Article

A Novel Methodology for Hydrocarbon Depth Prediction in Seabed Logging: Gaussian Process-Based Inverse Modeling of Electromagnetic Data

Hanita Daud [1], Muhammad Naeim Mohd Aris [1,*] , Khairul Arifin Mohd Noh [2,*] and Sarat Chandra Dass [3]

[1] Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS,
    Seri Iskandar 32610, Perak, Malaysia; hanita_daud@utp.edu.my
[2] Department of Geosciences, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak, Malaysia
[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 (M.N.M.A.); khairula.nmoh@utp.edu.my (K.A.M.N.)

Abstract: Seabed logging (SBL) is an application of electromagnetic (EM) waves for detecting potential marine hydrocarbon-saturated reservoirs reliant on a source–receiver system. One of the concerns in modeling and inversion of the EM data is associated with the need for realistic representation of complex geo-electrical models. Concurrently, the corresponding algorithms of forward modeling should be robustly efficient with low computational effort for repeated use of the inversion. This work proposes a new inversion methodology which consists of two frameworks, namely Gaussian process (GP), which allows a greater flexibility in modeling a variety of EM responses, and gradient descent (GD) for finding the best minimizer (i.e., hydrocarbon depth). Computer simulation technology (CST), which uses finite element (FE), was exploited to generate prior EM responses for the GP to evaluate EM profiles at "untried" depths. Then, GD was used to minimize the mean squared error (MSE) where GP acts as its forward model. Acquiring EM responses using mesh-based algorithms is a time-consuming task. Thus, this work compared the time taken by the CST and GP in evaluating the EM profiles. For the accuracy and performance, the GP model was compared with EM responses modeled by the FE, and percentage error between the estimate and "untried" computer input was calculated. The results indicate that GP-based inverse modeling can efficiently predict the hydrocarbon depth in the SBL.

Keywords: seabed logging; electromagnetic data; hydrocarbon depth; inverse modeling; Gaussian process; gradient descent; computer simulation technology

1. Introduction

Seabed logging (SBL) is an application of the marine controlled-source electromagnetic (CSEM) surveying technique for discovering and characterizing high-resistive bodies (e.g., hydrocarbon-saturated reservoirs) remotely in the deep marine environment based on contrasts of electrical resistivity of the subsurface. Its capability of identifying the fluid content inside the potential reservoirs is undoubted. This emerging application uses a powerful electric dipole transmitter to continuously emit low-frequency electromagnetic (EM) waves, normally between 0.1 and 10 Hz [1], in all directions through the seawater column and subsurface beneath the seabed. The SBL exploits the fact that hydrocarbon-saturated reservoirs have higher electrical resistivity (approximately 30–500 Ωm) than its surroundings, such as seawater (approximately 0.5–2.0 Ωm) and sedimentary rocks (approximately 1.0–2.0 Ωm) [2]. The conductivities (i.e., inversely proportional to resistivities) of hydrocarbon-filled reservoirs can reach about 0.01–0.1 S/m or lower in certain depths underneath the seabed. The main component of an SBL is the utilization of a towed EM source known as a horizontal electric dipole (HED) and stationary EM receivers for

recording the returned EM signals [1]. Studies related to the commercial SBL application can be found in [1–9]. Figure 1 is adopted from [10] to depict the illustration of the SBL application in the offshore environment.
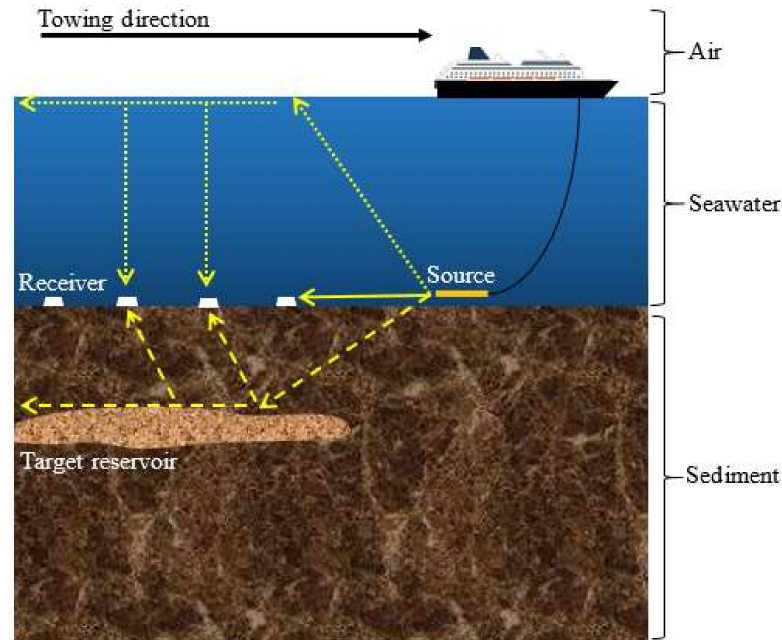

Figure 1. Illustration of the seabed logging (SBL) application in the deep marine environment. The source is towed by a ship and the electromagnetic (EM) receivers are placed on the seafloor beforehand. The target reservoir is in the sedimentary rock. The low-frequency EM waves travel in all directions (straight arrow: direct wave; dotted arrow: air wave; dashed arrow: guided wave).

Numerical modeling is a crucial task in geophysics problems since it provides prior knowledge and information of the potential reservoirs. Many powerful computational modeling techniques have been exploited for processing the EM responses, such as finite element (FE), finite difference (FD), finite volume (FV), integral equations (IE), etc. These numerical techniques are capable of modeling the subsurface realistically in the marine environment by defining grids or meshes. According to [11], the FE technique is highly preferred due to its ability to support unstructured meshes. These meshes are able to provide a quality geological representation by placing small grid elements in the areas where better resolution is highly needed, such as in large conductivity contrast areas [11]. However, as the consequence, the mesh-based algorithms require a very lengthy computational time to solve the linear system; in addition, these forward modeling techniques involve very complicated deterministic mathematical equations. According to [12,13], the most time-consuming task in the algorithms is solving the integral and multiple linear equations, not to mention for the inversion scheme that requires multiple EM forward solutions [13]. Here, the corresponding forward modeling should be robust and capable of operating the EM simulation rapidly, especially for repeated use in the large number of iterations involved in the inversion scheme as well as for multiple source–receiver positions in the SBL application [14].

To address these problems, this work aimed to develop a novel Gaussian process-based inversion methodology to provide meaningful information where the hydrocarbon potentially resides underneath the seabed. For the optimization procedure, this inverse modeling utilized the gradient descent (GD) method to find the best estimate (i.e., hydrocarbon depth) based on the forward Gaussian process (GP) regression model. As mentioned before, GP acts as the alternative forward modeling technique as it is capable of being a surrogate model for the complicated mathematical model that requires high computational time to solve [15]. Researchers in [16] also stated that GP is an elegant and simple approach,

where it performs well in modeling nonlinear problems. Mean squared error (MSE) was used as the objective function (i.e., loss function) in the optimization. The main focus of this work was to find the depth of hydrocarbon of observational EM responses, whose hydrocarbon layer was assumed to be present but with an unknown depth. This work is novel in two ways. First, it presents the integration of a stochastic process (the GP regression model acts as the EM forward model) and an indirect search algorithm (the iterative GD method as the acquirer of minimizer) in one inversion methodological flow. Second, applying this proposed inverse modeling for predicting the input parameter of the potential high-resistive bodies (e.g., hydrocarbon depth) is a new procedure or approach in SBL data processing. This attempt could be very beneficial to de-risk hydrocarbon exploration in the offshore environment. Details of the GP and GD approaches are discussed in the next section.

2. Mathematical Background

Gaussian process (GP) and gradient descent (GD) algorithms are commonly exploited in machine learning (ML) problems. In order to seek the most favorable fairness between the performance of the modeling and the need of low computational time, considering the strengths and integrating these two mathematical approaches could be very advantageous to any applications that struggle with the same concern. Here, Figure 2 shows the common ML process to predict outputs given a set of inputs, its corresponding outputs, and desired/query inputs. It represents the general idea on which this work is focusing. The proposed inversion methodology is believed capable of providing meaningful and useful information to the SBL application with an acceptable accuracy and low computational efforts. In this section, the background and examples from the literature of GP and GD approaches are thoroughly discussed in Sections 2.1 and 2.2, respectively.
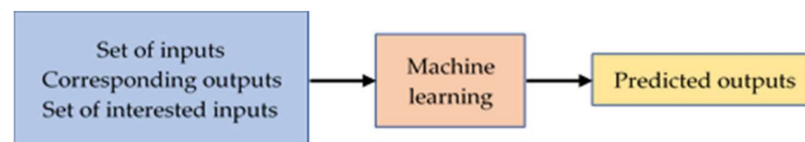

Figure 2. Machine learning (ML) process. ML utilizes sets of inputs, the corresponding outputs, and interested inputs to yield the predictive outputs.

