Multivariate Gaussian Process Regression for Evaluating Electromagnetic Profile in Screening Process of Seabed Logging Application

Multivariate Gaussian Process Regression for Evaluating Electromagnetic Profile in Screening Process of Seabed Logging Application

Muhammad Naeim Mohd Aris[1][0000-0001-6481-735X], Hanita Daud[2][0000-0003-1377-4606], Khairul Arifin Mohd Noh[3] and Sarat Chandra Dass[4][0000-0002-7376-0436]

[1,2,3] Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia
[4] Heriot-Watt University Malaysia, 62200 Putrajaya, Malaysia
[1]muhammad_16000991@utp.edu.my [2]hanita_daud@utp.edu.my
[3]khairula.nmoh@utp.edu.my [4]s.dass@hw.ac.uk
Corresponding author: muhammad_16000991@utp.edu.my

Abstract. Computer simulation is an important task for reservoir modeling in screening process of seabed logging. Information acquired from the computer simulation could provide reliable information of electromagnetic (EM) profile and subsurface underneath the seabed. However, the computer simulation could be a time-consuming task in the screening process due to its intricate mathematical equations. In this paper, a predictive model based on Gaussian process regression (GPR) is used to provide information of EM profile at various observations with low time consumption. Multivariate GPR model is developed based on computer simulation outputs. Normalized magnitude versus offset plots are analyzed to eliminate data from any undesired wave interaction. Root mean square error and coefficient of variation between the GPR model and the computer simulation outputs at untried observations are computed. On average, the resulting error was 0.0352 and the coefficient of variation was less than 0.5%. This indicates the multivariate GPR model is well-fitted and capable of evaluating EM profile with low processing-time.

Keywords: Multivariate Regression, Gaussian Process, Seabed Logging.

1 Introduction

Seabed logging (SBL) is an application of controlled-source electromagnetic (CSEM) surveying technique in marine environment to remotely detect and characterize hydrocarbon-filled reservoirs beneath the seafloor. This promising application exploits the fact that hydrocarbon reservoirs have higher electrical resistivity (approximately 30 – 500 Ωm) than its surroundings such as seawater (0.5 – 2.0 Ωm) and sedimentary rocks (1.0 – 2.0 Ωm) [1]. SBL survey is based on the use of an array seabed receivers as the EM signal recorders and a moving source called as horizontal electric dipole (HED) transmitter. The source is normally elevated at 30 – 50 m from the seafloor and it continuously emits low frequency, typically from 0.1Hz to 10.0Hz, from differ-

ent positions through the seawater and the subsurface [2]. Generally, offshore hydrocarbon-saturated reservoirs are embedded within conductive seabed geological formations [3]. Based on the property of electrical resistivity, electromagnetic (EM)-based surveying techniques can provide crucial information that complement seismic surveying technique to characterize hydrocarbon-filled reservoirs beneath the seafloor. In the SBL, screening process prior could identify whether the exploration regions are suitable for marine CSEM mapping or not. In other words, it can determine the applicability of the SBL given the potential target areas. Screening process is done based on synthetic forward modelling codes by computing and simulating synthetic EM datasets. Although this procedure could be disputed due to the CSEM synthetic modelling can be poorly constrained, but according to the [4], the study could indicate the limitation of the SBL application towards the "EM-friendly" regions.

Computer simulation for screening process uses powerful numerical modeling techniques to compute the synthetic EM profile. In geophysics applications, finite-element method (FEM) is commonly employed in the marine CSEM forward modeling due to its powerful capability of modeling offshore reservoirs compared to the other numerical methods. However, solving the linear equations system using FEM algorithms can become a time-consuming task [5]. Researchers in [6] mentioned that evaluating the integrals and solving the linear equations in the FEM require high computational time. In addition, only a few realizations of tried observations of the computer simulation are capable of being acquired after the very lengthy computational time [7]. Addressing these problems, calibration of Gaussian process regression (GPR) and computer simulation is proposed in this research work to evaluate the EM profile at all possible observations in the screening process of the SBL application with low computational time. Multivariate GPR models are developed based on prior information generated through computer simulation technology (CST) outputs. In the context of this research work, the source-receiver separation distance (i.e., offset), depth of hydrocarbon layer from the seabed surface (i.e., thickness of overburden) and the transmission frequency are the multivariate independent variables used corresponding to the magnitude of E-field as the dependent variable. The main goal of this paper is to extend the use of GPR in the SBL and determine its viability as a predictive intelligent computational method to evaluate the EM profile in multivariate distribution. The literature of GPR methodology is briefly discussed in next section.

