This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Performance Evaluation of Semantic Kriging: A Euclidean Vector Analysis Approach

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

***Abstract*—Prediction of spatial attributes in geospatial data repositories is indispensable in the field of remote sensing and geographic information system. The *semantic kriging* (*SemK*) approach semantically captures the domain knowledge of the terrain in terms of local spatial features for spatial attribute prediction. It produces better results than *ordinary kriging* and other prediction methods. This letter focuses on the theoretical and empirical analyses of the SemK. A *Euclidean* vector analysis approach is adopted to theoretically prove the efficacy of SemK in capturing semantic knowledge.**

***Index Terms*—Geographic information system (GIS), kriging, prediction, semantic kriging (SemK).**

I. Introduction

SPATIAL data repositories often consist of missing and erroneous spatial attributes. In this scenario, the prediction and forecasting of these attributes with better accuracy is a major challenge in the domain of *remote sensing* and *geographic information system*. It becomes critical in many spatial analyses, such as climatological/weather analysis, where several spatial attributes are involved. Among different statistical and machine learning-based prediction methods, *kriging* [1], [2] is the most popular regression-based interpolation technique, aiming at the mean square prediction error minimization. It models the spatial autocorrelation in terms of *semivariances*, which is the function of *Euclidean distance* in 2-D space. However, in case of meteorological attributes, particularly for the weather parameters, the association between the random fields between any pair of sample points does not solely depend on their Euclidean distance but also on their representative spatial features or the land covers. According to the *U.S. Environmental Protection Agency*,[1] the geofeatures, such as *building*, *road surface*, *water body*, etc., influence weather attributes, particularly the *land surface temperature* (*LST*), significantly. Hengl *et al.* [3] reported the LST of a location to be the function of its representative land cover. However, this knowledge is not incorporated into the prediction method. In our previous work [4], the proposed *semantic kriging* (*SemK*) blends this semantic knowledge of the terrain into the interpolation process for better estimation of the prediction attributes (PAs). The "semantics" of a sample point is considered as the additional knowledge of the terrain [5]. Thus, in this scenario, no multivariate kriging methods [e.g., *co-kriging* (*CK*)] will be suitable. The SemK extends the most popular univariate kriging method [2], i.e., *ordinary kriging* (*OK*), with the land cover information, which is modeled as the semantics of the sample points. A weighted ontology is considered to capture the semantic knowledge and the correlation between the surrounding spatial features. Ontology [6], [7] is a tool to capture the domain knowledge by analyzing the semantic relationships between the concepts. The SemK combines this ontology-based semantic association with the Euclidean distance-based relatedness (estimated by OK) of the sample points. The mathematical models for the semantics and spatial correlation, theoretical error analysis, and efficiency analysis in terms of entropy have been formalized in [4]. A case study [4] with actual LST data proves that SemK outperforms OK and several other popular methods. This letter presents a theoretical as well as an empirical analysis of SemK to prove its efficacy in handling the semantic knowledge for spatial prediction, verify its functionality, and establish its relationship with OK.

A. Objectives

This letter focuses on the theoretical analysis and performance evaluation of SemK. It investigates the capability of SemK in incorporating the semantic domain knowledge into the prediction method. It also validates the impact of the ontology, and its granularity, to emphasize the benefits of SemK over OK. Although OK has been considered as the base scheme for SemK, the similar theoretical analysis can be carried out with any other univariate kriging method. The formal proofs based on the *Euclidean* vector analysis approach are presented with four *lemmas* and a *proposition*, each of which exhibits important characteristics of SemK. The outlines of the lemmas and the proposition are given as follows:

- SemK assigns more weightage to the semantically similar and spatially important interpolating point than other points (Lemma 1).
- The amount of domain knowledge captured by SemK over OK is the angular difference between the weight vector produced by SemK and OK. In other words, the domain knowledge is correctly modeled by SemK (Lemma 2).
- The weight vector assigned by SemK is closer to the optimal weight vector produced by OK. Thus, SemK performs better than OK (Lemma 3).
- A less detailed ontology will eventually reduce SemK to OK (Lemma 4).
- The misclassification of the sample points in the ontology can be detected by SemK (Proposition 1).

This letter is organized in five sections. Section II describes the SemK interpolation method [4]. Section III presents the theoretical analysis of SemK. The empirical analysis of the proposed lemmas is presented in Section IV. Finally, the conclusion is drawn in Section V.

Manuscript received June 4, 2014; revised October 13, 2014; accepted November 4, 2014.