2.1. Gaussian Process

Mathematically, the Gaussian process (GP) can be defined as an infinite collection of random variables where any of the finite subsets can be treated as Gaussian (i.e., normal) distribution [17,18]. A GP model is fully expressed or specified by a mean function $m(x)$ in a vector and a covariance $k(x, x')$ in the form of a matrix. GP is a probabilistic and nonparametric method for fitting any functional forms based on observations of the domain. This method differs from most of the other black-box identification approaches. GP does not try to approximate the model by trying to fit the parameters of the chosen basis function, but it searches the relationship among the given knowledge or information. This stochastic approach can provide predictive mean and variance. The mean represents the most likely outputs, whereas the variance plays a role as the uncertainty's quantifier. The magnitude of variance is incorporated into the GP prediction and inference. It does reflect the amount and quality of the data distribution, and it is an essential numerical feature when it comes to differentiating the GP regression from any other computational intelligence methods.

GP regression is a well-known technique and has been adopted in many scientific and engineering fields, such as machine learning, computer experiments, geostatistics, electronics, etc. Its convenient properties in modeling make GP successful in demonstrating outstanding performances in numerous applications. Researchers in [19] employed GP regression for predicting the concentration of ozone in an area of Bourgas, Bulgaria. The prediction was based on the measurements of the concentration of ozone, phenol, sulphur

dioxide, and benzene in the air (hourly). Researchers in [20] studied the capability of GP in estimating load-bearing pile capacity. This research work compared the performance of GP prediction with other approaches, and GP worked very well in predicting the pile capacity. In a wireless networking application, researchers in [21] inferred that, besides resulting in low-cost outcomes, GP was also capable of obtaining low complex proximity reports for a received-signal-strength (RSS) threshold. Next, GP was also used by researchers in [22] to explain the uncertainty in estimating the state of health (SOH) of lithium-ion batteries. Additionally, according to researchers in [23], GP can successfully decrease the complexity of the computer vision with a better performance of reconstructions.

In [24], GP was hybridized with particle swarm optimization (PSO) to do advanced prediction of roadway broken road zone based on data collected in China. The study inferred that the hybrid model showed the best performance with higher variance accounted for (VAF), higher coefficient of determination (R^2), and lower root mean squared error (RMSE) compared to the other methods such as multiple linear regression (MLR) and artificial neural network (ANN). For geophysics problems, researchers in [25] mentioned that GP was able to yield reliable information of permeability and porosity in well analysis. GP was also used by researchers in [26] to predict the presence of a thin hydrocarbon layer by utilizing its predictive variance for quantifying the uncertainties of EM responses. Furthermore, GP was exploited for seismic signal detection as well. According to [27], the rate of detection when using GP was higher compared to other methods and the rate of false alarm was lower especially for the weak seismic signals. Researchers in [28] proved that GP was successfully capable of being exploited in processing EM responses by calibrating the stochastic and the computer experiment models for magnitude versus offset (MVO) analysis.

2.2. Gradient Descent

Gradient descent (GD) is a popular iterative process used to update the model parameters of a function by minimizing the objective function (i.e., loss function) based on a first-derivative (i.e., slope) basis. The derivative of a function is fully employed to determine whether to increase or decrease the model parameters (i.e., direction to move). Moreover, finding the slope can also tell how big of a step should be taken to reach the optimal point. Conceptually, this approach optimizes the function by iteratively moving in the negative direction of the gradient, also known as steepest descent. The smoothness of the moving process to reach the minimum depends on the size of steps, which in the GD method is called the learning rate, i.e., α. A higher learning rate can extend over a larger function area, but there is a risk of overshooting the optimal point. Meanwhile, a smaller learning rate is more precise since the slope is frequently re-calculated. However, if it is not wisely parameterized, the optimization needs more time to converge and obtain the minima. Thus, choosing the suitable step size implies a high efficiency of the optimization algorithm.

This well-known method has been employed and exercised in many scientific applications, mostly in optimization problems such as in machine learning [29]. Researchers in [30] exploited GD to establish an operation model of nodes for the classification of heart conditions based on a deep neural network. Furthermore, researchers in [31] adopted the GD method to produce phase-only computer-generated holograms (CGHs). This study inferred that the GD required low time to optimize the phase-only CGH distribution and provided higher-precision images compared to the Gerchberg–Saxton (GS) algorithm. Next, researchers in [32] used the GD method to optimize the vector centers of the consequent layer functions and receptive field matrices in a neuro-fuzzy model based on the standard criterion of mean square error. Furthermore, GD was also used by [33] for minimizing the equation of error used in the study by updating the W_1 and W_2 matrices in order to solve the predicting student performance (PSP) problem. Researchers in [34] also utilized stochastic GD as one of the machine learning methods to be compared for the inversion of a stochastic skin optical model. The researchers also used multi-layer perceptron (MLP),

linear support vector regressor (SVR), convolutional neural network (CNN), and lasso regressor in the comparative study.

3. Methodology

This work aimed to predict the depth of hydrocarbon in the seabed logging (SBL) application by using a novel methodology, namely Gaussian process (GP)-based inverse modeling. Here, four subsections are discussed: 3.1. Preprocessing computer outputs, 3.2. Evaluating electromagnetic (EM) profile using the GP, 3.3. Predicting depth of hydrocarbon, and 3.4. Estimate and GP model validation.

3.1. Preprocessing Computer Outputs

Computer simulation technology (CST) software, which uses finite element (FE), is exploited in this work to model and evaluate prior information of EM responses (e.g., magnitudes of electric (E)-field) in the SBL for the proposed GP-based inversion methodology. CST is part of SIMULIA, a Dassault Systèmes brand, sourced from Darmstadt, Germany. According to [35], CST Studio Suite® exclusively uses the FE method as the EM simulation solver for low-frequency application. Nine SBL models with the same input properties, except the hydrocarbon depth, were simulated using the CST software. Details of the SBL data acquisition are presented in Appendix A. For demonstration, one SBL model replicated by the CST is depicted in Figure A1. The SBL models, known as target models, were replicated with three background layers such as air, seawater, and sediment. The nine different "tried" depths were between 200 m and 1000 m from the seafloor with an increment of 100 m each. All the computer input properties and parameters used in the SBL modeling are indicated in Table 1.

Table 1. Input properties and parameters of the seabed logging (SBL) models.

Input Property or Parameter	Numerical Values in Unit
Electrical conductivity (Air)	1.0×10^{-11} S/m
Electrical conductivity (Seawater)	1.63 S/m
Electrical conductivity (Sediment)	1.00 S/m
Electrical conductivity (Hydrocarbon)	2.0×10^{-3} S/m
Thickness of layer (Air)	300 m
Thickness of layer (Seawater)	1000 m
Total thickness of layer (Sediment)	3500 m
Thickness of layer (Hydrocarbon)	200 m
Transmission frequency	0.125 Hz
Current strength	1250 A
Transmitter length	270 m
Depth of elevation (Transmitter)	30 m
Length of SBL models	20,000 m
Width of SBL models	20,000 m
Height of SBL models	5000 m

As discussed by researchers in [28], interpretation of the SBL data relies on an assumption that the normalized magnitude versus offset (MVO) between the target responses (i.e., with presence of hydrocarbon) and reference responses (i.e., without presence of hydrocarbon) can pre-identify the presence of high-resistive bodies. The idea is that the magnitudes of target response are divided by the magnitudes of reference response along the offsets (i.e., source–receiver separation distances). Ratio of the normalization should be greater than one, which is associated with the existence of the potential reservoirs. Thus, all the generated CST outputs were preprocessed before being used for the forward GP modeling. By considering the ratio measurements and the symmetrical property of the developed SBL models, we chose nine different datasets with 73 datapoints per set, which are from an approximate distance of 11,840.80 m to an approximate distance of 19,004.97 m, for the data processing presented in the next subsection.

3.2. Evaluating Electromagnetic Profile Using the Gaussian Process

This work uses Gaussian process machine learning (GPML) MATLAB codes for the GP regression, previously written by Rasmussen and Nickisch, which can be found in [36]. Note that the MATLAB version used by this work is R2019b. As mentioned before, the Gaussian process (GP) acts as the alternative approach to the numerical forward modeling in the SBL application for evaluating EM profiles at various "untried" hydrocarbon depths. Suppose all nine CST datasets, $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_9 = \{(x_i, y_i)\}_{i=1}^N$, where $N = 73$, are treated as the training datasets for the GP modeling. It is also possible to simply write the d-dimensional input variable vector and its corresponding output, which is magnitude of E-field vector, as $x_i \in X$ and $y_i \in Y$, respectively. Generally, in this work, X is the N-by-d matrix and Y is the N-by-1 column vector, where $x \in \mathbb{R}^{d=2}$ and $y \in \mathbb{R}^{d=1}$. In other words, $x = (s, h)$, where s and h represent the offset (i.e., source–receiver separation distance) and the depth of hydrocarbon, respectively.

