

**© 2016 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.**

# Constrained Sampling of 2.5D Probabilistic Maps for Augmented Inference

Lei Shi, Jaime Valls Miro, Teng Zhang, Teresa Vidal-Calleja, Liye Sun, Gamini Dissanayake

**Abstract**—This work exploits modeling spatial correlation in 2.5D data using Gaussian Processes (GPs), and produces constrained sampling realizations on these models to improve certainty in the predictions by means of integrating additional sparse information. Data organized in 2.5D such as elevation and thickness maps has been extensively studied in the fields of robotics and geostatistics. These maps are typically represented as a probabilistic 2D grid that stores an estimated value (height or thickness) for each cell. With the increasing popularity and deployment of robotic devices for infrastructure inspection, 2.5D data becomes a common interpretation of the condition of the target being inspected. Modeling the spatial dependencies and making inferences on new grid locations is a common task that has been addressed using GPs, but inference results on locations which are weakly correlated with the training data are generally not sufficiently informative and distinctly uncertain. The predictive capability of the proposed framework, which is applicable to any 2.5D data, is demonstrated with field inspection data from pipelines. Specifically, sparse and complementary measurements from alternative sensing modalities have been incorporated into the model to predict in more detail local thickness conditions where GP training data is limited. The output of this work aims to probabilistically present variations of the target in the case that both accuracy and reasonable diversity are of significant interest.

## I. INTRODUCTION

Two-and-a-half dimensional (2.5D) mapping has been widely used in robotics and geostatistics to represent the target in a compact but informative manner. A 2.5D map is a 2D grid where each cell stores an estimated value. For example, digital elevation maps store the height of the terrain [1] and thickness maps reflect the remaining material in pipe walls [2]. Other than containing the values of interest, 2.5D maps can also imply how these values are spatially correlated. Studying the spatial correlation has the potential of adding considerable value to analyzing the unorganized measurements alone.

In the last decades, increasing numbers of robotic devices and autonomous systems have been developed in the water industry for the purpose of pipeline inspection to better understand the current condition of buried pipes [3]. However, the way of efficiently and effectively exploiting the available data needs to be given greater consideration due to the cost and physical constraints of deploying these systems.

An example of the task to be addressed in this research is illustrated in Figure 1. On a buried pipeline, excavations

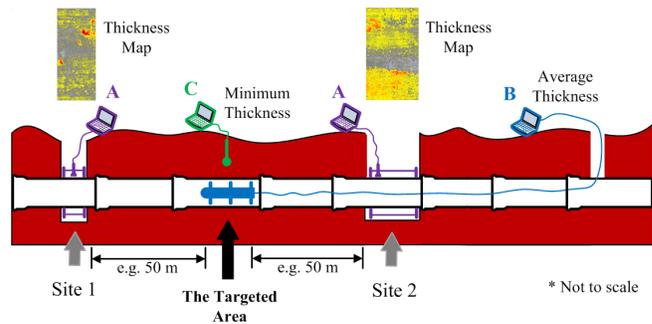


Fig. 1: An example of the targeted 2.5D data prediction task

are carried out on sites 1 and 2, and local inspections on these sites give details of the pipe condition in the form of 2.5D remaining wall thickness maps. The targeted area is far away from either inspected site and cannot be excavated, so the task is to predict the pipe condition at the targeted area with the greatest possible detail. There could be no specific information on the targeted area, but sometimes limited information is available from other sources. In the case shown in Figure 1, for the whole area of interest, an in-pipe robot *B* provides the average thickness value and an intelligent soil monitoring system *C* provides the minimum thickness value based on its knowledge on environmental factors. The framework proposed in this research allows the prediction of the targeted area in the form of a thickness map based on inspection outcomes from sites 1 and 2, with or without extra information provided by system *B* and/or *C*.

In our previous work, we have proposed a GP framework for modeling 2.5D data and designed a covariance function specifically for modeling spatial correlation in pipe wall thickness data [2]. To extend our previous work, this research paper reports the development of various mathematical strategies for integrating different constraints in the realizations. The proposed constrained simulation strategies, as per the block diagram shown in Figure 2, is applicable to any 2.5D data. The rest of this paper is arranged as follows. Section II provides a review of related work regarding spatial data modeling and realization, and the background of the application. Section III discusses in detail the underlying approach of modeling 2.5D data with GPs and the strategies of integrating various constraints. The experimental setup is described in Section IV. The experimental results are presented in Section V with analysis and discussion. Section VI concludes the paper.