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

Digital Object Identifier 10.1109/LGRS.2014.2387373

[1] http://www.epa.gov/hiri/resources/pdf/BasicsCompendium.pdf

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

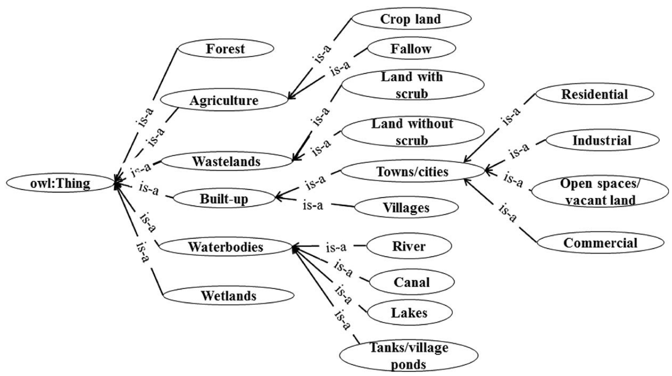

Fig. 1. Spatial feature ontology for LST prediction [4].

II. Overview of SemK

The SemK proposed in [4] blends the semantics and the correlation between the surrounding spatial features into the interpolation process. It maps the traditional *covariance* to the higher dimension by extending the popular kriging method, i.e., OK. According to the Tobler's law of spatial proximity [8], the correlated and the semantically similar features will have more importance than the distant one. The semantics and the spatial correlations are evaluated by organizing the spatial features into an ontology hierarchy, constructed with all possible features within the region of interest (RoI). The ontology hierarchy of *Kolkata*, a metropolitan city in India (central coordinate: 22.567° N, 88.367° E), is depicted in Fig. 1 [4]. It consists of the spatial land use/land cover classes, such as built-up, agriculture, forest, water bodies, etc. This hierarchy is constructed using *"is-a"* (*hyponym*) relation.

All the sample points are mapped to the most appropriate representative *leaf* features in the ontology hierarchy. The association between a pair of leaf features in the hierarchy is evaluated with two parameters, the *spatial importance* and the *semantic similarity* [4]. These two parameters modify and map the standard covariance measure into the higher dimension.

1) Spatial Importance: The spatial importance between each pair of leaf features in the ontology can be measured by the correlation analysis between them, with reference to the PAs. The RoI is divided into k number of nonoverlapping zones (R_k) such that $\bigcup_{i=1}^{k} R_k = \text{RoI}$. The k pairs of sample points are chosen within a predefined distance d from the whole RoI, such that each pair is retrieved from each of the k zones. The pairwise correlation analysis is carried out to evaluate the association between the leaf features in the ontology with the k pairs of sample points. This correlation score between a pair of sample points is termed as spatial importance. The spatial importance between the ith and the jth features in the ontology, i.e., SI_{ij}, is given as follows:

$$\begin{aligned}\text{SI}_{ij} &= \text{Corr}_{\text{PA}}(f_i, f_j)\\ &= \frac{\sum_{m=1}^{k}(Z(f_{i_m}) - \overline{Z(f_i)})(Z(f_{j_m}) - \overline{Z(f_j)})}{\sqrt{\sum\limits_{m=1}^{k}(Z(f_{i_m}) - \overline{Z(f_i)})^2 \sum\limits_{m=1}^{k}(Z(f_{j_m}) - \overline{Z(f_j)})^2}}\end{aligned}$$

where $Z(f_{p_q})$ represents the random field value of the qth sample point, representing the feature f_p; $\overline{Z(f_p)}$ represents the average of the random field values of the feature f_p over k sample points. For all the interpolating points with respect to the prediction point, an $[\mathbf{N} \times 1]$ vector is formed, given as $[\mathbf{SI}_{0i}]^T_{\mathbf{N}\times 1} = [\text{SI}_{01}\text{SI}_{02}\cdots\text{SI}_{0N}]$. Similarly, for N interpolating points, an $[\mathbf{N} \times \mathbf{N}]$ symmetric matrix is formed, termed as $[\mathbf{SI}_{ij}]_{\mathbf{N}\times\mathbf{N}}$.

TABLE I
Example Scenario: Specifications of the Prediction Point x_0 and Three Interpolating Points x_i, x_j, and x_k

Sample point	Representative feature	A_h	A_{SI}	A_{SS}	Assigned weight
x_0	f_0	0	1	1	–
x_i	f_i	d_{0i}	SI_{0i}	SS_{0i}	w_i
x_j	f_j	d_{0j}	SI_{0j}	SS_{0j}	w_j
x_k	f_k	d_{0k}	SI_{0k}	SS_{0k}	w_k

2) Semantic Similarity: The semantic similarity between any two sample points or their representative features in the ontology is measured using the modified *context resemblance* method [9]. The semantic similarity between the ith and jth features in the ontology, i.e., SS_{ij}, is given as follows:

$$\text{SS}_{ij} = \frac{\frac{m_i}{|f_i|} + \frac{m_j}{|f_j|}}{2}$$

where $|f_i|$ and $|f_j|$ are the total number of nodes in the ith and jth feature paths respectively, and m_i and m_j (where $m_i = m_j$) are the number of features matching in their paths. For this metric, an $[\mathbf{N} \times 1]$ vector is formed with reference to the interpolation point and is given as $[\mathbf{SS}_{0i}]^T_{\mathbf{N}\times 1} = [\text{SS}_{01}\text{SS}_{02}\cdots\text{SS}_{0N}]$. Furthermore, for all the interpolating points, an $[\mathbf{N} \times \mathbf{N}]$ symmetric matrix is formed, denoted by $[\mathbf{SS}_{ij}]_{\mathbf{N}\times\mathbf{N}}$.