In order to evaluate or estimate the EM profiles at various "untried" depth inputs conforming to the testing input, $x_* = (s_*, h_*)$, the posterior distribution of the testing output y_* is defined in Equation (1). This work's interest was to predict the depth parameter of hydrocarbon along the offsets from 11,840.80 m to 19,004.97 m. Thus, here, only the testing input variable of hydrocarbon depth, h_*, is varied accordingly, while the s_* remains unchanged. Note that the interested hydrocarbon depths in the GP regression are from 200 m to 1000 m with an increment of 20 m each (including the "tried" depths).

$$(y_* | X, Y, x) \sim G(m_*, \kappa_*), \tag{1}$$

where the predictive mean, m_*, and the predictive variance, κ_*, are expressed in Equations (2) and (3), respectively.

$$m_* = K(X, x_*)^T \left[K(X, X) + \sigma_N^2 I_N \right]^{-1} Y, \tag{2}$$

and

$$\kappa_* = k(x_*, x_*) - K(X, x_*)^T \left[K(X, X) + \sigma_N^2 I_N \right]^{-1} K(X, x_*). \tag{3}$$

Based on Equations (2) and (3), $K(X, X)$ is the covariance matrix with a dimension of N-by-N amongst the X, $K(X, x_*)$ is the covariance matrix with a dimension of N-by-1 between the X and x_*, $k(x_*, x_*)$ is the covariance of the x_*, I_N is an N-dimensional unit matrix, and $\sigma_N^2 I_N$ is the noise covariance matrix. All these covariance matrices are composed of squared exponential (SE) terms, where each element of the matrices is computed using Equation (4) as follows:

$$k(x, x') = \sigma_f^2 \left[exp\left(-\frac{|x - x'|^2}{2\ell^2} \right) \right], \tag{4}$$

where σ_f represents the signal variance and ℓ represents the isotropic length scale. These two parameters are known as the hyperparameters of the SE covariance function, $\theta = \{\sigma_f, \ell\}$. SE is a very famous covariance function used in GP regression. According to [37], GP with SE is infinitely differentiable. This feature is very important for the derivative-based approach.

Here, it must be noted that all the computations of Equations (1)–(4) require the estimate of θ. These hyperparameters are numerically estimated by minimizing the negative log marginal likelihood (NLML) using a gradient-based optimizer. The NLML equation is defined in Equation (5) as follows:

$$NLML = -log[p(Y|X, \theta)] = \frac{N \log(2\pi)}{2} + \frac{\log |K(X, X) + \sigma_N^2 I_N|}{2} + \frac{Y^T \left[K(X, X) + \sigma_N^2 I_N \right]^{-1} Y}{2}. \tag{5}$$

In Equation (5), a logarithmic identifier is used to simplify the integral expression involved in the marginal likelihood method [25]. Once the hyperparameters are successfully estimated, the estimation of the GP regression is then defined using Equations (2) and (3). The m_* is a mean value of the estimation that gives the best estimate of y_* (i.e., magnitude of E-field at x_*), and the κ_* is the predictive variance of the testing input in terms of two standard deviations (95% confidence interval), which reflects the reliability of the predictive mean. These two equations are the basis in the forward GP modeling.

3.3. Predicting Depth of Hydrocarbon

The fitted GP model is denoted as $Y_{fit}(s,h)$, where h here includes all "tried" and "untried" depths. This means that, for every h (i.e., 200 m–1000 m with a gap of 20 m), the GP gives estimates of the EM profile, $Y_{fit}(s,h)$, at distances from 11,840.80 m to 19,004.97 m. Assume that the observational dataset $Y_{obs}(s,\blacksquare)$ is observed at $N = 73$ distance points and the hydrocarbon resides underneath the seabed but with "unknown" depth. To find out the depth of hydrocarbon of the observational dataset Y_{obs}, the least square criterion is formed between the Y_{fit} and Y_{obs}. Mean squared error (MSE) is used as the loss function in the optimization algorithm. In general, MSE measures the average amount of squared differences between Y_{fit} and Y_{obs}. The loss function $L(h_j)$ is expressed in Equation (6) as follows:

$$\{L(h_j)\}_{j=1}^{M} = \frac{1}{N} \sum_{i=1}^{N} \left[Y_{fit}(s_i, h_j) - Y_{obs}(s_i) \right]^2, \tag{6}$$

where $L(h_j)\,|\,j = 1, 2, \ldots, M$ and $M = 41$.

Equation (6) is a function of h. The loss function is minimized over the h with the minimizer serving as an estimate of the hydrocarbon depth. By using Equation (6), the $L(h_j)$ between the Y_{fit} and Y_{obs} are calculated. Here, the least $L(h_j)$ is then identified. As for now, the minimizer of the least $L(h_j)$ serves as the best estimate of hydrocarbon depth for Y_{obs}. Thus, in this work, the respective minimizer is treated as the initial guess of depth. Then, the iteration index j is set as zero and the initial guess is denoted as h_0.

Next, for the optimization using the GD method, the first-partial derivative of the $L(h_j)$ with respect to h_j is required. The partial derivative of the loss function is defined in the equation below. The derivation of Equation (7) is provided in Appendix B (in Equation (A1)):

$$L'(h_j) = \frac{2}{N} \left(\sum_{i=1}^{N} \left[Y_{fit}(s_i, h_j) - Y_{obs}(s_i) \right] \right) \cdot Y'_{fit}(s_i, h_j), \tag{7}$$

where Y'_{fit} is the derivative of the GP using the SE covariance function with respect to h_j. It must be noted here that the forward GP model is defined by estimations composed by the predictive mean as shown in Equation (2). To find out the Y'_{fit}, the predictive mean m_* is differentiated with respect to h_j as follows:

$$m'_* = K'(X, x_*)^T \left[K(X,X) + \sigma_N^2 I_N \right]^{-1} Y, \tag{8}$$

where $K'(X, x_*)$ is the derivative of covariance function between X and x_* with respect to h_j. As mentioned earlier, each element of the covariance matrix is computed using the SE covariance function (Equation (4)). Thus, the derivative of SE between X and x_* with respect to h_j is defined in order to find the $K'(X, x_*)$, as expressed in Equation (9). Its derivation is provided in Appendix B (in Equation (A2)):

$$k'(x, x_*) = \sigma_f^2 \cdot \left(\frac{|x - x_*|}{\ell^2} \right) \cdot exp \left(-\frac{|x - x_*|^2}{2\ell^2} \right). \tag{9}$$

Next, the following step is finding the correction or update of h_j. It is denoted as h_{j+1}. By referring to the GD algorithm, the equation of h_{j+1} is expressed as follows:

$$h_{j+1} = h_j - \alpha \cdot L'(h_j), \tag{10}$$

where α is the learning rate, which governs the step size of the optimization. Normally, as a start, a geometric sequence with a common ratio of 10 is used to indicate which minimizer can give a significant optimization.

After getting the h_{j+1}, the EM profile at h_{j+1}, denoted as $Y_{fit}(s_i, h_{j+1})$, is then estimated using the GP regression by defining the h_{j+1} as its testing input and the GP modeling steps are repeated. Next, the new loss function for h_{j+1} between Y_{fit} and Y_{obs} is calculated using Equation (11) as follows:

$$L(h_{j+1}) = \frac{1}{N} \sum_{i=1}^{N} \left[Y_{fit}(s_i, h_{j+1}) - Y_{obs}(s_i) \right]^2. \tag{11}$$

The tolerance threshold ε, which acts as the stopping criterion for the iterations, is defined for verifying the $L(h_{j+1})$ before deciding the optimal hydrocarbon depth for the Y_{obs}. In the context of this work, it was decided that the changes between each subsequent minimizer or depth had to be less than 0.05. As presented in Section 4, the specific limit leads to acceptable results. Equation (12) shows the equation of relative percentage change (RPC) used in this work:

$$RPC_{j+1} = \frac{|h_{j+1} - h_j|}{|h_j|} \times 100\%. \tag{12}$$

The next conditions used in this work are the following: the maximum number of iterations is 100 and the loss function (i.e., MSE) must converge. In other words, $L(h_{j+1})$ must be always lower than $L(h_j)$. Each iteration must satisfy these three threshold conditions. The optimal depth of hydrocarbon for Y_{obs} is believed to be successfully predicted if the process does not meet the conditions. Otherwise, it is returned to the step of evaluating the $L'(h_j)$ onwards by defining a new iteration index, $j = j + 1$. All the flows are repeated until the ε is neglected, then the iterations stop. The methodological flow presented in this section is shown in Figure 3.

3.4. Estimate and GP Model Validation

Error measurement is very important to determine the reliability of estimations. In this work, for estimate validation, the percentage error between the computer input (i.e., hydrocarbon depth used in the computer simulation) and the depth predicted by the GP-based inverse modeling was calculated using Equation (13). Meanwhile, for GP model validation, three error validators were used in this work, namely mean absolute deviation (MAD), mean absolute percentage error (MAPE), and root mean squared error (RMSE). The errors of EM profiles provided by the GP model were calculated by assuming the EM responses modeled by the FE (through the CST software) as the true EM values. All equations of the MAD, MAPE, and RMSE are defined in Equations (14)–(16), respectively, as follows:

$$Percentage\ Error = \left| \frac{h_{true} - h_{estimate}}{h_{true}} \right| \times 100\%, \tag{13}$$

$$MAD = \frac{1}{N} \sum |y_{FE} - y_{GP}|, \tag{14}$$

$$MAPE = \frac{1}{N} \sum \frac{|y_{FE} - y_{GP}|}{|y_{FE}|} \times 100\%, \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N} \sum (y_{FE} - y_{GP})^2}, \tag{16}$$

