



Article Stochastic Process-Based Inversion of Electromagnetic Data for Hydrocarbon Resistivity Estimation in Seabed Logging

Muhammad Naeim Mohd Aris^{1,*}, Hanita Daud¹, Khairul Arifin Mohd Noh² and Sarat Chandra Dass³

- ¹ Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia; hanita_daud@utp.edu.my
- ² Department of Geosciences, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia; khairula.nmoh@utp.edu.my
- ³ School of Mathematical and Computer Sciences, Heriot-Watt University Malaysia, Putrajaya 62200, Malaysia; s.dass@hw.ac.uk
- * Correspondence: muhammad_16000991@utp.edu.my

Abstract: This work proposes a stochastic process-based inversion to estimate hydrocarbon resistivity based on multifrequency electromagnetic (EM) data. Currently, mesh-based algorithms are used for processing the EM responses which cause high time-consuming and unable to quantify uncertainty. Gaussian process (GP) is utilized as the alternative forward modeling approach to evaluate the EM profiles with uncertainty quantification. For the optimization, gradient descent is used to find the optimum by minimizing its loss function. The prior EM profiles are evaluated using finite element (FE) through computer simulation technology (CST) software. For validation purposes, mean squared deviation and its root between EM profiles evaluated by the GP and FE at the unobserved resistivities are computed. Time taken for the GP and CST to evaluate the EM profiles is compared, and absolute error between the estimate and its simulation input is also computed. All the resulting deviations were significantly small, and the GP took lesser time to evaluate the EM profiles compared to the software. The observational datasets also lied within the 95% confidence interval (CI) where the resistivity inputs were estimated by the proposed inversion. This indicates the stochastic process-based inversion can effectively estimate the hydrocarbon resistivity in the seabed logging.

Keywords: stochastic process; Gaussian process; seabed logging; electromagnetic data; gradient descent; inversion; optimization; hydrocarbon resistivity

1. Introduction

Determining high resistive structures, such as hydrocarbon-saturated reservoirs without wells drilling, is of crucial significance in oil and gas industry. In this context, the traditional surveying technique, such as seismic method, has been employed. The method utilizes sound waves to map the geological subsurface, but, according to Reference [1], the seismic results can be uncertain. Therefore, researchers in Reference [2] introduced seabed logging (SBL) to remotely determine the offshore hydrocarbon-saturated reservoirs and become the complementary tool to the seismic method. Marine controlled-source electromagnetic (CSEM) technique (also known as SBL) employs low-frequency electromagnetic (EM) energy sourced by a very powerful electric dipole, towed by a ship, to characterize high resistive structures (e.g., hydrocarbon) beneath the seabed [3]. This frequency-domain technique relies on a fact that hydrocarbon has higher electrical resistivity than its surroundings. Electrical resistivity of the hydrocarbon can approximately reach 30 to 500 Ω m, while the resistivities of its surroundings, such as seawater and sediment, are approximately 0.5 to 2.0 Ω m and 1.0 to 2.0 Ω m, respectively. The property of the electrical resistivity (inversely proportional to conductivity) plays an important role for the hydrocarbon exploration. Main references of the SBL can be found in References [1-12].



Citation: 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* 2021, 9, 935. https://doi.org/10.3390/ math9090935

Academic Editor: Vlad Stefan Barbu

Received: 9 February 2021 Accepted: 8 April 2021 Published: 23 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

Besides capable of detecting the hydrocarbon-saturated reservoirs remotely (i.e., without well drilling), the numerical forward modeling methods used for processing the CSEM data are also very powerful and reliable. Practically, the forward modeling is reliant on an assumption that the property of the resistivity is known. Then, a few processes take place in the forward solution, such as simulating the data based on the model, comparing the simulation with the measured data, and updating the model. The surveying CSEM technique usually employs finite element (FE) or finite difference (FD) methods for the dataprocessing [13]. FE method is preferable due to its ability to support unstructured meshes in order to represent the complex subsurface realistically. Note that these deterministic numerical methods are mesh-based algorithms which involve complicated integrations and linear system. Even though they are reliable, high computational efforts, in terms of cost and time, are needed especially for solving the integrals and the linear equations [14]. In addition, inverse modeling is also very useful, especially to recover the physical property (e.g., resistivity) of hydrocarbon; however, according to researchers in Reference [14], the inversion scheme needs multiple forward solutions, hence high computational time being required to solve the inversion scheme. Here, the corresponding forward modeling should be able to do the simulation rapidly, especially for the repeated use of number of iterations. SBL applications generate large volume of noisy data captured by all receivers located on the seafloor. Researchers in Reference [1] stated that the CSEM data-processing has become a challenging task to geophysicists. Therefore, quantification of uncertainty in processing the CSEM data is very important in the SBL. The use of the differential equation (DE) methods alone in the forward modeling is not capable of providing the quantification of the uncertainty.

To address these problems, we propose a stochastic process-based inversion methodology to estimate the electrical resistivity of hydrocarbon layer based on the multifrequency SBL data. Gaussian process (GP) regression is a stochastic and nonparametric method which is employed as the alternative forward modeling method that is able to provide the uncertainty quantification. According to researchers in Reference [15], this elegant method can perform very well in the nonlinear modeling problems. In the inverse modeling, gradient descent (GD) method is utilized for the optimization purposes. We use mean squared error (MSE) as the objective function (i.e., loss function) of the optimization. GD acts as the minimizer finder by obtaining the optimal electrical resistivity of hydrocarbon of the observational dataset. The novelty of this work can be described in three ways: (i) making full use of the stochastic process as the EM forward modeling by utilizing knowledge from computer simulation to reduce the computational effort involved in the modeling, (ii) integrating the stochastic approach and the indirect search algorithm in one inversion methodology to recover the hydrocarbon resistivity property, and (iii) applying the proposed inversion to find the resistivity of the hydrocarbon in the SBL application based on the multifrequency marine CSEM data with uncertainty justification. This work could be very advantageous to the petroleum industry to reduce the computational efforts involved in the ad hoc CSEM modeling. The literature reviews on the available methods used in processing the CSEM data are discussed in Section 2.

2. Available Methods Used in Processing CSEM Data

For decades, controlled-source electromagnetic (CSEM) technology has been widely employed in many land-based geophysical explorations [13]. Rising in oil and gas demand has triggered a commercial interest to apply the CSEM technique in ocean-based subsurface exploration [16]. The subsurface could be very complex due to a lateral variation of the subsurface conductivity beneath the seafloor (i.e., sea-bottom sediments). In the electromagnetic (EM) modeling, Maxwell's equations are solved in a discretized form to compute the propagation of the EM fields. In a 3-D scenario, the subsurface can be discretized using either regular brick, hexahedral, or tetrahedron elements.

