Modelling 1-D synthetic seabed logging data for thin hydrocarbon detection

Citation for published version:
Aris, MNM, Daud, H, Noh, KAM, Dass, SC & Mukhtar, SM 2020, Modelling 1-D synthetic seabed logging data for thin hydrocarbon detection: An application of Gaussian process. in *Proceedings of the 27th National Symposium on Mathematical Sciences (SKSM27).*, 090002, AIP Conference Proceedings, vol. 2266, AIP Publishing, 27th National Symposium on Mathematical Sciences 2019, Bangi, Selangor, Malaysia, 26/11/19. https://doi.org/10.1063/5.0018105

Digital Object Identifier (DOI):
10.1063/5.0018105

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Published In:
Proceedings of the 27th National Symposium on Mathematical Sciences (SKSM27)

Modelling 1-D Synthetic Seabed Logging Data for Thin Hydrocarbon Detection: An Application of Gaussian Process

Muhammad Naeim Mohd Aris[1, a)], Hanita Daud[1, b)], Khairul Arifin Mohd Noh[2, c)], Sarat Chandra Dass[3, d)] and Siti Mariam Mukhtar[1, e)]

[1]*Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia*

[2]*Department of Geosciences, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia*

[3]*School of Mathematical and Computer Sciences, Heriot-Watt University Malaysia, 62200 Putrajaya, Wilayah Persekutuan, Malaysia*

a)Corresponding author: muhammad_naeim@yahoo.com
b)hanita_daud@utp.edu.my
c)kharula.nmoh@utp.edu.my
d)s.dass@hw.ac.uk
e)mariamukhtar_91@yahoo.com

Abstract. Seabed Logging (SBL) is a technique that employs high-powered electric dipole source to emit electromagnetic (EM) signal to detect hydrocarbon (HC) reservoirs beneath the seabed. This application is based on electrical resistivity contrasts between target reservoirs and its surrounding. SBL analysis can become a challenging task when the target reservoirs are thin, and the contrasts of resistivity are not very significant. As HC reservoirs are getting thinner, target responses and reference (non-HC) responses are difficult to be distinguished. Addressing this problem, we propose a simple statistical method, Gaussian Process (GP), to model one-dimensional (1-D) SBL data with uncertainties quantification. In this paper, Computer Simulation Technology (CST) software was used to replicate SBL models with five different thicknesses of HC. Some characteristics of the SBL models such as seawater depth, reservoir thickness and reservoir depth were imitated as the case study of Troll West Oil Province, North Sea. We developed 1-D forward GP model for all the SBL responses. Both modelled responses, target and reference, were compared and mean percentage differences between the responses were then calculated. For every comparison, confidence bars for each modelled response were observed to confirm the existence of thin HC. For model validation, root mean square errors (RMSEs) between modelled and generated (CST software) data were calculated. The confidence intervals revealed that the target and reference responses are distinguishable for all HC thicknesses, and the calculated RMSEs showed that GP is reliable to be applied in SBL data to provide uncertainties quantification.

INTRODUCTION

Nowadays, an innovative application of Controlled-Source Electromagnetic (CSEM) technique to explore hydrocarbon (HC) reservoirs in deep water environment has gained a lot of attention in petroleum industry. This application is known as Seabed Logging (SBL). The reliability of this application in measuring the sub-surface electrical resistivity beneath the seabed has successfully been approved [1]. In marine CSEM, HC-filled reservoirs have effectively been characterized due to the fact that HC has higher electrical resistivity than the surroundings (e.g. saline water formations and sedimentary rocks). Typically, it is known to have a few tens of Ohm meter of electrical resistivity (30-500 Ohm-m), whereas seawater and sediment are about 0.5-2.0 Ohm-m and 1.0-2.0 Ohm-m, respectively [2].

The most important basis in SBL application is the use of Horizontal Electric Dipole (HED) transmitter as the source, and electromagnetic (EM) receivers, as the magnetic and electric sensors [3]. The acquisition method has commonly been performed in inline tows with horizontal dipole [4]. The EM energy is continuously emitted from the source along the survey, whereas the EM receivers, placed on the seabed, record the returned signals. Marine CSEM technique employs low-frequency EM wave during the survey. According to [5], the frequency of the EM signal can be between 0.01 Hz and 10 Hz, and it can travel up to 10 km offset. The purpose of exercising low-frequency EM wave is to obtain higher wavelength as the function of offset distance [2]. The representation of the SBL survey in deep offshore environment is depicted in Fig. 1.

