

“© 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.”

# Gaussian Process As a Benchmark for Optimal Sensor Placement Strategy

Nalika Ulapane

*Swinburne University of Technology &  
Peter MacCallum Cancer Centre*  
Melbourne, Australia.  
nalika.ulapane@petermac.org  
nalika.ulapane@gmail.com

Karthick Thiyagarajan

*iPipes Lab, UTS Robotics Institute*  
*University of Technology Sydney*  
Sydney, Australia.  
Karthick.Thiyagarajan@uts.edu.au

Sarath Kodagoda

*iPipes Lab, UTS Robotics Institute*  
*University of Technology Sydney*  
Sydney, Australia.  
Sarath.Kodagoda@uts.edu.au

**Abstract**—Optimal sensor placement is an important problem to look at. This problem becomes all the more relevant nowadays due to advancements in infrastructure monitoring robotic technologies including underground sensing. While there are multiple ways to solve optimal sensor placement problems, one of the most generic methods available is Bayesian Optimization and its variants. In this paper, we present a simple benchmark-like formulation for exploiting Gaussian Process uncertainty for sensor placement to measure a scalar field.

**Index Terms**—Bayesian Optimization; Gaussian Process; Optimization; Pulsed Eddy Current; Sensor Placement; Scalar Field; Sensor; Smart Pipes; Spatial Measurement; 2D Environment.

## I. INTRODUCTION

Within the task of measuring scalar fields, optimal sensor placement, or computing subsequent best, or subsequent most informative location(s) for sensor placement, is an important question to look at. This question has become especially relevant due to the modern advancements of robotic capabilities available to support sensing tasks [1], including underground measurements [2]–[6]. An approach that is classically used to solve such sensor placement problems is Bayesian Optimization and its extensions [7].

However, Bayesian Optimization approaches come with the burden of computational complexity [7], [8], especially due to some Gaussian Process (GP) [9] computations that have to be made. This burden is especially worrisome in real time computations for robotic deployments. Therefore, it becomes of interest to assess alternative approaches that can optimize sensor placement in computationally simpler ways.

When it comes to assessing such alternative approaches though, one would at some stage have to compare the alternatives against Bayesian Optimization methods. This comparison becomes necessary because Bayesian Optimization approaches, despite their computational burdens, can be argued to be a powerful benchmark means [7] to solve many a design of experiments problem, including sensor placement problems.

Having that premise set, with an intention of analyzing different sensor placement strategies in future, we present in this paper a simplified yet generic formulation for exploiting Gaussian Process coupled with the Squared Exponential kernel [9], describing a benchmark method for sensor placement in a

2D environment. The formulation can be extended to 3D environments and higher dimensions as well. We also demonstrate the performance of the benchmark through a Pulsed Eddy Current sensing experiment targeted at underground reinforced concrete sewer inspection [10].

## II. PROBLEM FORMULATION

We are given a bounded 2D plane, on which any point can be given by the coordinate pair  $(x, y)$  where  $x \in \mathbb{R}$  and  $y \in \mathbb{R}$ . The plane can have a ‘ $k$ ’ number of boundaries. The boundaries are known. With respect to those boundaries, the plane can be written as a region that the coordinates  $(x, y)$  satisfy the inequalities:  $b_1(x, y) \geq 0, b_2(x, y) \geq 0, \dots, b_k(x, y) \geq 0$ . Here  $b_1, b_2, \dots, b_k$  are functions that perform  $\mathbb{R}^2 \rightarrow \mathbb{R}$ , and these functions can be nonlinear for generality.

Our bounded 2D plane carries a scalar field. Said more specifically, every point  $(x, y)$  in our plane will have a measurable and static (i.e., non time-varying) scalar quantity given by  $g = f(x, y)$ . Here,  $g \in \mathbb{R}$ , and  $f$  is a smooth continuous function such that  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ . But, the function  $f$  is unknown.

Now, we are given a fixed set  $X_n$  containing an ‘ $n$ ’ number of coordinates taken from our bounded plane.  $X_n$  can be written as in (1). It must be assumed that the coordinates in  $X_n$  are sufficiently dense so that between any two adjacent coordinates the function  $f$  remains approximately smooth. We are then asked to measure  $G_n$ —the set given in (2) that contains the scalar values corresponding to the coordinates in  $X_n$ .

$$X_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (1)$$

$$G_n = \{g_1, g_2, \dots, g_n\} \quad (2)$$

We are given one sensor to take measurements. Measurement-taking will have to be iterative. At every iteration, the sensor can take a single measurement by reaching any point  $(x_i, y_i) \in X_n$ . We are thus required to iterate to cover the space taking one measurement at a time.