2 Statistical Background – Gaussian Process Regression (GPR)

Gaussian process regression has successfully been applied in numerous applications especially in computer experiments and machine learning (ML). This intelligent computational method becomes very famous due to its flexibility and its nature of being a probabilistic and non-parametric model [8]. According to [9], Gaussian process regression (GPR) has convenient properties for many tasks of modeling especially for the use of analyzing data in statistics application and ML. This method provides predictions through Bayesian inference [10]. It utilizes the prior knowledge of a condition associ-

ated to an event to describe the probability of the event. Bayesian approach gives an advantage by providing significant and interpretable information on the subject and group levels [11]. As in the GPR, the inference begins with a prior distribution and knowledge is updated as data points are being observed, then a posterior distribution over function is yielded as the output.

Studies related to calibration of statistical methodology and computer experiments can be found in [12–16] The details of the use of GPR in computer experiments also was thoroughly explained in [7]. Learning regression functions from given datasets is an essential task in ML and computer vision in many applications. According to [17], kernel machines such as Support Vector Machines (SVM) and GPR have shown an outstanding performance in learning non-linear mappings for big-dimensional data and low-dimensional as well. However, the regularization of parameters in SVM needs cross validation [18], while GPR learns the parameters without the process. The other advantage is, GPR is a probabilistic approach where it provides fully predictive distribution. Ideally, GPR also is capable of measuring confidence of the predictions. It can provide the predictive variance and show error bars along the predictions. If the predictive distribution spreads widely over the values range, it indicates that the model is uncertain [10].

In other application, GPR also has been utilized in optimizing the received-signal-strength (RSS) threshold to generate useful proximity reports [19]. The complexity of computation from GPR-based RSS models was compared with the other RSS models. From the results, low-cost and low-complex proximity report was able to be generated through the use of GPR algorithms. Next, the viability of GPR in geophysics application has been proven by [20] in determining the reservoir porosity and permeability of the southern basin of the South Yellow Sea. From the findings, estimation of reservoir porosity and permeability using GPR consumed lower computational time than using radial-basis-function-neural-network (RBFNN), back-propagation-neural-network (BPNN) and generalized-regression-neural-network (GRNN). Researchers in [21] and [22] applied GPR as the alternative forward modeling technique in inversion methodology to estimate hydrocarbon depth and resistivity in the SBL, respectively.

3 Data and Method

Subsections 3.1, 3.2 and 3.3 explain synthetic data acquisition using the computer simulation technology (CST) software, methodology of the multivariate Gaussian process regression (GPR) and validation of the GPR model, respectively.

3.1 Synthetic Data Acquisition

As mentioned earlier, the CST software is used to replicate the SBL models and acquire the synthetic EM datasets (i.e., SBL data). The CST software offers CST EM Studio (EMS) module as low-frequency solver to discretize Maxwell's equations. In this research work, two types of SBL models were replicated and modeled which are SBL model with presence of hydrocarbon layer (i.e., target model) and SBL model with all sediment half-space beneath the seabed surface (i.e., reference model). The

target model consists of four horizontal layers which are air, seawater, hydrocarbon and sediment, while the reference model has no hydrocarbon layer. These synthetic SBL models are parameterized based on the literature. Illustration of these two models is depicted in Fig. 1.

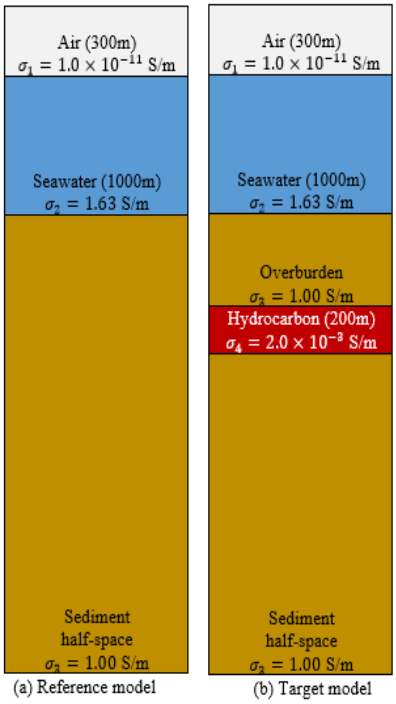

Fig. 1. (a) SBL model without hydrocarbon layer, (b) SBL model with hydrocarbon layer. Sediment above the hydrocarbon layer is known as overburden.

From Fig. 1, air and seawater layers for both target and reference models have similar thicknesses which are 300m and 1000m, respectively. The thickness of the overburden layer indicates the depth of hydrocarbon layer from the seafloor. Note that the CST software designs the simulation models in three-dimensional (3D) (length×width×height: 20km×20km×5km). In order to understand the EM profile with various possibilities of marine environment, various depths of hydrocarbon are tested (1000m – 2000m with gap of 200m each) and thickness of hydrocarbon layer replicated in the target models is fixed at 200m with volume of 80km^3 (length×width×height: 20km×20km×0.2km). A 270m-long transmitter is elevated at 30m from the seabed and positioned at the center of both the simulation models with multiple low-frequencies emitted from the source (0.125Hz, 0.25Hz, 0.375Hz and 0.5Hz). The current of the source is kept at 1250A. Receivers are positioned on the seabed and set to have gap of approximately 100m each along the length of the SBL

models. Note that static source-receiver combination is exercised in this research work along 20km of offset distance.

