Soil Organic Matter Estimation in Precision Agriculture using Wireless Sensor Networks

Linh Nguyen and Sarath Kodagoda
Centre for Autonomous Systems, Faculty of Engineering and Information Technology
University of Technology Sydney, Australia
Email: {vanlinh.nguyen,sarath.kodagoda}@uts.edu.au

***Abstract*—In order to achieve the ever increasing quantity and quality demands for agricultural products, technological innovations must be explored. In most cases, low cost and high quality sensing options are a top priority. This paper addresses the problem of predicting soil organic matter content in an agriculture field using information collected by a low-cost network of mobile, wireless and noisy sensors that can take discrete measurements in the environment. In this context, it is proposed that the spatial phenomenon of organic matter in soil to be monitored is modeled using Gaussian processes. The proposed model then enables the wireless sensor network to estimate the soil organic matter field at all unobserved locations of interest. The estimated values at predicted locations are highly comparable to those at corresponding points on a realistic image that is aerially taken by a very expensive and complex remote sensing system.**

I. Introduction

In agriculture production, precision farming is an emerging methodology that collects and processes intensive data and information on soil and crop conditions to make more efficient use of farm inputs such as fertilizers, herbicides, and pesticides. This leads to not only maximizing crop productivity and farm profitability but also minimizing environmental contamination [1]. Since cost of nitrogen fertilizer is relatively low and a small input can increase crop yields, many farmers tend to uniformly apply a large amount of nitrogen fertilizer to fields, resulting in potential for groundwater pollution [2]. Therefore, one of principal problems in precision agriculture is how to manage the nitrogen, which can also be supplied by mineralization of soil organic matter (SOM). In other words, there is a requirement to fully understand organic matter content and its spatial distribution in soil so that we can proportionally apply nitrogen fertilizer to the need in portions of the field, reducing over-application of the nitrogen fertilizer.

The most often utilized technique to observe the soil organic matter content is remote sensing, which gathers information about a phenomenon without making any physical contacts with it. There are two types of sensors in remote sensing systems, including passive and active. In monitoring soil and crop conditions, remote sensing is basically conducted from aerial and satellite platforms [3], and observed phenomena are represented by remotely sensed images [4]. Analyzing the observed images allows us to obtain spatial and spectral variations resulting from soil and crop characteristics. In the context of soil properties, SOM content has frequently been estimated from soil reflectance measurements by examining quantitative relationships between remotely sensed data and soil characteristics, focused on the reflective region of the spectrum (0.3 to 2.8 μm), with some relationships established from data in the thermal and microwave regions [5]. Recently, the work conducted by Bajwa *et al.* [6] demonstrates the potential of aerial visible/infrared (VIR) hyperspectral imagery for determining the SOM content, providing high spatial and spectral resolution.

Although remote sensing is considered as a promising approach to study organic matter content and its variability in soil, there still have several burdens that impede the adoption of this geographical technique for the nitrogen management. For instance, SOM content can be efficiently inferred from reflectance measurements if observations obtained in areas with moderate to high SOM levels, e.g. 10 to 15 g per kg [7] but low SOM levels since other soil factors may considerably affect the reflectance. Moreover, the reflectance based method is not really effective over large geographic areas owing to confounding impacts of nature such as moisture and underlying parent material [8], extensive plant canopy over a region [9] and variations in surface roughness [10] and vegetation [11]. Accuracy of estimating SOM content is questionable where surface features confuse spectral responses [8]. And cloud cover conditions probably influence quality of remotely sensed color photographs [12]. On the other hand, when considering small areas, the imagery must be of high spatial resolution. Such aerial or satellite images are either unavailable or fairly expensive [13]. More importantly, processing those high resolution imagery is computationally expensive, which restricts the feasibility of usage.

Recently, technological developments in micro-electro-mechanical systems and wireless communications, which involve substantial evolution in reducing the size and the cost of components, have led to the emergence of wireless sensor networks (WSN) that are increasingly useful in crucial applications in environmental monitoring [14]. WSN can be employed to enhance our understanding of environmental phenomena and direct natural resource management. It is proven that the networks of wireless sensors are very appealing and promising for supporting agriculture practices [15]. For instance, wireless sensor nodes are deployed in greenhouses and gardens [16] to gauge information of environmental parameters such as temperature, relative humidity and light intensity that significantly influence the development of the agricultural crops.

