorological parameters. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2015)* (pp. 4562–4565). Milan, Italy: IEEE.

Bhattacharjee, S., & Ghosh, S. K. (2017). *Semantic kriging* (Vol. *2*, pp. 1868–1879). Springer International Publishing.

Bhattacharjee, S., Mitra, P., & Ghosh, S. K. (2014). Spatial interpolation to predict missing attributes in GIS using semantic kriging. *IEEE Transactions on Geoscience and Remote Sensing*, *52*, 4771–4780.

Bhattacharjee, S., Prasad, R. R., Dwivedi, A., Dasgupta, A., & Ghosh, S. K. (2012). Ontology based framework for semantic resolution of geospatial query. In *2012 12th International Conference on Intelligent Systems Design and Applications (ISDA)* (pp. 437–442). Kochi, India: IEEE.

Carrera-Hernández, J., & Gaskin, S. (2007). Spatio temporal analysis of daily precipitation and temperature in the basin of Mexico. *Journal of Hydrology, 336*, 231–249.

Chen, B., Huang, B., Chen, L., & Xu, B. (2017). Spatially and temporally weighted regression: A novel method to produce continuous cloud-free Landsat imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *55*, 27–37.

Chen, J., Viatte, C., Hedelius, J. K., Jones, T., Franklin, J. E., Parker, H., … Wofsy, S. C. (2016). Differential column measurements using compact solar-tracking spectrometers. *Atmospheric Chemistry and Physics*, *16*, 8479–8498.

Chen, X., Li, W., Chen, J., Rao, Y., & Yamaguchi, Y. (2014). A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery. *Remote Sensing*, *6*, 2845–2863.

Colombi, A., De Michele, C., Pepe, M., Rampini, A., & Michele, C. D. (2007). Estimation of daily mean air temperature from MODIS LST in alpine areas. *EARSeL eProceedings*, *6*, 38–46.

Crisp, D., Pollock, H., Rosenberg, R., Chapsky, L., Lee, R., Oyafuso, F., … Wunch, D. (2016). The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO2) instrument and its radiometrically calibrated products. *Atmospheric Measurement Techniques*, *10*, 59–81.

Deng, M., Yang, W., Liu, Q., Jin, R., Xu, F., & Zhang, Y. (2018). Heterogeneous space–time artificial neural networks for space–time series prediction. *Transactions in GIS*, *22*, 183–201.

Dixon, B., & Uddameri, V. (2016). *GIS and geocomputation for water resource science and engineering*. John Wiley & Sons.

Fan, C., Rey, S. J., & Myint, S. W. (2017). Spatially filtered ridge regression (SFRR): A regression framework to understanding impacts of land cover patterns on urban climate. *Transactions in GIS*, *21*, 862–879.

Foster, M. P., & Evans, A. N. (2008). An evaluation of interpolation techniques for reconstructing ionospheric TEC maps. *IEEE Transactions on Geoscience and Remote Sensing*, *46*, 2153–2164.

Gerber, F., de Jong, R., Schaepman, M. E., Schaepman-Strub, G., & Furrer, R. (2018). Predicting missing values in spatio-temporal remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *56*, 2841–2853.

Heinle, L., & Chen, J. (2017). Automated enclosure and protection system for compact solar-tracking spectrometers. *Atmospheric Measurement Techniques Discussions*, *2017*, 1–21.

Hengl, T., Heuvelink, G. B., Tadić, M. P., & Pebesma, E. J. (2012). Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theoretical and Applied Climatology*, *107*, 265–277.

Heuvelink, G., & Griffith, D. A. (2010). Space-time geostatistics for geography: A case study of radiation monitoring across parts of Germany. *Geographical Analysis*, *42*, 161–179.

Hui, Z., Hu, Y., Yevenyo, Y. Z., & Yu, X. (2016). An improved morphological algorithm for filtering airborne LIDAR point cloud based on multi-level kriging interpolation. *Remote Sensing*, *8*, 35. https://doi.org/10.3390/rs8010035

Hulley, G., Veraverbeke, S., & Hook, S. (2014). Thermal-based techniques for land cover change detection using a new dynamic MODIS multispectral emissivity product (MOD21). *Remote Sensing of Environment*, *140*, 755–765.