The property of electrical resistivity is the key parameter in the SBL application. Relevant electrical conductivity (inverse of electrical resistivity) is parameterized for each layer of the simulation models. As of the prior knowledge, 24 EM datasets of the target models for 20km-range source-receiver separation distances, four different transmission frequencies (0.125Hz, 0.25Hz, 0.375Hz, 0.5Hz) and six different tested hydrocarbon depths (1000m – 2000m with an increment of 200m) are generated using the CST software, while four EM datasets of the reference models (for 20km-range offsets and four different frequencies) are generated as well. The multivariable of the target models is portrayed in Fig. 2.

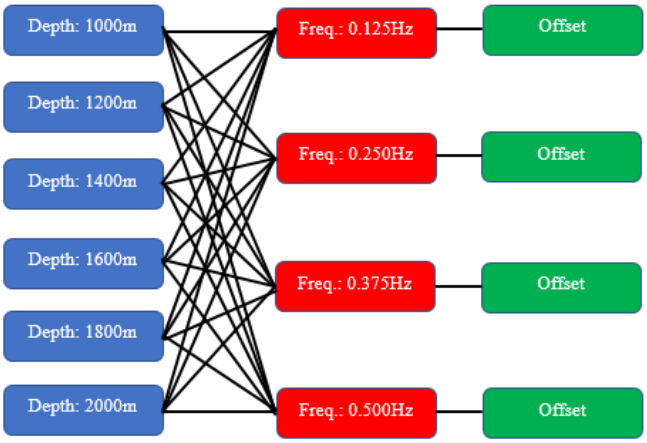

Fig. 2. Multivariable of the target model (hydrocarbon depth×frequency×offset: 6×4×1). For the reference model, the variable of hydrocarbon depth is ignored.

3.2 Developing Multivariate Gaussian Process Regression (GPR) Model

Gaussian process (GP) can be defined as an infinite collection of random variables where any of the finite subsets is normally distributed as well [23–25]. In the context of this research work, zero-mean function $m(x) = 0$ is assumed as the prior since it is a linear combination of random variables that is treated as Gaussian distribution (i.e., normal distribution). 24 training sets of n_{train} different input specifications are assumed to have the multivariate input, x, at spatial locations $x \in \mathbb{R}^{n_{train} \times n_d}$ and the output variable, $y \in \mathbb{R}$, represents as the magnitude of electric field (E-field). There are 62 data points for every EM dataset. In this research work, 1488 input-output specifications of the target model (i.e., n_{train}) are observed with spatial dimension $n_d = 3$ multivariate consisting offset, hydrocarbon depth and frequency. Thus, $x = (p, q, r)$ where $p_a|a = 1,2,\dots,62$, $q_b|b = 1,2,\dots,6$, $r_c|c = 1,\dots,4$ denote the offset, depth and frequency, respectively. The GPR on function f is then expressed in (1).

$$f(x) \sim GP(0, k(x, x')) \tag{1}$$

Mean function in the GPR represents the expected value of the function f at the input point, while covariance function defines the correlation between the points. Squared exponential (SE) covariance function, as shown in (2), is exercised in this research work.

$$k(x, x') = \sigma_f^2 exp\left(-\frac{d^2}{2}\right) \tag{2}$$

where

$$d = \frac{|x - x'|}{\ell} \tag{3}$$

There are two hyper-parameters involved in (3) which are signal variance, σ_f and characteristic-length scale, ℓ. In this research work, these hyper-parameters are optimized by minimizing negative marginal log-likelihood as expressed in (4).

$$logp(y|X, \theta) = \frac{-[nlog(2\pi) + \log|K| + y^T K^{-1} y]}{2} \tag{4}$$

where $\theta = \{\sigma_f, \ell\}$ and K is a covariance matrix of the training datasets obtained by using (2). Logarithmic identifier is used to simplify the integral expression involved in marginal likelihood equation. In order to get the expected function value f^* which is represented by m^*, prior joint Gaussian distribution given the test input point $x^* \in \mathbb{R}^{n_{test} \times n_d}$ (untried observation) is expressed in (5).

$$\begin{bmatrix} m \\ m^* \end{bmatrix} \sim G\left(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X,X^*) \\ K(X,X^*)^T & K(X^*,X^*) \end{bmatrix}\right) \tag{5}$$

where σ_n denotes the signal noise that is normally distributed, $K(X,X)$ is K, $K(X,X^*)$ represents the covariance matrix of the training-testing data and $K(X^*,X^*)$ is the covariance matrix of the testing-testing data. For the test input points, GPR model is used to evaluate EM profile for $n_{test} = 62 \times 15 \times 3 = 2790$ different test input specifications (offset×depth×frequency). The untried observations of the hydrocarbon depth are from 1050m to 1950m with a gap of 50m (the prior information is excluded) and for the transmission frequencies are 0.1875Hz, 0.3125Hz and 0.4375Hz.