Based on measurements gathered by the large-scale WSN, Langendoen *et al.* in [17] designed an optimal control system that can be utilized to adjust environmental quantities for the purpose of obtaining better production yields and minimizing use of resources. Furthermore, the WSN have been used to track animals. In [18], Butler *et al.* proposed a moving virtual fence method to control cow herd, based on a wireless system. To respond requirements to constantly monitor the conditions of individual animals, a WSN based system is designed to generally monitor animal health and locate any animals that are sick and can infect the others [19]. In the context of soil science, a farm based network of wireless sensors has been developed to assess soil moisture and soil temperature as demonstrated in [20].

In fact, not only do these systems provide a virtual connection with the physical field in general, the WSN can be utilized for developing optimal strategies for crop production. In [21], Hokozono *et al.* have employed the sensed data to study variability of environmental effects, which then influences the conversion from conventional to organic and sustainable crop production. Furthermore, real time information from the fields gathered by the WSN is really helpful for farmers to minimize potential risks in crop production by controlling their production strategies at any time, without using a tractor or any other vehicles to collect each sampling point [22]. More particularly, in addition to collecting the data, combining the measurements with a model, a wireless sensor network is also competent to estimate and predict the spatial phenomenon at unobserved locations. This interesting attribute enables the WSN to create a continuous surface by employing the set of measurements collected at discrete points to interpolate the physical field at unobserved locations. The more number of predicted points is, the more accurate the predictions of the resulting surface are as compared with the remotely sensed image.

In order to enhance the accuracy of the predicted field, it is essential to accurately model the spatial phenomena. Usually, the physical processes are described by deterministic and data-driven models [23]. The prime disadvantage of the deterministic model is that it requires model parameters and initial conditions to be known in advance. Furthermore, model complexity and various interactions in the deterministic models that are difficult to model tilt the balance in favor of data-driven approaches. In this work, it is particularly proposed to consider the Gaussian process data-driven model [24]–[26] to statistically model spatial fields. The use of a Gaussian process (GP) allows prediction of the environmental phenomena of interest effectively at any unobserved points.

Upon analysis above, it can be clearly seen that the use of remote sensing technique to monitor and estimate SOM content is costly, complicated and particularly impractical in areas with significant vegetation and litter cover. As a consequence, in this work we propose to utilize the low-cost WSN to discretely take SOM measurements at predefined locations and then use the GP to statistically predict the SOM field at the rest of space from the observations available. The resulting prediction surfaces of the SOM content at studied areas are highly comparable to the imagery obtained by the aerial or satellite platforms.

The structure of the paper is arranged as follows. Section II introduces wireless sensor networks for monitoring the SOM content and dataset that is used to conduct the experiments. The spatial field model and the interpolation technique are also presented in this section. Section III describes the experiments and discusses the results before conclusions are delineated in Section IV.

II. Materials and Methods

In this section, we first present structure of a wireless sensor network, a data set and we then discuss the spatial prediction approach utilized in this work. For simplicity, we define notations as follows. Let $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ denote the set of real and nonnegative real numbers. The Euclidean distance function is defined by $\|\cdot\|$. Let $\mathbb{E}$ denote the operator of the expectation and $\mathrm{tr}(\cdot)$ denote trace of a matrix. Other notations will be explained as and when they occur.

A. Wireless Sensor Network and Dataset

1) Wireless Sensor Network: A wireless sensor network is specifically composed of multiple autonomous, small size, low cost, low power and multifunctional sensor nodes. Each node can communicate untethered in short distances. These tiny sensor nodes could be equipped with various types of sensing devices such as temperature, humidity, chemical, thermal, acoustic and optical sensors. Therefore, by positioning the individual sensors inside or very close to the phenomenon, the sensor nodes not only measure it but also transmit the data to a central node that is also known as the base station or the sink. A unique feature of sensor nodes is that each is embedded with an on-board processor. In addition to controlling all activities on the board, the processor is responsible for locally conducting simple pre-computation of the raw measurements before sending the required or partially processed data to the sink. The pre-processing aims to enhance the energy conservation and reduce communicating time.