All authors are with the Centre for Autonomous Systems at Faculty of Engineering and IT, University of Technology, Sydney NSW 2007, Australia. {Lei.Shi-1 | Jaime.VallsMiro | Teresa.VidalCalleja | Gamini.Dissanayake}@uts.edu.au, {Teng.Zhang | Liye.Sun}@student.uts.edu.au

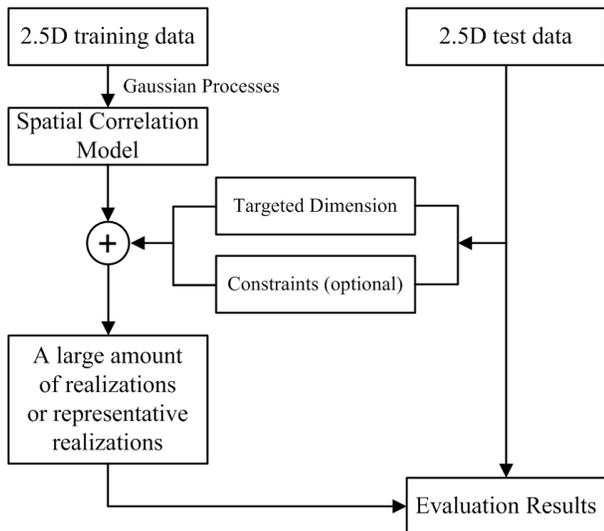


Fig. 2: Block diagram of the developed framework

## II. RELATED WORK

Modeling spatial dependencies for 2.5D data has been extensively studied in the areas of robotics and geostatistics. Mathematically, the problem can be described as a random field, which is a collection of random variables of the form  $\{y_{\mathbf{x}}, \mathbf{x} \in \mathbf{R}^d\}$ , where  $y_{\mathbf{x}}$  is the quantity measured at the position  $\mathbf{x}$  [4]. In some research domains, random fields are also known as spatial processes that are defined for modeling spatially arranged measurements and patterns [5], [6]. Random fields can be statistically specified by mean and covariance [4], [5]. When the mean is modeled as a constant which is a common practice, depending on the assumptions on the covariance, there are: stationary, isotropic and anisotropic random fields [4], [5]. A more specific type of random field being studied is Gaussian Random Fields [7], which is also known as Gaussian Spatial Processes [5] or Gaussian Processes [8], [9]. In this case the random variables jointly have a multivariate Gaussian distribution [5]. GPs can be generalized to any finite dimensions of random variables, and the distribution over a finite subset of a GP is also a multivariate Gaussian distribution.

In robotics, a typical solution in 2.5D terrain and surface modeling is to employ GPs [10]–[12]. In all these works, the utilization varies with the properties of the applications, the structure of the model and the usage of the correlation information. In our previous work, GPs and relevant probabilistic tools have been exploited for thickness mapping and fusion for applications on pipeline condition assessment [13], [14].

The modeling and realization process is an established approach in many research disciplines [15], [16]. In this study we propose performing GP realization instead of inference, because when the target locations are far away from the locations containing training data, realization results become a more suitable interpretation of the target to be predicted. Moreover, in the case that the global location information of either the training or the test data is unavailable, inference

becomes impossible and realization would be the only way of predicting the condition of the area of interest.

Generating realizations of a GP is also known as simulation or sampling; in the context of this paper, we use the three terms: sampling, realization and simulation interchangeably. A conventional sampling algorithm is Cholesky decomposition, and there exist more efficient approaches for large-scale tasks [5], [6]. In terms of incorporating constraints, methods on sampling a multivariate Gaussian distribution under constraints have been mathematically described in literature. Common examples include simulating truncated or conditional multivariate Gaussian distribution [16]–[18]. However, mean-constraint and a combination of different types of constraints are rarely discussed or evaluated in the applications of 2.5D data.

When assessing the condition of buried pipelines, a common approach to extrapolate the information obtained from local inspections into uninspected regions involves inspecting in detail a small amount of pipe segments (one or several), summarizing the extreme values such as depth of pitting, and carrying out an Extreme Value Analysis (EVA). This data-driven statistical procedure does not require full understanding of the complicated mechanism behind corrosion. However, it appears to the authors that the potential of condition assessment is not fully exploited by analyzing only extreme values, as modeling the spatial correlation in the local assessment outcomes (in the form of 2.5D thickness map) can potentially facilitate more reliable failure-related analysis. For example, a large amount of probabilistically distributed realizations as representation of an uninspected area, in the form of wall thickness maps at the same resolution of the local inspection outcome, are supposed to be more suitable for subsequent structural analysis [2], [15].