According to Reference [13], the electric and magnetic fields within each element can be approximated by either linear or higher order polynomial functions. Researchers in

Reference [17] used 3-D finite element (FE) method to compute the EM responses from the reservoir that consists of a set of resistive hydrocarbons separated by conductors in homogeneous and inhomogeneous environments. The resulting computational experiments showed that the presence of the resistive gaps gave a significant impact on the EM fields recorded by the receivers. Researchers in Reference [13] employed linear edge-based FE method for modeling the 3-D CSEM data in anisotropic conductive medium. Besides that, researchers in Reference [18] introduced a new method for processing CSEM data based on redatuming the observational data from the actual receivers into the virtual receivers. To compute the EM fields in the virtual receivers, Stratton-Chu type integral transformation was used in the research study. Next, 2.5-D FE method was used by researchers in Reference [19] as its forward modeling code for marine CSEM application in Niger Delta. Adaptive FE method has been known as a robust numerical modeling approach in complex environment. It adaptively refines the FE mesh using error estimator of a priori to produce high quality of EM responses for all the refinements though it caused high computational time. Researchers in Reference [20] used FE, namely linear edge FE method, for their computational implementation of the CSEM application. Other than that, parallel edge-based tool for geophysical EM modeling (PETGEM) was used by researchers in Reference [21] for 3-D CSEM forward modeling by employing Nédélec edge finite element method (EFEM). The EFEM can support the unstructured tetrahedral meshes and offers a good compromise between its accuracy and the number of degrees of freedom. To our knowledge, most of the existing numerical EM modeling techniques are mesh-based algorithms. These differential equation (DE) methods involve very large computational domain to satisfy to the corresponding boundary condition in the modeling. Other than the DE methods, some literature also reported the use of integral equation (IE) method for the EM modeling. However, according to researchers in Reference [22], the matrix system involved in the IE method needs significantly high computational effort in terms of time and memory.

3. Background of the Proposed Methodology

The proposed inversion methodology consists of two frameworks, which are Gaussian process (GP) as the forward modeling method, and gradient descent (GD) used for the optimization involved in the inversion scheme.

3.1. Stochastic Process–Gaussian Process

Gaussian process (GP) is a random function which consists of a collection of random variables that are treated as Normal (i.e., gaussian) distribution [23,24]. A model of GP is defined by a vector of its mean function, m(x), and a matrix composed by its covariance function, k(x, x'). The distribution can be expressed as $f(x) \sim G(m(x), k(x, x'))$. GP is a probabilistic and nonparametric method which capable of fitting functional forms reliant on the domain observations. According to Reference [25], this stochastic approach has a property that differs from most of the other black-box identification approaches. It searches for the correlations among the known knowledge rather than approximating the model by trying to fit the basis function parameters. GP can provide the predictive mean (i.e., most likely output) and predictive variance (i.e., confidence interval (CI)) through the Bayesian inference. It places a prior on the function space f without parameterizing the function and then generates posterior distribution over the function. Researchers in Reference [26] stated that a GP model can be utilized as a surrogate model for intricate mathematical model that needs high computational time to solve.

GP has been used in many applications due to its flexibility, such as in the engineering fields, geo-statistics, electronics, machine learning, etc. Researchers in Reference [27] mentioned that the GP is capable of decreasing the complexity of the computer vision with a finer performance of reconstructions. In addition, researchers in Reference [28] compared the performance of an improved GP model with other methods, such as multiple linear regression (MLR) and artificial neural network (ANN), in predicting roadway broken road zone. The results showed that the hybrid model gave the best performance in terms of

coefficient of determination (R^2), root mean squared error (RMSE), and variance accounted for (VAF). In geophysics field, researchers in Reference [29] stated that GP can provide reliable porosity and permeability information to the well analysis. Besides that, GP was also fully utilized by researchers in Reference [30] in the seismic signal detection. The study inferred that the GP was able to produce higher rate of detection and lower rate of false alarm, especially for the weak seismic signals. Next, researchers in Reference [31] proposed an improved methodology for active learning in the GP regression to predict the black-box functions. The proposed method revealed better RMSE with a fewer number of points compared to the other famous methods. Researchers in Reference [32] also enhanced the GP regression for estimating the state-of-health (SOH), and the results showed that the proposed method was capable of reducing the computational complexity with better estimation accuracy.

3.2. Gradient Descent

Gradient descent (GD) is a famous iterative method used for updating parameters of a function. GD updates the model by minimizing its objective function based on a first derivative (i.e., slope) property. Determining the slope can inform the direction of the model parameters (whether to decrease or increase), and it also could tell the step size that should be taken by the model to reach the optimum point. GD is also known as steepest descent method, where it optimizes the objective function by moving in negative direction of the slope iteratively. In the GD method, the step size is called as learning rate, α . A smaller α can learn the function area precisely since the slope is frequently recalculated. But, if the α is not carefully parameterized, the optimization may require high processing time to converge and reach the optimum point. Meanwhile, a bigger α can extend over larger area of the function but overshooting the optimum point may occur. Thus, sensible parameterization of the step size implies the efficiency of the GD optimization.

GD is a well-known method that is commonly used in optimization problems, such as in the applications of machine learning. Researchers in Reference [33] exercised the GD for the purpose of optimizing the vector centers of the consequent layer functions and receptive field matrices in a neuro-fuzzy model. This study relied on the mean squared error (MSE) criterion as its standard of the optimization. Besides that, researchers in Reference [34] employed the GD method to establish nodes operation model for classifying heart conditions based on a deep learning network. Next, researchers in Reference [35] also used GD method to identify the aperture shape in the direct aperture optimization (DAO). Furthermore, this optimization method was also utilized by researchers in Reference [36] to solve the predicting student performance (PSP) by minimizing the error equation, which is its loss function. The study updated the matrices of W_1 and W_2 in the optimization scheme. In addition, researchers in Reference [37] employed GD to find the optimal parameters of impedance control in an iterative learning control algorithm for the force/position hybrid control to control the trajectory of the hexa-robot. The results revealed that the proposed approach can successfully control the robot as expected. Researchers in Reference [38] also used GD to obtain phase-only computer-generated holograms (CGHs). The study concluded that the GD method needed low time to optimize the distribution of the phaseonly CGHs and obtained higher precision images compared to the results obtained when using the Gerchberg-Saxton (GS) algorithm.

4. Methodology

In Section 4, five subsections are discussed, which are, Section 4.1. Developing synthetic seabed logging (SBL) model, Section 4.2. Preprocessing computer simulation outputs, Section 4.3. Developing two-dimensional (2-D) Gaussian process (GP) models, Section 4.4. Electrical resistivity estimation, and Section 4.5. Validation of estimate and GP model.

4.1. Developing Synthetic Seabed Logging Model

A synthetic seabed logging (SBL) application is replicated using the computer simulation technology (CST) software. According to Reference [39], CST Studio Suite[®] exclusively employs finite element (FE) method as the solver of the electromagnetic (EM) simulation in the low-frequency applications. The synthetic model has the same input parameters, except for the conductivity (inversely proportional to the resistivity) of the hydrocarbon layer and the frequency of the EM signal. The SBL model is replicated with three different background layers, such as air, seawater, and sediment, with an isotropic hydrocarbon layer inside, as depicted in Figure 1.





Based on Figure 1, sediment layer located above the target layer (i.e., hydrocarbon) is called as overburden, whereas sediment below the target layer is called as under burden. The thickness of overburden (i.e., depth of hydrocarbon) is fixed at 500 m from the seafloor. To make it simpler, all the input parameters of the simulation are summarized in Table 1. Based on Table 1, five different transmission frequencies (i.e., multifrequency) and five resistivities of isotropic target (e.g., 90, 180, 270, 360, and 450 Ω m) are simulated in the SBL modeling. It means that for every frequency, five different EM profiles are generated and these profiles act as the training CST datasets for the data-processing purposes. Note that the set-up is fixed for the case of static source-receiver separation distances. The source (i.e., transmitter) is placed in the seawater at the center of the model with coordinate of (height = 1270 m, width = 10,000 m, length = 10,000 m), and the inline receivers are located along the seafloor. Hence, all the generated CST datasets are symmetrical from the model center to the right and left boundaries.