Given a task like this, typically, we would go and measure ‘ $g$ ’ values at every given coordinate. However, suppose measuring at all coordinates is considered tedious, and we would like to do fewer measurements, but still get a good enough

job done. Achieving that objective is the problem we address in this paper. Our problem can be expressed mathematically as follows.

Given a fixed set of coordinates  $X_n$ , we would like to identify an optimal subset  $X_m$ , where  $X_m \subset X_n$  and obtain a corresponding set of measurements  $G_m$ , where  $G_m \subset G_n$ , such that the underlying unknown continuous scalar function  $f$  that does the mapping  $g = f(x, y)$ , can be estimated. The estimate of  $f$  must converge under some convergence criteria we define.

Solving the aforesaid problem essentially addresses a version of the optimal sensor placement task in 2D. Moreover, obtaining a converged estimate of  $f$  enables finding the scalar values at coordinates in the set  $X_n \cap X_m^c$  through prediction and without measurement (here,  $X_m^c$  is the complement set of  $X_m$ ). In the subsequent sections, we formulate a solution for this problem using Gaussian Process (GP) [9]. Our formulation can also be used as a benchmark to be compared against to evaluate other methods.

### III. METHODOLOGY

This section presents the GP-based formulation to accomplish the sensing task described in Section II. An iterative method for GP maximum uncertainty-driven, information maximization-based, active measurement-taking is proposed. Examples are available to how GP maximum uncertainty criteria has been exploited in different contexts [11], [12].

Given any set of prior measurements, this method calculates the next best, or next most informative point to take a measurement. The next best point is considered where the maximum GP uncertainty (explained later) results from the remaining unmeasured points. Sequential steps to be followed are summarized herein.

**Step 1, Conduct prior measurements:** Select a set  $X_m$  (where  $X_m \subset X_n$ ) that contains an ‘ $m$ ’ number of ‘uniformly’ or ‘evenly’ distributed points across the given region to sense (note: select the number  $m$  such that  $m/n \times 100\% \approx 50\%$ ). Then conduct the measurements corresponding to  $X_m$ , and construct the corresponding set  $G_m$ . Also construct the corresponding unmeasured coordinate set  $X_*$  where  $X_* = X_n \cap X_m^c$ . Now arrange  $X_m$ ,  $G_m$  and  $X_*$  to be corresponding matrices as shown in (3), (4) and (5). The notation  $[\dots]^T$  henceforth denotes matrix transpose.

$$X_m = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \end{bmatrix}^T \quad (3)$$

$$G_m = [g_1 \quad g_2 \quad \dots \quad g_m]^T \quad (4)$$

$$X_* = \begin{bmatrix} x_*^{(1)} & x_*^{(2)} & \dots \\ y_*^{(1)} & y_*^{(2)} & \dots \end{bmatrix}^T \quad (5)$$

Moreover, the notation  $X_m^{(i)}$  and  $X_*^{(j)}$  will henceforth denote any generic  $i^{\text{th}}$  and  $j^{\text{th}}$  row of two matrices  $X_m$  and  $X_*$  respectively.  $X_m^{(i)}$  and  $X_*^{(j)}$  will be considered vectors in  $\mathbb{R}^2$ . This notation is useful for (6).

**Step 2, Estimate Hyper-Parameters:** The objective of the second step is to estimate and save the hyper-parameter set  $\theta^{(l)} = \{\alpha^{(l)}, \eta^{(l)}, \sigma^{(l)}\}$ , where  $\theta^{(l)} \in \mathbb{R}^3$ . The superscript  $(l)$  denotes the number of iterations the measurement-taking has gone for. The value of  $(l)$  will be one when entering **Step 2** soon after completing **Step 1**, i.e., taking the prior measurements. On subsequent returns to **Step 2** (as described later), the value of  $(l)$  will increment.

This hyper-parameter set in  $\mathbb{R}^3$  results by considering the Squared Exponential kernel and Gaussian likelihood in GP formulation [9] (presenting full GP formulation is not done in this paper, only vital steps are shown, readers are advised to refer to works like [9] for complete derivations). Taking the squared exponential kernel becomes appropriate for our task due to our underlying condition that the function ‘ $f$ ’ is smooth and continuous (recall from Section II).