## III. APPROACH

This section provides the detailed methodology for modeling spatially correlated 2.5D data using GPs, followed by a description on different strategies for constrained sampling.

### A. Gaussian processes

In this section GPs will be briefly introduced in the aspects of problem identification, probabilistic modeling, parameter estimation, inference and realization.

Consider  $n$  thickness-location pairs  $\mathcal{D}$  defined as:

$$\mathcal{D} = \{(y_1, \mathbf{x}_1), (y_2, \mathbf{x}_2), \dots, (y_n, \mathbf{x}_n)\}, \quad (1)$$

where  $\mathbf{x}_i \in X$  is the position in  $\mathbf{R}^d$  from which the thickness measurements  $y_i \in Y$  was taken, and  $d = 2$  in the case of 2.5D data. The data set  $\mathcal{D}$  is assumed to be drawn from a noisy process

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \text{ where } \varepsilon_i \sim \mathcal{N}(0, \sigma_n^2), \quad (2)$$

where noise  $\varepsilon_i$  follows independent and identically distributed (i.i.d.) zero-mean Gaussian with variance  $\sigma_n^2$ . GPs learn the distribution  $p(\mathbf{f}|X, \mathcal{D})$  from  $\mathcal{D}$  and have the capability of inferring  $p(\mathbf{f}|X^*, \mathcal{D})$  for arbitrary location  $X^*$ .

Having specified the mean and covariance functions, and the hyper-parameter set  $\theta$  (parameters for the mean function, the kernel and the noise), parameter estimation can be achieved through maximizing the likelihood function as described in equation 3.

$$\log p(\mathbf{y}|X) = -\frac{1}{2}(\mathbf{y} - \mathbf{m}(X))^\top K_y^{-1}(\mathbf{y} - \mathbf{m}(X)) - \frac{1}{2}\log|K_y| - \frac{n}{2}\log 2\pi, \quad (3)$$

where  $\mathbf{m}$  and  $K$  are mean and covariance functions respectively, and  $K_y = K(X, X) + \sigma_n^2 I$  denotes the joint prior distribution covariance of the function at positions  $X$ . The variance of the noise  $\sigma_n^2$  is another parameter to be learned.

Inference at query locations  $X^*$  can be performed by calculating the predicted mean  $\mu_P$  and covariance  $\Sigma_P$ .

$$\mu_P = \mathbf{m}(X^*) + K(X^*, X)K_y^{-1}(\mathbf{y} - \mathbf{m}(X)) \quad (4)$$

$$\Sigma_P = K(X^*, X^*) - K(X^*, X)K_y^{-1}K(X^*, X)^\top \quad (5)$$

The matrix  $K(X^*, X)$  is obtained from the covariance function  $K$  and it indicates the cross-correlation between the function at  $X^*$  and the training inputs  $X$ .

The test data itself at query locations  $X^*$  follows the multivariate Gaussian distribution  $\mathcal{N}(\mathbf{m}(X^*), K(X^*, X^*) + \sigma_n^2 I)$ , and the learned model can be simulated on this distribution.

### B. Range-constrained simulation

In the case that the measurement of interest  $y$  is subject to range constraint, the multivariate Gaussian distribution  $Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$  can be considered as conditioned on  $LB \leq Y \leq UB$ , where  $LB$  and  $UB$  are lower and upper bounds. The new distribution is a truncated multivariate Gaussian distribution. The inequalities for lower and upper bounds can be described in an alternative way as  $MY \leq N$  for the ease of formulation, by defining

$$M = \begin{bmatrix} I \\ -I \end{bmatrix}, \quad N = \begin{bmatrix} UB \\ -LB \end{bmatrix}, \quad (6)$$

where  $I$  is the identity matrix.

In terms of implementation, efficient ways of sampling truncated Gaussian distribution have been discussed in [17].

### C. Mean-constrained simulation

Mean constraint can be mathematically defined as  $Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$  conditioned on  $\frac{1}{n}\sum_{i=1}^n y_i$  being fixed. The condition is equivalent to  $\sum_{i=1}^n y_i = c$  where  $c$  is a constant.