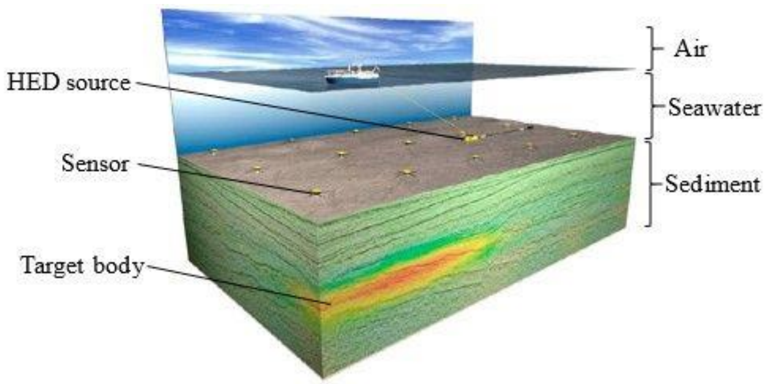


FIGURE 1. Representation of SBL survey

In real-field environment, the magnitudes of electric field (E-field) acquired from SBL survey decrease as the thickness of HC reservoirs decrease [6]. Due to this property, the magnitudes of E-field acquired from thin resistive reservoirs (target responses) and non-HC regions (reference responses) are indistinguishable and hardly to be interpreted. Without any uncertainty quantifications, misinterpretation may occur in the processing of one-dimensional (1-D) SBL data. As mentioned by [7 – 9], processing the data numerically and analytically has become a challenging task to many geophysicists.

Therefore, we propose a simple and flexible statistical method, called as Gaussian Process (GP), to model synthetic 1-D SBL data (consist of magnitude of E-field and offset distance) with uncertainties quantification. In this paper, some characteristics of the offshore environment such as the seawater depth, thickness of the reservoir and reservoir depth are imitated as discussed by [10]. Referring to [10], data in the case study were acquired from Troll West Oil Province, located in the northern part of the North Sea. The survey was done in offshore environment with seawater depth of 310 m to 350 m, and the target reservoir was located at the depth of 1460 m with thickness of approximately 22 m to 26 m. Thus, based on this information, SBL models are replicated to generate data, and then, GP is used to model the 1-D SBL data and quantify the uncertainties.

Although better representation of real models can be provided by higher dimensional modelling, 1-D analysis can give notable contribution to the exploration especially when the acquired datasets are insufficient to analyze in multi-dimensional interpretation. By using GP, we try to fully utilize the synthetic 1-D SBL data to provide some useful insights to the SBL application. The way on how the proposed method could help in making decision for HC exploration is thoroughly explained in the next sections.

STATISTICAL BACKGROUND

Theoretically, Gaussian Process (GP) is a finite collection of random variables such that for every collection has a multivariate Gaussian distribution. GP is a distribution over function. This probabilistic and non-parametric approach has gained immense interest to many areas of scientific and engineering applications such as geo-statistics, machine learning, electronics, etc. [11]. According to [12], excellent performance has been shown by GP regression in plenty of applications. GP approach follows Bayesian interpretation settings where over-fitting in the modelling can be avoided [11].

Ideally, GP is defined by a mean function, $m(s)$, and a covariance function, $k(s, s')$. The collection of random variables is distributed according to GP function, $(f(s_1), f(s_2), \dots, f(s_n)) \sim G\ (\mu, \Sigma)$, where

$$\mu = (m(s_1), m(s_2), \dots, m(s_n)) \tag{1}$$

$$\Sigma = (k(s, s'))_{n \times n} \quad (2)$$

Uncertainties quantification provided by GP is the main reason why this approach is very famous. [13] stated that GP is capable of providing a full predictive distribution. It can provide the variance measured as a description of uncertainties in terms of $\pm$ two standard deviation. Thus, from this advantage, we attempt to utilize GP in SBL application since there are a lot of possible risks that may happen while interpreting and processing the acquired 1-D SBL data. The details of the stepwise processes can be referred in [14 – 15].

METHODOLOGY

This section is separated into four sub-sections, which are: (i) seabed logging (SBL) modelling using Computer Simulation Technology (CST) software; (ii) one-dimensional (1-D) SBL data acquisition; (iii) statistical analysis – Gaussian Process; (iv) calculating mean percentage difference and model validation using root mean square error (RMSE). Details are explained as below.