Input Parameters	Numerical Values in Unit
Conductivity (Air)	$1.00 imes 10^{-11}{ m Sm^{-1}}$
Conductivity (Seawater)	$1.63 \mathrm{S} \mathrm{m}^{-1}$
Conductivity (Sediment)	$1.00 \mathrm{S} \mathrm{m}^{-1}$
Resistivity (Hydrocarbon)	90 Ωm 180 Ωm 270 Ωm 360 Ωm 450 Ωm
Frequency	0.0625 Hz 0.1250 Hz 0.2500 Hz 0.3750 Hz 0.5000 Hz
Hydrocarbon thickness	200 m
Current strength	1250 A
Length of source	270 m
Source depth (from seawater surface)	970 m

Table 1. Input parameters of the computer simulation.

4.2. Preprocessing Computer Simulation Outputs

As discussed in Reference [40], the SBL data interpretation is based on an assumption that the existence of target bodies can be preidentified by normalizing the magnitude versus offset (MVO) between the target responses (with hydrocarbon) and the reference responses (without hydrocarbon). The normalized MVO can be calculated by dividing the target responses with the reference responses. The existence of the potential target bodies can be detected if the ratio of normalization is greater than 1.0.

The same input parameters are used for replicating the reference model, as indicated in Table 1 and presented in Section 4.1, but the hydrocarbon information is not included. After generating the multifrequency reference model and preprocessing the electromagnetic (EM) profiles by considering the normalized MVO and symmetricity of the model, only 93 datapoints per dataset are used for the GP modeling (discussed in the next subsection), which are taken from source-receiver separation distances of 10,845.77–20,000 m.

4.3. Developing Two-Dimensional Gaussian Process Models

Gaussian process machine learning (GPML) MATLAB codes are used for the Gaussian process (GP) regression. The GPML codes can be found in Reference [41]. As mentioned earlier, the GP regression is employed to evaluate the EM profiles at various unobserved resistivities of hydrocarbon. For every frequency, all five CST datasets, as expressed in Equation (1), are treated as the GP training datasets.

$$D_1, D_2, \dots, D_5 = \{ (x_j, y_j) \}_{i=1}^M,$$
(1)

where M = 93. The *d*-dimensional input vector and its corresponding output vector can be simply written as $x_j \in X$ and $y_j \in Y$, respectively. X indicates *M*-by-*d* matrix (i.e., $x \in \mathbb{R}^{d=2}$), and Y indicates *M*-by-1 column vector (i.e., $y \in \mathbb{R}^{d=1}$). In this work, the input is bivariate, x = (a, b), where *a* stands for the source-receiver separation distance (i.e., offset), and *b* is the electrical resistivity of the hydrocarbon layer, whereas the corresponding output *y* is the magnitude of electric field (E-field). To evaluate the EM profiles at unobserved resistivity inputs conforming to the testing input $x_* = (a_*, b_*)$, the posterior distribution is defined as in Equation (2):

$$(y_*|X,Y,x) \sim G(\mu,\kappa),\tag{2}$$

where μ and κ are the predictive mean and variance, respectively. These two terms are defined in Equations (3) and (4). Since the resistivity of hydrocarbon is the only parameter that is varied, thus $a_* = a$ (i.e., remains unchanged). One hundred and sixteen unobserved target resistivities, which are 93–447 Ω m with an increment of 3 Ω m each (excluding the training input points), are set as the desired information. Therefore, for every frequency, there are entirely 121 EM profiles (including the training datasets) evaluated in the GP model.

$$\mu = K(X, x_*)^T K_{\varepsilon}(X, X)^{-1} Y,$$
(3)

and

$$\kappa = k(x_*, x_*) - K(X, x_*)^T K_{\varepsilon}(X, X)^{-1} K(X, x_*),$$
(4)

where

$$K_{\varepsilon}(X,X) = K(X,X) + \sigma_M^2 I_M.$$
(5)

Based on Equations (3)–(5), K(X, X) is an *M*-by-*M* covariance matrix amongst the X, $\sigma_M^2 I_M$ is an *M*-dimensional noise covariance matrix, $K(X, x_*)$ is an *M*-by-1 covariance matrix between the *X* and x_* , and $k(x_*, x_*)$ is a covariance matrix of x_* . Note that, in the context of this work, all these covariance matrices are composed of the squared exponential (SE) covariance function, where Equation (6) is utilized to compute every element of the matrices.

1

$$k(x,x') = \sigma_f^2 \left[exp\left(\frac{-|x-x'|^2}{2\ell^2}\right) \right],\tag{6}$$

where hyperparameters $\theta = \{\ell, \sigma_f\}$. ℓ and σ_f represent the isotropic length scale and the signal variance, respectively. Note that the SE covariance function is used to govern the GP realizations smoothness which is dictated by the computer simulation outputs. According to Reference [42], GP regression with the SE covariance function is infinitely differentiable. This is an advantageous for any inversion scheme that relies on the derivative basis such as the gradient descent (GD) method.

Here, Equations (2)–(6) need the computation of the hyperparameters θ . In the GP regression, this θ is computed by minimization of the negative log marginal likelihood (NLML) using the gradient-based approach. The equation of NLML is expressed in Equation (7):

$$log[p(Y|X,\theta)] = \frac{1}{2} \Big[Mlog(2\pi) + log|K_{\varepsilon}(X,X)| + Y^T[K_{\varepsilon}(X,X)]^{-1}Y \Big].$$
(7)

Researchers in Reference [29] mentioned that the logarithmic identifier used in Equation (7) is to simplify the expression of integrals involved in the marginal likelihood method. Once the θ is successfully computed, the GP predictive distribution is defined by using Equations (3) and (4). Here, the μ is the best estimate of y_* (i.e., predictive mean) which represents the magnitude of E-field at the x_* , and the κ represents the predictive variance of the x_* in terms of 95% confidence interval (CI) (±two standard deviations). This feature expresses the reliability of the μ .

4.4. Electrical Resistivity Estimation

It must be noted here that the developed GP model is denoted as $Y_{fit}(a, b)$ where all the EM profiles are evaluated at all training and testing input parameters. For every frequency, $Y_{fit}(a, b)$ consists of 121 EM profiles observed at 93 points of offset from 10,845.77 m to 20,000 m with 121 different resistivities of hydrocarbon (90–450 Ω m, with a gap of 3 Ω m). Assume that the observational dataset $Y_{obs}(a, \blacksquare)$ is observed at M = 93 points of offset, and

there is a potential target layer resides 500 m underneath the seafloor with "unknown" resistivity. Then, the least square criterion is formed between the Y_{fit} and the Y_{obs} in order to determine the target electrical resistivity of the Y_{obs} . This work uses mean squared error (MSE) as the objective function of the optimization scheme. Generally, MSE calculates the mean amount of squared differences between the Y_{fit} and Y_{obs} . The objective function $J(b_i)$ is defined in Equation (8).

$$\{J(b_i)\}_{i=1}^{N} = \frac{1}{M} \sum_{j=1}^{M=93} \Big[Y_{fit}(a_j, b_j) - Y_{obs}(a_j, \blacksquare) \Big]^2,$$
(8)

where N = 121. The objective function $J(b_i)$ is a function of b. Thus, Equation (8) is minimized over the b, and the minimizer indicates the estimate of the target resistivity. For every frequency, all the computed $J(b_i)$ are then plotted in order to observe the trend of the function. From the trend, a random initial guess of target resistivity can be specified. Since this is a minimization procedure, the decreasing trend of the function is the key idea of choosing the suitable initial guess. To make the process runs efficiently, a point closes to the minimum zero-gradient point is selected as the initial guess, and the iteration index i is set to zero. Here, the initial minimizer is denoted as b_0 .