where h_{true} is the CST depth input, $h_{estimate}$ is the depth estimated by the GP-based inverse modeling, y_{FE} is the EM response modeled by FE method, y_{GP} is the EM response evaluated by the forward GP modeling, and N is the total number of observations.
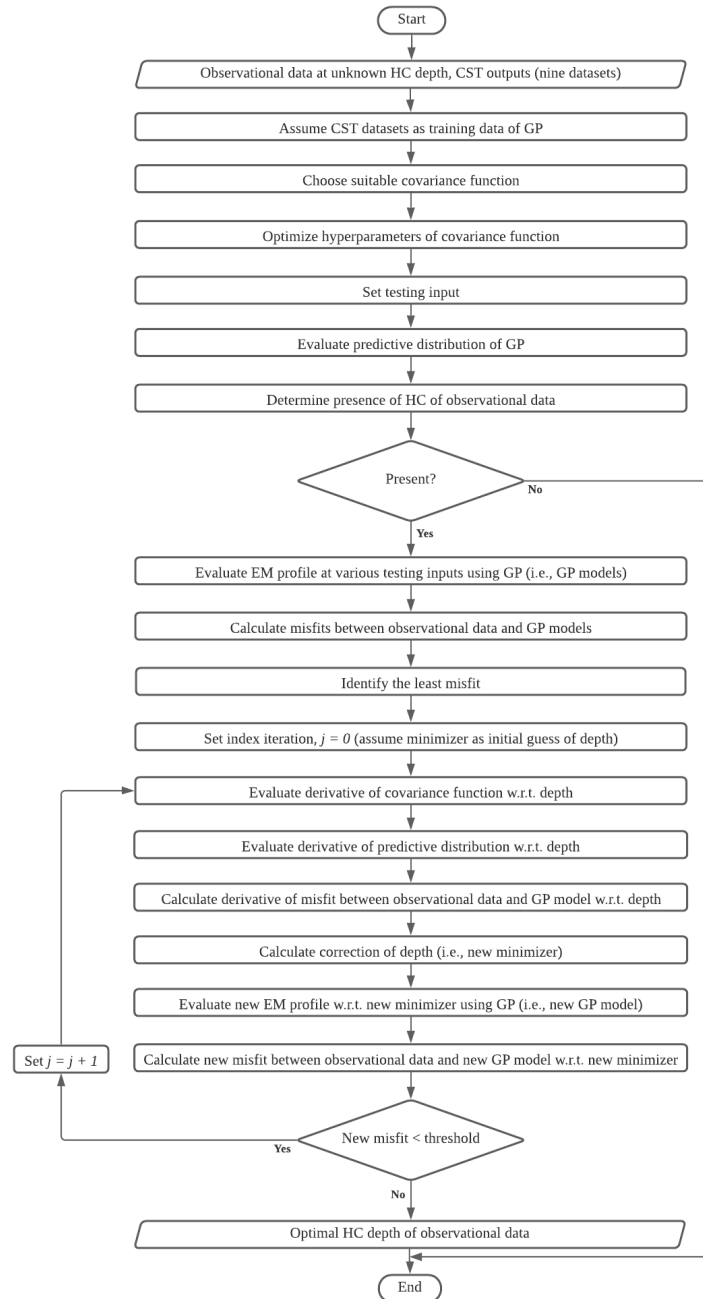

Figure 3. Methodological flow of Gaussian process (GP)-based inverse modeling. Note that all the seabed logging (SBL) models replicated by the computer simulation technology (CST) software are target models (i.e., with hydrocarbon layer). Hence, determining the presence of hydrocarbon is not discussed in this work. HC stands for hydrocarbon and w.r.t. represents with respect to.

4. Results and Discussion

All prior electromagnetic (EM) responses were generated using the computer simulation technology (CST) software. Nine seabed logging (SBL) models were parameterized with the same input properties, except for its depth of hydrocarbon, which were 200 m–1000 m with a gap of 100 m each. It must be noted that these EM responses were evaluated using the finite element (FE) method, which was used exclusively by the CST software

for the low-frequency applications. In this work, only magnitude of electric (E)-field was considered as the corresponding output. Figure 4 shows the magnitude versus offset plot for all nine SBL datasets.
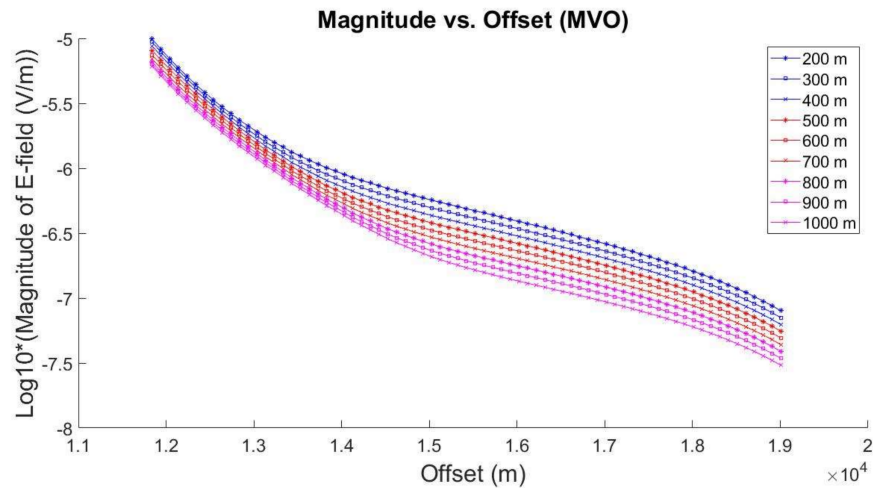
Figure 4. Magnitude of electric (E-)field versus offset for nine different depths of hydrocarbon from 200 m to 1000 m with an increment of 100 m each.

Based on Figure 4, the *x*-axis represents the offset distances from approximately 11,840.80 m to 19,004.97 m and the *y*-axis represents the magnitude of E-field with a logarithmic scale. This figure reveals that the CST was capable of simulating the SBL application in the deep marine environment. The generated EM responses behave similarly to the real behavior of EM signals in media with various conductivities. The EM signals travel through a high-resistive body with less attenuation. Thus, a model with higher hydrocarbon depth generates a much smaller magnitude of E-field due to the factor of far EM signal propagation. Next, these EM responses were fully utilized by the GP regression in order to provide more EM profiles at "untried" depths. The contour plot of the two-dimensional (2D) GP model is depicted in Figure 5.
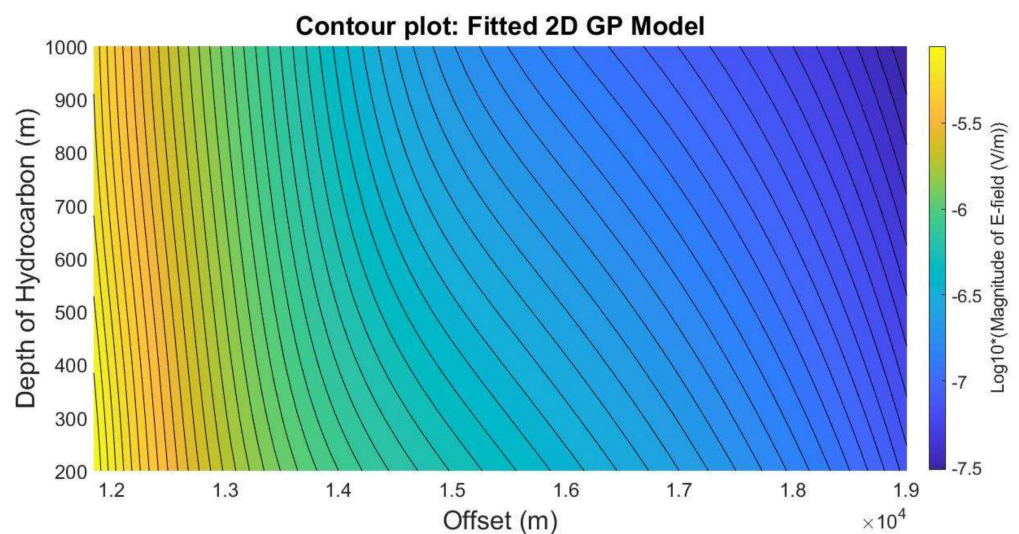
Figure 5. Two-dimensional (2D) forward Gaussian process (GP) model with 41 different electromagnetic (EM) datasets with depths from 200 m to 1000 m with an increment of 20 m.

Based on Figure 5, the *x*-axis is the offset distance and the *y*-axis is the 41 depths of hydrocarbon from 200 m to 1000 m, while the magnitudes of the E-field are represented

in color codes. This forward GP model is a combination of EM profiles at "tried" and "untried" hydrocarbon depths. The purpose of this GP modeling was to evaluate as many EM responses as possible at various depths of hydrocarbon for the inversion scheme, for which generating all these EM responses using the CST software could take high computational time. Simulating and generating all these 41 datasets using the CST could require around 3 h and 30 min, whereas the GP modeling only took 45 min for the prior knowledge acquisition and less than 1 min for developing the forward GP model and evaluating 41 profiles.

In this work, three random observational datasets were generated separately using the CST software where the depths of hydrocarbon were assumed to be "unknown", i.e., 350, 650, and 950 m. Here, the goal was to find the best estimates of hydrocarbon depths of the observational datasets with the smallest MSE values. In other words, a small MSE value indicates that the EM profile evaluated by the forward GP model gives the most likely values for the observational dataset, thereby giving the best depth information. As discussed before, the step size of the optimization was determined by the learning rate. For this work, three different learning rates were used for every inversion scheme. The first scheme (1st observational dataset) used 0.051 as its learning rate, whereas the second (2nd observational dataset) and third (3rd observational dataset) schemes used 0.087 and 0.081, respectively. Details of how this work determined the suitable learning rates are provided in Appendix C. Plots of MSE versus learning rate and iteration number versus learning rate for every observational dataset are depicted in Figures 2–6 and A1. After identifying the learning rate, optimization using the gradient descent was executed. Thus, three MSE versus hydrocarbon depth plots are plotted as shown in Figures 6–8.