SBL Modelling Using CST Software

Computer Simulation Technology (CST) software is capable of replicating SBL survey to generate synthetic SBL data as suggested by [16 – 17]. Yahya et al. (2012) [18] mentioned that this software can be used to solve any applications of low-frequency problem. CST software uses Finite Integration Technique (FIT) to discretize the Maxwell's equations involved in the data interpretation to probe the resistivity contrasts. FIT solves the equations in a finite calculation domain which is in grid cell where the mesh element can be in hexahedral or tetrahedral [19]. In this paper, two types of SBL model are replicated by the CST software, which are model without presence of HC (reference model) and model with different thicknesses of HC (target model). The illustrations of the models are represented in Fig. 2(a) and 2(b), respectively.

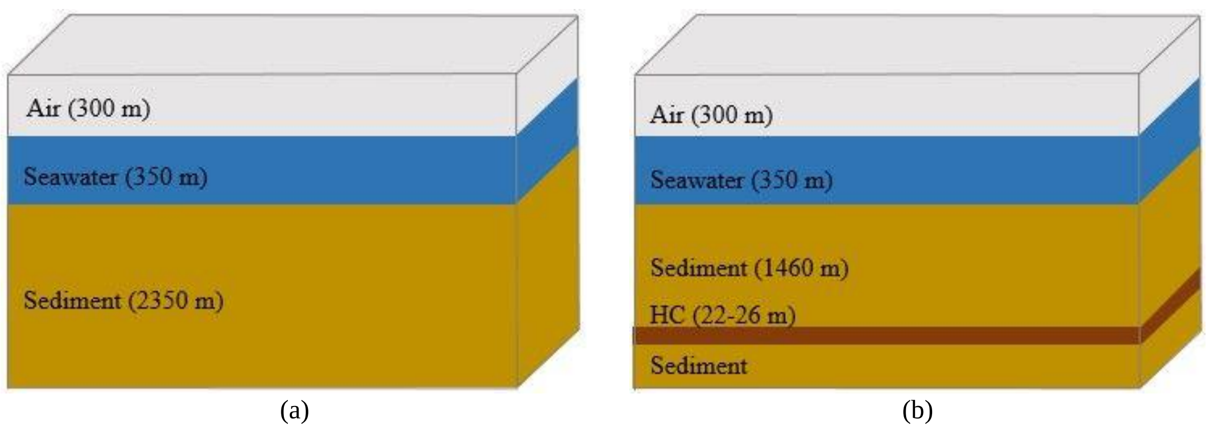


FIGURE 2. (a) SBL model without HC reservoir (reference model) (b) SBL model with HC reservoir (target model)

As mentioned earlier, the seawater depth, HC thickness and HC depth are imitated as in the case study of Troll West Oil Province presented by [10]. The air and seawater layers for both models are kept at 300 m and 350 m, respectively. From Fig. 2(b), the HC reservoir is located at 1460 m from the seabed surface. As the main concern of this work, the thickness of HC is varied from 22 m to 26 m with an increment of 1 m. The resistive body is designed to have 10x5 km^2. The length and the width of both models are fixed at 10 km, whereas the height of the models is 3 km. The indication and details of every layer in each replicated model are tabulated in Table 1.

TABLE 1. Indication of each SBL model replicated in this research work

SBL Model	Thickness (m)				
	Air	Seawater	Sediment		Hydrocarbon (HC)
			Overburden	Under Burden	
Reference Model: 1	300	350	2350		No HC
Target Model: 2	300	350	1460	868	22
Target Model: 3	300	350	1460	867	23
Target Model: 4	300	350	1460	866	24
Target Model: 5	300	350	1460	865	25
Target Model: 6	300	350	1460	864	26

The simulation is carried out with 270 m long HED transmitter. The source is located at the center of the model and 30 m above the seabed. The current and frequency used are fixed at 1250 A and 0.125 Hz, respectively. These settings are found to be reliable and suitable for synthetic SBL survey as exercised by [17], [18], [20], [21]. In real environment, each geophysical structure has its own physical properties. For this paper, each layer is specified with unique relative permittivity, electrical conductivity and thermal conductivity. The physical properties of all layers involved in the models are referred as in [17] and tabulated in Table 2.

