buildings

MDPI

Article

Prediction of Unstable Hydrodynamic Forces on Submerged Structures under the Water Surface Using a Data-Driven Modeling Approach

Van My Nguyen [1], Hoang Nam Phan [1,*] and Thanh Hoang Phan [2]

[1] Faculty of Road and Bridge Engineering, The University of Danang—University of Science and Technology, Da Nang 550000, Vietnam
[2] School of Mechanical Engineering, Pusan National University, Busan 46241, Korea
* Correspondence: phnam@dut.udn.vn

Abstract: Catastrophic failures of partially or fully submerged structures, e.g., offshore platforms, hydrokinetic turbine blades, bridge decks, etc., due to the dynamic impact of free surface flows such as waves or floods have revealed the need to evaluate their reliability. In this respect, an accurate estimation of hydrodynamic forces and their relationship to instability in structures is required. The computational fluid dynamics (CFD) solver is known as a powerful tool to identify dynamic characteristics of flow; however, it commonly consumes a huge computational cost, especially in cases of re-simulations needed. In this paper, an efficient surrogate model based on the Gaussian process is developed to rapidly predict the nonlinear hydrodynamic pressure coefficients on submerged bodies near the water surface. For this purpose, a CFD model is first developed, which is based on a two-dimensional incompressible Navier–Stokes solver incorporating free surface treatment and turbulent flow models. Then, an experimental design is adopted to generate initial training samples considering the effect of the submerged body shape ratio and flow Re number. Surrogate models of hydrodynamic pressure coefficients and their instability based on Gaussian process modeling are established using the outcome from the CFD simulations, where optimal trend and correlation functions are also investigated. Once surrogate models are obtained, the mean and oscillation amplitudes of hydrodynamic pressure coefficients on a submerged rectangular body, which represents the shape of most civil structures, can be rapidly predicted without the attempt at re-simulation. The findings can be practically applied in rapidly assessing hydrodynamic forces and their instability of existing submerged civil structures or in designing new structures, where a suitable shape ratio should be adopted to avoid flow-induced instability of hydrodynamic forces.

Keywords: submerged structure; free surface flow; hydrodynamic force; computational fluid dynamics; surrogate model; Gaussian process; data-driven approach

Citation: Nguyen, V.M.; Phan, H.N.; Phan, T.H. Prediction of Unstable Hydrodynamic Forces on Submerged Structures under the Water Surface Using a Data-Driven Modeling Approach. *Buildings* **2022**, *12*, 1683. https://doi.org/10.3390/buildings12101683

Academic Editors: Dongming Li and Zechuan Yu

Received: 14 September 2022
Accepted: 8 October 2022
Published: 13 October 2022

1. Introduction

Past failures of partially or fully submerged civil structures, e.g., offshore platforms, hydrokinetic turbine blades, bridge decks, etc., due to the dynamic impact caused by floods, waves or tsunamis have revealed the significant need for evaluating the reliability of these structures under hydrodynamic forces [1–3]. Regarding this issue, one of the challenges is the accurate evaluation of multiphase flows and their impact on submerged structures. This is essential not only for optimum designs of the newly designed structures but also for the estimation of the degree of risk for existing ones.

Large-scale structures under free surface flows commonly require time-consuming experimental analyses. In addition, flow characteristics, e.g., flow patterns, velocities and pressure fields, are uncertain and nonlinear; hence, resulting in a very high cost and time required for many experimental setups in order to observe accurate flow characteristics [4,5].

While analytical models for hydrodynamic forces on submerged bodies have been developed and incorporated into design codes, they are often overestimated and with a degree of error [6–8]. An alternative is the development of numerical models for free surface flows and their impact on structures; this can reduce much effort in terms of cost and time as compared with experimental models [9]. The numerical models are mainly based on computational fluid dynamic (CFD) approaches that solve Navier-Stokes equations along with the treatment of the free surface. However, due to the uncertain flow characteristics, varying with each particular case, the design of these numerical simulations becomes complicated and requires numerous computing resources, especially in cases of the large number of samples considered in a reliability analysis [10,11].