The modified *covariance matrix* ($\mathbf{C}'$) and the *distance matrix* ($\mathbf{D}'$) for SemK [4], based on the spatial importance and the semantic similarity, are given as follows:

$$\mathbf{C}' = \overset{\mathbf{C}}{\underset{([\mathbf{SI_{ij}}]\circ[\mathbf{SS_{ij}}])}{-.-.-}} \quad \mathbf{D}' = \overset{\mathbf{D}}{\underset{([\mathbf{SI_{0i}}]\circ[\mathbf{SS_{0i}}])}{-.-.-}}$$

where "$\circ$" and "$-.-.-$" denote the *Hadamard product* and *Hadamard division* between matrices, respectively. In both $\mathbf{C}$ and $\mathbf{D}$, the covariance is calculated with respect to the Euclidean distance in 2-D space [4]. The weight matrix $\mathbf{W}'$ of SemK is given as

$$\mathbf{W}' = \left[\overset{\mathbf{C}}{\underset{([\mathbf{SI}_{ij}]\circ[\mathbf{SS}_{ij}])}{-.-.-}}\right]^{-1}\left[\left[\overset{\mathbf{D}}{\underset{([\mathbf{SI}_{0i}]\circ[\mathbf{SS}_{0i}])}{-.-.-}}\right] - \lambda\mathbf{1}\right]$$

where λ is the *Lagrange multiplier* of SemK.

III. Analysis of SemK

This section presents a Euclidean vector analysis-based performance evaluation of SemK. It argues that the sample points representing the semantically similar and spatially correlated features should have more impact on the prediction point than the less similar and loosely correlated sample points. The lemmas and the proposition are presented to establish the characteristics of SemK. An example scenario (refer to Table I) is considered with an interpolation/prediction point x_0 and three interpolating points x_i, x_j, and x_k, along with three supporting parameters, namely, Euclidean distance (A_h), spatial importance (A_{SI}), and semantic similarity (A_{SS}). The parameters are measured with respect to the interpolation point x_0. *Assigned weight* w' is computed by SemK for each interpolating point, with respect to these supporting attributes.

Lemma 1: Between any pair of interpolating points with the same Euclidean distance from the prediction point, the point that represents more similar feature with that of the prediction point will be assigned more weightage by SemK than the other one.

$$= \frac{(1-\lambda_{\mathrm{OK}})(\frac{1}{\mathrm{SIS}_{0j}} - \lambda_{\mathrm{SemK}}) + (1-\lambda_{\mathrm{OK}})(\frac{1}{\mathrm{SIS}_{0i}} - \lambda_{\mathrm{SemK}})}{\sqrt{(1-\lambda_{\mathrm{OK}})^2 + (1-\lambda_{\mathrm{OK}})^2}\sqrt{\left(\frac{1}{\mathrm{SIS}_{0j}} - \lambda_{\mathrm{SemK}}\right)^2 + \left(\frac{1}{\mathrm{SIS}_{0i}} - \lambda_{\mathrm{SemK}}\right)^2}} = \frac{1}{\sqrt{2}}\left(\frac{\mathbb{SIS}_{0i} + \mathbb{SIS}_{0j}}{\sqrt{\mathbb{SIS}_{0i}^2 + \mathbb{SIS}_{0j}^2}}\right)$$

Proof: Consider any pair of the interpolating points x_i and x_j from Table I, and let the Euclidean distances from the interpolation point be the same for both x_i and x_j, i.e., $d_{0i} = d_{0j}$. Let f_i be more semantically similar and correlated with f_0 than f_j. According to the hierarchical ontology property and the Tobler's law of spatial proximity [8], $(\mathrm{SI}_{0i} * \mathrm{SS}_{0i}) > (\mathrm{SI}_{0j} * \mathrm{SS}_{0j}) \Rightarrow \mathrm{SIS}_{0i} > \mathrm{SIS}_{0j}$, where $(\mathrm{SI}_{mn} * \mathrm{SS}_{mn})$ is referred to as SIS_{mn}. We need to prove $w'_i > w'_j$. The modified covariance matrix $\mathbf{C}'$ and the distance matrix $\mathbf{D}'$ of SemK (defined in [4]) are given as follows:

$$\mathbf{C}' = \begin{bmatrix} \frac{\gamma(d_{ii})}{1*1} & \frac{\gamma(d_{ij})}{\mathrm{SIS}_{ij}} \\ \frac{\gamma(d_{ji})}{\mathrm{SIS}_{ji}} & \frac{\gamma(d_{jj})}{1*1} \end{bmatrix} \qquad \mathbf{D}' = \begin{bmatrix} \frac{\gamma(d_{0i})}{\mathrm{SIS}_{0i}} \\ \frac{\gamma(d_{0j})}{\mathrm{SIS}_{0j}} \end{bmatrix}.$$