By referring to Figures 6–8, the *x*-axis is the depth of hydrocarbon and the *y*-axis is the misfit, which is the MSE. The blue points are the 41 MSE values between the GP model from depths 200 m to 1000 m and the observational data. The red points indicate the optimized MSEs with respect to the updated minimizers. Every optimization must satisfy the convergence condition. Thus, in order to monitor the convergence status, this work illustrates MSE versus number of iterations plots for every observational dataset, which are depicted in Figures 9–11. In these figures, the *x*-axis and *y*-axis represent the number of iterations and MSE, respectively. Based on these figures, even though the first inversion scheme needed a higher number of iterations than the second and third ones, all the iterations were smaller than the defined maximum number, which was 100. Moreover, all the MSEs converged where the optimizations stopped at the lowest MSE values. To make it easy to interpret, the main results are presented in Table 2.
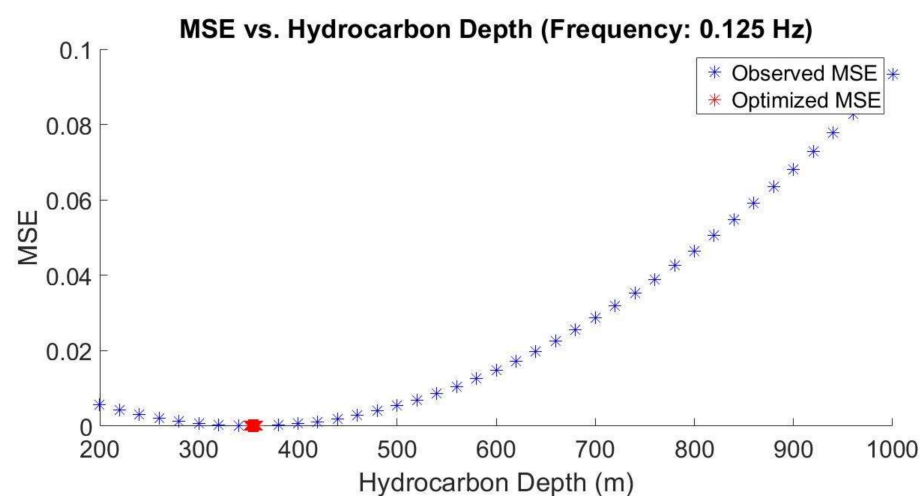

Figure 6. Mean squared error (MSE) versus hydrocarbon depth plot (1st observational dataset).
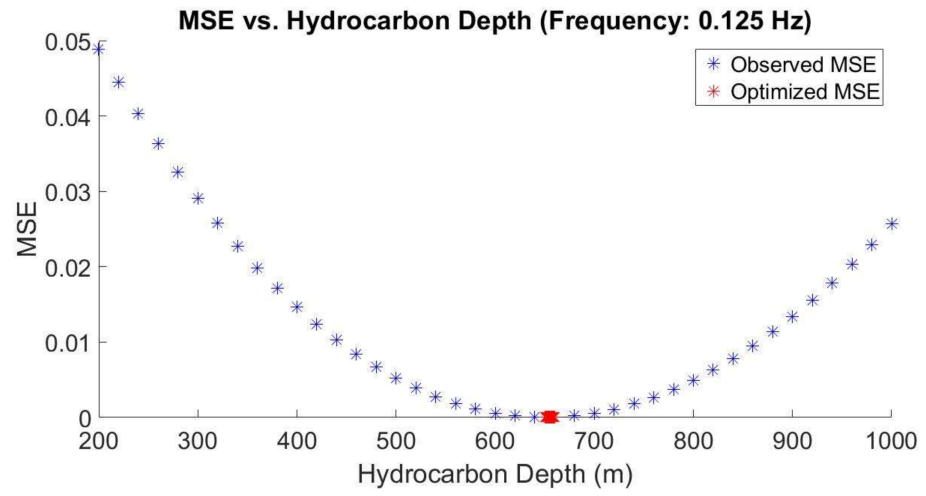
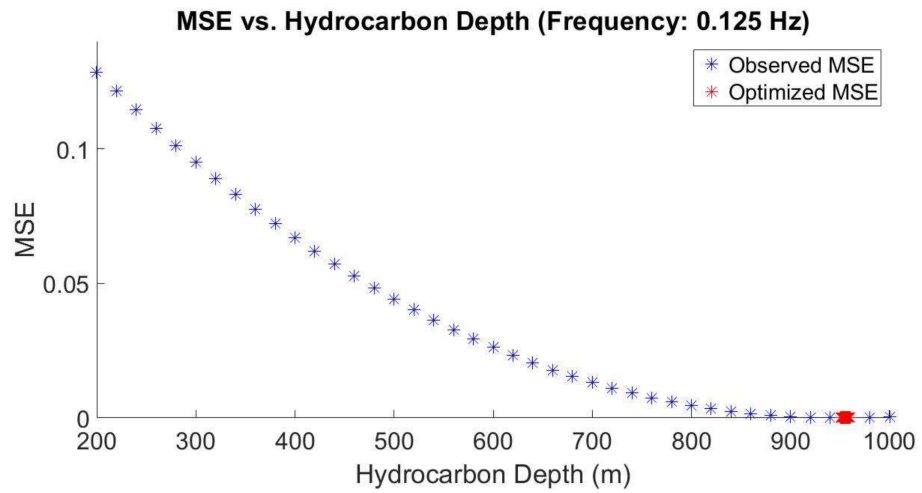**Figure 7.** Mean squared error (MSE) versus hydrocarbon depth plot (2nd observational dataset).

Figure 8. Mean squared error (MSE) versus hydrocarbon depth plot (3rd observational dataset).
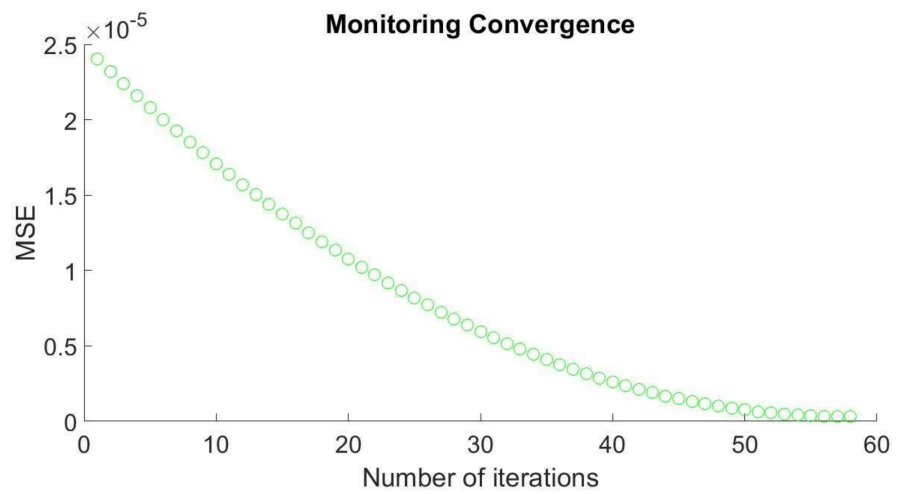

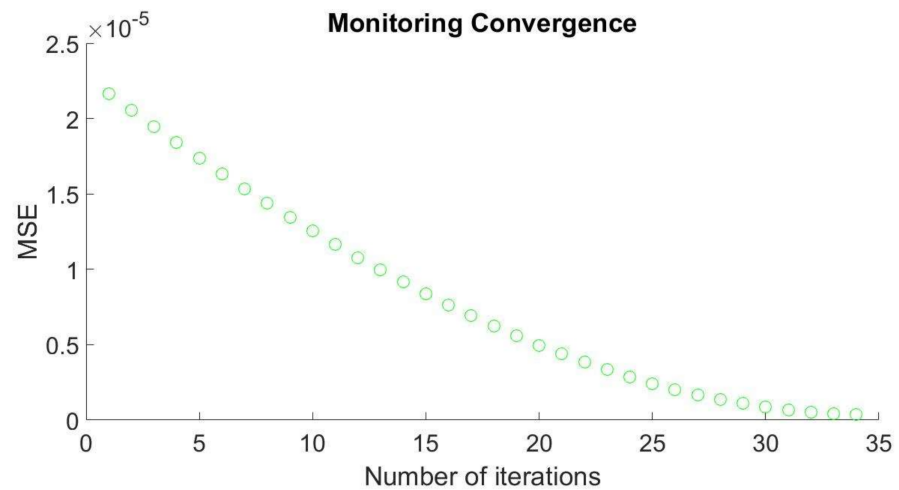**Figure 9.** Convergence test (1st observational dataset).