Here, the posterior conditional Gaussian distribution given X, y and x^* is expressed in (6).

$$p(m^*|X, y, x^*) \sim G(m^*|\mu^*, \kappa^*) \tag{6}$$

where

$$\mu^* = K(X,X^*)^T [K(X,X) + \sigma_n^2 I]^{-1} y \tag{7}$$

and

$$\kappa^* = K(X^*, X^*) - K(X, X^*)^T [K + \sigma_n^2 I]^{-1} K(X, X^*) \quad (8)$$

The GPR outputs are defined in terms of predictive mean μ^* and predictive variance κ^*. The predictive mean is the EM profile evaluated by the GPR at the untried observations, while the predictive variance represents the uncertainty quantification of the predictive distribution in terms of ± two standard deviations (95% confidence interval).

3.3 Electromagnetic Profile Validation

This subsection briefly discusses the GPR model validation using root mean square error (RMSE) and coefficient of variation (CV) in percentage. The purpose of this subsection is to quantify the differences between the true EM values (generated through the CST) and the estimates (the multivariate GPR model). The RMSE and CV are calculated based on (9) and (10), respectively.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{z}\sum \left(y_i - y_i'\right)^2} \quad (9)$$

and

$$CV = \frac{RMSE}{|\mu_{y'}|} \times 100\% \quad (10)$$

where z is number of data, y is the true value (i.e., CST output), y' is the estimate evaluated by the multivariate GPR and $\mu_{y'}$ is the average of the estimates.

4 Results and Discussion

Normalized magnitude versus offset (MVO) was plotted for all the EM datasets of the simulation models for the purpose of identifying EM data acquired from insignificant waves contribution in the computer simulation. Note that the normalized MVO was calculated by dividing the EM profile from the target model with the EM profile from the reference model. The normalized MVO plot is capable of determining the anomalies underneath the seafloor. If the calculated value is approaching to one, it means that the EM profile of the target model apparently looks like the EM profile of the reference model. This indicate that the surveyed area may absolutely have low certainty of the presence of hydrocarbon-saturated reservoir underneath the seafloor.

Note that all the simulation models were modeled with 20km source-receiver separation distance and the source was located at the center of the simulation models. Thus, all the CST outputs were symmetry. The EM signal traveled at equal separation distances from the source location which was at coordinate of (x = 10000, y = 10000) to the left (0km) and right (20km) of the model boundaries. Therefore, only half of the CST outputs (10km – 20km) were considered for the visualization and GPR modeling

purposes in order to make it more interpretable. The normalized MVO plots for all the EM datasets are depicted in Fig. 3.