To calculate hyper-parameters following GP formulation [9], we construct matrix  $\Sigma_m$  such that  $\Sigma_m = K(X_m, X_m) + \sigma^2 I$ . Here,  $\sigma \in \mathbb{R}$  is a hyper-parameter to be estimated,  $K(X_m, X_m) \in \mathbb{R}^{m \times m}$ , and  $I$  is an identity matrix of corresponding size. A generic formulation for  $X_a \subset X_n$ ;  $X_b \subset X_n$ ;  $a, b \in \mathbb{Z}^+$ , and  $a \neq b$  becomes  $K(X_a, X_b) \in \mathbb{R}^{a \times b}$ . A generic element of  $K(X_a, X_b)$  is given by  $K(X_a, X_b)_{i,j} = k(X_a^{(i)}, X_b^{(j)})$ . Here,  $k(X_a^{(i)}, X_b^{(j)})$  is the squared exponential kernel given by (6) where  $\|\dots\|$  denotes the Euclidean Norm and  $\alpha, \eta \in \mathbb{R}$  are hyper-parameters to be estimated.

$$k(X_a^{(i)}, X_b^{(j)}) = \alpha^2 \exp\left(-\frac{1}{2\eta^2} \|X_a^{(i)} - X_b^{(j)}\|^2\right) \quad (6)$$

We then proceed to derive maximum likelihood estimation formulation [9]. This results in the negative log marginal likelihood given by (7) where  $|*|$  denotes the matrix determinant.

$$-\log[p(G_m|X_m, \theta)] = \frac{1}{2} G_m^T \Sigma_m^{-1} G_m + \frac{1}{2} \log |\Sigma_m| + \frac{m}{2} \log(2\pi). \quad (7)$$

After expressing  $-\log[p(G_m|X_m, \theta)]$ , we can then estimate hyper-parameters by minimizing the negative log marginal likelihood as in (8). A way to initialize the hyper-parameters to perform this optimization is presented in [13].

$$\theta^{(l)} = \arg \min_{\theta} (-\log[p(G_m|X_m, \theta)]) \quad (8)$$

**Step 3, Check Convergence:** This step prescribes a convergence, or a stopping criteria for the sensing task. To define a stopping criteria, in every iteration  $(l)$ , we make a prediction on the full coordinate set  $X_n$  based on the measurement sets  $X_m$  and  $G_m$  available at that iteration. This means by following GP prediction [9] we perform:  $G_n^* | G_m, X_m, X_n \sim \mathcal{N}(\mu_n^*, \Sigma_n^*)$ , such that  $\mu_n^*$  and  $\Sigma_n^*$  are given by (9) and (10). Then, for the  $(l)^{\text{th}}$  iteration we calculate  $\gamma^{(l)}$  as shown in (11) where  $|*|$  denotes absolute value.  $\gamma^{(l)}$  here is the measure of mean percentage value of  $2 \times$  prediction uncertainty with respect to prediction mean. As we collect more and more measurements, i.e., as our  $X_m$  set gets larger in every iteration  $(l)$ , we would expect  $\gamma^{(l)}$  to decrease. Therefore, we define the stopping criteria to be  $\gamma^{(l)} \leq \gamma_c$  condition

be satisfied for  $N_{it}$  consecutive iterations. Here,  $\gamma_c$  will be some small positive number we define. Similarly,  $N_{it}$  will be a number of consecutive iterations we define. This means, as long as  $\gamma^{(l)} > \gamma_c$  prevails, the iterative measurement-taking will continue. It should also be noted that for the way we calculate  $\gamma^{(l)}$  in (11), we do not wish any  $\mu_n^{*(i)}$  to lie close to zero. Therefore, we recommend adding a bias constant to the measured  $G_m$  set prior to doing GP in a way that all values in  $G_m$  become positive and the minimum value in  $G_m$  becomes larger than one.

$$\mu_n^* = K(X_n, X_m)\Sigma_m^{-1}G_m \quad (9)$$

$$\Sigma_n^* = K(X_n, X_n) + \sigma^2 I - K(X_n, X_m)\Sigma_m^{-1}K(X_m, X_n) \quad (10)$$

$$\gamma^{(l)} = 100\% \times \frac{2}{n} \sum_{i=1}^n \frac{\sqrt{(\Sigma_n^*)_{i,i}}}{|\mu_n^{*(i)}|} \quad (11)$$