With the development of computer science, besides various computer models which have been adopted in many fields of CFD [12], machine learning techniques have been developed and widely used to predict multiphase flow characteristics [13–15], in which, data-driven techniques, e.g., neural network [13], support vector machine [14] and Gaussian process [15], have been implemented to build surrogate models for the prediction of the multiphase flow pattern, as well as the hydrodynamic pressure distribution. These surrogate models, which are used when an outcome of interest cannot be easily measured or computed, can efficiently reduce the computational effort and rapidly estimate flow characteristics in the context of uncertainty treatments and reliability analyses. The first two methods (i.e., neural network and support vector machine) commonly require an adequate dataset for a reliable prediction, which depends on the number of input and output parameters, while the Gaussian process regression makes it possible to predict the model response with little observed data. In addition, the Gaussian process offers a flexible kernel method for regression due to various available trends and correlation functions. Therefore, in many complex problems, this technique is more suitable and efficient in terms of reducing computational costs [11,16].

The objective of this paper is to develop surrogate models, which are based on Gaussian process modeling, to rapidly predict nonlinear hydrodynamic pressure coefficients and their instability effect on submerged bodies. As a case study, rectangular submerged bodies near the water surface are considered. This type of shape is standard and represents the shape of many engineering applications such as bridge decks, offshore platforms and hydrokinetic turbine blades. Firstly, a modeling approach of the flow passing a submerged body is presented based on a two-dimensional incompressible Navier–Stokes solver. The free surface is treated using the volume of fluid method and the effect of the turbulent flow is also considered by using the shear stress transport turbulence model. Then, an appropriate experimental design is used to generate initial training samples considering the effect of the aspect ratio of the submerged body and the Re number of the flow. Surrogate models of hydrodynamic pressure coefficients based on the Gaussian process modeling are established using the outcome from the CFD simulations, where optimal trend and correlation functions are also investigated. Once surrogate models are obtained, the mean and oscillation amplitudes of hydrodynamic pressure coefficients of the free surface flow on a submerged cylinder with an arbitrary aspect ratio can be accurately and rapidly predicted without the need for attempt at re-simulation. The findings from the work can be practically applied in rapidly assessing hydrodynamic force and its instability effect on existing submerged civil structures, or in designing new structures, where a suitable shape ratio range is recommended to avoid the detrimental effects of flow-induced instability from hydrodynamic forces.

2. Numerical Model of Free Surface Flow

In this study, a two-dimensional (2D) incompressible Reynolds-averaged Navier-Stokes (RANS) homogeneous two-phase mixture model is adopted to simulate the nonlinear interactions of a submerged cylinder beneath the free surface [9,17]. The model solves the mixture continuity and momentum equations to obtain the mean flow velocity and pressure fields. The RANS model is closed by including a turbulence model to predict

fluctuating velocity components; thus, the shear stress transport $k-\omega$ model is adopted. In addition, to capture complex free surface behaviors and non-linear hydrodynamic pressure coefficients, the interface between the water and air phases is numerically treated using the volume-of-fluid (VOF) method [18,19]. These equations can be described in the Cartesian,

$$\frac{\partial u_j}{\partial x_j} = 0 \tag{1}$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho_m}\frac{\partial p}{\partial x_i} + g_i + \mu_m \frac{\partial^2 u_i}{\partial x_j} - \frac{\partial \overline{u_i' u_j'}}{\partial x_j}, \tag{2}$$

$$\frac{\partial \alpha_g}{\partial t} + \frac{\partial (\alpha_g u_j)}{\partial x_j} = 0, \tag{3}$$

Here, the subscripts $i,\ j = 1,\ 2$ represent two directions of x and y in the computational domain, respectively; t is the computation time; p denotes the pressure, u and u' are the mean and the fluctuating velocities; g is the gravitational term; and ρ_m and μ_m are the mixture density and viscosity, respectively.

The interface position is numerically treated via the phasic volume fraction of the gas phase, α_g. Here, the pure gas phase is obtained in the case of α_g equals 1.0, while the pure water phase is obtained in the case of α_g equals 0.0. The interfaces with a limited thickness between two phases are identified by the values in the range from 0.0 to 1.0. The properties of the mixture phase at the interface are predicted by a function of the volume of fraction of the individual phase, an example, for the calculation of the mixture values for density and mixture valuables

$$\rho_m = \alpha_g \rho_g + (1-\alpha_g)\rho_w, \tag{4}$$

$$\mu_m = \alpha_g \mu_g + (1-\alpha_g)\mu_w, \tag{5}$$

where ρ_w, μ_w and ρ_g, μ_g are the density and viscosity of the individual water and gas phases, respectively.