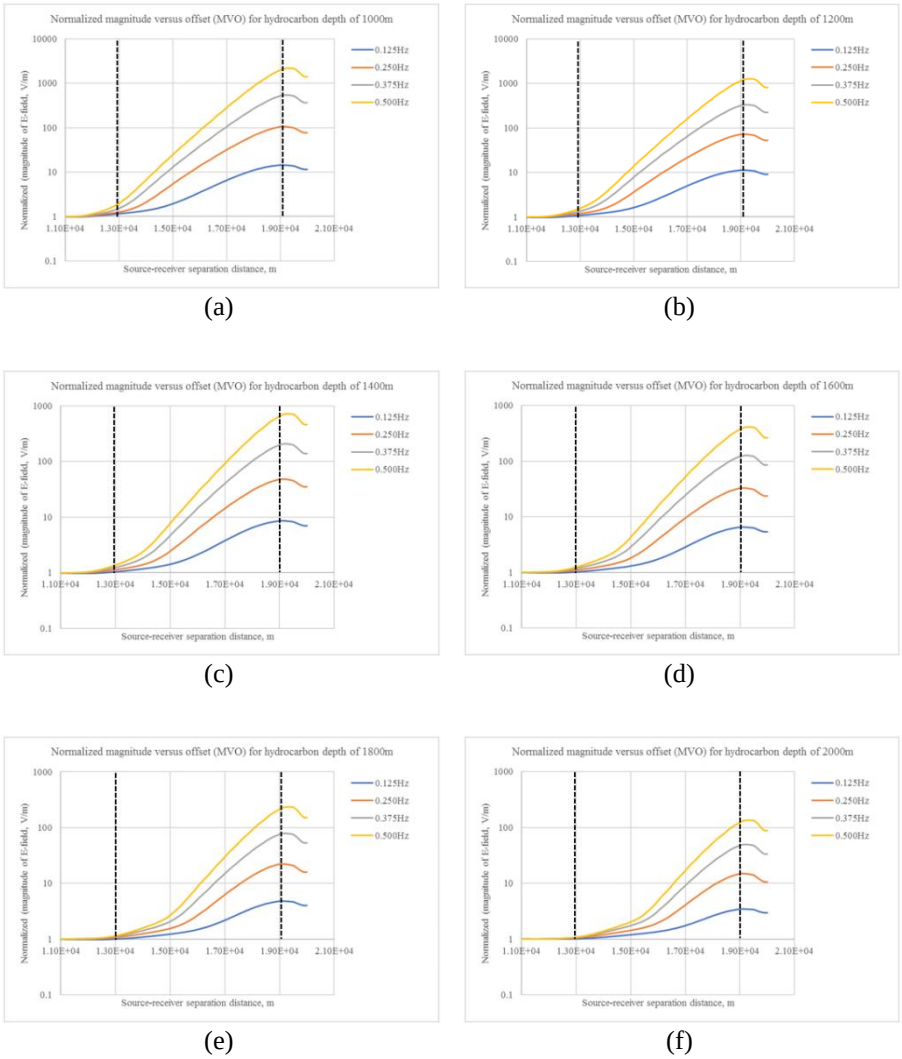

Fig. 3. Normalized magnitude versus offset (MVO); (a) 1000m, (b) 1200m, (c) 1400m, (d) 1600m, (e) 1800m, (f) 2000m.

Based on Fig. 4, there are six different plots where the x-axis represents the source-receiver separation distance (i.e., offset) and the y-axis is the log10 of magnitude of E-field. These plots were grouped based on the depth of hydrocarbon layer. For every hydrocarbon depth, four prior multi-frequencies were exercised in the computer simulation models. Every curve defines normalized magnitude at different transmission

frequency (0.125Hz, 0.25Hz, 0.375Hz and 0.5Hz). It is known that this research work replicated simulation target models with hydrocarbon layer positioned at different depths. Thus, from the normalized MVO plots, the anomalies are present where at certain offsets. The normalized EM profiles become irrelevant at short and long source-receiver separation distances. At short offset which is approximately less than 13km (<~3km from the source), the normalized magnitude values are less than one. Theoretically, The EM responses of the target model should be higher than responses from the reference model. Thus, this research work assumed insignificant wave contribution (e.g., direct wave) occurred at the short offset distances. Referring to Fig. 3, at long offset (>~9km from the source), the normalized magnitudes suddenly decrease after gradually increase from the center of the models due to presence of noise at long source-receiver separation distances. Therefore, this research work inferred to only process the EM data points at offsets of approximately 13km to 19km (~3 – 9km from the source), as shown in Fig. 4.

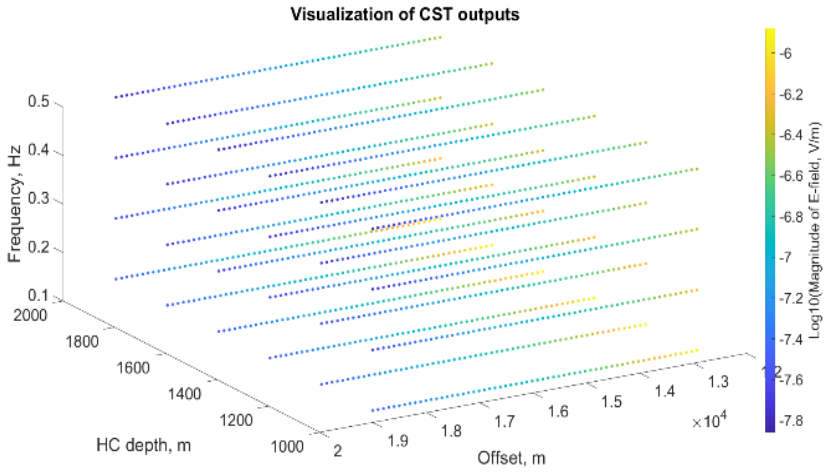

Fig. 4. EM profiles used as the prior knowledge of the GPR model. The colors represent the magnitude of E-field in logarithmic scale.

Fig. 4 shows the visualization of all the CST outputs utilized for the GPR. There are three different axes which represent source-receiver separation distance, depth of hydrocarbon (HC) and transmission frequency. Legend provided in the figure indicates the magnitude of E-field in logarithmic scale with base 10. Based on the color codes, the EM profile exponentially decreases from the smallest offset distance to the biggest distance. This proves that the CST software is capable of providing relevant EM profile to the GPR. The CST consumed approximately 18 minutes to generate four different EM datasets with different frequencies for each of the hydrocarbon depth. Thus, in total, the time for the CST software to generate the 24 EM profiles of the target models was approximately 108 minutes (~1 hour 48 minutes). For the comparison purposes, the multivariate GPR model is depicted in Fig. 5.