Jiang, J., & Tian, G. (2010). Analysis of the impact of land use/land cover change on land surface temperature with remote sensing. *Procedia Environmental Sciences*, *2*, 571–575.

Li, J. (2008). *A review of spatial interpolation methods for environmental scientists (Record 2008/23)*. Canberra, Australia: Geoscience Australia.

Li, J., & Heap, A. D. (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics*, *6*, 228–241.

Liu, H., & Weng, Q. (2009). Scaling effect on the relationship between landscape pattern and land surface temperature. *Photogrammetric Engineering & Remote Sensing*, *75*, 291–304.

Lussana, C., Tveito, O., & Uboldi, F. (2018). Three-dimensional spatial interpolation of 2 m temperature over Norway. *Quarterly Journal of the Royal Meteorological Society*, *144*, 344–364.

Mendiratta, N., Kumar, R. S., & Rao, K. S. (2008). *Standards for bio-geo database* (Technical Report 1). Natural Resources Data Management System (NRDMS) Division, New Delhi.

Metz, M., Andreo, V., & Neteler, M. (2017). A new fully gap-free time series of land surface temperature from modis lst data. *Remote Sensing*, *9*. Retrieved from http://www.mdpi.com/2072-4292/9/12/1333

Oguz, H. et al. (2013). LST calculator: A program for retrieving land surface temperature from Landsat TM/ETM+ imagery. *Environmental Engineering and Management Journal*, *12*, 549–555.

Romaguera, M., Vaughan, R., Ettema, J., Izquierdo-Verdiguier, E., Hecker, C., & van der Meer, F. (2018). Detecting geothermal anomalies and evaluating LST geothermal component by combining thermal remote sensing time series and land surface model data. *Remote Sensing of Environment*, *204*, 534–552.

Shiode, N., & Shiode, S. (2011). Street-level spatial interpolation using network-based IDW and ordinary kriging. *Transactions in GIS*, *15*, 457–477.

Snepvangers, J., Heuvelink, G., & Huisman, J. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*, *112*, 253–271.

Spadavecchia, L., & Williams, M. (2009). Can spatio-temporal geostatistical methods improve high resolution regionalisation of meteorological variables? *Agricultural and Forest Meteorology*, *149*, 1105–1117.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*, 234–240.

Tran, D. X., Pla, F., Latorre-Carmona, P., Myint, S. W., Caetano, M., & Kieu, H. V. (2017). Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS Journal of Photogrammetry and Remote Sensing*, *124*, 119–132.

Trigo, I. F., Dacamara, C. C., Viterbo, P., Roujean, J.-L., Olesen, F., Barroso, C., … Geiger, B. (2011). The satellite application facility for land surface analysis. *International Journal of Remote Sensing*, *32*, 2725–2744.

Tsendbazar, N.-E., De Bruin, S., Fritz, S., & Herold, M. (2015). Spatial accuracy assessment and integration of global land cover datasets. *Remote Sensing*, *7*, 15804–15821.

Wagner, H. H. (2003). Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. *Ecology*, *84*, 1045–1057.

Weiss, D. J., Atkinson, P. M., Bhatt, S., Mappin, B., Hay, S. I., & Gething, P. W. (2014). An effective approach for gap-filling continental scale remotely sensed time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 106–118.

Xiong, Y., Huang, S., Chen, F., Ye, H., Wang, C., & Zhu, C. (2012). The impacts of rapid urbanization on the thermal environment: A remote sensing study of Guangzhou, South China. *Remote Sensing*, *4*, 2033–2056.

Xu, C.-D., Wang, J.-F., Hu, M.-G., & Li, Q.-X. (2013). Interpolation of missing temperature data at meteorological stations using P-BSHADE. *Journal of Climate*, *26*, 7452–7463.

Yang, D., Gu, C., Dong, Z., Jirutitijaroen, P., Chen, N., & Walsh, W. M. (2013). Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging. *Renewable Energy*, *60*, 235–245.

Yang, J., Wang, Y., & August, P. (2004). Estimation of land surface temperature using spatial interpolation and satellite-derived surface emissivity. *Journal of Environmental Informatics*, *4*, 37–44.