Consider an invertible  $n \times n$  matrix  $B$  with the last row being the vector  $\vec{1} = [1, 1, \dots, 1]$ ,  $\hat{Y} = BY$  transforms  $Y$  into a new coordinate and  $\hat{Y} \sim \mathcal{N}(\hat{\mu}_Y, \hat{\Sigma}_Y)$ , where  $\hat{\mu}_Y = B\mu_Y$ ,  $\hat{\Sigma}_Y = B\Sigma_Y B^\top$  which can also be written in the form of block matrices

$$\hat{\mu}_Y = \begin{bmatrix} \mu_U \\ \mu_D \end{bmatrix}, \quad \hat{\Sigma}_Y = \begin{bmatrix} \Sigma_{UU} & \Sigma_{UD} \\ \Sigma_{DU} & \Sigma_{DD} \end{bmatrix} \quad (7)$$

After transformation,  $\hat{y}_n = c$ , and  $Z$  defined as  $[\hat{y}_1, \dots, \hat{y}_{n-1}]^\top$  forms a conditional multivariate Gaussian distribution  $(Z|\hat{y}_n = c) \sim \mathcal{N}(\mu_Z, \Sigma_Z)$ , where

$$\mu_Z = \mu_U + \Sigma_{UD}\Sigma_{DD}^{-1}(c - \mu_D) \quad (8)$$

$$\Sigma_Z = \Sigma_{UU} - \Sigma_{UD}\Sigma_{DD}^{-1}\Sigma_{DU} \quad (9)$$

Therefore mean-constrained sampling of  $Y$  can be achieved indirectly through unconstrained sampling of  $Z$ .

### D. Mean and range constrained simulation

Sampling  $Y$  constrained on both the mean and the range can be performed indirectly through linearly constrained sampling of  $Z$ .

As discussed previously, given the transformation  $Y = B^{-1}\hat{Y}$ , the range constraint  $MY \leq N$  can be written as

$$MB^{-1} \begin{bmatrix} Z \\ c \end{bmatrix} \leq N \quad (10)$$

Define  $MB^{-1} = G = [G_1 \quad G_2]$ , the range constraint on  $Y$  becomes the linear constraint on  $Z$ :  $G_1 Z \leq N - G_2 c$ .

### E. Minimum and range constrained Simulation

The minimum value itself introduces a condition and also serves as a refined lower range limit. Consider  $Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$  conditioned on  $y_n = g$ , where  $g$  is the minimum value in  $Y$ . Define  $L = [y_1, \dots, y_{n-1}]^\top$ , and the conditional Gaussian distribution  $(L|y_n = g) \sim \mathcal{N}(\mu_L, \Sigma_L)$  can be derived in a similar way as shown in equations (7), (8) and (9). The new range constraint can be applied following the similar way as shown in equation (10) with new lower bound value  $g$  and one less dimension.

However, the above statement is based on  $y_n = g$  which implies the  $n^{\text{th}}$  random variable is of the minimum value. Due to the fact that the location associate with the minimum value is actually unknown, samples generated subject to the minimum and range constraints need to be drawn from a finite mixture multivariate Gaussian distribution. The simulation process is equivalent to sampling conditional distributions conditioned on uniformly distributed locations of the minimum value. Therefore mean and range constrained sampling of  $Y$  can be achieved indirectly through range constrained sampling of a finite variation of  $L$ .

### F. Minimum, mean and range constrained simulation

Consider  $Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$  conditioned on both  $y_n = g$  and  $\sum_{i=1}^n y_i = c$ . Find an invertible  $n \times n$  matrix  $\tilde{B}$  with the last row being the vector  $\vec{1} = [1, 1, \dots, 1]$ , and the second last row being the vector  $[0, \dots, 0, 1]$ .  $\tilde{Y} = \tilde{B}Y$  transforms  $Y$  into a new coordinate and  $\tilde{Y} \sim \mathcal{N}(\tilde{\mu}_Y, \tilde{\Sigma}_Y)$ , where  $\tilde{\mu}_Y = \tilde{B}\mu_Y$ ,  $\tilde{\Sigma}_Y = \tilde{B}\Sigma_Y\tilde{B}^\top$ . Define  $R = [y_1, \dots, y_{n-2}]^\top$ , and the conditional distribution  $(R|y_n = c, y_{n-1} = g) \sim \mathcal{N}(\mu_R, \Sigma_R)$  can be derived in a similar way as shown in equations (7), (8) and (9). The range constraint can be applied following the similar way as shown in equation (10). Consider the unknown locations of the minimum value, mean-and-range-constrained sampling of  $Y$  can be achieved indirectly through linearly constrained sampling of a finite variation of  $R$ .