After specifying the initial guess, the first-partial derivative of the $J(b_i)$ is defined with respect to b_i , as derived in Equation (9).

$$J(b_{i}) = \frac{1}{M} \sum_{j=1}^{M} \left[Y_{fit} - Y_{obs} \right]^{2},$$

$$J'(b_{i}) = \frac{1}{M} \frac{\partial}{\partial b_{i}} \left[\sum_{j=1}^{M} \left[Y_{fit} - Y_{obs} \right]^{2} \right],$$

$$I'(b_{i}) = \frac{2}{M} \sum_{j=1}^{M} \left[Y_{fit} - Y_{obs} \right] \frac{\partial}{\partial b_{i}} \left[Y_{fit} - Y_{obs} \right],$$

$$J'(b_{i}) = \frac{2}{M} \sum_{j=1}^{M} \left[Y_{fit} - Y_{obs} \right] \frac{\partial}{\partial b_{i}} Y_{fit},$$

$$J'(b_{i}) = \frac{2}{M} \left(\sum_{j=1}^{M} \left[Y_{fit} - Y_{obs} \right] \right) Y'_{fit},$$
(9)

where Y'_{fit} is the first derivative of the GP model with respect to b_i . As mentioned before, the GP model is composed by the predictive mean μ , shown in Equation (3). Thus, the derivative of the μ with respect to the b_i is written in Equation (10).

$$\mu' = K'(X, x_*)^T K_{\varepsilon}(X, X)^{-1} Y,$$
(10)

where $K'(X, x_*)$ is the derivative of the covariance matrix between X and x_* with respect to the x_* (specifically with respect to b_i). In order to find the $K'(X, x_*)$, the derivative of the SE covariance function with respect to x_* is derived. The equation is shown in Equation (11):

$$k(x, x_{*}) = \sigma_{f}^{2} \left[exp\left(\frac{-|x-x_{*}|^{2}}{2\ell^{2}}\right) \right],$$

$$k'(x, x_{*}) = \sigma_{f}^{2} \frac{\partial}{\partial x_{*}} \left[exp\left(\frac{-|x-x_{*}|^{2}}{2\ell^{2}}\right) \right],$$

$$k'(x, x_{*}) = \sigma_{f}^{2} exp\left(\frac{-|x-x_{*}|^{2}}{2\ell^{2}}\right) \left(\frac{2|x-x_{*}|}{2\ell^{2}}\right),$$

$$k'(x, x_{*}) = \sigma_{f}^{2} \left(\frac{|x-x_{*}|}{\ell^{2}}\right) exp\left(-\frac{|x-x_{*}|^{2}}{2\ell^{2}}\right).$$
(11)

The next step is finding the update of the b_i . Based on the GD method, the update can be computed by using Equation (12).

$$b_{i+1} = b_i - \alpha \cdot J'(b_i), \tag{12}$$

where α is the learning rate. To gain more knowledge about the learning rate, in the context of this work, positive values of 10, 1, 0.1, 0.01, 0.001, and 0.0001 are used as the indicator of which range of α can give the significant optimization. For the update, the changes between the subsequent update (i.e., minimizer) and the total number of iterations are set to be at most 0.05% and 100, respectively. Equation (13) is used to calculate the subsequent changes of each updated minimizer:

$$C_{i+1} = \left| \frac{b_{i+1} - b_i}{b_i} \right| \times 100\%.$$
(13)

For every iteration, GP regression is utilized to evaluate the EM profile at every update. The new objective function between the Y_{fit} and Y_{obs} is computed using Equation (14):

$$J(b_{i+1}) = \frac{1}{M} \sum \left[Y_{fit}(a, b_{i+1}) - Y_{obs}(a, \blacksquare) \right]^2.$$
(14)

Besides identifying the subsequent changes of the update and the number of iterations, the new objective function $J(b_{i+1})$ for every iteration also needs to be verified. One criterion is added as the threshold where the objective function must always converge. In other words, $J(b_{i+1})$ must be always smaller than $J(b_i)$. Each iteration must satisfy these stopping criteria to continue the optimization process. The flow is repeated until the stopping criteria (i.e., threshold) are neglected, then the iteration stops, and the optimal hydrocarbon resistivity for the Y_{obs} is concluded. The methodological flow of the proposed inversion is depicted in Figure 2.

4.5. Validation of Estimate and Gaussian Process Model

To determine the reliability of the estimates of hydrocarbon resistivity obtained by the inversion methodology, this work calculates the absolute error between the estimate and the computer simulation input. The equation of the absolute error is shown in Equation (15).

$$Absolute \ Error = |b_{true} - b_{estimate}|, \tag{15}$$

where b_{true} indicates the simulation input, and $b_{estimate}$ is the target resistivity estimated by the stochastic process-based inversion. Furthermore, the Gaussian process (GP) model is also validated by calculating the mean squared deviation (MSD), and root mean squared deviation (RMSD) between the electromagnetic (EM) profile evaluated by the GP model and EM response modeled by the finite element (FE) method through the computer simulation technology (CST) software at the unobserved input points. Note that the EM responses modeled by the FE are assumed as the true EM values. The equations of the MSD and RMSD are formulated in Equations (16) and (17), respectively.

$$MSD = \frac{1}{M} \sum (y_{FE} - y_{GP})^2,$$
 (16)

and

$$RMSD = \sqrt{MSD} = \sqrt{\frac{1}{M}\sum(y_{FE} - y_{GP})^2},$$
(17)

where y_{FE} and y_{GP} represent the EM response modeled by the FE and GP methods, respectively.



Figure 2. Methodological flow of the proposed methodology. In the context of this work, all the simulation models are models with presence of the hydrocarbon. Thus, the step of determining the target layer is not discussed in this work. w.r.t. indicates with respect to.

5. Algorithms of the Proposed Methodology

Algorithm 1 shows the steps involved in the forward Gaussian process (GP) modeling, whereas the process of the inversion scheme for each frequency developed in this work is shown in Algorithm 2.

Algorithm 1 Forward Gaussian Process (GP) Modeling

Input:

Training input: Source-receiver separation distances (i.e., offsets) and hydrocarbon resistivities Training output: Magnitude of electric field (E-field)

Steps:

- 1. Optimize the hyperparameters
- 2. Define the testing inputs (i.e., unobserved hydrocarbon resistivities)
- 3. Compute the covariance matrices
- 4. Evaluate the GP predictive distribution (i.e., predictive mean and variance)

Output:

Two-dimensional (2-D) GP model which consists of various electromagnetic (EM) profiles (i.e., magnitude of electric field) at the unobserved hydrocarbon resistivities

Algorithm 2 Stochastic process-based Inversion

Input:

Two-dimensional (2-D) GP models at various hydrocarbon resistivities (observed and unobserved)

Observational dataset with "unknown' resistivity of isotropic hydrocarbon **Steps:**

- 1. Compute the objective function (i.e., mean squared errors (MSEs)) between the 2-D GP model and the observational dataset
- 2. Plot all the MSEs
- 3. Define the initial guess (i.e., initial minimizer) and its respective MSE
- 4. Set index iteration i = 0
- 5. Compute the derivative of the covariance function when i = 0
- 6. Define the learning rate
- 7. Define the maximum number of iterations (equal to 100) and the maximum percentage change (equal to 0.05%)
- 8. Set index iteration i = 1
- 9. Evaluate the derivative of the GP predictive mean for the previous minimizer
- 10. Compute the derivative of the MSE for the previous minimizer
- 11. Calculate the new minimizer when i = 1
- 12. Define new testing point for the GP regression based on the new minimizer
- 13. Evaluate the new predictive mean (i.e., new GP model) based on the new minimizer
- 14. Compute the new MSE between the new GP model and the observational dataset
- 15. Calculate the subsequent change between the new minimizer and the previous minimizer
- 16. Return to step 8 by setting new index iteration i = i + 1 if the *i* is less than 100, the percentage change is less than 0.05%, and the new MSE is smaller than the previous MSE. If not, the iteration stops

Output:

The optimal electrical resistivity of hydrocarbon of the observational dataset

6. Results and Discussion

It must be noted here that all the training datasets were evaluated using the finite element (FE) method, which was exclusively employed by the computer simulation technology (CST) software to solve the low-frequency problems. All the multifrequency computer simulation datasets are depicted in Figure 3.



