

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Multivariate versus Univariate Sensor Selection for Spatial Field Estimation

1<sup>st</sup> Linh Nguyen

*School of Engineering, IT and Physical Sciences*  
*Federation University Australia*  
VIC, Australia  
l.nguyen@federation.edu.au

3<sup>rd</sup> Nalika Ulapane

*School of Health Sciences*  
*Swinburne University of Technology*  
VIC, Australia  
<https://orcid.org/0000-0003-3432-3943>

2<sup>nd</sup> Karthick Thiyagarajan

*UTS Robotics Institute*  
*University of Technology Sydney*  
NSW, Australia  
karthick.thiyagarajan@uts.edu.au

4<sup>th</sup> Sarath Kodagoda

*UTS Robotics Institute*  
*University of Technology Sydney*  
NSW, Australia  
sarath.kodagoda@uts.edu.au

**Abstract**—The paper discusses the sensor selection problem in estimating spatial fields. It is demonstrated that selecting a subset of sensors depends on modelling spatial processes. It is first proposed to exploit Gaussian process (GP) to model a univariate spatial field and multivariate GP (MGP) to jointly represent multivariate spatial phenomena. A Matérn cross-covariance function is employed in the MGP model to guarantee its cross-covariance matrices to be positive semi-definite. We then consider two corresponding *univariate* and *multivariate sensor selection* problems in effectively monitoring multiple spatial random fields. The sensor selection approaches were implemented in the real-world experiments and their performances were compared. Difference of results obtained by the univariate and multivariate sensor selection techniques is insignificant; that is, either of the methods can be efficiently used in practice.

**Index Terms**—Sensor selection, multivariate, univariate, spatial fields, Gaussian process, multimodal sensing, multivariable

## I. INTRODUCTION

Sensor selection for estimating spatial fields, which we call spatial sensor selection, is a fundamental but critical problem in various monitoring applications [1]. In practice, spatial sensors, monitoring spatial phenomena, can be excessively deployed in a sensing task partly due to their low cost, which may lead to an over-sampling issue in a system. The over-sampling problem is defined as multiple sensors deployed maximally recording quite similar data causing redundancy in sensor measurements. The redundancy is operationally expensive since it requires more memory, computation, communication, maintenance and energy resources while redundant data does not contribute any additional information to a sensing system. Therefore, selecting the most informative sensors out of all possible ones is an important problem to address.

The critical crux in the spatial sensor selection is that selected sensors are able to not only effectively observe spatial phenomena but also efficiently predict the fields at any unmeasured locations of interest [2]. In other words, selected sensors are equipped by a spatial process model to represent

spatial fields and to predict the fields at unobserved locations. The argument leads to the spatial sensor selection problem that selects the most informative sensors through minimizing prediction uncertainties at predicted locations. Thus, spatial modelling is important in the spatial sensor selection formulation.

Single or univariate spatial process, e.g. indoor temperature, can be mathematically modelled through a statistical model such as Gaussian process (GP) [3]. And, the sensor selection problem in monitoring a univariate spatial field, which is also called the *univariate sensor selection*, has been well studied [4]. In cases there exist multiple or multivariate spatial processes in the same environment, e.g. indoor temperature and humidity, they may have cross-correlation [5]. If those spatial phenomena are simultaneously observed, presence of their cross-correlation may influence on results of the spatial sensor selection for monitoring multivariate spatial fields, which here we define as the *multivariate sensor selection*. Since, to the best of our knowledge, the multivariate sensor selection has not been well studied yet, does it outperform the univariate sensor selection? In this paper, we discuss comparison of spatial prediction results correspondingly obtained by both the approaches.

We first exploit GP to model a univariate spatial field. To represent multivariate spatial phenomena we employ *multivariate Gaussian process* (MGP) [6] with a Matérn cross-covariance function that was proved to produce valid cross-correlation between multiple spatial processes [7]. The spatial sensor selection approaches were implemented in the real-life experiments where the obtained results demonstrate their effectiveness. The remaining of the paper is arranged as follows. We introduce the univariate and multivariate models in Section II before discussing the spatial sensor selection problems in Section III. Section IV summarizes our experimental results with discussions before conclusions are drawn in Section V.

## II. UNIVARIATE AND MULTIVARIATE MODELS

In order to formulate the multivariate and univariate sensor selection problems for finding the most informative sensor nodes in efficiently estimating spatial fields, this section introduces the corresponding multivariate and univariate models that can be employed to represent spatial processes given a limited number of sensor measurements and then predict them at unmeasured locations of interest.