By carefully engineering the communication topology, a sensor node can communicate with others or a base station based on a routing structure. The wireless communication technology widely utilized in sensor networks is the ZigBee standard. ZigBee is a suite of high-level communication protocols that uses small, low-power digital radios based on the IEEE 802.15.4 standard for wireless area networks [28]. In a small-scale network, each node directly transmit its data to the sink, which is called single hop communication. Nevertheless, the single hop transmission is inefficient in a large-scale network, where transmission energy expense is exponential of a transmitting distance. Hence, the multihop communication in which the data is transmitted to sensor nodes' neighbors in multiple times before reaching the sink is practically feasible. A typical multihop wireless sensor network architecture is demonstrated in Fig. 1. On the other hand, Fig. 1 also illustrates another efficient solution for communication in a large-scale

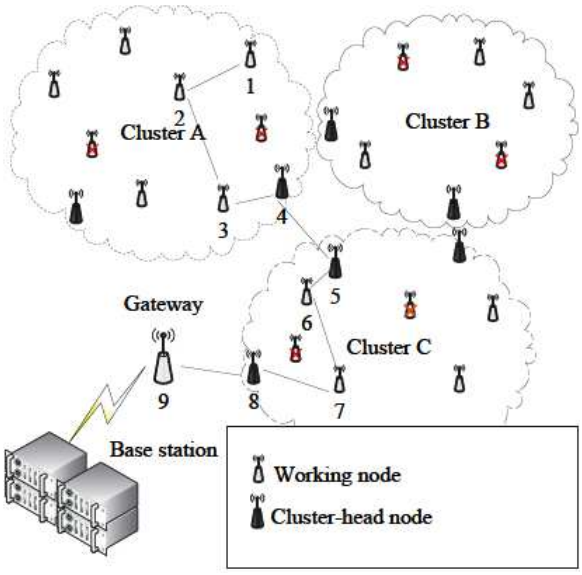

Fig. 1: Wireless sensor network structure [27].

network. In this configuration, the network is organized by clusters; and each cluster-head node aggregates data from all the sensors within its cluster and transmits to the sink.

After gathering measurements from all sensor nodes, the base station performs computations and fuses the data before making decision about the phenomenon.

2) Dataset: In order to illustrate our proposed approach as compared with the remote sensing technique, we conducted experiments using published data sets that were collected from a real-world field in Benton county, Indiana, USA [29]. In the work [29], a hyperintensive aerial photograph of the field was taken by a digital camera from an airplane flying at a height of 1219 m. After analyzing the raw data, imagery of soil organic matter contents calculated in percentage were created. For the purpose of comparisons, in this work, we suppose that sensors can take the soil organic matter content measurements at locations on imagery maps published in [29].

B. Spatial Field Model for Soil Organic Matter

In this section, we introduce the dominant concepts and properties on the spatial field model that are used in this paper. We refer the interested readers to [26] for further details.

Consider the spatial field of interest $\mathcal{Q} \subset \mathbb{R}^d$, we let spatial locations within $\mathcal{Q}$ denote as $v = (v_1^T, v_2^T, ..., v_n^T)^T \in \mathbb{R}^{dn}$. The data consists of one measurement taken at each observed location in v. Let a random vector $y(v)$ denoted by $y(v) = (y(v_1), y(v_2), ..., y(v_n))^T \in \mathbb{R}^n$ describe a vector of measurements. In this study, it is supposed that $v_i, i = 1, ..., n$ varies continuously through $\mathcal{Q}$. The spatial field model is a summation of a large scale component, a random field and a noise. The noise is supposed to be independent and identically distributed (i.i.d.). Hence, the model is defined by

$$y(v_i) = X(v_i)\beta + \xi(v_i) + \varepsilon(v_i), \tag{1}$$

where

- $X(v_i)\beta$ is the expectation of $y(v_i)$, which is also referred to as a spatial trend function;
- $\xi(v_i) \sim \mathcal{N}(0, cov(v_i, v_j))$ is a Gaussian process that will be presented in the following;
- $\varepsilon(v_i)$ is a noise with a zero mean and an unknown variance τ^2.

The expectation of $y(v_i)$ in the model (1) is frequently derived through a polynomial regression model, for example a constant, first, or second order polynomial function. Here, $X(v_i)$ is given by $X(v_i) = (1, X_1(v_i), ..., X_{p-1}(v_i)) \in \mathbb{R}^p$, a spatially referenced non-random variable (known as covariate) at location v_i. And $\beta = (\beta_0, \beta_1, ..., \beta_{p-1})^T$ is an unknown vector of mean parameters. For instance, it is assumed that $v_i \in \mathbb{R}^2$, that is $v_i = (v_{i1}, v_{i2})$, the second order polynomial expectation is dependent on the coordinates of a sensing location, specified by

$$X(v_i)\beta = \beta_0 + \beta_1 v_{i1} + \beta_2 v_{i2} + \beta_3 v_{i1}^2 + \beta_4 v_{i2}^2 + \beta_5 v_{i1} v_{i2}. \tag{2}$$

In this case, $X(v_i) = (1, v_{i1}, v_{i2}, v_{i1}^2, v_{i2}^2, v_{i1}v_{i2})$ and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$.