The numerical discretization of the equation system in the generally structured grid is based on the finite volume method with a pressure-based solver. For time discretization, the first-order implicit method is applied. The first-order upwind scheme is adopted for both the convective and viscous terms and the advection equation is approximated using the implicit compressive scheme. The main reason behind using first-order and implicit schemes is to obtain better convergence than high-order and explicit schemes, especially for strongly deformable free surfaces with breaking wave phenomena. All simulation results are performed using the Ansys Fluent software [20].

3. Surrogate Model of Hydrodynamic Pressure Coefficients

3.1. Basic Formulations

Gaussian process regression uses a set of observed training data to predict spatially correlated data, which postulates a combination of a functional basis and departure in the following form [21],

$$Y_i = \sum_{j=1}^{p} \beta_j f_j(\boldsymbol{x}^{(i)}) + Z(\boldsymbol{x}^{(i)}) \text{ with } i = 0, \ldots, m, \tag{6}$$

where the first term is the unknown multivariate polynomial function, $\boldsymbol{f} = \left\{ f_j(\boldsymbol{x}^{(i)}) \right\}$ with $j = 1, \ldots, p$, called the trend, and $Z(\mathbf{x})$ is the realization of the Gaussian process having zero mean and variance σ^2; $Z(\boldsymbol{x})$ is expressed as

$$Cov\left[Z(\boldsymbol{x}), Z(\boldsymbol{x}')\right] = \sigma^2 R(\boldsymbol{x} - \boldsymbol{x}', \ell), \tag{7}$$

where σ^2 is the variance of the Gaussian process, whereas R is the correlation function which is the function of the difference $\boldsymbol{x} - \boldsymbol{x}'$ and scale parameters ℓ ($\ell_i > 0, i = 1, \ldots, n$).

Several correlation functions are proposed, e.g., the exponential (Equation (8)), Gaussian (Equation (9)), and Matérn-3/2 (Equation (10)),

$$R(\boldsymbol{x}-\boldsymbol{x}',\ell)=\sum_{i=1}^{n}e^{-1/\ell_i|x_i-x_i'|}, \tag{8}$$

$$R(\boldsymbol{x}-\boldsymbol{x}',\ell)=\sum_{i=1}^{n}e^{-\frac{1}{2}\left(\frac{|x_i-x_i'|}{\ell_i}\right)^2}, \tag{9}$$

$$R(\boldsymbol{x}-\boldsymbol{x}',\ell)=\sum_{i=1}^{n}\left(1+\frac{\sqrt{3}}{\ell_i}|x_i-x_i'|\right)e^{-\sqrt{3}/\ell_i|x_i-x_i'|} \tag{10}$$

The vector of the prediction $\hat{\boldsymbol{Y}}_0$ and the true response $\boldsymbol{Y}=\{Y_i\}$ with $i=1,\ldots,m$ is normally distributed,

$$\begin{Bmatrix}\hat{\boldsymbol{Y}}_0\\ \boldsymbol{Y}\end{Bmatrix}\sim \mathrm{N}_{1+m}\begin{Bmatrix}\boldsymbol{f}_0^T\boldsymbol{\beta}\\ \mathbf{F}\boldsymbol{\beta}\end{Bmatrix},\sigma^2\begin{bmatrix}1 & \boldsymbol{r}_0^T\\ \boldsymbol{r}_0 & \mathbf{R}\end{bmatrix}, \tag{11}$$

where $\boldsymbol{f}_0$ is the vector of regression models evaluated at $x^{(0)}$, $\mathbf{F}$ is the regression matrix, $\boldsymbol{r}_0$ is the vector of cross-correlations between the point $x^{(0)}$, $r_{0i}=R\left(x^{(0)}-x^{(i)},\ell\right)$, $\mathbf{R}$ is the correlation matrix of the true response, $R_{ij}=\left(x^{(i)}-x^{(j)},\ell\right)$.