TABLE 2. Physical properties of each layer

Layer	Relative Permittivity, (*F/m*)	Electrical Conductivity, (*S/m*)	Thermal Conductivity, (*W/mK*)
Air	1	1×10^{-11}	0.024
Seawater	80	1.630	0.593
Hydrocarbon	4	0.002	0.492
Sediment	30	1.000	2.000

It should be noted that electrical conductivity is the inverse of electrical resistivity [6]. From Table 2, HC has lower electrical conductivity than its surroundings (seawater and sediment). This implies that HC reservoir replicated in this research work is parameterized with a reliable electrical resistivity such in real-field environment.

1-D SBL Data Acquisition

Six responses from different SBL models are acquired in this paper. The data are generated for total offset distance of 10 km. The source is located at the center of the model. This means that the EM signals travel equidistant from the center to 5 km offset on the left and 5 km offset on the right of the model. The generated data from 0 km to 5 km is identical to the data from 5 km to 10 km, but with different direction. Thus, in this work, only data generated from 5 km to 10 km offset are considered to be analyzed for data processing purposes. Logarithmic scale with base 10 is applied to the magnitudes of E-field to make the data interpretable since the values of the acquired data are extremely small.

Statistical Analysis – Gaussian Process

As mentioned earlier, Gaussian Process (GP) is fully defined by a mean function, $m(s)$, and covariance function, $k(s,s')$. Zero mean function, $m(s)=0$, is used in this work as the prior mean. Prior zero-mean function does not imply zero-mean predictive distribution and exercising a simple mean function as the prior make the model easy to interpret. Thus, a GP on function, f, is written as $f(s) \sim GP(0, k(s,s'))$. For every thickness of HC, let $Y(s_i)$ be the CST outputs which are the magnitudes of E-field at $n=60$ different specification of inputs, and offset distance, s_i , where $i=1,2,\ldots,60$.

Squared exponential (SE) is chosen as the covariance function used in this research work. SE covariance function is very famous, flexible and infinitely differentiable. The equation of SE covariance function is as follows,

$$k(s,s') = \sigma_f^2 \, exp\left(\frac{-1}{2l^2}|s-s'|^2\right) \tag{3}$$

where σ_f is the signal variance which identifies the function variance from the mean, and l is the characteristic-length scale which describes how smooth the function is. These hyper-parameters control the smoothness of the function over the distribution. In the GP model, these hyper-parameters are properly estimated by numerically minimizing the negative marginal log-likelihood. Next, the predictive equation for predicting the magnitude of E-field at m unobserved inputs, s^*, is given as below.

$$\hat{Y}(s^*) = \sum_{i=1}^{m} \hat{\beta}_i f_i(s^*) + K_*^T K^{-1}(Y - Y_0) \tag{4}$$

where K is the covariance matrix, K_* is the column vector (correlation between unobserved and observed data points), $\hat{\beta}_i$ is the estimate of β_i yielded from Bayesian methodology, Y is the column vector of CST outputs, and Y_0 is the column vector of $\hat{Y}_0(s_i)$. Equation (4) is the main equation in GP prediction. Details of this methodology in SBL application can be referred in numerical example discussed in [22].

Calculating Mean Percentage Difference and Model Validation Using RMSE

Mean percentage difference is calculated to see the difference between target response and reference response at each input point. Mean percentage difference can become a monitoring procedure to determine whether the data acquired from the CST software agree with the real-field data behavior or not. The equation of the percentage difference is referred as in [17].

$$Mean\ percentage\ difference = \frac{\sum\left[\frac{(y_i(HC) - y_i(No\_HC))}{y_i(HC)} \times 100\%\right]}{k} \tag{5}$$

where $y_i(HC)$ is the target responses (magnitudes of E-field from model with HC), $y_i(No\_HC)$ is the reference responses (magnitudes of E-field form model without presence of HC), and k is the total number of data points.

Next, the purpose of model validation is to determine the reliability of the developed 1-D forward GP models whether it can fit well in the 1-D SBL data or not. Root mean square error (RMSE) is calculated between data modelled by the GP and data generated through the CST software (measured by FIT). Small values of RMSE indicate the developed GP models are fit well, and information provided by the GP is reliable to be used to confirm the presence of HC by quantifying the uncertainties.

$$RMSE = \sqrt{\left(\frac{1}{k}\sum (y_i - y_i^*)^2\right)} \tag{6}$$

where y_i is the magnitudes of E-field generated through the CST software (measured by FIT), y_i^* is the magnitudes of E-field modelled by the GP, and k is the total number of data.