From the definition of SemK, the weight matrix $\mathbf{W}'$ is given as follows: $\mathbf{C}'^{-1}[\mathbf{D}' - \lambda_{\mathrm{SemK}}\mathbf{1}]$. Considering the constraint of SemK, i.e., $\mathbf{1}^T\mathbf{W}' = 1$, $[\mathbf{D}' - \lambda_{\mathrm{SemK}}\mathbf{1}]$ can be written as $\mathbb{D}'$, where $\mathbb{D}'_{0i} = ((\gamma(d_{0i})/\mathrm{SIS}_{0i}) - \lambda_{\mathrm{SemK}})$. For this scenario, $\mathbf{W}'$ is modified as $\mathbf{C}'^{-1}\mathbb{D}$ and is given as follows:

$$\mathbf{W}' = \begin{bmatrix} w_i^{\mathrm{SemK}} \\ \\ w_j^{\mathrm{SemK}} \end{bmatrix} = \frac{1}{K}\begin{bmatrix} -\gamma(d_{jj}) & \frac{\gamma(d_{ij})}{\mathrm{SIS}_{ij}} \\ \frac{\gamma(d_{ji})}{\mathrm{SIS}_{ji}} & -\gamma(d_{ii}) \end{bmatrix}\begin{bmatrix} \mathbb{D}'_{0i} \\ \mathbb{D}'_{0j} \end{bmatrix}$$
$$= \frac{1}{K}\begin{bmatrix} \frac{\gamma(d_{ij})*\mathbb{D}'_{0j}}{\mathrm{SIS}_{ij}} \\ \frac{\gamma(d_{ij})*\mathbb{D}'_{0i}}{\mathrm{SIS}_{ij}} \end{bmatrix}$$

where $K = -\gamma(d_{ii})\gamma(d_{jj}) + (\gamma(d_{ij})^2/(\mathrm{SI}_{ij} * \mathrm{SS}_{ij})^2)$. From the definition of OK, $\gamma(d_{ii}) = \gamma(d_{jj}) = 0$ (as both represent *self-covariance* with respect to Euclidean distance), and $d_{ij} = d_{ji}$. From the definition of SemK, $\mathrm{SI}_{ij} = \mathrm{SI}_{ji}$ and $\mathrm{SS}_{ij} = \mathrm{SS}_{ji} \Rightarrow \mathrm{SIS}_{ij} = \mathrm{SIS}_{ji}$. Thus, K is modified as $\gamma(d_{ij})^2/(\mathrm{SIS}_{ij})^2$. Hence, the definition of $\mathbb{D}'$ implies that $w_i^{\mathrm{SemK}} > w_j^{\mathrm{SemK}}$. This concludes the proof.

Lemma 2: The domain knowledge captured by SemK, in terms of semantic similarity and the spatial importance for all the interpolating points, is reflected in the semantic weights assigned by SemK. This additional knowledge of SemK over OK can be represented by the angular difference of the weight vectors of OK and SemK.

Proof: For proving this lemma, we need to prove that the angular difference between the weight vectors of OK ($\overrightarrow{\mathbf{W}_{\mathrm{OK}}}$) and SemK ($\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}}$), which represents the amount of change incorporated to the semantic weight vector of SemK over OK, is exactly equal to the angular difference between the captured knowledge by SemK in terms of the semantics and that by OK. Thus, the formal argument is as follows: $\theta_{(\overrightarrow{\mathbf{Change}},(1-\lambda_{\mathrm{OK}})\overrightarrow{\mathbf{1}})} = \theta_{(\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}})}$.