Gaussian process: A Gaussian process (GP) is a very popular non-parametric Bayesian technique for modeling spatially correlated data. Initially known as *kriging*, the technique has its roots in geostatistics where it is mainly used for estimation of mineral resources [30]. The Gaussian processes (GPs) extend multivariate Gaussian distributions over a finite vector space to function space of infinite dimensionality.

Consider a spatial location $v_i \in \mathbb{R}^d$, a random variable $z(v_i)$ at v_i is modeled as a GP and written as

$$z(v_i) \sim \mathcal{GP}(\mu(v_i), cov(v_i, v_j)), \tag{3}$$

where $v_i, v_j \in \mathbb{R}^d$ are the inputs. $\mu(v_i)$ is a mean function and $cov(v_i, v_j)$ is a covariance function, often called a kernel function. These functions are defined as

$$\mu(v_i) = \mathbb{E}[z(v_i)],$$
$$cov(v_i, v_j) = \mathbb{E}[(z(v_i) - \mu(v_i))(z(v_j) - \mu(v_j))]$$

A spatial GP is stationary if $cov(v_i, v_j) = cov(v_i - v_j)$. That is, the covariance depends only on the vector difference between v_i and v_j. Furthermore, if $cov(v_i, v_j) = cov(\|v_i - v_j\|)$, the stationary process is isotropic. Hence, the covariance between a pair of variables of $z(v_i)$ at any two locations is only dependent on the distance between them.

The covariance function is a vital ingredient in a GP. In fact, there is a practical family of parametric covariance functions proposed in [31]. For example, one of the frequently used kernel functions is squared exponential, that is,

$$cov(v_i, v_j) = \sigma^2 \exp\left(-\frac{\|v_i - v_j\|^2}{2\phi^2}\right), \tag{4}$$

where

- σ^2 is the marginal variance (also known as the maximum allowable covariance);
- ϕ is the range parameter (also called the length scale) that is referred to as the reduction rate of the correlation between $z(v_i)$ and $z(v_j)$ when $\|v_i - v_j\|$ increases.

The parameters of σ^2 and ϕ can be varied; therefore, these parameters are referred to as hyperparameters since they correspond to the hyperparameters in neural networks.

C. Spatial Inference

After introducing the spatial field model, we now delineate the regression technique, which is utilized to predict continuous quantities of the physical process.

Consider a data set of n observations $\mathcal{D} = \{(v_i, y_i)|i = 1, ..., n\}$ collected by the wireless sensor network, where v_i is a location vector of dimension d and y_i is a scalar value of noise corrupted output. The corresponding vector of noise-free observations is referred to as $z = (z(v_1), z(v_2), ..., z(v_n))^T \in \mathbb{R}^n$. As discussed in Section II-B, the prior z can be described as

$$z \sim \mathcal{N}(\mu, \Sigma_{zz}), \tag{5}$$

where $\mu \in \mathbb{R}^n$ is the mean vector obtained by $\mu_i = \mu(v_i)$, and Σ_{zz} is an $n \times n$ covariance matrix whose elements can be computed by $\Sigma_{zz}[i, j] = cov(v_i, v_j)$. By the use of the spatial field model presented in (1), the mean value at each v_i can be obtained by

$$\mu_i = X(v_i)\beta.$$

The advantage of the GP formulation is that the combination of the prior and noise can be implemented exactly by matrix operations [32]. Therefore, the noisy observations can be normally distributed as

$$y \sim \mathcal{N}(z, \tau^2 I), \tag{6}$$

where τ^2 is a noise variance and I is an $n \times n$ identity matrix. Note that the GP models and all formulas are always conditional on the corresponding locations. In the following, the explicit conditioning on the matrix v will always be neglected.

Given the observations, the objective of probabilistic regression is to compute the prediction of the real values $z_* = z(v_*)$ at m interested points v_*. In [25], Rasmussen *et al.* demonstrated that the GP has a marginalization property, which implies that the joint distribution on random variables at v and v_* is Gaussian, specified by

$$\begin{bmatrix} y \\ z_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} X(v)\beta \\ X(v_*)\beta \end{bmatrix}, \begin{bmatrix} \Sigma_{zz} + \tau^2 I & \Sigma_{zz_*} \\ \Sigma_{z_* z} & \Sigma_{z_* z_*} \end{bmatrix} \right), \tag{7}$$

where $X(v)$ and $X(v_*)$ are $n \times p$ and $m \times p$ matrices of covariates, respectively. Then $X(v)\beta$ and $X(v_*)\beta$ are the mean vectors of y and z_*. $\Sigma_{z_* z_*}$ is the covariance matrix of z_*. $\Sigma_{zz_*} (= \Sigma_{z_* z}^T)$ is the cross-covariance matrix between y and z_*.