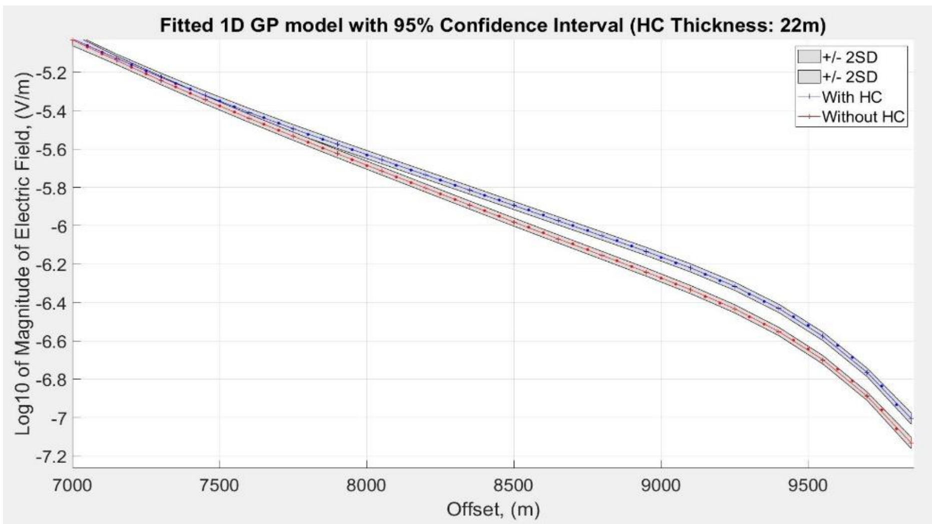
RESULTS AND DISCUSSION

The mean percentage differences between responses of model with hydrocarbon layer (target model) and without hydrocarbon layer (reference model) are tabulated in Table 3. Based on Table 3, the mean percentage difference of SBL model 1 and 2 is the lowest compared to the other models. This is because, as the thickness of HC is getting thinner, the values of the response (magnitudes of E-field) between model with and without HC are getting nearer and become very hard to distinguish. Thus, it implies that the data generated through the CST software are in an agreement with the behavior of real-field data.

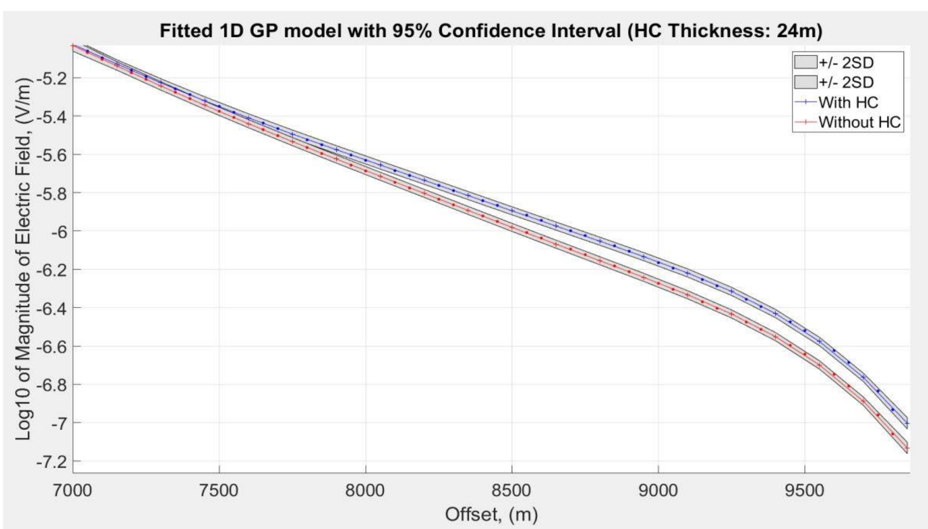
TABLE 3. Mean percentage difference between target model for each HC thickness and response model

SBL Model Comparison	Mean Percentage Difference (%)
1 and 2	17.0460
1 and 3	17.0703
1 and 4	17.0750
1 and 5	17.0831
1 and 6	17.0913