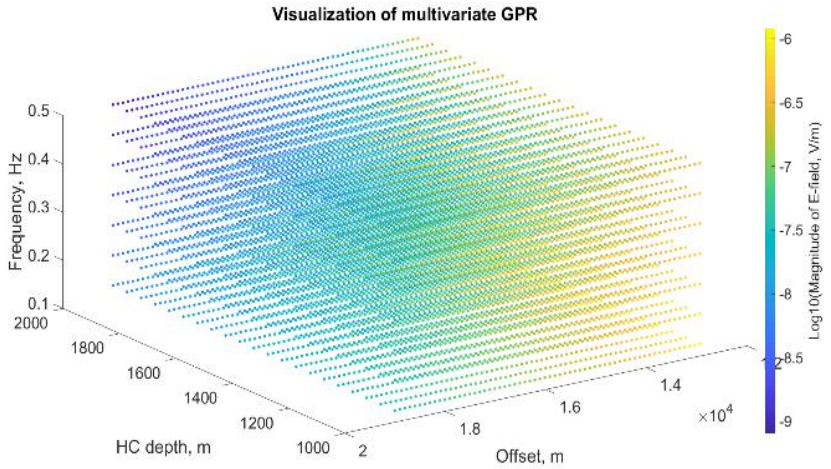

Fig. 5. Multivariate GPR model. The colors represent the magnitude of E-field in logarithmic scale.

Note that GPR algorithms used in this research work can be found in study presented by [9]. The multivariate GPR model consisted prior knowledge from the CST outputs and the posterior distribution (i.e., estimates) at the untried observations. In total, Fig. 5 presents $62 \times 21 \times 7 = 9114$ different EM data points. From the figure, there are seven tested frequencies, 21 hydrocarbon depths and 62 source-receiver separation distance points. GPR took about approximately less than five minutes to evaluate 9114 EM data points. The difference between the time taken of the computer simulation and GPR to generate the EM profiles is very significant. Computer simulation could consume approximately 378 minutes (~6 hours 18 minutes) in order to generate EM profiles with the 21 depths of hydrocarbon and four frequencies. Furthermore, the capability of the GPR to provide such multivariable input points with huge number of outputs (big input-output pair combination) with low time consumption should be highlighted as well. Besides, this GPR modeling also provides the uncertainty quantification along the predictions. The resulting predictive variances revealed very small values. This means that the multivariate GPR is able to fit the synthetic EM datasets very well in order to produce new information.

The EM profiles estimated by the multivariate GPR were validated and compared with the CST outputs at the untried observations. Simulations for target models at the untried observations were modeled and run separately for the purpose of the GPR model validation. Root mean square error (RMSE) and coefficient of variation in percentage (CV) were computed. Low RMSE and CV indicate the GPR is able to evaluate reliable EM profile in the screening process of the SBL application. Fig. 6 shows the comparison of MVO plots between the CST outputs as the true values and the estimates (i.e., GPR model) at untried observations.