### A. Univariate Spatial Prediction

In univariate scenarios, we only consider a single spatial random phenomenon, e.g. temperature [8]. It is assumed that  $n$  sensors positioned at locations  $\mathbf{s} = (\mathbf{s}_1^T, \dots, \mathbf{s}_n^T)^T \in \mathbb{R}^{n \times d}$ , where  $d$  is dimension of the environment, are used to measure that spatial process at their locations. A measurement  $y_i$  collected by an  $i^{\text{th}}$  sensor can be mathematically modelled by

$$y_i = \mathbf{x}(\mathbf{s}_i)\beta + w(\mathbf{s}_i) + \varepsilon(\mathbf{s}_i), \quad (1)$$

where  $\mathbf{x}(\mathbf{s}_i)$  is a vector of the spatially referenced covariates while  $\beta$  is a vector of the trend parameters. The  $\mathbf{x}(\mathbf{s}_i)\beta$  presents trend of the spatial phenomenon at locations  $\mathbf{s}_i$ . The spatial trend can be constant, first or second orders and chosen by users. For instance, if we select the second order trend, then  $\mathbf{x}(\mathbf{s}_i)$  and  $\beta$  are a  $1 \times 6$  and a  $6 \times 1$  vectors, respectively.  $w(\mathbf{s}_i)$  is a latent random function that is proposed to be represented by a zero-centred Gaussian process (GP) while  $\varepsilon(\mathbf{s}_i)$  is a normally distributed independent and identically distributed noise with a zero mean and an unknown variance  $\tau^2$ .

Let  $\mathbf{y} = (y_1, \dots, y_n)^T$ ,  $\mathbf{w}(\mathbf{s}) = (w(\mathbf{s}_1), \dots, w(\mathbf{s}_n))^T$  and  $\mathbf{x}(\mathbf{s}) = (\mathbf{x}(\mathbf{s}_1)^T, \dots, \mathbf{x}(\mathbf{s}_n)^T)^T$ . If it is assumed that there are  $p$  unobserved locations of interest in the environment  $\mathbf{z} = (\mathbf{z}_1^T, \dots, \mathbf{z}_p^T)^T \in \mathbb{R}^{p \times d}$ , then their corresponding variables can be predicted through a posterior normal distribution where its posterior mean and covariance matrix can be computed by,

$$\mu_{\mathbf{z}|\mathbf{s}} = \mathbf{x}(\mathbf{z})\beta + \Sigma_{\mathbf{z}\mathbf{s}}(\Sigma + \tau^2 I)^{-1}(\mathbf{y} - \mathbf{x}(\mathbf{s})\beta), \quad (2)$$

$$\Sigma_{\mathbf{z}|\mathbf{s}} = \Sigma_{\mathbf{z}\mathbf{z}} - \Sigma_{\mathbf{z}\mathbf{s}}(\Sigma + \tau^2 I)^{-1}\Sigma_{\mathbf{s}\mathbf{z}}^T, \quad (3)$$

where  $\mathbf{x}(\mathbf{z})$  is a covariate matrix corresponding to  $\mathbf{z}$  while  $I$  is an  $n \times n$  identity matrix. Both  $\Sigma$  and  $\Sigma_{\mathbf{z}\mathbf{z}}$  are covariance matrices of random variables at locations  $\mathbf{s}$  and  $\mathbf{z}$ , respectively.  $\Sigma_{\mathbf{z}\mathbf{s}}$  is an  $p \times n$  cross-covariance matrix between random variables of a single spatial process presented at  $\mathbf{z}$  and  $\mathbf{s}$ . It is noted that the covariance and cross-covariance matrices  $\Sigma$ ,  $\Sigma_{\mathbf{z}\mathbf{z}}$  and  $\Sigma_{\mathbf{z}\mathbf{s}}$  can be calculated by the use of a covariance function such as the squared exponential function [2], where its hyperparameters can be estimated by the maximum likelihood method [9] given collected sensor measurements.

### B. Multivariate Spatial Prediction

In some practical scenarios, several spatial phenomena are monitored simultaneously [5]. Though each spatially distributed process can be modelled individually, cross-correlation between them is ubiquitous [10]. In those cases, we expect to model multiple spatial fields jointly.