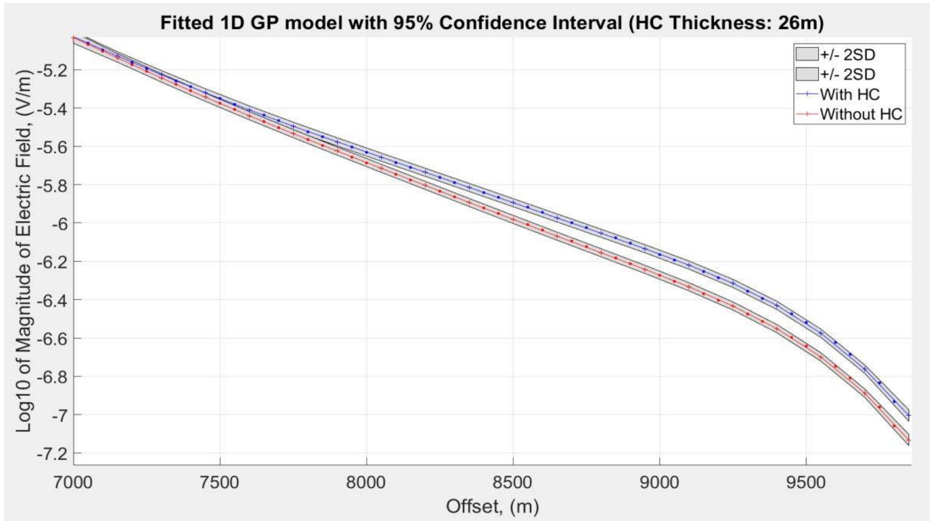
Next, we proceed with the thin HC detection. As mentioned earlier, each response was modelled by using GP, and for every thickness, model with HC (target model) was compared with model without HC (reference model) in order to confirm the presence of the thin HC reservoir. Due to limited spaces, only three figures are provided in this paper. Figure 3(a), 3(b) and 3(c) are the modelled responses for comparison between target models (with thicknesses of 22, 24, 26 m) and reference model.



(a)



(b)



(c)

FIGURE 3. (a) 1-D forward GP model between target responses with 22 m thick HC and reference responses; (b) 1-D forward GP model between target responses with 24 m thick HC and reference responses; (c) 1-D forward GP model between target responses with 26 m thick HC and reference responses

Based on figures above, the x-axis is the offset distance from 7 km to 10 km, and the y-axis is the log10 of magnitude of E-field. HC stated in the legend denotes hydrocarbon. The grey-shaded regions in these figures are the 95% confidence interval provided by the GP which acts as the uncertainty measurement. From these figures, we can see that the confidence bars between the target responses and reference responses are not overlapping each other. This means that the data points between the target and reference models are certainly non-dispersed from the $\pm$ two standard deviation. This analysis reveals that the presence of the thin HC (22-26 m) is able to be detected and confirmed by providing the uncertainties quantification.

Next is for model validation. All data points from SBL model 1 to SBL model 6 were validated by using RMSE. As mentioned earlier, data generated through CST software were measured and solved using FIT. Thus, this sub-section is purposely done to determine the reliability of the developed model by evaluating the differences (errors) between data generated through the CST software and data modelled by the GP. Assume that data measured by the FIT are the true values. The calculated RMSEs are tabulated in Table 4.

TABLE 4. RMSEs between modelled datasets and data generated through CST software

SBL Model	RMSE
1	0.012315001
2	0.012685111
3	0.012698648
4	0.012860642
5	0.012347879
6	0.012339854

From this table, we can see that all values of the calculated RMSE are very small and at average of 0.0125. This means that the modelled responses (using GP) and responses measured using the FIT have no significant differences. The RMSE values were indirectly influenced by the chosen hyper-parameters involved in (3) as well. The negative marginal log-likelihood for each data fitting was well minimized to obtain the best modelled data sets. Based on the RMSEs tabulated in Table 4, it implies that GP is capable of evaluating the 1-D SBL data with much simpler ways and equations.