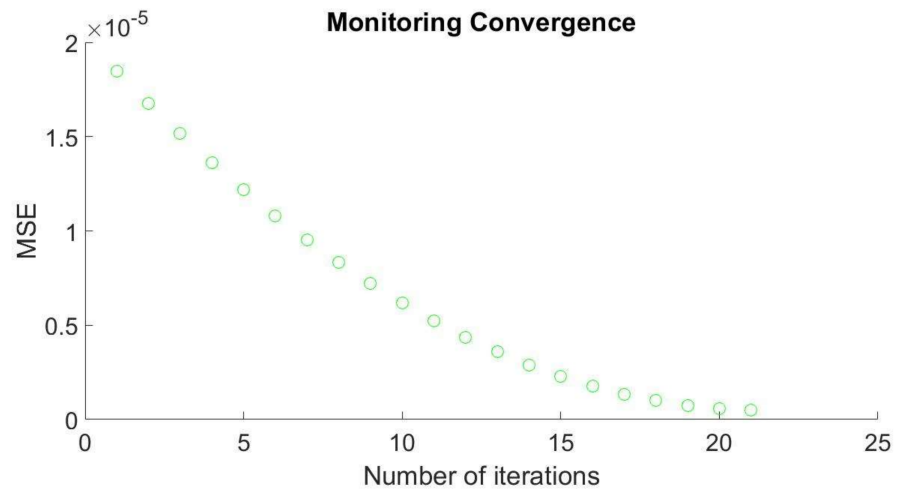**Figure 10.** Convergence test (2nd observational dataset).

Figure 11. Convergence test (3rd observational dataset).

Table 2. The main Gaussian process (GP)-based inverse modeling results (all three observational datasets).

Dataset	Index	Iteration Number	Estimate (m)	MSE	RPC (%)
1	h_{58}	58	349.92	0.00000033	0.04977145
2	h_{34}	34	650.20	0.00000039	0.04978711
3	h_{21}	21	950.22	0.00000051	0.04991896

Based on Table 2, the first observational dataset required 58 iterations to yield the best estimate of hydrocarbon depth, which was 349.92 m. The second observational dataset needed 34 iterations to get its best estimate, which was 650.20 m. In contrast, only 21 iterations were needed by the third observational dataset to produce the estimate of 950.22 m. Moreover, only 114.047 s, 73.760 s, and 60.556 s were taken by the GP-based inversion methodology to reach the estimates of 349.92 m, 650.20 m, and 950.22 m, respectively. From the table, the calculated MSE values were also very small. This means that the deviations between its true EM values and the predicted EM profiles were very low. All the final RPCs were also lesser than 0.05 as compared to what was defined in the methodology. In addition, in the context of this work, the reliability of the GP regression in evaluating the EM profiles was investigated as well by considering data modeled by the FE method as the true EM values. Thus, MADs, MAPEs, and RMSEs between the EM responses modeled by FE and EM evaluated by the GP (at depths 200–1000 m with a gap

of 100 m each) were calculated. The results are presented in Table 3. Furthermore, to find out the accuracy and performance of the GP-based inverse modeling in predicting the hydrocarbon depth, percentage errors between the true depths (i.e., computer input) and the estimates were calculated as well. The percentage errors are indicated in Table 4.

Table 3. Validation of the Gaussian process (GP) model.

Hydrocarbon Depth (m)	MAD	MAPE (%)	RMSE
200	1.39×10^{-9}	0.0849	2.63×10^{-9}
300	1.84×10^{-9}	0.1027	4.25×10^{-9}
400	1.48×10^{-9}	0.1035	3.26×10^{-9}
500	1.64×10^{-9}	0.1144	3.49×10^{-9}
600	1.68×10^{-9}	0.1168	4.28×10^{-9}
700	1.21×10^{-9}	0.1270	2.33×10^{-9}
800	1.37×10^{-9}	0.1204	3.57×10^{-9}
900	1.85×10^{-9}	0.1346	5.79×10^{-9}
1000	9.64×10^{-9}	0.1212	1.81×10^{-9}
Average	1.49×10^{-9}	0.1140	3.49×10^{-9}

Table 4. Percentage errors of hydrocarbon depth prediction.

True Depth (m)	Estimate (m)	Percentage Error (%)
350.0000	349.9172	0.0237
650.0000	650.2015	0.0310
950.0000	950.2225	0.0234

Based on Table 3, all the MADs, MAPEs, and RMSEs for depths 200 m to 1000 m with 100 m gap each were very small, which were 1.49×10^{-9} (on average), 0.114% (on average), and 3.49×10^{-9} (on average), respectively. As seen in Table 4, the percentage errors were very low as well, which was 0.0260% on average. All the results revealed that GP regression was capable of evaluating the EM profiles at various depths of hydrocarbon accurately and with low time consumption. The main goal was also achieved in that the GP-based inversion methodology can predict the depth of hydrocarbon in the SBL application efficiently.

5. Conclusions

This work proposed Gaussian process (GP)-based inverse modeling to predict the depth of hydrocarbon in the seabed logging (SBL) application. The capability of computational modeling and inversion of the SBL data with good accuracy and low computational effort is one of the main concerns in electromagnetic (EM) problems. Thus, this work used machine learning-based approaches, namely GP, as the forward modeling of EM profiles and the gradient descent (GD) method as the optimizer for finding the input parameter of the hydrocarbon (e.g., depth) in the SBL data processing. This work utilized three datasets generated through the computer simulation technology (CST) software with random "unknown" hydrocarbon depths as the observational responses. Based on the results, GP regression can provide accurate EM profiles at various depths of hydrocarbon with low computational time (45 min for the prior data acquisition) and less than 2 min for the GP-based inversion methodology. We believed that if the forward modeling performed by the GP could be done rapidly for the repeated use, the inverse modeling (by the GD) would not significantly affect the total computational time of the inversion methodology. Next, the resulting mean absolute deviations (MADs), mean absolute percentage errors (MAPEs), and root mean squared errors (RMSEs) between the GP model and the data modeled by finite element (FE) method were very small. Furthermore, the GP-based inverse modeling could also predict the depth of hydrocarbon for each observation dataset accurately. All the calculated percentage errors between the estimates (by the GP-based

inversion methodology) and the true depth values (computer input) were very small as well. Therefore, it can be concluded that the GP-based inverse modeling was successfully developed to predict the depth of hydrocarbon in the SBL. This attempt could help marine EM data processing, especially in order to de-risk hydrocarbon exploration in the deep marine environment.

Appendix A. Data Acquisition Using CST Software

Typical synthetic seabed logging (SBL) models were replicated using the computer simulation technology (CST) software. The input parameters can be seen in Table 1. The first layer was air, followed by seawater and sediment. Based on Figure A1, the isotropic hydrocarbon layer was positioned in the sediment layer at 200 m from the seafloor (for demonstration purposes, only one model is shown here). The source was placed at the center of the model and elevated at 30 m from the seafloor (coordinate: x = 10,000 m, y = 10,000 m, z = 1270 m), and inline receivers were located on the seafloor (i.e., static source–receiver separation distances). All the generated electromagnetic (EM) responses were symmetrical from the center to the left and right boundaries.
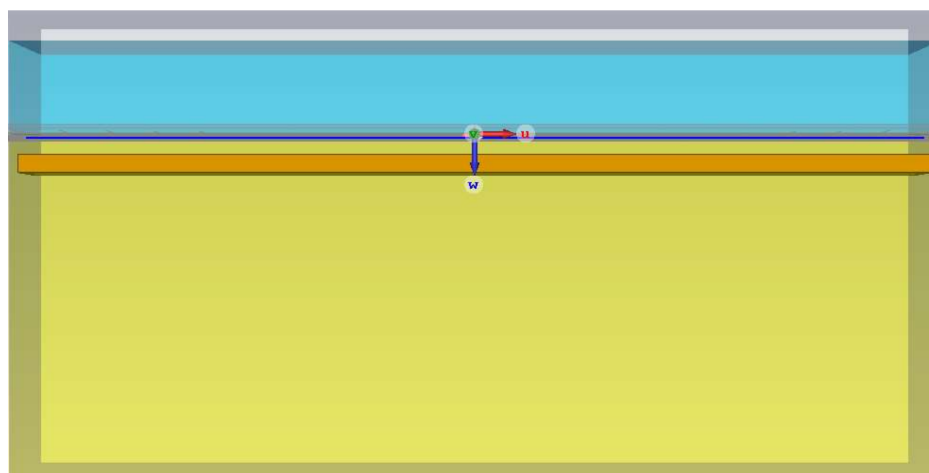

Figure A1. Setup of the seabed logging (SBL) model in the computer simulation technology (CST) software. The depth of hydrocarbon is 200 m from the seafloor.