Figure 3. Magnitude of electric field (E-field) versus offset for five different hydrocarbon resistivities (90, 180, 270, 360, and 450 Ω m) for five different transmission frequencies: (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.

Based on Figure 3, the *x*- and *y*-axes indicate the source-receiver separation distance and magnitude of E-field in logarithmic scale, respectively. In this work, we observed 93 points of offset, which were from 10,845.77 m to 20,000 m. In addition, we tested multiple transmission frequencies to identify which frequency can give the best estimation of the hydrocarbon electrical resistivity based on the proposed inversion methodology given specific model conditions. To make the figure interpretable, the zoom-in scaled version of Figure 3 at far source-receiver separation distances is provided and depicted in Figure 4.



Figure 4. Cont.



Figure 4. The zoom-in scaled of magnitude of electric field (E-field) versus offset at far source-receiver separation distances for five different transmission frequencies: (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.

From Figure 4, we can see that, when using higher frequency (i.e., 0.5 Hz), the electromagnetic (EM) signals attenuate rapidly along the lower electrical resistivity of hydrocarbon (i.e., 90 Ω m). Lower frequency gave better resolution of magnitude of E-field at far offset distances especially when the hydrocarbon resistivities were higher than 90 Ω m. This is because low-frequency of EM signals less attenuate along the high-resistive medium. We believed that the property of multiple frequencies of EM signal in the seabed logging could give unique clarification to each estimation made by the stochastic process-based inversion methodology.

We utilized all the computer simulation datasets to be used as the training datapoints in the Gaussian process (GP) modeling. For every frequency, we developed a two-dimensional (2-D) GP model which consisted of EM profiles at all training and testing hydrocarbon resistivities. All the contour plots of the 2-D GP model are shown in Figure 5. Based on Figure 5, the *x*- and *y*-axes indicate the 93 points of offset and 121 target resistivities, respectively. The log 10 of magnitude of E-field is represented in the color codes. We developed these multifrequency 2-D GP models to evaluate as many EM profiles as possible by fully utilizing the computer simulation datasets. This information is very useful for the inversion scheme. Note that generating all the multifrequency 121 EM profiles using the CST software could consume high computational time, which is approximately 10 h. Meanwhile, the GP modeling only need around 25 min for the prior data acquisition using the computer simulation and approximately less than 1 min to develop all the 2-D GP models at the unobserved inputs. This proves that, in the context of this study, the forward GP modeling can provide more "untried" information with lower computational time compared to the time taken by the mesh-based modeling.

In order to validate the reliability of each of the 2-D GP models, we calculated the mean squared deviation (MSD) and root mean squared deviation (RMSD) between the GP models and the EM responses modeled by the finite element (FE) at eight random unobserved hydrocarbon resistivities chosen in this research work which were 120 Ω m, 150 Ω m, 210 Ω m, 240 Ω m, 300 Ω m, 330 Ω m, 390 Ω m, and 420 Ω m. The resulting MSDs and RMSDs for frequencies of 0.0625, 0.125, 0.25, 0.375, and 0.5 Hz are tabulated in Tables 2 and 3, respectively.



Figure 5. Two-dimensional (2-D) Gaussian process (GP) model with 121 different electromagnetic (EM) profiles for multiple transmission frequencies; (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.

Table 2. Mean squared deviation (MSD) results.

Target Resistivity	0.0625 Hz	0.1250 Hz	0.2500 Hz	0.3750 Hz	0.5000 Hz
120 Ωm	$3.78 imes10^{-6}$	$8.45 imes 10^{-6}$	$1.94 imes 10^{-5}$	$3.00 imes 10^{-5}$	$3.92 imes 10^{-5}$
150 Ωm	$1.12 imes 10^{-6}$	$2.45 imes10^{-6}$	$5.69 imes10^{-6}$	$8.61 imes10^{-6}$	$1.14 imes10^{-5}$
210 Ωm	$2.55 imes10^{-7}$	$5.48 imes10^{-7}$	$1.32 imes 10^{-6}$	$1.96 imes10^{-6}$	$2.69 imes10^{-6}$
240 Ωm	$2.05 imes10^{-7}$	$4.28 imes10^{-7}$	$1.03 imes10^{-6}$	$1.43 imes10^{-6}$	$2.02 imes 10^{-6}$
300 Ωm	$1.47 imes10^{-7}$	$3.06 imes10^{-7}$	$7.42 imes10^{-7}$	$1.05 imes 10^{-6}$	$1.48 imes 10^{-6}$
330 Ωm	$1.64 imes10^{-7}$	$3.40 imes10^{-7}$	$8.17 imes10^{-7}$	$1.12 imes 10^{-6}$	$1.59 imes10^{-6}$
390 Ωm	$2.32 imes10^{-7}$	$4.87 imes10^{-7}$	$1.13 imes10^{-6}$	$1.70 imes10^{-6}$	$2.31 imes10^{-6}$
420 Ωm	$3.80 imes 10^{-7}$	$8.34 imes 10^{-7}$	$1.94 imes 10^{-6}$	$3.06 imes 10^{-6}$	$4.11 imes 10^{-6}$

Based on Tables 2 and 3, we can see that all the MSD and RMSD values for every frequency are very small. It shows that the EM profiles evaluated by the GP regression are significantly close to its true values (i.e., EM responses modeled by the FE method). This implies that the GP regression is capable of evaluating multifrequency EM profiles at various unobserved input parameters accurately with low time consumption.

Target Resistivity	0.0625 Hz	0.1250 Hz	0.2500 Hz	0.3750 Hz	0.5000 Hz
120 Ωm	$1.94 imes 10^{-3}$	$2.91 imes 10^{-3}$	$4.41 imes 10^{-3}$	$5.48 imes 10^{-3}$	$6.26 imes 10^{-3}$
150 Ωm	$1.06 imes10^{-3}$	$1.57 imes10^{-3}$	$2.38 imes10^{-3}$	$2.93 imes10^{-3}$	$3.38 imes10^{-3}$
210 Ωm	$5.05 imes10^{-4}$	$7.40 imes10^{-4}$	$1.15 imes 10^{-3}$	$1.40 imes 10^{-3}$	$1.64 imes 10^{-3}$
240 Ωm	$4.53 imes10^{-4}$	$6.54 imes10^{-4}$	$1.02 imes 10^{-3}$	$1.20 imes 10^{-3}$	$1.42 imes 10^{-3}$
300 Ωm	$3.84 imes10^{-4}$	$5.53 imes10^{-4}$	$8.62 imes10^{-4}$	$1.02 imes 10^{-3}$	$1.22 imes 10^{-3}$
330 Ωm	$4.05 imes10^{-4}$	$5.83 imes10^{-4}$	$9.04 imes10^{-4}$	$1.06 imes 10^{-3}$	$1.26 imes 10^{-3}$
390 Ωm	$4.82 imes 10^{-4}$	$6.98 imes10^{-4}$	$1.06 imes 10^{-3}$	$1.30 imes10^{-3}$	$1.52 imes 10^{-3}$
420 Ωm	$6.17 imes10^{-4}$	$9.13 imes10^{-4}$	$1.39 imes 10^{-3}$	$1.75 imes 10^{-3}$	$2.03 imes 10^{-3}$

Table 3. Root mean squared deviation (RMSD) results.