It is assumed that we have  $m$  different types of sensors to monitor  $m$  spatial phenomena that the different types of sensors are co-located.  $n$  sensors for each type are positioned at locations  $\mathbf{s}$ , and all the collected measurements are mathematically modelled by

$$\mathbf{Y} = \mathbf{X}(\mathbf{s})\beta + \mathbf{W}(\mathbf{s}) + \varepsilon(\mathbf{s}), \quad (4)$$

where  $\mathbf{Y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_m^T)^T$ ,  $\beta = (\beta_1^T, \dots, \beta_m^T)^T$  and

$$\mathbf{X}(\mathbf{s}) = \begin{bmatrix} \mathbf{x}_1(\mathbf{s}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_2(\mathbf{s}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}_m(\mathbf{s}) \end{bmatrix}.$$

The measurement noises are now modelled by a multivariate normal distribution  $\varepsilon \sim MVN(\mathbf{0}, \Psi)$ , where the dispersion matrix  $\Psi$  is arranged by

$$\Psi = \begin{bmatrix} \tau_1^2 I & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \tau_2^2 I & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \tau_m^2 I \end{bmatrix} \in \mathbb{R}^{mn \times mn}$$

$\mathbf{W}(\mathbf{s}) = (\mathbf{w}^T(\mathbf{s}_1) \cdots \mathbf{w}^T(\mathbf{s}_n))^T$ , which is jointly modelled by a MVP  $\mathbf{W}(\mathbf{s}) \sim MGP(\mathbf{0}, \Sigma_m)$ . The multivariate cross-covariance matrix  $\Sigma_m$  is given by

$$\Sigma_m = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1m} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{m1} & \mathbf{C}_{m2} & \cdots & \mathbf{C}_{mm} \end{bmatrix} \in \mathbb{R}^{mn \times mn},$$

where  $\mathbf{C}_{ii}$  is a univariate covariance matrix computed from only a single  $i^{\text{th}}$  spatial field.  $\mathbf{C}_{ij}$  is a cross-covariance matrix calculated between any two  $i^{\text{th}}$  and  $j^{\text{th}}$  spatial phenomena, where  $1 \leq i \neq j \leq m$ . Computing the univariate covariance matrix  $\mathbf{C}_{ii}$  is widely known [2]; however, working out the cross-covariance matrix  $\mathbf{C}_{ij}$  is not completely straightforward due to its ill-famed requirement of positive semi-definite. To this end, Gneiting *et al.* in [7] proposed a colocated correlation coefficient  $\rho_{ij}$  to account for cross-correlation between spatial process components. That is, the cross-covariance matrix can be calculated as follows,

$$\mathbf{C}_{ij} = \mathbf{C}_{ji} = \rho_{ij} \text{COV}(\mathbf{h} | \theta),$$

where  $\text{COV}(\mathbf{h} | \theta)$  is a covariance function,  $\theta$  is a vector of its hyperparameters and  $\mathbf{h} \in \mathbb{R}^d$  is a separation vector between any two locations  $\mathbf{s}$ . The authors of [7] proposed to use the Matérn covariance function and proved a valid  $\rho_{12}$  for two spatial processes so that a cross-covariance matrix  $\mathbf{C}_{ij}$  is guaranteed to be positive semi-definite, which demonstrates in the following theorem.