Appendix B. Derivation of First-Partial Derivative

Equation (A1) shows the derivation of the first-partial derivative of the mean squared error (MSE) with respect to h_j.

$$
\begin{aligned}
L(h_j) &= \frac{1}{N}\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]^2 \\
\frac{\partial}{\partial h_j}L(h_j) &= \frac{\partial}{\partial h_j}\left[\frac{1}{N}\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]^2\right] \\
\frac{\partial}{\partial h_j}L(h_j) &= \frac{1}{N}\cdot\frac{\partial}{\partial h_j}\left[\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]^2\right] \\
\frac{\partial}{\partial h_j}L(h_j) &= \frac{2}{N}\cdot\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]\cdot\frac{\partial}{\partial h_j}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right] \\
\frac{\partial}{\partial h_j}L(h_j) &= \frac{2}{N}\cdot\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]\cdot\frac{\partial}{\partial h_j}Y_{fit}(s_i,h_j) \\
L'(h_j) &= \frac{2}{N}\left(\sum_{i=1}^{N}\left[Y_{fit}(s_i,h_j)-Y_{obs}(s_i)\right]\right)\cdot Y'_{fit}(s_i,h_j)
\end{aligned}
\tag{A1}
$$

Equation (A2) shows the derivation of the derivative of the squared exponential (SE) covariance function with respect to the testing point, x_*.

$$
\begin{aligned}
k(x,x_*) &= \sigma_f^2\left[exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\right] \\
\frac{\partial}{\partial x_*}k(x,x_*) &= \frac{\partial}{\partial x_*}\left[\sigma_f^2\left[exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\right]\right] \\
\frac{\partial}{\partial x_*}k(x,x_*) &= \sigma_f^2\cdot\frac{\partial}{\partial x_*}\left[exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\right] \\
\frac{\partial}{\partial x_*}k(x,x_*) &= \sigma_f^2\cdot exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\cdot\frac{\partial}{\partial x_*}\left[-\frac{|x-x_*|^2}{2\ell^2}\right] \\
\frac{\partial}{\partial x_*}k(x,x_*) &= \sigma_f^2\cdot exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\cdot\left(\frac{-1}{2\ell^2}\right)\cdot\frac{\partial}{\partial x_*}|x-x_*|^2 \\
\frac{\partial}{\partial x_*}k(x,x_*) &= \sigma_f^2\cdot exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)\cdot\left(\frac{-1}{2\ell^2}\right)\cdot -2|x-x_*| \\
k'(x,x_*) &= \sigma_f^2\cdot\left(\frac{|x-x_*|}{\ell^2}\right)\cdot exp\left(-\frac{|x-x_*|^2}{2\ell^2}\right)
\end{aligned}
\tag{A2}
$$

Appendix C. Identifying Learning Rate (α) of Gradient Descent

Geometric sequences with a common ratio of 10 were tested first in order to indicate which learning rates would give the significant optimization. From here, other values of learning rate were examined as well to find the smallest mean squared error (MSE) that could be optimized. Based on Figure 2, Figure 4, and Figure 6 3.30×10^{-7}, 3.90×10^{-7}, and 5.10×10^{-7} were the smallest optimized MSEs for all observational data. Thus, based on the MSEs, the size of the learning rates was narrowed down. Next, the number of iterations was then considered as plotted in Figure 3, Figure 5, and Figure 7. From here, learning rates that gave the smallest MSE with the lowest iteration number and followed the relative percentage change (RPC) condition were chosen for the Gaussian process (GP)-based inverse modeling.
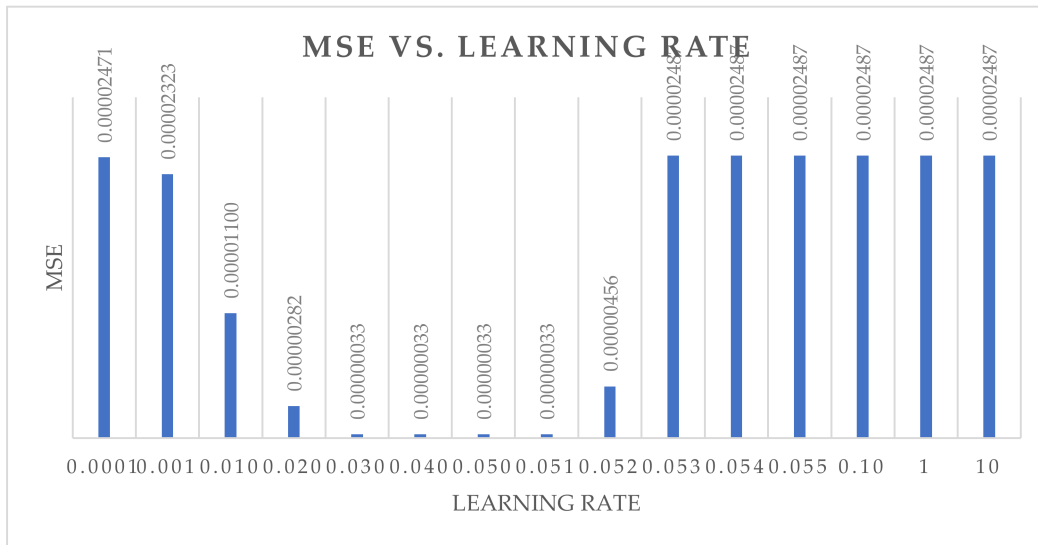
Figure 2. Bar plot of mean squared error (MSE) versus learning rate (1st observational dataset).
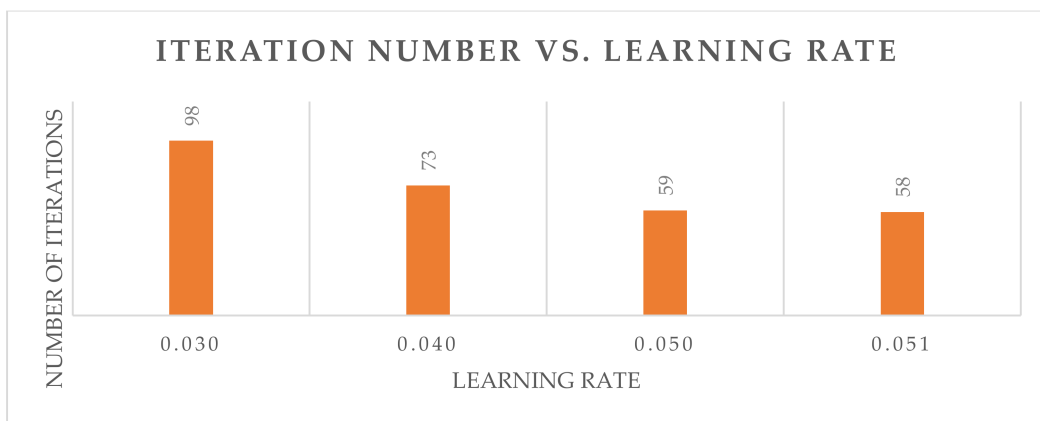

Figure 3. Bar plot of iteration number versus learning rate for the smallest mean squared error (MSE) (1st observational dataset).
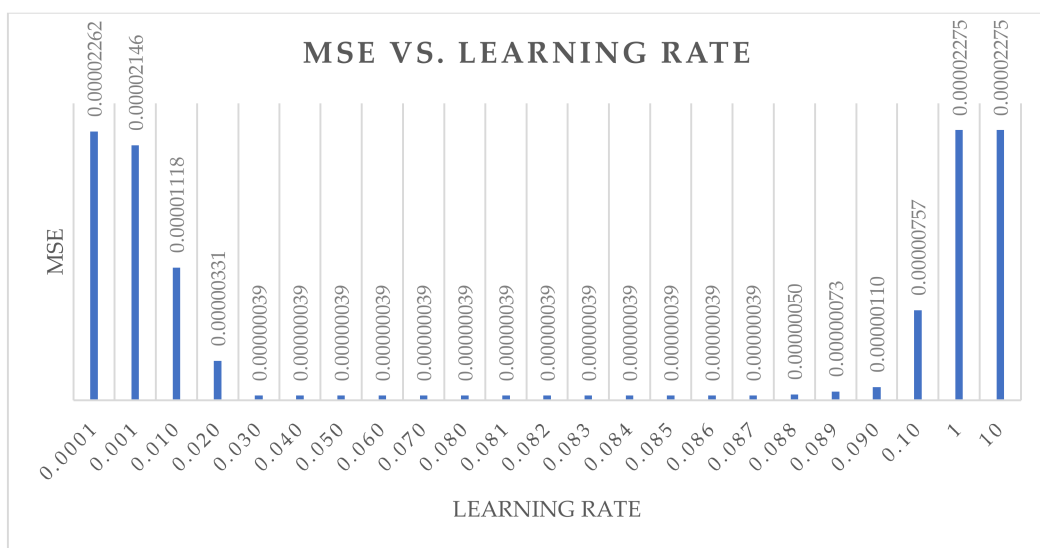

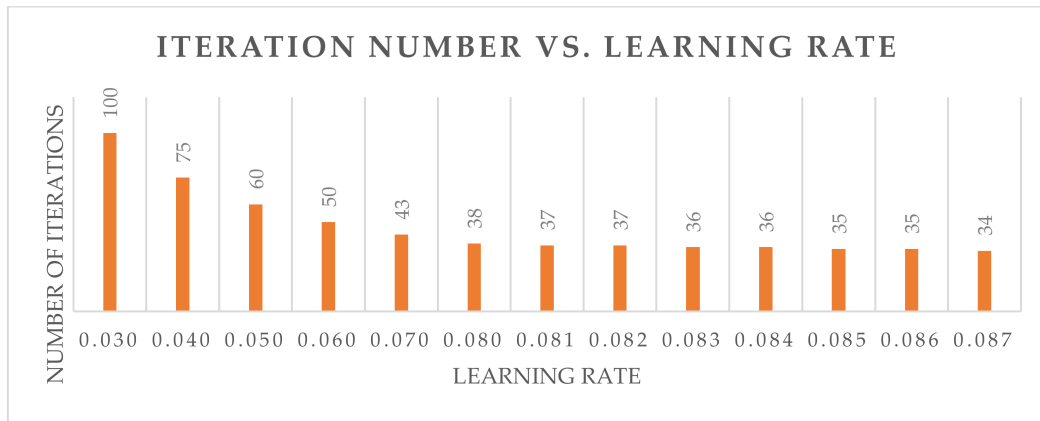**Figure 4.** Bar plot of mean squared error (MSE) versus learning rate (2nd observational dataset).