#### Step 4, Compute and Perform Next Best Measurement:

This step is performed for as long as the convergence criteria in **Step 3** is not satisfied. To compute the coordinate for the next best measurement, we make a prediction on the unmeasured coordinate set  $X_*$  based on the measurement sets  $X_m$  and  $G_m$  available at iteration ( $l$ ). This means by following GP prediction [9] we perform:  $G_*|G_m, X_m, X_* \sim \mathcal{N}(\mu_*, \Sigma_*)$ , such that  $\mu_*$  and  $\Sigma_*$  are given by (12) and (13). When  $\mu_*$  and  $\Sigma_*$  are computed, we can then determine the coordinate with highest uncertainty in iteration ( $l$ ) by solving (14). The resulting coordinate  $(x^{(l)}, y^{(l)})$  is deemed as the next best point to take a measurement according to our maximum uncertainty criteria. Once the point  $(x^{(l)}, y^{(l)})$  is determined, the corresponding measurement  $g^{(l)}$  is taken. We then add point  $(x^{(l)}, y^{(l)})$  to the matrix  $X_m$  and  $g^{(l)}$  to the matrix  $G_m$ , and return to **Step 2** thereby closing the loop of the algorithm to continue iterating.

$$\mu_* = K(X_*, X_m)\Sigma_m^{-1}G_m \quad (12)$$

$$\Sigma_* = K(X_*, X_*) + \sigma^2 I - K(X_*, X_m)\Sigma_m^{-1}K(X_m, X_*) \quad (13)$$

$$(x^{(l)}, y^{(l)}) = \underset{(x^{(i)}, y^{(i)}) \in X_*}{\arg \max} (\Sigma_*)_{i,i} \quad (14)$$

## IV. RESULTS AND CONCLUSION

Results from an example where the proposed method was applied is shown. The example comes from [10]. A Pulsed Eddy Current (PEC) sensor is used to locate reinforcement rods (i.e., rebars). This application is related to underground reinforced concrete sewer infrastructure inspection. The PEC sensor is allowed to move across a planar surface above the rebars. Signal intensity is indicative of the sensor moving above the rebar (see Fig. 1(a)). This exercise has 117 (i.e.,  $13 \times 9$ ) uniformly placed square grid locations. The sensor can take a maximum of 117 measurements—one measurement per grid location. Fig. 1(a) shows the case where all 117 measurements have been taken. Fig. 1(b) shows the measurements completed at the point of convergence by following the method proposed in this paper. White squares are indicative of unmeasured points. Fig. 1(c) shows the completed map

by filling the blank spaces of Fig. 1(b) via GP regression. The effectiveness of the method is evidenced by there being hardly any visible difference in Fig. 1(c) from Fig. 1(a). Fig. 1(d) shows the percentage error between the GP predictions in Fig. 1(c) and the corresponding actual measurements in Fig. 1(a). The maximum percentage error does not exceed 5%. Fig. 2 shows how the convergence measure  $\gamma^{(l)}$  has varied along the iterations ( $l$ ). Since there are 117 maximum allowable measurements in this example, the prior measurements (i.e., 50%, recall from **Step 1**) accounted for 59 points. After those measurements, only 40 more measurements were required (notice from Fig. 2—convergence reached in the 41<sup>st</sup> iteration) to reach convergence at the convergence parameters set to  $\gamma_c = 10\%$  and  $N_{it} = 10$ . Thus can be demonstrated the capability of the proposed method. Future work can explore different convergence criteria, different next best point calculation criteria, and use the proposed method as a benchmark to be compared against when evaluating other methods for similar purposes, and also larger data sets [14].

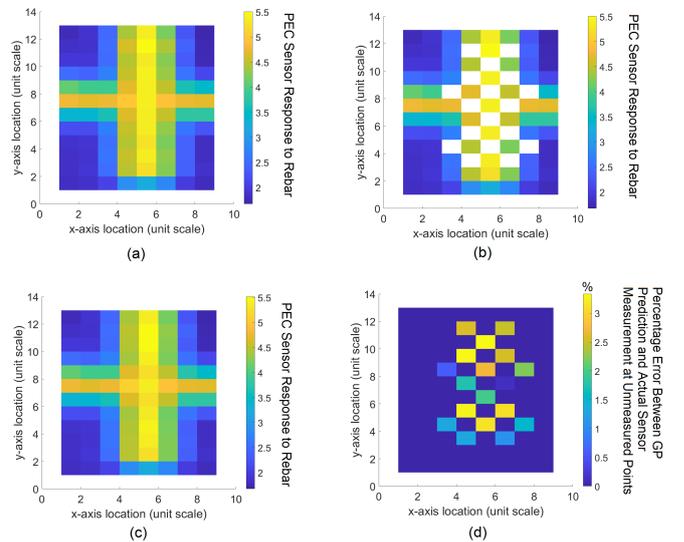


Fig. 1. Experimental results from a PEC sensor measurement task to locate sewer reinforcement rods (rebars): (a) Reference map with all sensor measurements; (b) Measurements completed at convergence of the proposed method (white squares indicate unmeasured points); (c) Measured map from the proposed method filled with GP predictions on the unmeasured points; (d) Percentage error between GP predictions and actual measurements of the unmeasured points.