In probabilistic terms, the conditional distribution at predicted positions of v_* given y is derived as follows.

$$\mu_{z_*|y} = X(v_*)\beta + \Sigma_{z_* z}(\Sigma_{zz} + \tau^2 I)^{-1}(y - X(v)\beta), \tag{8}$$

and

$$\Sigma_{z_*|y} = \Sigma_{z_* z_*} - \Sigma_{z_* z}(\Sigma_{zz} + \tau^2 I)^{-1}\Sigma_{zz_*}, \tag{9}$$

where $\mu_{z_*|y}$ and $\Sigma_{z_*|y}$ are posterior mean vector and covariance matrix of z_*, given y. As a consequence, using observations at locations in set v, quantities at unobserved locations, v_*, can be predicted. Nonetheless, in order to practically implement the full inference, all of the mean parameters β and hyperparameters σ^2, ϕ, and τ^2 are required to be known; hence the estimations are primarily discussed in the next subsection.

D. Parameter Estimation

Let $\theta = (\sigma^2, \phi, \tau^2) \in \mathbb{R}^3_{\geq 0}$ denote a hyperparameter vector. The mean parameters β and hyperparameters θ that are hereafter called model parameters of the spatial field model can be estimated by utilizing generalized least squares technique [24] and the maximum likelihood approach [26]. In the following, a recursive algorithm for estimating the mean parameters β and hyperparameters θ is delineated. Rewriting the marginal distribution of $y(v)$ given model parameters yields

$$y(v)|\sigma^2, \phi, \tau^2, \beta \sim \mathcal{N}(X(v)\beta, \Sigma_{zz} + \tau^2 I), \tag{10}$$

For the sake of simplicity, it is denoted $\Sigma = \Sigma_{zz} + \tau^2 I$.

First, in the best linear unbiased estimator [24], β can be obtained by minimizing the function

$$f(\beta) = (y(v) - X(v)\beta)^T \Sigma^{-1} (y(v) - X(v)\beta).$$

If given θ, i.e. Σ is known, the estimated β can be specified by

$$\hat{\beta} = (X(v)^T \Sigma^{-1} X(v))^{-1} X(v)^T \Sigma^{-1} y(v). \tag{11}$$

Second, from (10) the log-likelihood function can be obtained by

$$\mathcal{L}(\theta, \beta) = -\frac{1}{2}\{(y(v) - X(v)\beta)^T \Sigma^{-1} (y(v) - X(v)\beta) + \\ + \log\det(\Sigma) + n\log(2\pi)\}. \tag{12}$$

By substituting $\hat{\beta}$ into the log-likelihood function and numerically optimizing this function with respect to σ^2, ϕ and τ^2, the estimated $\hat{\theta}$ can be obtained. Eventually, $\hat{\beta}$ can be computed by the back substitution of $\hat{\theta}$.

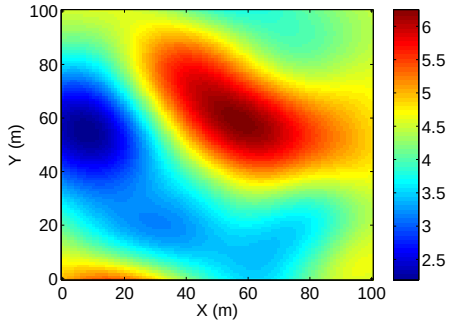

Fig. 2: The true field of the soil organic matter content. Percentage of the soil organic matter content is shown in color bar.