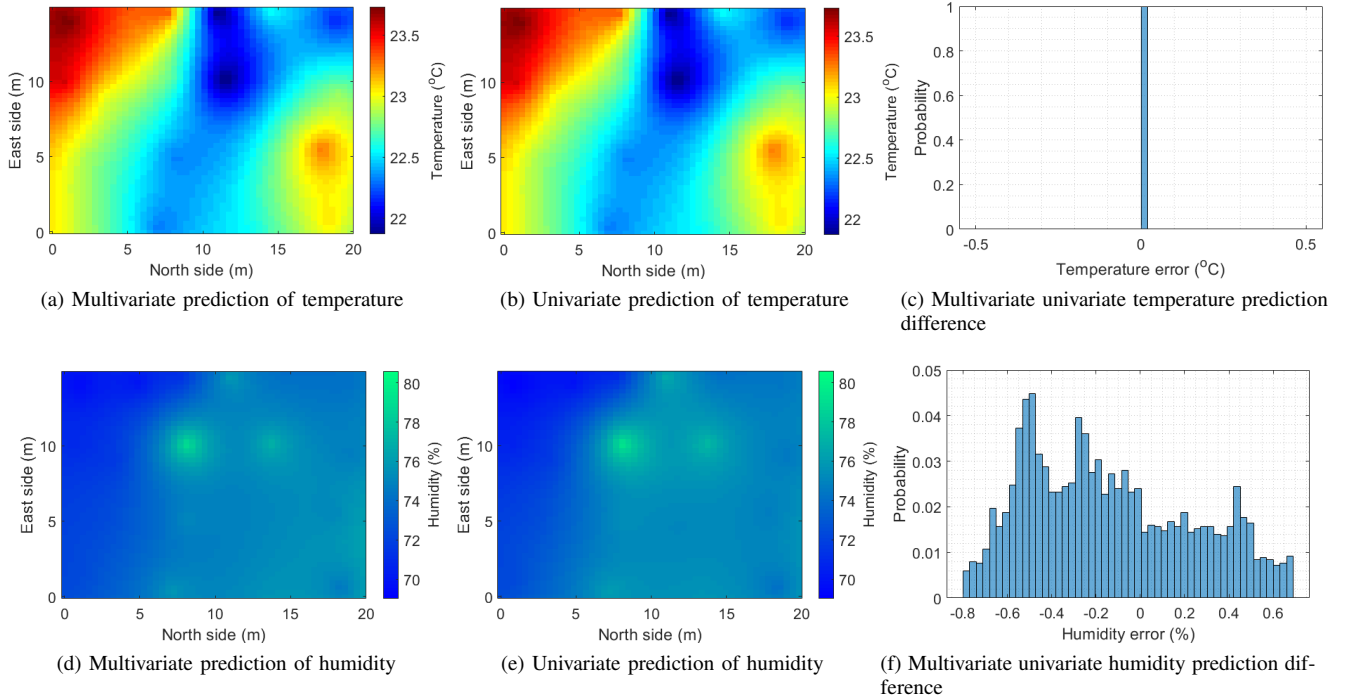


Fig. 1: Comparison of multivariate versus univariate prediction of temperature (top row) and humidity (bottom row). Each prediction is obtained by the 20 measurements collected by the 20 sensors.

*Theorem 1:* [7] If  $\nu_{12} = \frac{1}{2}(\nu_1 + \nu_2)$  and  $\kappa_{12} \geq \max(\kappa_1, \kappa_2)$ , the collocated correlation coefficient  $\rho_{12}$  is computed by

$$\rho_{12} = \frac{1}{2} \left( \frac{\kappa_1}{\kappa_{12}} \right)^{\nu_1} \left( \frac{\kappa_2}{\kappa_{12}} \right)^{\nu_2} \frac{\sqrt{\Gamma(\nu_1 + 0.5d)\Gamma(\nu_2 + 0.5d)}}{\sqrt{\Gamma(\nu_1)\Gamma(\nu_2)}} \times \frac{\Gamma(0.5(\nu_1 + \nu_2))}{\Gamma(0.5(\nu_1 + \nu_2 + d))}, \quad (5)$$

where  $\nu_i$  and  $\kappa_i$  are the Matérn smoothness and spatial scale parameters of a component process, respectively, while  $\Gamma(\cdot)$  denotes a gamma function.

Similar to the univariate modelling presented in Section II-A, multiple spatial phenomena at unmeasured locations  $\mathbf{z}$  can also be predicted by the use of (2) and (3) with some updates. Both  $\mathbf{y}$  and  $\mathbf{x}(\mathbf{s})$  are replaced by  $\mathbf{Y}$  and  $\mathbf{X}(\mathbf{s})$ , respectively.  $\Sigma$  is replaced by  $\Sigma_m$  while  $\tau^2 I$  is altered by  $\Psi$ .

$$\mathbf{X}(\mathbf{z}) = \begin{bmatrix} \mathbf{x}_1(\mathbf{z}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_2(\mathbf{z}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}_m(\mathbf{z}) \end{bmatrix}.$$

Both a covariance matrix  $\Sigma_{\mathbf{z}\mathbf{z}} \in \mathbb{R}^{mp \times mp}$  and a cross-covariance matrix  $\Sigma_{\mathbf{z}\mathbf{s}} \in \mathbb{R}^{mp \times mn}$  can be computed as  $\Sigma_m$ . It is noted that these covariance and cross-covariance matrices are accounted for multiple spatial fields.