Figure 5. Bar plot of iteration number versus learning rate for the smallest mean squared error (MSE) (2nd observational dataset).
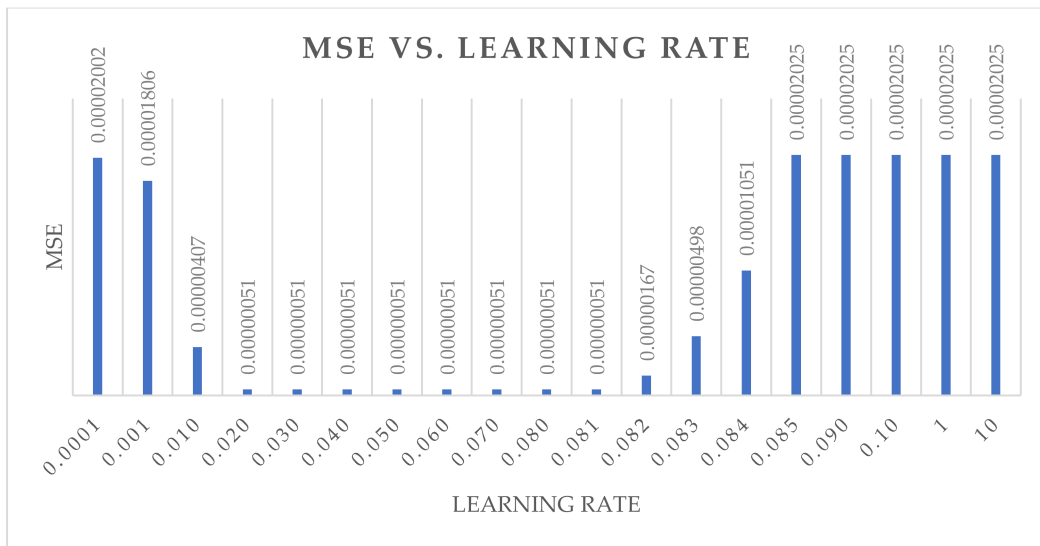

Figure 6. Bar plot of mean squared error (MSE) versus learning rate (3rd observational dataset).
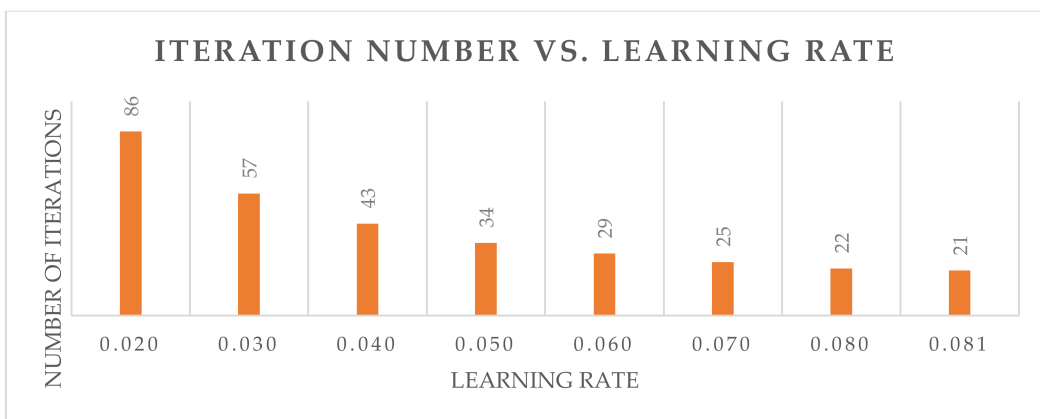

Figure 7. Bar plot of iteration number versus learning rate for the smallest mean squared error (MSE) (3rd observational dataset).

References

1. 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.

2. Ellingsrud, S.; Eidesmo, T.; Johansen, S. Remote sensing of hydrocarbon layers by seabed logging: Results from a cruise offshore Angola. *Lead. Edge* **2002**, *21*, 972–982. [CrossRef]

3. Hesthammer, J.; Boulaenko, M. The offshore EM challenges. *First Break* **2005**, *23*, 59–66. [CrossRef]

4. Carazzone, J.J.; Burtz, O.M.; Green, K.E.; Pavlov, D.A. Three dimensional imaging of marine controlled source EM data. *SEG Expand. Abstr.* **2005**, *24*, 575.

5. Srnka, L.J.; Carazzone, J.J.; Ephron, M.S.; Eriksen, E.A. Remote reservoir resistivity mapping. *Lead. Edge* **2006**, *25*, 972–975. [CrossRef]

6. Constable, S.; Srnka, L.J. An introduction to marine controlled-source electromagnetic methods for hydrocarbon exploration. *Geophysics* **2007**, *72*, WA3–WA12. [CrossRef]

7. Um, E.S.; Alumbaugh, D.L. On the physics of the marine controlled-source electromagnetic method. *Geophysics* **2007**, *72*, W13–WA26. [CrossRef]

8. Andréis, D.; MacGregor, L.M. Controlled-source electromagnetic sounding in shallow water: Principles and applications. *Geophysics* **2008**, *73*, F21–F32. [CrossRef]

9. Zhdanov, M.S. Electromagnetic geophysics: Notes from the past and the road ahead. *Geophysics* **2010**, *75*, A49–A66. [CrossRef]

10. Mohd Aris, M.N.; Daud, H.; Dass, S.C.; Mohd Noh, K.A. Gaussian Process Methodology for Multi-Frequency Marine Controlled-Source Electromagnetic Profile Estimation in Isotropic Medium. *Processes* **2019**, *7*, 661. [CrossRef]

11. Puzyrev, V.; Koldan, J.; de la Puente, J.; Houzeaux, G.; Vázquez, M.; Cela, J.M. A parallel finite-element method for three-dimensional controlled-source electromagnetic forward modelling. *Geophys. J. Int.* **2013**, *193*, 678–693. [CrossRef]

12. Li, Y.; Key, K. 2D marine controlled-source electromagnetic modeling: Part 1—An adaptive finite-element algorithm. *Geophysics* **2007**, *75*, WA51–WA62. [CrossRef]

13. 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]

14. Daeung, Y.; Micheal, S.Z.; Hongzhu, C.; Alexander, G. A hybrid finite difference and integral equation method for modeling and inversion of marine CSEM data. In Proceedings of the SEG 2015 New Orleans Annual Meeting, New Orleans, LA, USA, 18–23 October 2015; pp. 823–827.

15. Harari, O.; Steinberg, D.M. Optimal designs for Gaussian process models|via spectral decomposition. *J. Stat. Plan. Inference* **2014**, *154*, 87–101. [CrossRef]

16. 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]

17. 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]

18. Kong, D.; Chen, Y.; Li, N. Gaussian process regression for tool wear prediction. *Mech. Syst. Signal. Process.* **2018**, *104*, 556–574. [CrossRef]

19. Petelin, D.; Grancharova, A.; Kochigan, J. Evolving Gaussian Process models for prediction of Ozone concentration in the air. *Simul. Model. Pract. Theory* **2013**, *33*, 68–80. [CrossRef]

20. Pal, M.; Deswal, S. Modelling pile capacity using Gaussian process regression. *Comput. Geotech.* **2010**, *37*, 942–947. [CrossRef]

21. Yin, F.; Zhao, Y.; Gunnarsson, F.; Gustafsson, F. Received-Signal-Strength Threshold Optimization Using Gaussian Processes. *IEEE Trans. Signal. Process.* **2017**, *65*, 2164–2177. [CrossRef]

22. 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. *Microelectron. Reliab.* **2013**, *53*, 832–839. [CrossRef]

23. 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]

24. 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]

25. 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]

26. Mohd Aris, M.N.; Daud, H.; Mohd Noh, K.A.; Dass, S.C.; Mukhtar, S.M. Modelling 1-D synthetic seabed logging data for thin hydrocarbon detection: An application of Gaussian process. In Proceedings of the National Symposium of Mathematical Sciences, Bangi, Malaysia, 26–27 November 2019.

27. Li, J.; He, M.; Cui, G.; Wang, X.; Wang, W.; Wang, J. A Novel Method of Seismic Signal Detection Using Waveform Features. *Appl. Sci.* **2020**, *10*, 2919. [CrossRef]

28. 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]

29. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

30. 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]

31. Liu, S.; Takaki, Y. Optimization of Phase-Only Computer-Generated Holograms Based on the Gradient Descent Method. *Appl. Sci.* **2020**, *10*, 4283. [CrossRef]

32. 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]

33. 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]

34. Annala, L.; Ayramo, S.; Polonen, I. Comparison of Machine Learning Methods in Stochastic Skin Optical Model Inversion. *Appl. Sci.* **2020**, *10*, 7097. [CrossRef]

35. 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).

36. Documentation for GPML Matlab Code Version 4.2. Available online: http://www.gaussianprocess.org/gpml/code/matlab/doc/ (accessed on 9 September 2020).

37. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.