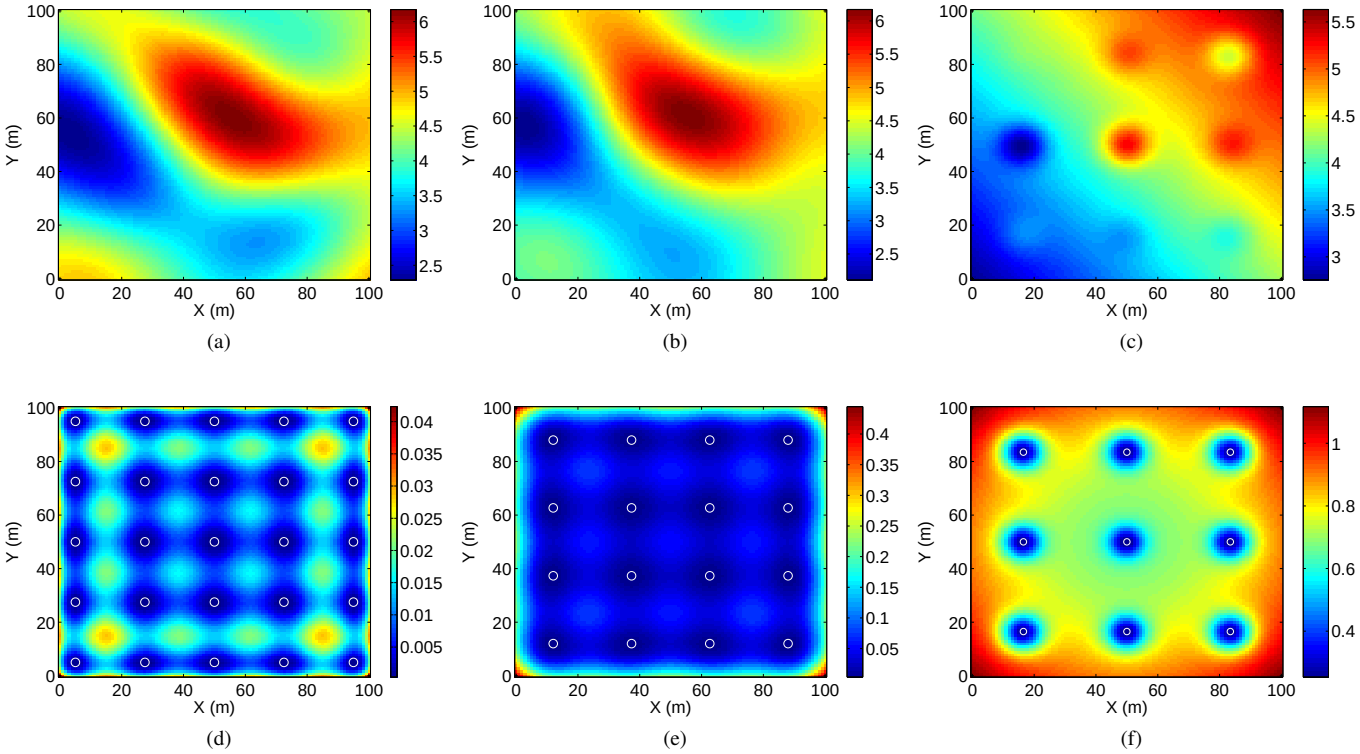

Fig. 3: The predicted fields and the predicted error variances of the SOM content using (a) and (d) 25, (b) and (e) 16, and (c) and (f) 9 sensors. The positions of sensor nodes are illustrated by white circles.

Notice that in order to optimize the log-likelihood function, the partial derivative can be specified by

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = -\frac{1}{2} tr\left((\alpha\alpha^T - \Sigma^{-1}) \frac{\partial \Sigma}{\partial \theta_i} \right),$$

where $\alpha = \Sigma^{-1}(y(v) - X(v)\beta)$, and θ_i is σ^2, ϕ and τ^2.

III. Results and Discussions

In this section, we provide experimental performances of our proposed approach on predicting the soil organic matter content for whole space of interest using a specific number of measurements collected by a wireless sensor network. As described in Section II-A2, the ground truth of the soil organic matter content in area of 100 m $\times$ 100 m was reconstructed as shown in Fig. 2. And then a network of wireless sensors was deployed by a grid in the selected area. In the illustrated experiments, 25, 16 and 9 sensing nodes were positioned at white circles in Figures 3d, 3e and 3f, respectively.

All the sensors make observations and transmit them to the sink via a specific routing tree. Then the base station estimates the mean parameters and hyperparameters for the Gaussian process model of the soil organic matter content. Based on the learned model, the estimated values of soil organic matter field at all unobserved locations of interest can be effectively predicted. In the implementations, we carried out the resulting predictions of means and error variances for whole space of 100 m $\times$ 100 m area. Note that the experiments were implemented in two dimensional environments.