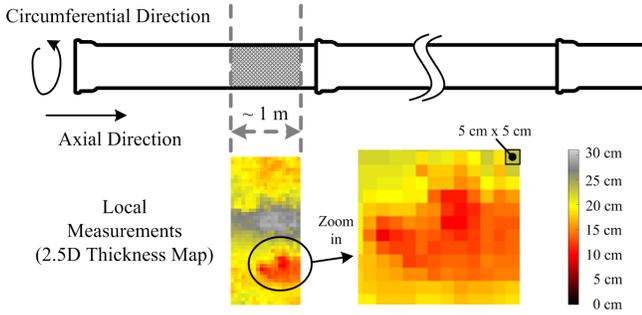


Fig. 3: The remaining wall thickness map of a pipe segment

### G. Covariance function for pipe wall thickness modeling

The above-stated modeling and simulation framework is applicable to any 2.5D data. Note that in a specific application, a model selection is necessary as there is no single mean and covariance function setup that fits all tasks. In our previous work, an anisotropic composite covariance function as described in equation 11 has been proposed and evaluated to better model the wall thickness maps of buried pipes [2].

$$K(X, X^*) = \sigma_1^2 \sigma_2^2 \exp\left(-\frac{2}{l_1^2} \sin^2\left(\frac{\pi d_c}{p}\right) - \frac{d_a}{l_2}\right) + \sigma_3^2 \sigma_4^2 \exp\left(-\frac{2}{l_3^2} \sin^2\left(\frac{\pi d_c}{p}\right)\right) \left(1 + \frac{d_a}{2\alpha l_4^2}\right)^{-\alpha}, \quad (11)$$

where  $d_a$  and  $d_c$  are absolute values of axial and circumferential distances respectively.  $\sigma_i$ ,  $l_i$  and  $p$  are parameters of the model. Please refer to our previous work for more details of the kernel function design [2].

## IV. EXPERIMENTAL SETUP

The evaluation is conducted on the thickness maps of 12 pipe segments, each approximately 1 meter long (ID: S1 to S12). These thickness maps have a spatial resolution of  $5\text{cm} \times 5\text{cm}$  as shown in Figure 3, such that each grid (cell) contains the average thickness value of the area it covers. More information on the ground-truth production can be found in our previous work [2], [19].

Developing a robust objective metric to measure the similarity between correlated 2.5D data (e.g. two thickness maps) remains a challenge in the community. In this research, following the convention, Root-Mean-Square Error (RMSE) for 2.5D data has been employed to evaluate the goodness of prediction. RMSE conveniently has the same unit as the quantity being measured, but it does not explicitly compare structural correlation. A smaller RMSE indicates smaller point-to-point discrepancy in general, thus two identical thickness maps have a reported RMSE of 0.

## V. RESULTS AND DISCUSSION

### A. Inference and realization

Figure 4 shows an example of the GP inference results on thickness map S5 in the form of test output mean and the associated uncertainty. The prediction is a flat surface associated with relatively uniform uncertainty values. Due to

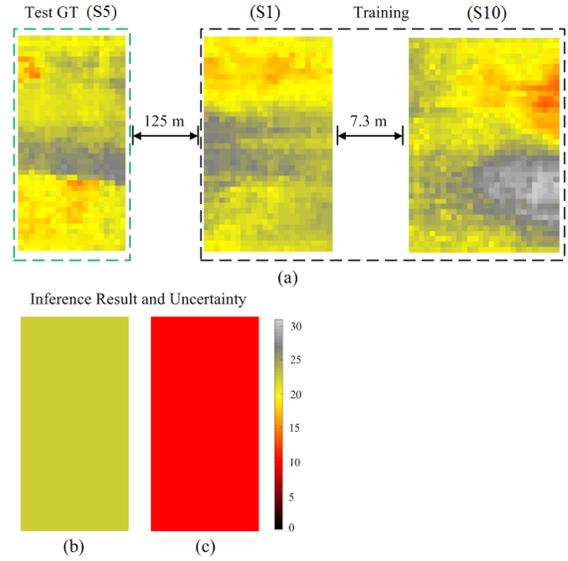


Fig. 4: (a) the ground-truth of thickness map S5, S1 and S10, (b) GP inference results as test output mean on S5, and (c) the associated uncertainty

the lack of spatially close training data, the inference result may not be a suitable target for further analysis. In worse cases, the unavailability of the global location information makes GP inference unfeasible. Unfortunately, in some applications, neither local training data nor accurate location information can be obtained cost-efficiently. Therefore, instead of performing inference, in these situations, studying the realizations generated from a GP model becomes more applicable.