For N interpolating points, $\overrightarrow{\mathbf{Change}}$ is a Euclidean vector, given as $[\mathrm{Change}_1\mathrm{Change}_2 \cdots \mathrm{Change}_N]^T$, where Change_i is the amount of semantic knowledge captured by SemK for the ith interpolating point, over OK. Since only SemK can capture this knowledge with respect to the spatial importance and the semantic similarity of the surrounding features, the amount of change captured by SemK for the ith interpolating point is given as follows:

$$\mathrm{Change}_i = \overrightarrow{\mathbf{SIS_{iN}}} \circledast (\overrightarrow{\mathbf{SIS_{0N}}} - \overrightarrow{\lambda_{\mathrm{SemK}}\mathbf{1}}).$$

Formally, $\overrightarrow{\mathbf{SIS_{iN}}}$ and $\overrightarrow{\mathbf{SIS_{0N}}}$ denote the ith row of the matrix $[([\mathbf{SI}_{ij}]_{\mathbf{N}\times\mathbf{N}} \circ [\mathbf{SS}_{ij}]_{\mathbf{N}\times\mathbf{N}})^{-h}]^{-1}$ and the matrix $([\mathbf{SI}_{0i}]_{\mathbf{N}\times 1} \circ [\mathbf{SS}_{0i}]_{\mathbf{N}\times 1})^{-h}$, respectively; $\circ$ is the *Hadamard product*; $\circledast$ denotes the *dot product*, and $-h$ denotes the *Hadamard inverse*. For any two interpolating points x_i and x_j, as specified in Table I, the $\overrightarrow{\mathbf{Change}}$ matrix is given as follows:

$$\mathbf{Change} = \begin{bmatrix} \mathrm{Change}_i \\ \mathrm{Change}_j \end{bmatrix}$$
$$= [[[\mathbf{SI}_{ij}] \circ [\mathbf{SS}_{ij}]]^{-h}]^{-1}_{2\times 2}[[[\mathbf{SI}_{0i}] \circ [\mathbf{SS}_{0i}]]^{-h}_{2\times 1} - [\lambda_{\mathrm{SemK}}\mathbf{1}]_{2\times 1}]$$
$$= \frac{1}{K_{change}}\begin{bmatrix} \frac{1}{\mathrm{SIS}_{ij}*\mathbb{SIS}_{0j}} \\ \frac{1}{\mathrm{SIS}_{ij}*\mathbb{SIS}_{0i}} \end{bmatrix}.$$

Here, $\overrightarrow{\mathbf{1}}$ represents the semantic knowledge vector of OK, $\mathbb{SIS}^{-h} = (\mathbf{SIS}^{-h} - \lambda_{\mathrm{SemK}}\mathbf{1})$. In case of semantic knowledge, $d_{0i} = d_{0j}$. Considering this constraint and normalizing the distance matrices for OK and SemK, respectively, the weight matrices of OK ($\mathbf{W}_{\mathrm{OK}}$) and SemK ($\mathbf{W}'_{\mathrm{SemK}}$) are given as follows:

$$\mathbf{W}_{\mathrm{OK}} = \frac{1}{K}\begin{bmatrix} \gamma(d_{ij}) * (1-\lambda_{\mathrm{OK}}) \\ \gamma(d_{ij}) * (1-\lambda_{\mathrm{OK}}) \end{bmatrix}$$
$$\mathbf{W}'_{\mathrm{SemK}} = \frac{1}{K'}\begin{bmatrix} \frac{\gamma(d_{ij})}{\mathrm{SIS}_{ij}} * (\frac{1}{\mathrm{SIS}_{0j}} - \lambda_{\mathrm{SemK}}) \\ \frac{\gamma(d_{ij})}{\mathrm{SIS}_{ij}} * (\frac{1}{\mathrm{SIS}_{0i}} - \lambda_{\mathrm{SemK}}) \end{bmatrix}.$$

In terms of semantics of the surrounding spatial features, the angular difference between the incorporated knowledge by the SemK, with respect to OK is given as follows:

$$\theta_{(\overrightarrow{\mathbf{Change}},1-\lambda_{\mathrm{OK}}\overrightarrow{\mathbf{1}})} = Cos^{-1}\left(\frac{1}{\sqrt{2}}\frac{\mathbb{SIS}_{0i} + \mathbb{SIS}_{0j}}{\sqrt{\mathbb{SIS}_{0i}^2 + \mathbb{SIS}_{0j}^2}}\right).$$

From Table I, assuming $d_{0i} = d_{0j}$ for the semantic knowledge, the angular difference between $\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}}$ and $\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}}$ is given by $\theta_{(\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}})}$. Now, $\mathrm{Cos}_{(\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}})}\theta$, shown at the top of the page, which implies $\theta_{(\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}})} = \mathrm{Cos}^{-1}\frac{1}{\sqrt{2}}((\mathbb{SIS}_{0i} + \mathbb{SIS}_{0j})/\sqrt{\mathbb{SIS}_{0i}^2 + \mathbb{SIS}_{0j}^2}))$. Hence, $\theta_{(\overrightarrow{\mathbf{Change}},(1-\lambda_{\mathrm{OK}})\overrightarrow{\mathbf{1}})} = \theta_{(\overrightarrow{\mathbf{W}^{\mathrm{sem}}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'^{\mathrm{sem}}_{\mathrm{SemK}}})}$. This concludes the proof.