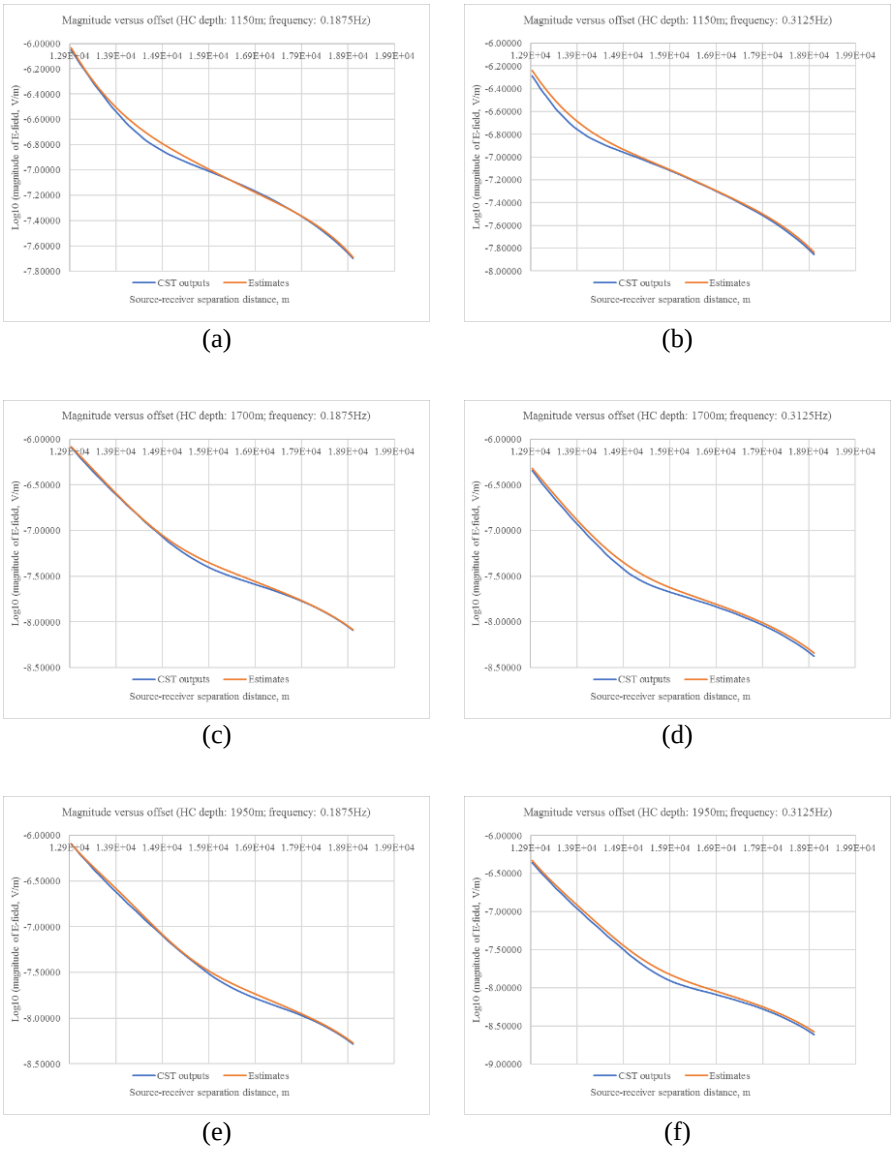

Fig. 6. Comparison of magnitude versus offset (MVO) plots between CST outputs and GPR at random untried observation; (a) depth: 1150m, frequency: 0.1875Hz, (b) depth: 1150m, frequency: 0.3125Hz, (c) depth: 1700m, frequency: 0.1875Hz, (d) depth: 1700m, frequency: 0.3125Hz, (e) depth: 1950m, frequency: 0.1875Hz, (f) depth: 1950m, frequency: 0.3125Hz.

The random untried hydrocarbon depths were 1150m, 1700m and 1950m, while 0.1875Hz and 0.3125Hz for the frequency. Based on Fig. 6, the x-axis represents the offset and the y-axis represents the log10 of the magnitude of E-field. The red curve

denotes the estimates, while the blue curve represents the CST outputs. From the figure, the comparisons look like no significant differences. This inference was strongly strengthened by the numerical proof of the computed RMSE and CV. All the RMSE and CV values for the comparisons shown in Fig. 6 are tabulated in Table 1.

Table 1. The resulting RMSE and CV at the untried observations.

HC depth (m)	Frequency (Hz)	RMSE	CV (%)
1150	0.1875	0.0289	0.4152
	0.3125	0.0333	0.4686
1700	0.1875	0.0268	0.3699
	0.3125	0.0450	0.5985
1950	0.1875	0.0277	0.3769
	0.3125	0.0493	0.6430

Based on Table 1, all the resulting RMSE and CV values are very small. In average, the RMSE was 0.0352 which was less than 0.04, while the average of the CV was 0.48% which was less than 0.5%. This indicates that the GPR model is well-fitted and capable of providing such beneficial information of EM profile at untried observations in the screening process of the SBL application.

5 Conclusion

Multivariate Gaussian process regression (GPR) methodology was proposed in this research work to evaluate electromagnetic (EM) profile in the screening process of the SBL application. The GPR has proved that this methodology is able to estimate the huge number of input-output pair specifications of EM profile with low time consumption compared to the computer simulation. The error measurements (i.e., root mean square error and coefficient of variation) also revealed very small values and percentages which were less than 0.04 and 0.5%, respectively. This means that the multivariate GPR model for the untried observations was successfully developed with deviation from the computer simulation outputs. This attempt could help geophysicists to understand and pre-determine the effect of the setup of the seabed logging (SBL) application such as the transmission multi-frequency used during the survey and the geometry configurations such as hydrocarbon depth towards the EM profile behaviors before in-depth analysis. Besides de-risking the hydrocarbon exploration in the SBL, this attempt also will be very useful for the higher-dimensional modeling that faces insufficient information. Therefore, GPR could be very helpful processing tool for the screening process of the SBL to evaluate EM profile at all possible observations with low time consumption.

Acknowledgments

The authors would like to thank to those who have contributed to this research work. This research work is sponsored by Yayasan Universiti Teknologi PETRONAS-Fundamental Research Grant (YUTP-FRG) (cost center: 015LC0-055).