### C. Comparison of Multivariate and Univariate Modelling

In order to demonstrate how well the univariate and multivariate modelling can model and then predict spatial processes,

we conducted the experiments with the temperature and humidity sensors to monitor the indoor climate. There were 20 temperature and 20 humidity sensors utilized in the experiments. The experiment set-up and the data collection will be discussed in detail in Section IV. In this section, we present difference of two predictions, for each spatial phenomenon of temperature and humidity, in the whole experimented environment, obtained by the univariate and multivariate models. In other words, given the 20 temperature measurements and the 20 humidity measurements, we built two univariate models as presented in Section II-A separately. Each built model was then employed to predict its corresponding spatial process at the 2500 unmeasured locations on a grid uniformly distributed in the space. The obtained results are visualized in Figures 1b for the temperature and 1e for the humidity, respectively. Moreover, we also used the 40 temperature and humidity measurements to jointly build a multivariate model as presented in Section II-B, which was then exploited to concomitantly estimate the temperature and humidity at the similar 2500 unobserved locations. The predicted results are depicted in Figures 1a and 1d correspondingly. Qualitatively, a pair of Figures 1a and 1b (likewise 1d and 1e) are highly comparable. More importantly, to quantify difference in each pair, we computed difference of the predictions at each of the 2500 locations and summarized all the differences in histograms as shown in Figures 1c and 1f, for the temperature and humidity, respectively. Interestingly, there is no difference between the predictions obtained by the multivariate and univariate models

in the temperature scenario. However, the difference occurs in the humidity example though the different values are very small, less than 1%, quite uniformly ranging from -0.8% to 0.7%.

### III. SPATIAL SENSOR SELECTION

Spatial sensor selection is a fundamental problem in various applications [1]. The critical crux of a sensor selection problem is to select the most informative subset of sensor nodes out of all possible ones, which can be utilized to efficiently predict spatial fields at any unmeasured positions of interest. In equivalent words, the sensor selection problem requires an incorporated spatial modelling representation such as the univariate or multivariate models introduced in Section II. It is noticed that in multivariate scenarios, different types of sensors are assumed to be colocated. For instance, to monitor temperature and humidity spatial processes, two temperature and humidity sensors are embedded on a single device and deployed at one location.

#### A. Problem and Solution

In order to effectively predict spatial phenomena observed by selected sensors as expectation of the spatial sensor selection problem, uncertainty at prediction results must be minimized. It is required to minimize predicted variances at unmeasured locations of interest given measurements gathered by selected sensors. If the multivariate and univariate models are employed, a predicted covariance matrix can be calculated by (3) for a single field or a similar calculation for a multivariate as discussed in Section II-B. Then all predicted variances of predicted locations  $\mathbf{z}$  lie along a diagonal of the covariance matrix. Therefore, if  $\mathcal{P}$  denotes a set of all possible sensors and  $\mathcal{C}$  denotes a selected subset of sensors, then a spatial sensor selection problem can be formulated as an optimization problem as follows,

$$\mathcal{C}^{opt} = \underset{\mathcal{C} \subset \mathcal{P}}{\operatorname{argmin}} \operatorname{tr}(\Sigma_{\mathbf{z}|\mathcal{S}_{\mathcal{C}}}), \quad (6)$$

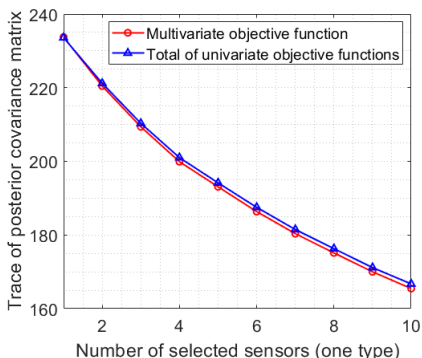


Fig. 2: Values of the objective functions in the sensor selection problems. The red solid curve was obtained from the multivariate temperature and humidity jointly while the blue solid curve was added from the univariate temperature and humidity separately.

where  $\operatorname{tr}(\Sigma_{\mathbf{z}|\mathcal{S}_{\mathcal{C}}})$  is trace of a covariance matrix  $\Sigma_{\mathbf{z}|\mathcal{S}_{\mathcal{C}}}$ , and  $\mathcal{S}_{\mathcal{C}}$  is locations of the selected sensors in the subset  $\mathcal{C}$ . Mathematically, it is quite straightforward to prove that (6) is NP-hard [2]. Although solving (6) in polynomial time is intractable, the spatial sensor selection optimization problem can be efficiently and near-optimally addressed by an approximate algorithm, e.g. greedy. Interested readers are referred to [2] for more detail about the algorithm.