In the context of this work, we generated three different multifrequency observation datasets separately using the CST software which the resistivities of the hydrocarbon were assumed to be "unknown", i.e., 100 Ω m, 200 Ω m, and 400 Ω m. To avoid biased, we purposely selected these "unknown" hydrocarbon resistivities since there was no information of the resistivities in the developed forward GP model. As mentioned before, the goal of this work was to find the best estimate of hydrocarbon resistivity of the observational dataset by minimizing the mean squared error (MSE). A small MSE indicates that the GP model gives the most likely EM responses for the observational dataset, thereby presenting the best information of the hydrocarbon resistivity. Here, we provide plots of MSE versus hydrocarbon resistivity for every frequency in Figure 6 (for 100 Ω m), Figure 7 (200 Ω m), and Figure 8 (400 Ω m).

By referring to Figures 6–8, the *x*- and *y*-axes indicate the hydrocarbon resistivities involved in the 2-D GP model and the computed MSE values, respectively. The blue points are the 121 MSEs at the respective hydrocarbon resistivities, whereas the red points represent the optimized MSEs with respect to the updated minimizers. We also monitored the status of the convergence for every optimization. Thus, Figures 9–11 show the plots of MSE versus number of iterations for every observational dataset and transmission frequency.

In Figures 9–11, the *x*- and *y*-axes indicate the number of iterations and the computed MSEs, respectively. Based on these figures, we can see that all the optimizations converged to the smallest MSE with less than 100 iterations (maximum iterations number). Here, the most important thing is all the optimizations converged and stopped at the smallest MSE values and approaching to zero. It agrees the main goal of optimization scheme which is to minimize the loss function. Next, in order to make all the figures obtained easier to interpret, we tabulate all the resulting estimates, the validation results (i.e., absolute error), and the learning rate used for every observational data and frequency in table below.



Figure 6. Cont.



Figure 6. Mean squared error (MSE) versus hydrocarbon resistivity for multiple transmission frequencies (Observational dataset 1 with "unknown" resistivity of 100 Ωm): (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.



Figure 7. Cont.



Figure 7. Mean squared error (MSE) versus hydrocarbon resistivity for multiple transmission frequencies (Observational dataset 2 with "unknown" resistivity of 200 Ωm): (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.



Figure 8. Mean squared error (MSE) versus hydrocarbon resistivity for multiple transmission frequencies (Observational dataset 3 with "unknown" resistivity of 400 Ωm): (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.



Figure 9. Mean squared error (MSE) versus number of iterations for multiple transmission frequencies (Observational dataset 1 with "unknown" resistivity of 100 Ω m): (a) 0.0625 Hz, (b) 0.125 Hz, (c) 0.25 Hz, (d) 0.375 Hz, and (e) 0.5 Hz.

Based on Table 4, all the number of iterations were less than 100 (i.e., the maximum number of iterations). Different learning rate was used for every frequency of the observational dataset. From the table, we can see all the computed MSE were very small. Moreover, for the frequency of 0.0625 Hz, only 72.289 s, 35.775 s, and 33.923 s were taken by the stochastic process-based inversion to estimate 104.03 Ω m, 201.45 Ω m, and 402.22 Ω m, respectively, whereas, for the frequency of 0.125 Hz, the inversion took 50.482 s, 44.682 s, and 46.848 s to estimate 102.72 Ω m, 198.08 Ω m, and 404.34 Ω m, respectively. For the frequency of 0.25 Hz, 63.797 s, 55.898 s, and 51.133 s were needed to estimate 102.17 Ω m, 197.51 Ω m, and 410.64 Ω m, respectively. Next, for the frequency of 0.375 Hz, the time taken by the proposed methodology to yield 102.01 Ω m, 197.93 Ω m, and 383.43 Ω m were 57.818 s, 42.565 s, and 49.087 s, respectively. The last one, for the frequency of 0.5 Hz, only 68.353 s, 47.105 s and 65.238 s was needed by the inversion methodology to produce estimates of 101.91 Ω m, 198.20 Ω m, and 386.18 Ω m, respectively. On average, the GP-inversion only took 52.333 s to obtain the optimal hydrocarbon resistivity. This indicates that the proposed inversion methodology could solve the forward and inverse modeling rapidly. In addition, to observe the trend of the hydrocarbon resistivity estimations based on the multifrequency property, we plotted the calculated absolute errors versus the true resistivity for every tested transmission frequency. The plot is depicted in Figure 12.



Figure 10. Mean squared error (MSE) versus number of iterations for multiple transmission frequencies (Observational dataset 2 with "unknown" resistivity of 200 Ω m): (a) 0.0625 Hz, (b) 0.125 Hz, (c) 0.25 Hz, (d) 0.375 Hz, and (e) 0.5 Hz.





Figure 11. Mean squared error (MSE) versus number of iterations for multiple transmission frequencies (Observational dataset 3 with "unknown" resistivity of 400 Ω m): (**a**) 0.0625 Hz, (**b**) 0.125 Hz, (**c**) 0.25 Hz, (**d**) 0.375 Hz, and (**e**) 0.5 Hz.

Dataset	Iterations Number	True Resistivity (Ωm)	Estimate (Ωm)	MSE	Absolute Error	
Frequency: 0.0625 Hz; Learning rate: 1.50						
1	83	100	104.03	$6.10 imes10^{-7}$	4.03	
2	38	200	201.45	$1.30 imes10^{-7}$	1.45	
3	35	400	402.22	$1.70 imes10^{-7}$	2.22	
	Frequency: 0.1250 Hz; Learning rate: 1.00					
1	58	100	102.72	$2.70 imes10^{-7}$	2.72	
2	49	200	198.08	$3.80 imes10^{-7}$	1.92	
3	54	400	404.34	$3.50 imes10^{-7}$	4.34	
	Frequency: 0.2500 Hz; Learning rate: 0.50					
1	76	100	102.17	$4.00 imes10^{-7}$	2.17	
2	63	200	197.51	$4.50 imes10^{-7}$	2.49	
3	60	400	410.64	$1.22 imes 10^{-6}$	10.64	
Frequency: 0.3750 Hz; Learning rate: 0.50						
1	67	100	102.01	$5.30 imes10^{-7}$	2.01	
2	48	200	197.93	$5.50 imes10^{-7}$	2.07	
3	55	400	383.43	$1.45 imes10^{-6}$	16.57	
Frequency: 0.5000 Hz; Learning rate: 0.20						
1	83	100	101.91	$8.10 imes10^{-7}$	1.91	
2	54	200	198.20	$8.30 imes10^{-7}$	1.80	
3	76	400	386.18	$1.17 imes10^{-7}$	13.82	

 Table 4. The main stochastic process-based inversion results.