Lemma 3: The weight vector generated by SemK is closer to the optimal weight vector than that produced by OK.

Proof: To prove this argument, it can be proved alternatively that the angular difference between the optimal weight vector with SemK is always less than the angular difference between the optimal weight vector with OK, i.e., $\theta_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})} > \theta_{(\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}$.

Let the optimal weight vector be $[w_i^{\mathrm{OPT}}w_j^{\mathrm{OPT}}]^T$ for any two interpolating points x_i and x_j. Let there be some extra domain knowledge (e.g., $\mathbf{U}$) required to be incorporated to get the optimal solution. Given Lemma 1, if $(w_i^{\mathrm{SemK}} - w_j^{\mathrm{SemK}}) \geq 0$,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

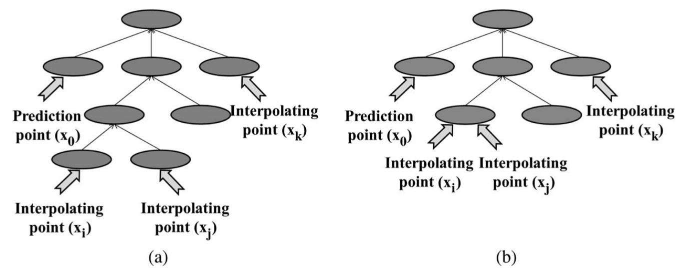

Fig. 2. Position of the sample points in the ontology. (a) Initial ontology. (b) Modified ontology.

then $(w_i^{\mathrm{OPT}} - w_j^{\mathrm{OPT}}) \geq (w_i^{\mathrm{SemK}} - w_j^{\mathrm{SemK}}) \geq 0$, and vice versa. In other words, as semantic similarity and spatial importance are inversely proportional to the Euclidean distance-based traditional covariance, $\mathbf{U}$ is also inversely proportional to that and takes some positive real values between (0, 1]. According to Lemma 2, $\mathrm{Cos}_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}\theta$ is given as $(((1/(\mathrm{SIS}_{0j}U_{0j})) - \lambda_{\mathrm{OPT}}) + ((1/(\mathrm{SIS}_{0i}U_{0i})) - \lambda_{\mathrm{OPT}}))/(\sqrt{2}(((1/(\mathrm{SIS}_{0j}U_{0j}))-\lambda_{\mathrm{OPT}})^2+((1/(\mathrm{SIS}_{0i}U_{0i}))-\lambda_{\mathrm{OPT}})^2))$.

Similarly, $\mathrm{Cos}_{(\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}\theta$ is given as $(((1/U_{0j}) - \lambda_{\mathrm{OPT}}) + ((1/U_{0i}) - \lambda_{\mathrm{OPT}}))/(\sqrt{2}(((1/U_{0j}) - \lambda_{\mathrm{OPT}})^2 + ((1/U_{0i}) - \lambda_{\mathrm{OPT}})^2))$.

As the semantic knowledge (SIS_{ij}) between any two sample points is inversely proportional to their traditional covariance and lies between (0, 1], therefore $\mathrm{Cos}_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}\theta < \mathrm{Cos}_{(\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}\theta \Rightarrow \theta_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})} > \theta_{(\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}},\overrightarrow{\mathbf{W}_{\mathrm{OPT}}})}$. This proves that the weight vector produced by SemK is closer to the optimal vector than that produced by OK.

Lemma 4: Less detailed ontology will eventually reduce SemK to OK.

Proof: Let us consider three interpolating points from the example scenario in Table I. In both ontologies in Fig. 2, the positions of the representative features of the sample points are identified. In the initial ontology [refer to Fig. 2(a)], the interpolating points x_i and x_j are pointing to two different specialized concepts in the hierarchy. In the modified ontology, e.g., the above specialized concepts are generalized to the same parent concept [refer to Fig. 2(b)]. Thus, both x_i and x_j are now represented by the same concept in the ontology hierarchy. For both x_i and x_j, with respect to the SemK process with general ontology (*SemK_mod*), the semantic properties have been changed to $\mathrm{SI}_{0i}^{\mathrm{mod}}$, $\mathrm{SS}_{0i}^{\mathrm{mod}}$ and $\mathrm{SI}_{0j}^{\mathrm{mod}}$, $\mathrm{SS}_{0j}^{\mathrm{mod}}$, with reference to the prediction point x_0. However, the semantic property of x_k remains unchanged. According to the property of hierarchical ontology, it can be verified that the more general concept has higher semantic similarity to others than the specialized concept in the same path [4]. For evaluating the spatial importance of the parent concept, let m be the number of specialized concepts mapped to its parent concept. Among the $(m * k)$ sample points of the parent concept, the first k sample points are chosen for the correlation study, which are spatially closer. It will eventually lead to a higher correlation score than any specialized concept (according to the Tobler's law of spatial proximity [8]). Thus, for any $p = i, j$ and $q = 0, k$, the following inequality holds:

$$1 > \mathrm{SIS}_{pq}^{\mathrm{mod}} > \mathrm{SIS}_{pq} > 0.$$

For both x_i and x_j, $w_i^{\mathrm{SemK}} < w_i^{\mathrm{mod}} < w_i^{\mathrm{OK}}$ and $w_j^{\mathrm{SemK}} < w_j^{\mathrm{mod}} < w_j^{\mathrm{OK}}$. Similarly, from Lemma 2, it can be proved that $\theta_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'_{SemK\_mod}})} < \theta_{(\overrightarrow{\mathbf{W}_{\mathrm{OK}}},\overrightarrow{\mathbf{W}'_{\mathrm{SemK}}})}$. This concludes the proof.

TABLE II
EMPIRICAL STUDY OF LEMMA 1

Point Type	Point No.	Representative Spatial feature in the ontology	Similarity score	Assigned weigh-tage
Prediction point	—	Commercial	—	—
Interpolating points	1	Land with scrub	0.194	0.128
	2	Industrial	0.525	0.357
	3	Residential	0.623	0.515

Proposition 1: If the sample points are misclassified in the ontology hierarchy (i.e., represented by a wrong concept in the hierarchy), SemK will be able to capture this fact.

Proof: If there is any misclassification of sample points in the ontology hierarchy, it affects the SemK process by generating erroneous semantic covariance values for the sample points. If any point is misclassified as a less similar concept in the ontology with respect to the interpolation point, its assigned weight is less than optimal, and *vice versa*.

The SemK is capable of identifying this kind of misclassification of the sample points in the ontology hierarchy by some preprocessing steps. To check this property, a set of dummy points has been introduced, each of them is represented by one distinct leaf feature in the ontology. These points are assumed to have the same Euclidean distance from the prediction point. Thus, the weight assigned by SemK to each of the dummy points will be the function of "semantic" properties (i.e., semantic similarity and spatial importance) only. Considering the given scenario in Table I with two interpolating points, i.e., x_i and x_j, the semantic weight assigned to the ith dummy point (d_i) is given as $w_{d_i(f_i)}^{\mathrm{SemK}} = ((1/\mathrm{SIS}_{0j})-\lambda_{\mathrm{SemK}})$; f_i is the representative feature of d_i. If the ith interpolating point x_i is actually represented by the feature f_i, the normalized w_i^{SemK} is given as $w_{i(f_i)}^{\mathrm{SemK}} = ((\gamma(d_{0j})/\mathrm{SIS}_{0j}) - \lambda_{\mathrm{SemK}})$. However, if it is misclassified and misrepresented by another feature f_k, the w_i^{SemK} is given as follows:

$$w_i^{\mathrm{SemK}} = \left(\frac{\gamma(d_{0j})}{\mathrm{SIS}_{0k}} - \lambda_{\mathrm{SemK}}\right).$$

Thus, the assigned weight to the interpolating point x_i is λ_{SemK} less than $\gamma(d_{0j})$ times of SIS_{0k}, instead of SIS_{0j}. Hence, it can be concluded that the sample point is misclassified as feature f_k in the ontology, instead of feature f_i.

IV. EMPIRICAL STUDY

An experimentation has been carried out for empirical analysis of the proposed lemmas with LST data. These data are generated by the processing of USGS[2] Landsat ETM+ satellite imagery. The spatial resolution of these data is 30m (for bands 1–5 and 7) and 60m (for band 6). The satellite data are processed to generate the LST data. *West Bengal*, a state in India, has been considered as the spatial RoI for the empirical study.

In order to prove Lemma 1, a scenario with three interpolating points (as per the description of Lemma 1) is considered, which are equidistant from the prediction point. Their semantic properties are also measured in the range of (0, 1]. It is observed from Table II that the interpolating point with a higher semantic knowledge score has been assigned higher weightage by SemK. The result supports Lemma 1.

[2]U.S. Geological Survey (USGS): http://www.usgs.gov/.