Fig. 3 demonstrate the posterior means and posterior variances of the soil organic matter content, predicted for whole studied area. While Figures 3a and 3d show the predicted results using 25 SOM observations, pairs of Figures 3b and 3e, 3c and 3f illustrate resulting means and variances using SOM measurements gauged by 16 and 9 sensor nodes, respectively. It can be apparently seen that the more numbers of sensing devices are, the more accurate the resulting predictions of the SOM content are. In equivalent words, when 25 SOM sensors are in use, as deployed in Fig. 3d, the snapshot of the surface of the SOM content predicted in whole space of 100 m $\times$ 100 m in Fig. 3a is very close to the real image that represents the SOM in the same area obtained by the remote sensing technique, shown in Fig. 2. Moreover, even when we experimented with only 16 measuring devices positioned at white circles in Fig. 3e, the predicted means of the SOM field demonstrated in Fig. 3b are highly comparable with the ground truth illustrated in Fig. 2. As expected, quality of prediction degrades with fewer number of sensors, for example as in Fig.3c. Nonetheless, patterns corresponding to the SOM content values in 3c are clearly classified as compared with those in Fig. 2. In the context of variances, it can be clearly seen that the accuracy of the predictions is dependent on numbers of sensors participating in sensing task. And, the prediction errors at locations in the range around the sensor nodes are trivial.

More importantly, to evaluate the quality of prediction in

the case studied we investigated the root mean square errors (RMSE) of the predicted field at M spatial locations of interest, which are based on RMSE $= \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\mu_{z|y}[i] - z[i]\right)^2}$, where z is a vector of the values actually observed, and $\mu_{z|y}$ is a vector of predicted means at interested positions given observations y. It can be clearly seen in Fig. 4 that the RMSE gradually reduce with increased number of observations. Thus, given a required accuracy of the predictions, projecting that value to the RMSE curve, a number of sensors and their locations can also be chosen for a network.

IV. Conclusions

The paper has presented a Gaussian process based inference approach to estimate the spatial soil organic matter content using measurements gathered by a wireless sensor network. A publicly available data set was used for validation. The SOM predicted by our low-cost GP based algorithm is highly comparable with expensive remote sensing system results. The proposed method has potential for applying to precision agriculture, where management of nitrogen levels are required. Our system also allows farmers to choose a number of sensing nodes, corresponding to their expected prediction accuracy. In future work, we will concentrate on finding optimal locations to deploy sensors.

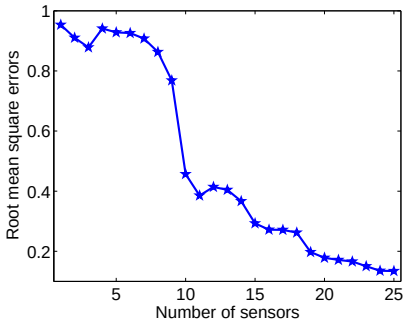

Fig. 4: Root mean square errors.

References

[1] T. Harmon, C. Kvien, D. Mulla, G. Hoggenboom, J. Judy, and J. Hook, "Precision agriculture scenario," in *Proc. NSF Workshop on Sensors for Environmental Observatories*, Baltimore, MD, USA, 2005.

[2] A. Schepers, *Comparison of GIS approaches that integrate soil and crop variables to delineate management zones for precision agriculture*. Masters Thesis, Department of Geography, University of Nebraska-Lincoln, 2002.

[3] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosystems engineering*, vol. 114, pp. 358–371, 2013.

[4] D. P. Paine and J. D. Kiser, *Aerial photography and image interpretation*. Wiley, 2012.

[5] M. Chellasamy, R. T. Zielinski, and M. H. Greve, "A multievidence approach for crop discrimination using multitemporal worldview-2 imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7(8), pp. 3491–3501, 2014.

[6] S. G. Bajwa and L. F. Tian, "Soil fertility characterization in agricultural fields using hyperspectral remote sensing," *Transactions of the ASAE*, vol. 48(6), pp. 2399–2406, 2005.

[7] D. G. Sullivan, J. N. Shaw, D. Rickman, P. L. Mask, and J. Luvall, "Using remote sensing data to evaluate surface soil properties in alabama ultisols," *Soil Science*, vol. 170, pp. 954–968, 2005.

[8] J. W. Hummel, K. A. Sudduth, and S. E. Hollinger, "Soil moisture and organic matter prediction of surface and subsurface soils using an nir soil sensor," *Computers and Electronics in Agriculture*, vol. 32, pp. 149–165, 2001.

[9] P. Kongapai, *Application of remote sensing and geographic information system for estimation of soil organic matter in Nakhon Pathom Province*. Master Thesis, Mahidol University, Thailand, 2007.