In this experiment, a large amount of realizations (i.e. a bag of realizations) are generated from the learned GP model to collectively represent any thickness map of the size of S5. Figure 5 presents five randomly selected realizations and their RMSE.

Studying each instance in the bag of realizations involves enormous effort, therefore selecting representative realizations makes the analysis process more efficient. For example, if the standard deviation of the thickness values in a realization is set to reflect the quality of the pipe segment, typical thickness maps representing different conditions can be selected from all realizations for the subsequent structural analysis. Figure 6 shows three typical realizations of S5 in terms of variations of thickness values.

Without any additional information, the bag of realizations is supposed to represent any thickness map of the size of S5, but not only the specific thickness map S5. Therefore it is expected that S5 being closely matched with one thickness map in the bag of realizations. Figure 7 provides the most and least similar thickness maps to the reference S5.

### B. Quantitative evaluation of realization results

This section presents the realization results evaluated with leave-one-out cross-validation. Given the twelve thickness maps, for each iteration, one map is retained for testing

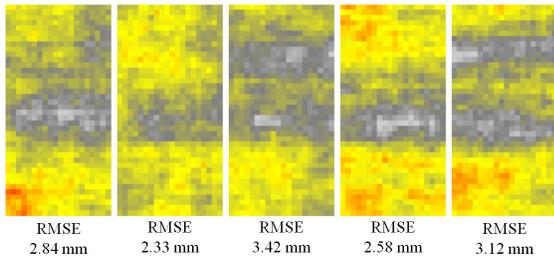


Fig. 5: Five randomly selected realizations of S5

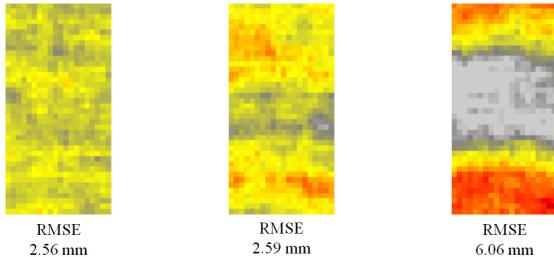


Fig. 6: Three typical (in terms of variations of thickness values) realizations of S5. From the left to the right: realizations with small, median and large thickness variations

while the eleven remaining maps are utilized in training. Then the learned Gaussian Process model is simulated in accord with the dimension of the test data and constraints (if any) derived from the test data, to generate a certain number of realizations which are then compared against the ground-truth of the test data.

Table I shows the evaluation results using RMSE as the evaluation metric, based on the average RMSE values on a hundred thousand realizations in each training-realization-comparison cycle. In 8 out of 12 cases and also in average, truncated sampling results outperform the unconstrained sampling. Due to the fact that the remaining pipe wall thickness is a non-negative value no greater than the nominal wall thickness, the lower and upper bounds constraints should always be applied in this application, thus the results from truncated sampling are set as the benchmark for subsequent experiments.

In Table I, RMSE of mean-constrained-and-truncated realizations shows superior results over truncated-only realizations in all tests. In Table II, RMSE of minimum-constrained-and-truncated realizations shows improved results over truncated-only realizations in 8 out of 12 tests and also in average. The results demonstrate that introducing accurate constraints enables the GP model to produce realizations which are more similar to the reference. However, using RMSE as the evaluation metric, the improvement by introducing the minimum value as a constraint is not as much as by incorporating the mean value as a constraint. The reason is that without knowing the location of the minimum value, realizations having the wrong minimum point locations may result in relatively larger point-to-point discrepancies, i.e. higher RMSE values.

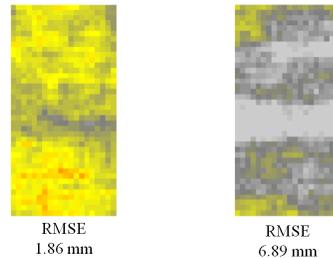


Fig. 7: The most (left) and least (right) similar realizations of S5, in terms of RMSE

Applying a combination of all constraints gives the results shown in the last column in Table II. It further outperforms the minimum-constrained-and-truncated realizations in all tests, as expected. However, as explained above, due to the unknown location of the minimum, in 11 out of 12 tests considering all constraints does not show RMSE reduction over mean-constrained-and-truncated sampling. However, in these cases the performance of incorporating all available constraints is close to mean-constrained-and-truncated sampling, except for in one case (S3).