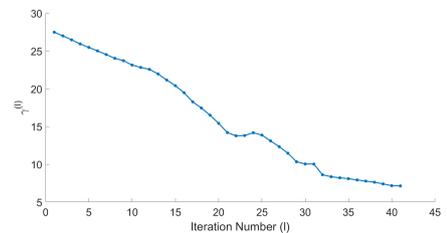


Fig. 2. The trend of  $\gamma^{(l)}$  reduction along with iterations for the example in Figure 1.

## REFERENCES

- [1] T. Glotzbach, D. Moreno-Salinas, A. Pascoal, and J. Aranda, "Optimal sensor placement for acoustic range-based underwater robot positioning," *IFAC Proceedings Volumes*, vol. 46, no. 33, pp. 215–220, 2013.
- [2] A. Gunatilake, L. Piyathilaka, A. Tran, V. K. Vishwanathan, K. Thiagarajan, and S. Kodagoda, "Stereo vision combined with laser profiling for mapping of pipeline internal defects," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11 926–11 934, 2021.
- [3] K. Thiagarajan, S. Kodagoda, R. Ranasinghe, D. Vitnage, and G. Iori, "Robust sensor suite combined with predictive analytics enabled anomaly detection model for smart monitoring of concrete sewer pipe surface moisture conditions," *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8232–8243, 2020.
- [4] J. Valls Miro, N. Ulapane, L. Shi, D. Hunt, and M. Behrens, "Robotic pipeline wall thickness evaluation for dense nondestructive testing inspection," *Journal of Field Robotics*, vol. 35, no. 8, pp. 1293–1310, 2018.
- [5] N. Ulapane, A. Alempijevic, T. Vidal-Calleja, J. V. Miro, J. Rudd, and M. Roubal, "Gaussian process for interpreting pulsed eddy current signals for ferromagnetic pipe profiling," in *2014 9th IEEE Conference on Industrial Electronics and Applications*. IEEE, 2014, pp. 1762–1767.
- [6] K. Thiagarajan, S. Kodagoda, R. Ranasinghe, D. Vitnage, and G. Iori, "Robust sensing suite for measuring temporal dynamics of surface temperature in sewers," *Scientific Reports*, vol. 8, no. 1, p. 16020, 2018.
- [7] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, and S. Venkatesh, "Bayesian optimization for adaptive experimental design: A review," *IEEE Access*, vol. 8, pp. 13 937–13 948, 2020.
- [8] K. Thiagarajan, S. Kodagoda, L. V. Nguyen, and S. Wickramanayake, "Gaussian markov random fields for localizing reinforcing bars in concrete infrastructure," in *Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC)*, J. Teizer, Ed. Taipei, Taiwan: International Association for Automation and Robotics in Construction (IAARC), July 2018, pp. 1052–1058.
- [9] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [10] N. Ulapane, S. Wickramanayake, and S. Kodagoda, "Pulsed eddy current sensing for condition assessment of reinforced concrete," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2019, pp. 1–6.
- [11] N. Ulapane, K. Thiagarajan, and S. Kodagoda, "System identification of static nonlinear elements: A unified approach of active learning, overfit avoidance, and model structure determination," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2020, pp. 1001–1006.
- [12] R. Chin, A. I. Maass, N. Ulapane, C. Manzie, I. Shames, D. Nešić, J. E. Rowe, and H. Nakada, "Active learning for linear parameter-varying system identification," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 989–994, 2020, 21th IFAC World Congress.
- [13] N. Ulapane, K. Thiagarajan, and S. Kodagoda, "Hyper-parameter initialization for squared exponential kernel-based gaussian process regression," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2020, pp. 1154–1159.
- [14] N. Ulapane, K. Thiagarajan, J. V. Miro, and S. Kodagoda, "Surface representation of pulsed eddy current sensor signals for improved ferromagnetic material thickness quantification," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 5413–5422, 2021.