[10] A. D. Matthias, A. Fimbres, E. E. Sano, D. F. Post, L. Accioly, A. K. Batchily, and L. G. Ferreira, "Surface roughness eff ects on soil albedo," *Soil Science Society of America Journal*, vol. 64, pp. 1035–1041, 2000.

[11] J. P. Walker, P. R. Houser, and G. R. Willgoose, "Active microwave remote sensing for soil moisture measurement: A field evaluation on using ers-2," *Hydrological Process*, vol. 1811, pp. 1975–1997, 2004.

[12] M. D. Nellis, K. P. Price, and D. Rundquist, "Remote sensing of cropland agriculture," *Papers in Natural Resources*, pp. 217–245, 2009.

[13] A. Bannari, K. S. Pacheco, H. McNairn, and K. Omari, "Estimating and mapping crop residues cover on agricultural lands using hyperspectral and ikonos data," *Remote Sensing of Environment*, vol. 104, pp. 447–459, 2006.

[14] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.

[15] L. Ruiz-Garcia, L. Lunadei, P. Barreiro, and I. Robla, "A review of wireless sensor technologies and applications in agriculture and food industry: state of the art and current trends," *Sensors*, vol. 9(6), pp. 4728–4750, 2009.

[16] K. Kim, J. Kim, K. Ban, E. Kim, and M. Jang, "U-it based greenhouse environment monitoring system," in *Proc. FTRA International Conference on Multimedia and Ubiquitous Engineering*, Crete, Greece, 2011, pp. 203–206.

[17] K. Langendoen, A. Baggio, and O. Visser, "Experiences from a pilot sensor network deployment in precision agriculture," in *Proc. International Parallel and Distributed Processing Symposium*, Rhodes Island, 2006, pp. 8–16.

[18] Z. Butler, P. Corke, R. Peterson, and D. Rus, "Virtual fences for controlling cows," in *Proc. IEEE International Conferece on Robotics and Automation*, New Orlean, USA, April 2004, pp. 4429–4436.

[19] D. Davcev and J. M. Gomez, *ICT Inovations*. Springer, 2009.

[20] P. Sikka, P. Corke, P. Valencia, C. Crossman, D. Swain, and G. Bishop-Hurley, "Wireless ad hoc sensor and actuator networks on the farm," in *Proc. International Conference on Information Processing in Sensor Networks*, Nashville, USA, 2006, pp. 492–499.

[21] S. Hokozono and K. Hayashi, "Variability in environmental impacts during conversion from conventional to organic farming: a comparison among three rice production systems," *Journal of Cleaner Production*, vol. 28, pp. 101–112, 2012.

[22] D. D. Wu, D. L. Olson, and J. R. Birge, "Risk management in cleaner production," *Journal of Cleaner Production*, vol. 53, pp. 1–6, 2013.

[23] R. Graham and J. Cortes, "Spatial statistics and distributed estimation by robotic sensor network," in *Proc. IEEE American Control Conference*, Baltimore, MD, USA, 2010, pp. 2422–2427.

[24] N. A. Cressie, *Statistics for spatial data*. Wiley, 1991.

[25] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. The MIT Press, Cambridge, Massachusetts, London, England, 2006.

[26] P. J. Diggle and P. J. Ribeiro, *Model-based geostatistics*. Springer, New York, USA, 2007.

[27] M. A. Alsheikh, S. Lin, D. Niyato, and H. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Communications Surveys and Tutorials*, vol. 16(4), pp. 1996–2018, 2014.

[28] M. Kuorilehto, M. Kohvakka, J. Suhonen, P. Hmlinen, M. Hnnikinen, and T. D. Hamalainen, *Ultra-low energy wireless sensor networks in practice: Theory, realization, and deployment*. John Wiley and Sons, 2007.

[29] D. Mulla, M. Beatty, and A. Sekely, "Evaluation of remote sensing and targeted soil sampling for variable rate application of nitrogen," in *Proc. 5th International Conference on Precision Agriculture*, America, 2001.

[30] P. Tahmasebi and M. Sahimi, "Enhancing multiple-point geostatistical modeling: 1. graph theory and pattern adjustment," *Water Resources Research*, vol. 52(3), pp. 2074–2098, 2016.

[31] J. P. Chilès and P. Delfiner, *Geostatistics: Modelling spatial uncertainty*. Wiley, 1999.

[32] C. M. Bishop, *Pattern recognition and machine learning*. Springer, New York, USA, 2006.