#### B. Selection Comparison

To show how the spatial sensor selection works and compare influence of the multivariate and univariate spatial modelling on sensor selection results, we continue the experiments we briefly mentioned in Section II-C. It is supposed that we expect to select 1 to 10 most informative sensors out of 20 potential ones in both the temperature and humidity cases. Given the measurements, we first built two univariate models and one multivariate model. We then run the greedy algorithm [2] to solve the optimization problem (6) three times, given the univariate temperature model, the univariate humidity model and the multivariate model, respectively. The values of the objective function in (6) corresponding to the selected subset of the sensors were summarized and are demonstrated in Fig. 2. More specifically, the red solid curve presents the objective function values when the algorithm selected 1 to 10 most informative sensor nodes given the multivariate model. Likewise, the blue curve is a sum of all the predicted variances in both the temperature and humidity fields, obtained by two separate algorithm runs for two univariate corresponding models. It can be clearly seen that difference between the optimization values under the multivariate and univariate modelling is trivial, slightly increased when a number of the selected sensors increase.

Moreover, locations of the 10 selected sensor nodes are illustrated in Fig. 3, where the blue circles are locations jointly selected by the multivariate sensor selection method for both the temperature and humidity sensors. In considerations of the univariate sensor selection, the 10 most informative temperature sensors are located at the red stars in Fig. 3a while those for measuring the humidity are positioned at the red triangles in Fig. 3b. In both the scenarios, the selections obtained by the multivariate and univariate algorithms are concurrent up to 9 out of 10 sensor nodes.

### IV. EXPERIMENTAL RESULTS

The section demonstrates in more detail how effectively a subset of sensor nodes selected by the multivariate and univariate sensor selection techniques can monitor spatial fields. We implemented the methods in the experiments using two networks of two different types of temperature and humidity sensors to observe the indoor climate at the Nanyang Technological University campus, Singapore.

#### A. Experimental Data Collection

The experiments were conducted at the room S2.1.B4.01, which is 19.80 m in length and 14.86 m in width, on the 25<sup>th</sup>

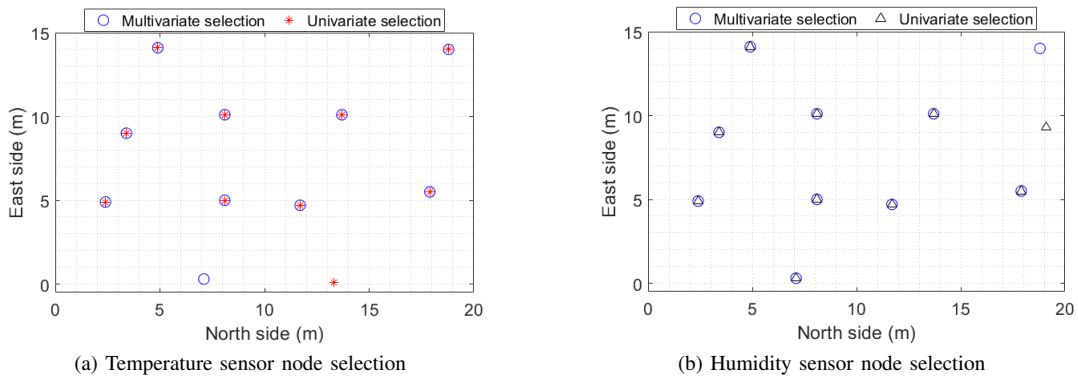


Fig. 3: Selecting the 10 most informative sensor nodes out of 20 potential ones, obtained by the multivariate sensor selection approach (blue circles) and the univariate sensor selection method (temperature sensors - red stars) - (humidity sensors - red triangles).