From Figure 12, the *x*- and *y*-axes indicate the true resistivity and absolute error, respectively. Based on this figure, we can infer that the frequency of 0.5 Hz gave better hydrocarbon resistivity estimation for the observational dataset 1 (true resistivity = $100 \Omega m$) with the absolute error of 1.91 compared to the other absolute errors when using 0.0625 Hz, 0.125 Hz, 0.25 Hz, and 0.5 Hz. But, when the true resistivity increased, the frequency of 0.0625 Hz gave smaller absolute errors, which were 1.45 for the observational dataset 2 (true resistivity = $200 \Omega m$) and 2.22 for the dataset 3 (true resistivity = $400 \Omega m$) compared

to the other absolute errors. This is because, in the seabed logging (SBL) applications, lower transmission frequency is used to get far propagation of the EM signals. This feature is normally employed when the location of the target bodies is deep from the seafloor. Besides that, lower frequency also attenuates slowly along the high-resistive bodies, such as hydrocarbon. Thus, far propagation of the EM signals beneath the seabed can be achieved when the embedded bodies have high electrical resistivity. Since the depth of hydrocarbon layer set in this work was fixed at 500 m from the seafloor, the frequency of 0.5 Hz gave the best resistivity estimation when the resistivity of the hydrocarbon layer was 100 Ω m. Meanwhile, when the hydrocarbon resistivity is 200 Ω m and above, the frequency of 0.0625 Hz gave the best estimation of the hydrocarbon resistivity.



Figure 12. Absolute error versus true resistivities of 100 Ω m, 200 Ω m, and 400 Ω m for multifrequency property.

Based on those results, its uncertainty quantifications produced by the GP modeling also were investigated. In the context of this work, the uncertainty quantifications for estimates of 101.91 Ω m (at frequency of 0.5 Hz), 201.45 Ω m (at frequency of 0.0625 Hz), and 402.22 Ω m (at frequency of 0.0625 Hz) were examined. EM profiles for these three estimates were evaluated by the GP regression separately, and the predictive variances for each profile were computed in order to determine whether the observational datasets lie within the 95% confidence interval (CI) of the respective GP model or not. If the observational datasets (with 100, 200, and 400 Ω m) fail to fit in the CI of the respective forward GP model, it means that the proposed methodology is unable to provide reliable resistivity information. For the estimate of 101.91 Ω m, only 1.08% of the EM responses over 93 datapoints per EM profile lied outside the 95% CI. Meanwhile, for the other two estimates, only 4.30% of the EM responses lied outside its respective CI for both estimates. By following the testing of significance for hypothesis testing, all the percentages were less than the 5% level of significance. Thus, it can be inferred that the observational datasets significantly lied within the 95% CI of the GP models where its hydrocarbon resistivity inputs were estimated by the stochastic process-based inversion.

7. Conclusions

We proposed stochastic process-based inversion to estimate the electrical resistivity of isotropic hydrocarbon layer based on multifrequency synthetic seabed logging (SBL) model. Five different transmission frequencies were tested which were 0.0625, 0.125, 0.25, 0.375, and 0.5 Hz to determine the effect of the multifrequency property of the SBL application in estimating the hydrocarbon resistivity by using the proposed inversion methodology.

We employed the stochastic process, namely Gaussian process (GP) regression, as the alternative forward modeling method to evaluate the electromagnetic (EM) profiles at various unobserved resistivities with uncertainty quantification by utilizing the prior knowledge generated through the computer simulation technology (CST) software. Based on the calculated deviations (presented in Tables 2 and 3), we proved that the GP was well-fitted to evaluate the EM profiles even at the unobserved resistivity inputs. In addition to that, GP also took significantly lower computational time (approximately 26 min including the prior data acquisition) to evaluate the 121 EM profiles compared to the time that could be taken by the software (approximately 10 h).

Next, for the inversion scheme, we proved that the proposed methodology can efficiently estimate the electrical resistivity of the hydrocarbon layer with low time consumption, with the average time of 52.333 s. The objective function also converged, and the optimization stopped at the smallest mean squared error (MSE). The computed absolute errors also agreed the nature or behavior of multifrequency EM data. For the case of the observational dataset 1, the stochastic process-based inversion was significantly capable of estimating the hydrocarbon resistivity when the transmission frequency used was 0.05 Hz, while, for the case of the observational datasets 2 and 3, frequency of 0.0625 Hz gave better hydrocarbon resistivity estimation. This is because lower frequency of EM signal attenuates slower along the medium with higher resistivity. The choice of the transmission frequency practiced in the SBL applications can be predetermined based on the prior knowledge of the hydrocarbon resistivity. In addition, we also fully utilized the uncertainty quantification feature (i.e., predictive variance) provided by the GP. By following the testing of significance for hypothesis testing, all the percentages of the observational datapoints outside the 95% confidence interval (CI) of the respective GP model were less than 5% level of significance. This showed the observational datasets significantly lied within the uncertainty quantification computed by the GP where the resistivity inputs were estimated by the inversion methodology. Therefore, the stochastic process-based inversion methodology can be fully utilized to effectively estimate the hydrocarbon resistivity in the SBL applications with low computational efforts. This attempt is very useful for the EM data-processing, especially to recover the electrical resistivity distribution of the hydrocarbon reservoir. Consequently, this attempt could de-risk the offshore hydrocarbon exploration.

Author Contributions: Conceptualization, H.D. and S.C.D.; methodology, S.C.D. and M.N.M.A.; software, H.D.; validation, S.C.D., K.A.M.N., and M.N.M.A.; writing—original draft preparation, M.N.M.A.; writing—review and editing, H.D., K.A.M.N., and M.N.M.A.; supervision, H.D., K.A.M.N., and S.C.D.; funding acquisition, H.D. and S.C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Yayasan Universiti Teknologi PETRONAS–Fundamental Research Grant (YUTP–FRG) with cost center of 015LC0-055.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented are openly available in Mendeley Data at doi:10.17632/kvtxc2w4sj.1. The data can be found here http://dx.doi.org/10.17632/kvtxc2w4sj.1.

Acknowledgments: The authors are grateful to Universiti Teknologi PETRONAS (UTP) for the facilities provided.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Ellingsrud, S.; Eidesmo, T.; Sinha, M.C.; MacGregor, L.M.; Constable, S.C. Remote sensing of hydrocarbon layers by sea bed logging (SBL): Results from a cruise offshore Angola. *Lead. Edge* **2002**, *20*, 972–982. [CrossRef]
- Eidesmo, T.; Ellingsrud, S.; MacGregor, L.M.; Constable, S.; Sinha, M.C.; Johansen, S.; Kong, F.N.; Westerdahl, H. Sea bed logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deep water areas. *First Break* 2002, 20, 144–152.