CONCLUSION

This study proposed a numerical methodology to confirm the existence of thin hydrocarbon reservoir in 1-D SBL data-processing by quantifying the uncertainties of E-field responses by using GP. In 1-D EM data interpretation, the presence of hydrocarbon was able to be detected, but, the uncertainties are still become the big concern in the hydrocarbon detection. GP-based analysis was proven to be reliable in thin hydrocarbon detection. The uncertainties quantification provided by the GP is very useful in order to provide a certain claim of the existence of thin hydrocarbon beneath the seabed. Based on the results, responses from SBL models with hydrocarbon and without hydrocarbon were able to certainly be distinguished by utilizing the confidence intervals provided by GP even though the hydrocarbon thicknesses are very thin (about 22-26 m). In addition, the reliability of the developed GP models also was proven where all the calculated RMSEs were very small. This means that GP can fit well the synthetic 1-D SBL data. This analysis could be an alternative methodology to SBL data interpretation before in-depth analysis is done.

ACKNOWLEDGEMENT

All the contributions are greatly acknowledged. This work was funded by Fundamental Research Grant Scheme (FRGS) from Ministry of Higher Education (MOHE) Malaysia with the cost center of 0153AB-I83 and Yayasan Universiti Teknologi PETRONAS-Fundamental Research Grant (YUTP-FRG) with the cost center of 015LC0-055.

REFERENCES

1. K. Feather, Scandinavian Oil and Gas Magazine, **5/6**, 37 (2007).
2. H. Daud, N. Yahya, V. Sagayan and M. Talib, Proceedings – 2011 IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2011, Penang, 2011, pp. 80–85.
3. T. Eidesmo, S. Ellingsrud, L. M. MacGregor, S. Constable, M. C. Sinha, S. Johansen, F. N. Kong and H. Westerdahl, First Break, **20**, 144 (2002).
4. K. Key, Geophysics, **74(2)**, F9 (2009).
5. S. A. Bakr, D. Pardo and T. Mannseth, Journal of Computer Physics, **255**, 456 (2013).
6. A. Kumar and J. E. Lie, Proceedings of 7th International Conference & Exposition on Petroleum Geophysics, Panjagutta, 2008, p. 93.
7. C. S. Cox, S. C. Constable, A. D. Chave and S. C. Webb, Nature, **320(6057)**, 52 (1986).
8. S. Ellingsrud, T. Eidesmo, T. Schaug-Pettersen and H. M. Pedersen, European Patent Office, EP 1 309 887 B1 (2000).
9. S. Ellingsrud, T. Eidesmo, M. C. Sinha, L. M. MacGregor and S. C. Constable, Leading Edge, **20(10)**, 972 (2002).
10. M. Zhdanov, M. Endo and J. Mattsson, in *Offshore Technology Conference*, (Offshore Technology Conference, Texas, 2015).
11. O. Harari and D. M. Steinberg, J. Stat. Plan. Inference, **154**, 87 (2014).
12. A. Schwaighofer and V. Tresp, in *Advances in Neural Information Processing Systems*, edited by S. Becker, S. Thrun and K. Obermayer (MIT Press, Canada, 2003).
13. L. L. T. Chan, Y. Liu and J. Chen, Industrial & Engineering Chemistry Research, **52(51)**, 18276 (2013).
14. M. N. Mohd Aris, H. Daud and S. C. Dass, Journal of Physics: Conf. Series, **1123**, 012025 (2018).
15. M. N. Mohd Aris, H. Daud, S. C. Dass and K. A. Mohd Noh, Processes, **7(10)**, 661 (2019).
16. H. Daud, V. A. Sagayan, R. Razali and M. Talib, AIP Conference Proceedings, **1605(1)**, 268 (2014).
17. S. M. Mukhtar, H. Daud and S. C. Dass, Procedia Computer Science, **72C**, 225 (2015).
18. N. Yahya, N. Nasir, M. N. Akhtar, M. Kashif, H. Mohd Zaid and A. Shafie, Journal of Electromagnetic Analysis and Applications, **4** (2012).
19. A. Canova, F. Dughiero, F. Fasolo, M. Forzan, F. Freschi, L. Giaccone and M. Repetto, IEEE Transactions on Magnetics, **45(3)**, 1855 (2009).
20. A. Ansari, A. Shafie and A. B. Said, International Journal of Computer Science, **9(3)**, 214 (2012).
21. M. Rauf, N. Yahya, T. E. Nyamasvisva, A. Ansari, A. Shafie and N. Nahar, Proceedings of 2014 IEEE Asia-Pacific Conference on Applied Electromagnetics, APACE 2014, Johor Bahru, 2015, pp. 99–102.
22. M. N. M. Aris, H. Daud, S. C. Dass and K. A. M. Noh, Journal of Environmental and Engineering Geophysics, **24(3)**, 399 (2019).