TABLE I: Cross-validation of realization results (Part I)

Test Map ID	Unconstrained Simul. Avg. RMSE (mm)	Truncated Simul. Avg. RMSE (mm)	Mean-cstr. and Truncated Simul. Avg. RMSE (mm)
S1	2.82	2.63	2.40
S2	5.32	5.03	4.05
S3	3.26	3.22	2.56
S4	3.14	3.33	2.26
S5	2.96	2.70	2.46
S6	3.77	4.06	2.11
S7	4.07	3.80	3.24
S8	3.43	3.22	2.95
S9	4.18	4.38	4.02
S10	3.42	3.47	3.21
S11	3.54	3.39	3.19
S12	3.05	2.82	2.63
<b>Avg.</b>	<b>3.58</b>	<b>3.50</b>	<b>2.92</b>

TABLE II: Cross-validation of realization results (Part II)

Test Map ID	Minimum-cstr. and Truncated Simul. Avg. RMSE (mm)	Mean, Minimum-cstr. and Truncated Simul. Avg. RMSE (mm)
S1	2.55	2.44
S2	4.67	4.57
S3	5.05	3.99
S4	2.29	2.25
S5	2.60	2.46
S6	2.29	2.12
S7	3.40	3.25
S8	3.42	2.95
S9	4.67	3.25
S10	3.63	3.27
S11	3.31	3.20
S12	2.79	2.62
<b>Avg.</b>	<b>3.39</b>	<b>3.03</b>

As typical examples, two reference thickness maps S3, S9 and their closest (in terms of RMSE) realizations un-

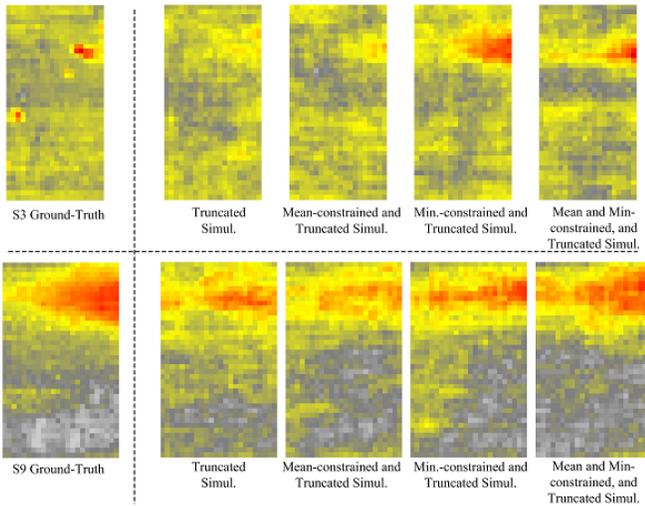


Fig. 8: Ground-truth data and their closest realizations under different sampling strategies for pipe segments S3 and S9

der different constrained sampling strategies is presented in Figure 8. For S3 which has small sections of obvious wall thinness, it can be visually verified that realizations without considering the minimum value tends to result in the lack of thickness variation. For S9, which in general is evenly thin on the upper half section and thick on the lower half section, the best prediction from mean-constrained-and-truncated simulation can be regarded as modest compared to the best prediction after introducing minimum value constraint.

The above results and discussion indicate that introducing constraints guides the sampling process towards generating realizations which satisfy all the actual constraints. However, realizations satisfying all constraints may also vary largely from each other. This is what is required in this particular application because without having access to the ground-truth, all realizations which capture the spatial correlation and satisfy all constraints are regarded as best-possible predictions of the target.

## VI. CONCLUSION

This paper presents an application of GPs to capture the spatial correlation of 2.5D data, and to sample the learned model aiming at generating predictions of the target areas. To further refine the prediction results when limited and sparse information of the target areas is available, constrained sampling incorporating three different types of constraints (range, average and minimum) and their combinations are discussed and corresponding mathematical strategies are presented. The proposed framework is broadly applicable to any 2.5D data. Using condition assessment of buried pipes as the specific application, the spatial correlation of local pipe thickness maps is modeled with GP. Based on the learned model, we proposed to use GP realizations to extrapolate the conditions of the same pipe cohort. The prediction performance is then evaluated using RMSE as the metric and experimental results show that with more constraints introduced, realizations of

this model as best-possible predictions of the target move from general to specific.