References

1. Gelius, L.J.: Multi-component processing of sea bed logging data. PIERS ONLINE 2, 589–593 (2006).
2. Chiadikobi, K.C., Chiaghanam, O.I., Omoboriowo, A.O., Etukudoh, M.V., Okafor, N.A.: Detection of hydrocarbon reservoirs using the controlled-source electromagnetic (CSEM) method in the 'Beta' field deep water offshore Niger Delta, Nigeria. International Journal of Science & Emerging Technologies 3(1), 7–18 (2012).
3. Persova, M.G., Soloveichik, Y.G., Domnikov, P.A., Vagin, D.V., Koshkina, Y.I.: Electromagnetic field analysis in the marine CSEM detection of homogeneous and inhomogeneous hydrocarbon 3D reservoirs. Journal of Applied Geophysics 119, 147–155 (2015).
4. Stefatos, A., Boulaenko, M., Hesthammer, J.: Marine CSEM technology performance in hydrocarbon exploration – limitations or opportunities. First break 27, 65–71 (2009).
5. Li, Y., Key, K.: 2D marine controlled-source electromagnetic modeling: Part 1—An adaptive finite-element algorithm. Geophysics 75, WA51–WA62 (2007).
6. Bakr, S.A., Pardo, D., Mannseth, T.: Domain decomposition Fourier finite element method for the simulation of 3D marine CSEM measurements, J. Comput. Phys. 255, 456–470 (2013).
7. Mohd Aris, M.N., Daud, H., Mohd Noh, K.A., Dass, S.C.: Model calibration of stochastic process and computer experiment for MVO analysis of multi-low-frequency electromagnetic data. Processes 8(5), 605 (2020).
8. Liu, D., Pang, J., Zhou, J., Peng, Y., Pecht, M.: Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression. Microelectronics Reliability 53, 832–839 (2013).
9. Rasmussen, C.E., Nickisch, H.: Gaussian processes for machine learning (GPML) toolbox. Journal of Machine Learning Research 11, 3011–3015 (2010).
10. Chan, L.L.T., Liu, Y., Chen, J.: Nonlinear system identification with selective recursive gaussian process model. Industrial & Engineering Chemistry Research 52, 18276–18286 (2013).
11. Turner, B.M., Forstmann, B.U., Wagenmakers, E.J., Brown, S.D., Sederberg, P.B., Steyvers, M.: A Bayesian framework for simultaneously modeling neural and behavioral data. Neuroimage 72, 193–206 (2013).
12. Sacks, J., Welch, W., Mitchell, T., Wyan, H.P.: Design analysis of computer Experiments. Statistical Science 4, 409–423 (1989).
13. Santner, T.J., Williams, B.J., Nolts, W.I.: The design and analysis of computer experiments. Springer-Verlag, Berlin Heidelberg (2003).
14. Higdon, D., Gattiker, J., Williams, B., Rightley, M.: Computer model calibration using high dimensional output. J. Amer. Statist. Assoc. 103, 570–583 (2008).
15. Hills, R.G.: Model validation: Model parameter and measurement uncertainty. Journal of Heat Transfer 128, 339–351 (2006).
16. Wang, S., Chen, W., Tsui, K.L.: Bayesian validation of computer models. Technometrics 51, 439–451 (2009).

17. Ranganathan, A., Yang, M.H.: Online sparse matrix Gaussian process regression and vision applications. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) European Conference on Computer Vision (ECCV 2008), vol. 5302, pp. 468–482. Springer, Berlin Heidelberg (2008).
18. Chu, W., Ghahramani, Z.: Gaussian processes for ordinal regression. Journal of Machine Learning Research 6, 1019–1041 (2005).
19. Yin, F., Zhao, Y., Gunnarsson, F., Gustafsson, F.: Received-signal-strength threshold optimization using Gaussian processes. IEEE Transactions on Signal Processing 65, 2164–2177 (2017).
20. Asante-Okyere, S., Shen, C., Ziggah, Y.Y., Rulegeya, M.M., Zhu, X.: Investigating the predictive performance of Gaussian process regression in evaluating reservoir porosity and permeability. Energies 11, 3261 (2018).
21. Daud, H., Mohd Aris, M.N., Mohd Noh, K.A., Dass, S.C.: A novel methodology for hydrocarbon depth prediction in seabed logging: Gaussian process-based inverse modeling of electromagnetic data. Applied Sciences 11(4), 1492 (2021).
22. Mohd Aris, M.N., Daud, H., Mohd Noh, K.A., Dass, S.C.: Stochastic process-based inversion of electromagnetic data for hydrocarbon resistivity estimation in seabed logging. Mathematics 9(9), 935 (2021).
23. Fang, D., Zhang, X., Yu, Q., Jin, T.C., Tian, L.: A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression. J. Clean. Prod. 173, 143–150 (2018).
24. Kong, D., Chen, Y., Li, N.: Gaussian process regression for tool wear prediction. Mech. Syst. Signal Process 104, 556–574 (2018).
25. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. The MIT Press, Cambridge (2006).