- Edwards, N. Marine controlled-source electromagnetics: Principles, methodologies, future commercial applications. *Surv. Geophys.* 2005, 26, 675–700. [CrossRef]
- 4. Hesthammer, J.; Boulaenko, M. The offshore EM challenges. First Break 2005, 23, 59–66. [CrossRef]
- 5. Carazzone, J.J.; Burtz, O.M.; Green, K.E.; Pavlov, D.A.; Xia, C. Three dimensional imaging of marine controlled source EM data. SEG Tech. Program Expand. Abstr. 2005, 24, 575.
- 6. Srnka, L.J.; Carazzone, J.J.; Ephron, M.S.; Eriksen, E.A. Remote reservoir resistivity mapping. *Lead. Edge* **2006**, *25*, 972–975. [CrossRef]
- Constable, S.; Srnka, L.J. An introduction to marine controlled-source electromagnetic methods for hydrocarbon exploration. *Geophysics* 2007, 72, WA3–WA12. [CrossRef]
- 8. Um, E.S.; Alumbaugh, D.L. On the physics of the marine controlled-source electromagnetic method. *Geophysics* 2007, 72, WA13–WA26. [CrossRef]
- Andréis, D.; MacGregor, L.M. Controlled-source electromagnetic sounding in shallow water: Principles and applications. *Geophysics* 2008, 73, F21–F32. [CrossRef]
- 10. Zhdanov, M.S. Electromagnetic geophysics: Notes from the past and the road ahead. Geophysics 2010, 75, A49–A66. [CrossRef]
- 11. Key, K. 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics* 2009, 74, F9–F20. [CrossRef]
- 12. Mittet, R.; Schaug-Pettersen, T. Shaping optimal transmitter waveforms for marine CSEM surveys. *Geophysics* 2008, 73, F97–F104. [CrossRef]
- Cai, H.; Xiong, B.; Han, M.; Zhdanov, M. 3D controlled-source electromagnetic modeling in anisotropic medium using edgebased-finite element method. *Comput. Geosci.* 2014, 73, 164–176. [CrossRef]
- 14. Bakr, S.A.; Pardo, D.; Mannseth, T. Domain decomposition Fourier finite element method for the simulation of 3D marine CSEM measurements. *J. Comput. Phys.* 2013, 255, 456–470. [CrossRef]
- 15. Chan, L.L.T.; Liu, Y.; Chen, J. Nonlinear system identification with selective recursive Gaussian process models. *Ind. Eng. Chem. Res.* **2013**, *52*, 18276–18286. [CrossRef]
- 16. Li, Y.; Key, K. 2D marine controlled-source electromagnetic modeling: Part 1—An adaptive finite-element algorithm. *Geophysics* **2007**, 72, WA51–WA62. [CrossRef]
- 17. 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. *J. Appl. Geophys.* **2015**, *119*, 147–155. [CrossRef]
- 18. Zhdanov, M.; Cai, H. Redatuming controlled-source electromagnetic data using Stratton-Chu type integral transformation. *J. Appl. Geophys.* **2016**, *126*, 1–12. [CrossRef]
- 19. Folorunso, A.F.; Li, Y. Seafloor bathymetry in deep and shallow water marine CSEM responses of Nigerian Niger Delta oil field: Effects and corrections. *J. Appl. Geophys.* **2015**, *123*, 194–210. [CrossRef]
- Castillo-Reyes, O.; Puente, J.; Modesto, D.; Puzyrev, V.; Cela, J.M. A parallel tool for numerical approximation of 3D electromagnetic surveys in geophysics. *Comput. Sist.* 2016, 20, 29–39. [CrossRef]
- 21. Castillo-Reyes, O.; Puente, J.; Cela, J.M. PETGEM: A parallel code for 3D CSEM forward modeling using edge finite elements. *Comput. Geosci.* 2018, 119, 123–136. [CrossRef]
- 22. Yoon, D.; Zhdanov, M.S.; Cai, H.; Gribenko, A. A hybrid finite difference and integral equation method for modeling and inversion of marine CSEM data. *SEG Tech. Program Expand. Abstr.* **2015**, 823–827. [CrossRef]
- 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.* **2018**, *173*, 143–150. [CrossRef]
- 24. Kong, D.; Chen, Y.; Li, N. Gaussian process regression for tool wear prediction. *Mech. Syst. Signal Process.* **2018**, 104, 556–574. [CrossRef]
- Mohd Aris, M.N.; Daud, H.; Dass, S.C.; Mohd Noh, K.A. Gaussian process methodology for multi-frequency marine controlledsource electromagnetic profile estimation in isotropic medium. *Processes* 2019, 7, 661. [CrossRef]
- 26. Harari, O.; Steinberg, D.M. Optimal designs for Gaussian process models via spectral decomposition. *J. Stat. Plan. Inference* **2014**, 154, 87–101. [CrossRef]
- Wang, H.; Gao, X.; Zhang, K.; Li, J. Fast single image super-resolution using sparse Gaussian process regression. *Signal Process*. 2017, 134, 52–62. [CrossRef]
- Yu, Z.; Shi, X.; Zhou, J.; Huang, R.; Gou, Y. Advanced prediction of roadway broken rock zone based on a novel hybrid soft computing model using Gaussian process and particle swarm optimization. *Appl. Sci.* 2020, 10, 6031. [CrossRef]
- 29. 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* **2018**, *11*, 3261. [CrossRef]
- 30. Li, J.; He, M.; Cui, G.; Wang, X.; Wang, W.; Wang, J.A. A novel method of seismic signal detection using waveform features. *Appl. Sci.* **2020**, *10*, 2919. [CrossRef]
- 31. Meka, R.; Alaeddin, A.; Oyama, S.; Langer, K. An active learning methodology for efficient estimation of expensive noisy black-box functions using Gaussian process regression. *IEEE Access* 2020, *8*, 111460–111474. [CrossRef]
- Xiang, M.; He, Y.; Zhang, H.; Zhang, C.; Wang, L.; Wang, C.; Sui, C. State-of health prognostic for lithium-ion batteries considering the limitations in measurements via maximal information entropy and collective sparse variational Gaussian process. *IEEE Access* 2020, *8*, 188199–188217. [CrossRef]

- Vlasenko, A.; Vlasenko, N.; Vynokurova, O.; Bodyanskiy, Y.; Peleshko, D. A novel ensemble neuro-fuzzy model for financial time series forecasting. *Data* 2019, 4, 126. [CrossRef]
- Yoo, H.; Han, S.; Chung, K. A frequency pattern mining model based on deep neural network for real-time classification of heart conditions. *Healthcare* 2020, *8*, 234. [CrossRef] [PubMed]
- 35. Zhang, L.; Zhang, P.; Yang, J.; Li, J.; Gui, Z. Aperture shape generation based on gradient descent with momentum. *IEEE Access* **2019**, *7*, 157623–157632. [CrossRef]
- Gomez-Pulido, J.A.; Duran-Dominguez, A.; Pajuelo-Holguera, F. Optimizing latent factors and collaborative filtering for students' performance prediction. *Appl. Sci.* 2020, 10, 5601. [CrossRef]
- 37. Ba-Phuc, H.; Cheng-Wei, W.; Yong-Lin, K. Force/position hybrid control for a hexa robot using gradient descent iterative learning control algorithm. *IEEE Access* 2019, *7*, 72329–72342.
- 38. Liu, S.; Takaki, Y. Optimization of phase-only computer-generated holograms based on the gradient descent method. *Appl. Sci.* **2020**, *10*, 4283. [CrossRef]
- Electromagnetic Simulation Solvers (CST Studio Suite). Available online: https://www.3ds.com/products-services/simulia/ products/cst-studio-suite/solvers/ (accessed on 2 November 2020).
- 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* 2020, *8*, 605. [CrossRef]
- 41. Documentation for GPML Matlab Code Version 4.2. Available online: http://www.gaussianprocess.org/gpml/code/matlab/ doc/ (accessed on 9 September 2020).
- 42. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006. Available online: http://196.189.45.87/bitstream/123456789/51320/1/17.pdf (accessed on 3 July 2020).