TABLE III
EMPIRICAL STUDY OF LEMMA 2

Scenario	Point No.	Representative spatial feature in the ontology	Similarity score	Assigned weightage
1	1	Land with scrub	0.194	0.128
	2	Industrial	0.525	0.357
	3	Residential	0.623	0.515
	Semantic angular difference		**Weight vectors' angular difference**	
	25.48°		25.48°	
2	1	Crop land	0.570	0.629
	2	Residential	0.180	0.073
	3	Wetlands	0.344	0.297
	Semantic angular difference		**Weight vectors' angular difference**	
	34.40°		34.40°	
3	1	River	0.218	0.120
	2	Commercial	0.623	0.657
	3	Forest	0.315	0.223
	Semantic angular difference		**Weight vectors' angular difference**	
	34.88°		34.88°	

TABLE IV
EMPIRICAL STUDY OF LEMMA 3

Scenario	**Angular difference with the optimal weight vector**	
	OK	***SemK***
1	49.95°	29.90°
2	34.40°	5.45°
3	29.09°	10.70°

TABLE V
EMPIRICAL STUDY OF LEMMA 4

Scenario	**Angular difference with the weight vector of *OK***	
	***SemK* with initial ontology**	***SemK* with modified ontology**
1	25.48°	24.19°
2	39.68°	30.22°
3	34.88°	29.47°

For the analysis of Lemma 2, three scenarios have been considered, corresponding to the specifications provided in Table I. The semantic vectors and the assigned weight vectors for each of the scenarios are compared (refer to Table III). It is evident from Table III that the angular difference between the semantic knowledge captured by OK and SemK is exactly equal to the angle between their weight vectors.

The scenario of Table I has been considered for the empirical analysis of Lemma 3. An additional knowledge of the terrain, "*elevation of the earth surface*" is considered for further improving the prediction, along with the semantic knowledge. These three knowledge (Euclidean distance, semantic property, and elevation) are considered to generate the optimal weight vector. The angular differences between the optimal weight vector with the SemK and OK, respectively, are computed and specified in Table IV. This result supports Lemma 3.

For the empirical analysis of Lemma 4, the ontology in Fig. 1 is modified and all the *level-4* concepts (i.e., spatial features) are generalized to their respective *level-3* parent concepts. The performance of SemK with both the initial and the modified ontologies are analysed in accordance with Lemma 4 (refer to Table V). It proves that the less detailed ontology will eventually reduce SemK to OK.

For the overall performance analysis, the proposed SemK is compared with other prediction and spatial interpolation methods. Seven other popular prediction techniques are considered for comparison, namely, simple spatial averaging (Average), multilayer perceptron (MLP), Bayesian network (BN), nearest neighbors (NN), inverse distance weighting (IDW), simple cokriging (CK), and ordinary kriging (OK). The predicted value is compared with the actual value (obtained from the satellite imagery), and different error metrics are evaluated. The performance of each of the methodologies is compared with two standard error metrics, namely, *mean absolute error* (*MAE*) and *root mean square error* (*RMSE*) [4]. The corresponding results are presented in Table VI. It is observed that the SemK outperforms most of the methodologies, in terms of accuracy in prediction.

TABLE VI
COMPARATIVE STUDY OF SEMK WITH POPULAR PREDICTION METHODS

Prediction Method	**Error in Prediction**	
	MAE	**RMSE**
Average	1.19	1.32
MLP	1.31	1.37
BN	0.88	1.17
NN	1.33	1.58
IDW	1.39	1.53
CK	0.97	1.08
OK	0.94	1.25
SemK	**0.69**	**0.58**

V. CONCLUSION

The SemK [4] is a regression-based interpolation method, belonging to the family of kriging. It incorporates the semantic domain knowledge of the terrain for more accurate estimation of the spatial attributes. This letter focuses on Euclidean vector analysis-based performance evaluation of SemK (both theoretically and empirically) to demonstrate its usability and advantages over other popular prediction methods. This may facilitate further development of spatiotemporal prediction and forecasting framework for weather attributes.

REFERENCES

[1] S. K. Das and R. Singh, "Performance of kriging-based soft classification on WiFS/IRS-1D image using ground hyperspectral signatures," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 453–457, Jul. 2009.

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

[3] T. Hengl, G. Heuvelink, M. Perec Tadi, and E. Pebesma, "Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images," *Theoretical Appl. Climatology*, vol. 107, no. 1/2, pp. 265–277, Jan. 2012.

[4] S. Bhattacharjee, P. Mitra, and S. K. Ghosh, "Spatial interpolation to predict missing attributes in GIS using semantic kriging," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4771–4780, Aug. 2014.

[5] S. Bhattacharjee and S. K. Ghosh, "Automatic resolution of semantic heterogeneity in GIS: An ontology based approach," in *Advanced Computing, Networking and Informatics.* Vol. 1. ser. Smart Innovation, Systems and Technologies, vol. 27, New York, NY, USA: Springer, 2014, pp. 585–591.

[6] S. Patil, S. Bhattacharjee, and S. K. Ghosh, "A spatial web crawler for discovering geo-servers and semantic referencing with spatial features," in *Distributed Computing and Internet Technology*, Lecture Notes in Computer Science vol. 8337. New York, NY, USA: Springer, 2014, pp. 68–78.

[7] 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.

[8] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Econ. Geography*, vol. 46, pp. 234–240, Jun. 1970.

[9] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, vol. 1. Cambridge, U.K.: Cambridge Univ. Press, 2008.