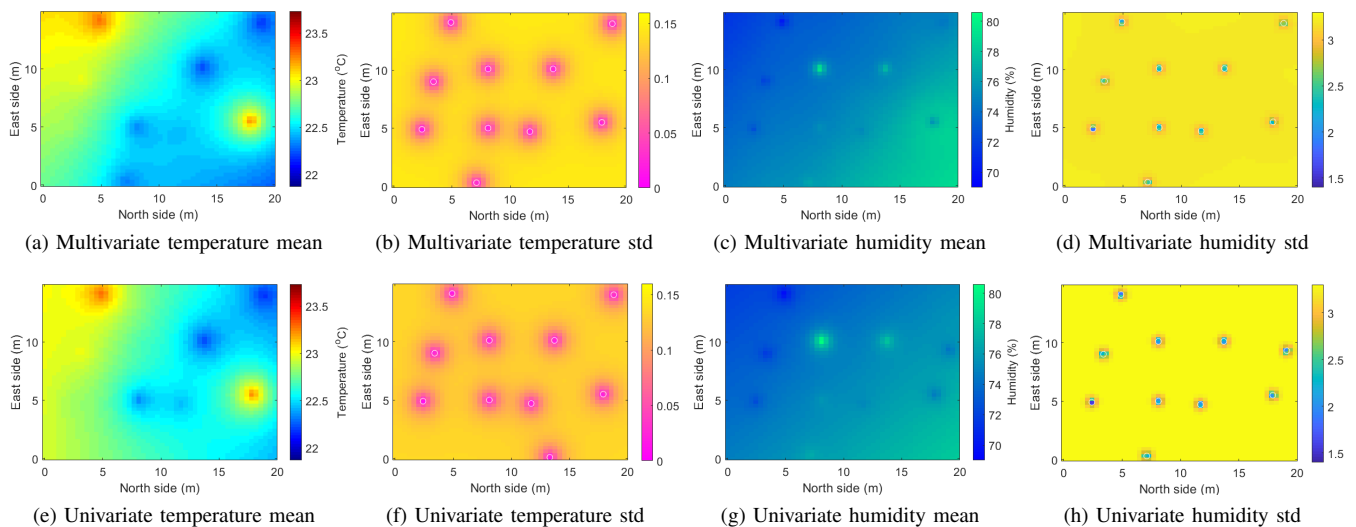


Fig. 4: Column-wise comparison of predicted temperature and humidity fields (mean and stand deviation (std)) obtained by the 10 most informative sensor nodes that are selected by the multivariate and univariate sensor selection approaches. The corresponding selected sensor nodes are demonstrated by the white circles.

of April, 2016. In the experiments, 20 Libelium temperature and 20 Libelium humidity sensors were utilized to observe the indoor temperature and relative humidity. Due to the Libelium sensor board design, a pair of the temperature and humidity sensors are integrated in the same device. That is, when we deployed the sensors, that pair was colocated. We formed the 40 sensors into two separate networks monitoring two separate spatial random fields. In each network, the sensors were wirelessly communicated and deployed at 20 predefined locations in the room. The collected measurements at each sensor were transmitted to a base station through a wireless router [11]. As the collected data can be used to evaluate human comfort in a building, we deliberately deployed all the sensors on the same plane at a sitting level.

### B. Results and Discussion

We now investigate to see how well the 10 most informative sensor nodes selected by the method as presented in Section III can monitor the spatial phenomena. Note that one may select 6, 12 or any number of sensors that suit to a particular application. For reasons of space, in this work, we demonstrate the results corresponding to the 10 selected nodes only.

Let us remind that there are two subsets of the 10 selected temperature (likewise, humidity) sensor nodes obtained by both the multivariate and univariate sensor selection approaches as discussed in Section III. In the multivariate scenario, since the 10 most informative sensor nodes were jointly selected by the multivariate sensor selection approach, we employed all the 10 corresponding temperature and 10 corresponding humidity measurements to build a multivariate model and then utilized the model to simultaneously predict two spatial processes at the 2500 unmeasured locations as

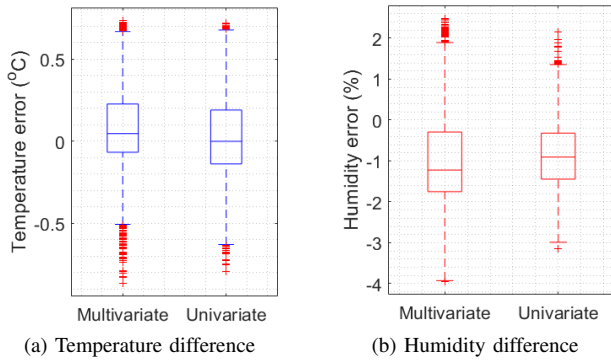


Fig. 5: Differences between the spatial field predictions obtained by all 20 sensors and 10 selected sensors that are obtained by the either multivariate or univariate sensor selection approaches.

introduced in Section II-C. The results of the predicted fields and variances are depicted in Figures 4a and 4b for the temperature and Figures 4c and 4d, respectively. In contrast, in the univariate scenarios, for each subset of the selected sensors, we developed a univariate model from the 10 corresponding measurements. That is, we have two univariate models, one for the temperature and one for the humidity. By the use of two the univariate models separately, we also independently predicted the spatial phenomena at the 2500 unobserved positions. The predicted results in the univariate scenario are shown in Figures 4e, 4f, 4g and 4h. In each column of Fig. 4, we can now compare the results corresponding obtained by the multivariate versus univariate sensor selection approaches. Visually, difference in each column is trivial. Furthermore, though only the 10 sensors were used to take the measurements, the prediction uncertainty, in both the scenarios, is quite low, which is highly practical in building monitoring applications [11].