## ACKNOWLEDGMENT

This work is an outcome from the Critical Pipes Project funded by Sydney Water, Water Research Foundation USA, Melbourne Water, Water Corporation(WA), UK Water Industry Research Ltd, South Australia Water, South East Water, Hunter Water, City West Water, Monash University, University of Technology Sydney and University of Newcastle.

## REFERENCES

- [1] S. Lacroix, I.-K. Jung, and A. Mallet, "Digital elevation map building with low altitude stereo imagery," *Robotics and Autonomous systems*, vol. 41, no. 2, pp. 119–127, 2002.
- [2] L. Shi, L. S. Sun, T. Vidal-Calleja, and J. Valls Miro, "Kernel-specific gaussian process for predicting pipe wall thickness maps," in *Australas. Conf. Robot. Autom. (ACRA)*, 2015.
- [3] J. M. Mirats Tur and W. Garthwaite, "Robotic devices for water main in-pipe inspection: A survey," *Journal of Field Robotics*, vol. 27, no. 4, pp. 491–508, 2010.
- [4] G. J. Lord, C. E. Powell, and T. Shardlow, *An Introduction to Computational Stochastic PDEs*. Cambridge University Press, 2014, ch. Random Fields, pp. 257–310.
- [5] D. P. Kroese and Z. I. Botev, "Spatial process generation," in *Lectures on Stochastic Geometry, Spatial Statistics and Random Fields, Volume II: Analysis, Modeling and Simulation of Complex Structures*, V. Schmidt, Ed. Springer-Verlag, 2013.
- [6] M. Schlather, A. Malinowski, P. J. Menck, M. Oesting, and K. Storkorb, "Analysis, simulation and prediction of multivariate random fields with package randomfields," *Journal of Statistical Software*, vol. 63, no. 8, pp. 1–25, 2015.
- [7] T. M. Davies and D. Bryant, "On circulant embedding for Gaussian random fields in  $\mathbb{R}^2$ ," *Journal of Statistical Software*, vol. 55, no. 9, pp. 1–21, 2013.
- [8] C. M. Bishop, *Pattern recognition and machine learning*, ser. Information science and statistics. New York: Springer, 2006, book Kernel Methods, pp. 291–324.
- [9] C. E. Rasmussen and C. K. I. Williams, *Gaussian Process for Machine Learning*. Cambridge: MA:MIT press, 2006, ch. Regression, pp. 7–32.
- [10] S. O'Callaghan and F. Ramos, "Gaussian process occupancy maps," *International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, 2012.
- [11] M. Smith, I. Posner, and P. Newman, "Efficient non-parametric surface representations using active sampling for push broom laser data," in *Robot. Sci. Syst. (RSS)*, 2010.
- [12] S. Vasudevan, F. Ramos, E. Nettleton, H. Durrant-Whyte, and A. Blair, "Gaussian process modeling of large scale terrain," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2009, pp. 1047–1053.
- [13] L. Sun, T. Vidal-Calleja, and J. V. Miro, "Bayesian fusion using conditionally independent submaps for high resolution 2.5 D mapping," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2015, pp. 3394–3400.
- [14] T. Vidal-Calleja, D. Su, F. De Bruijn, and J. V. Miro, "Learning spatial correlations for Bayesian fusion in pipe thickness mapping," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2014, pp. 683–690.
- [15] P. Spanos and B. Zeldin, "Monte Carlo treatment of random fields: a broad perspective," *Applied Mechanics Reviews*, vol. 51, no. 3, pp. 219–237, 1998.
- [16] G. Christakos, *Random Field Models in Earth Sciences*. Boston: Academic Press, 1992, ch. Simulation of Natural Processes, pp. 294 – 336.
- [17] C. P. Robert, "Simulation of truncated normal variables," *Statistics and computing*, vol. 5, no. 2, pp. 121–125, 1995.
- [18] Y. Hoffman and E. Ribak, "Constrained realizations of Gaussian fields—a simple algorithm," *The Astrophysical Journal*, vol. 380, pp. L5–L8, 1991.
- [19] J. V. Miro, J. Rajalingam, T. Vidal-Calleja, F. de Bruijn, R. Wood, D. Vitanage, N. Ulapane, B. Wijerathna, and D. Su, "A live test-bed for the advancement of condition assessment and failure prediction research on critical pipes," in *Leading-Edge Strategic Asset Manag. Conf. (LESAM13)*, 2013.