TABLE I: Quantiles of Prediction Errors

Spatial Fields	Quantiles		
	5%	50%	95%
Temperature (Multivariate)	-0.44	0.04	0.57
Temperature (Univariate)	-0.43	0.01	0.50
Humidity (Multivariate)	-2.47	-1.22	1.21
Humidity (Univariate)	-2.04	-0.90	0.86

More specifically, we quantified errors between the predictions obtained by the selected subsets of the sensor nodes and those obtained by all the 20 sensors. In other words, we computed errors between the predictions shown in Fig. 4a against Fig. 1a and Fig. 4c against Fig. 1d for the multivariate scenarios, and likewise Fig. 4e against Fig. 1b and Fig. 4g against Fig. 1e for the univariate scenarios. Those errors were summarized in the box plots as illustrated in Fig. 5. It can be clearly seen that the errors in both the scenarios are relatively small, which demonstrate effectiveness of the spatial sensor selection methods. Furthermore, we statistically summarize

quantiles of these prediction errors in Table I. The quantile results show that 90% of the temperature errors between the predictions obtained by the multivariate and univariate sensor selection techniques are within about  $0.5\text{ }^{\circ}\text{C}$  while those in the humidity scenarios range at most from  $-2.47\%$  to  $1.21\%$  though the univariate sensor selection occurs slightly better.

## V. CONCLUSIONS

Two spatial sensor selection problems utilizing univariate and multivariate models in effectively monitoring spatial random processes have been considered in the paper. Spatial modelling is required in the spatial sensor selection problems as it provides selected sensors with ability to predict spatial fields at unmeasured locations. Therefore, GP and MGP are exploited to model a univariate and multivariate spatial phenomena, respectively, which also allows selecting the most informative sensor nodes through minimizing prediction uncertainty. The efficient greedy algorithm [2] was used to solve the sensor selection optimization problems, where its implementation and the corresponding results obtained in the real-life experiments validated effectiveness of the selection methods. Comparison of the prediction results obtained by the univariate and multivariate sensor selection techniques was also made with insignificant difference shown.

## ACKNOWLEDGMENT

We would like to thank Prof. Guoqiang Hu at Nanyang Technological University, Singapore for providing the opportunity for the data collection.

## REFERENCES

- [1] L. Nguyen, N. Ulapane, and J. V. Miro, "Adaptive sampling for spatial prediction in environmental monitoring using wireless sensor networks: A review," in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2018, pp. 346–351.
- [2] L. V. Nguyen, G. Hu, and C. J. Spanos, "Efficient sensor deployments for spatio-temporal environmental monitoring," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 12, pp. 5306–5316, 2020.
- [3] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. The MIT Press, Cambridge, Massachusetts, London, England, 2006.
- [4] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.
- [5] G. Mastrantonio, G. Jona Lasinio, A. Pollice, G. Capotorti, L. Teodonio, G. Genova, and C. Blasi, "A hierarchical multivariate spatio-temporal model for clustered climate data with annual cycles," *Ann. Appl. Stat.*, vol. 13, no. 2, pp. 797–823, 06 2019.
- [6] A. O. Finley, S. Banerjee, A. R. Ek, and R. E. McRoberts, "Bayesian multivariate process modeling for prediction of forest attributes," *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 13(1), pp. 60–83, 2008.
- [7] T. Gneiting, W. Kleiber, and M. Schlather, "Matérn cross-covariance functions for multivariate random fields," *Journal of the American Statistical Association*, vol. 105, no. 491, pp. 1167–1177, 2010.
- [8] L. Nguyen, S. Kodagoda, R. Ranasinghe, and G. Dissanayake, "Mobile robotic sensors for environmental monitoring using gaussian markov random field," *Robotica*, p. 1–23, 2020.
- [9] N. A. C. Cressie, *Statistics for Spatial Data*. Wiley, 1993.
- [10] M. G. Genton and W. Kleiber, "Cross-covariance functions for multivariate geostatistics," *Statist. Sci.*, vol. 30, no. 2, pp. 147–163, 05 2015.
- [11] L. Nguyen, G. Hu, and C. J. Spanos, "Efficient spatio-temporal sensor deployments: A smart building application," in *2017 13th IEEE International Conference on Control Automation (ICCA)*, 2017, pp. 612–617.