The best linear unbiased predictor of the unknown quantity of interest y_0 is the Gaussian random variate $\hat{\boldsymbol{Y}}_0$ with mean and variance,

$$\mu_{\hat{Y}_0}=\boldsymbol{f}_0^T\hat{\boldsymbol{\beta}}+\boldsymbol{r}_0^T\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{F}\hat{\boldsymbol{\beta}}), \tag{12}$$

$$\sigma^2_{\hat{Y}_0}=\sigma^2\left(1-\boldsymbol{r}_0^T\mathbf{R}^{-1}\boldsymbol{r}_0+\left(\mathbf{F}^T\mathbf{R}^{-1}\boldsymbol{r}_0-\boldsymbol{f}_0\right)^T\left(\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\left(\mathbf{F}^T\mathbf{R}^{-1}\boldsymbol{r}_0-\boldsymbol{f}_0\right)\right), \tag{13}$$

where $\hat{\boldsymbol{\beta}}=\left(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{R}^{-1}\boldsymbol{y}$.

The maximum likelihood estimation technique is better suited for deriving estimators. Here, the likelihood of the observations $\boldsymbol{y}$ is defined concerning its multivariate normal distribution, which depends on $\boldsymbol{\beta},\sigma^2$, and ℓ,

$$L\left(\boldsymbol{y}\middle|\boldsymbol{\beta},\sigma^2,\ell\right)=\frac{1}{\left((2\pi\sigma^2)^m[\det\mathbf{R}(\ell)]\right)^{1/2}}\exp\left[-\frac{1}{2\sigma^2}\left(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta}^{\mathbf{T}}\right)\mathbf{R}(\ell)^{-1}(\boldsymbol{y}-\mathbf{F}\boldsymbol{\beta})\right]. \tag{14}$$

By maximizing the quantity described in Equation (14), the following analytical estimates of $\boldsymbol{\beta}$ and σ^2 that are functions of ℓ are obtained as

$$\hat{\boldsymbol{\beta}}(\ell)=\left(\mathbf{F}^T\mathbf{R}(\ell)^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^T\mathbf{R}(\ell)^{-1}\boldsymbol{y}, \tag{15}$$

$$\hat{\sigma^2}(\ell)=\frac{1}{m}\left(\boldsymbol{y}-\mathbf{F}\hat{\boldsymbol{\beta}}^{\mathbf{T}}\right)\mathbf{R}(\ell)^{-1}(\boldsymbol{y}-\mathbf{F}\hat{\boldsymbol{\beta}}). \tag{16}$$

By substituting these two solutions into Equation (14), its corresponding opposite log-likelihood reads

$$-\log L\left(\boldsymbol{y}\middle|\boldsymbol{\beta},\sigma^2,\ell\right)=\frac{m}{2}\log(\psi(\ell))+\frac{m}{2}(\log(2\pi)+1)\text{ with }\psi(\ell)=\hat{\sigma^2}(\ell)[\det\mathbf{R}(\ell)]^{\frac{1}{2}}, \tag{17}$$

and thus, the maximum likelihood estimate of ℓ is given as

$$\hat{\ell}=\underset{\ell}{\operatorname{argmin}}\psi(\ell). \tag{18}$$

3.2. Procedure of Surrogate Model-Based Hydrodynamic Pressure Coefficient Prediction

The overall procedure for the development of a Gaussian process based-surrogate model is as follows:

(i) In the first step, the most important variables and their distribution functions should be identified. An appropriate DOE is then conducted within the range of interest variables. As a result, several initial training samples are generated and corresponding CFD models are then built.

(ii) CFD simulations of the flow field are conducted for each combination of training conditions. The flow field characteristics, as well as hydrodynamic pressure coefficients, are obtained at each simulation. In this study, the mean and oscillation values of hydrodynamic pressure coefficients, i.e., drag, lift and moment, are considered as the model responses.

(iii) Once the training dataset has been established based on the DOE and the corresponding model responses, a surrogate model of the model response is built using the Gaussian process modeling incorporated in a Matlab-based software, Uqlab [22]. In this step, different trend and correlation functions that compose the Gaussian process are tested. An optimal surrogate model is finally obtained based on the error estimation from the cross-validation.

4. Case study of Submerged Bodies beneath the Water Surface

4.1. CFD Simulation and Design of Experiments

In this study, a rectangular shape body fully submerged beneath a free surface is selected, which represents the shape of most bridge deck or other civil structure components under the free surface flow. The computational domain and simulation conditions are adopted following the experimental work by Chu et al. [23], in which the problem of an open channel with a particular size is shown schematically in Figure 1. The rectangular body submerges in the water at a depth of h and a distance between the channel bed and the cylinder S. To reduce the computational cost, a planar symmetric numerical model is used under the main assumption that there are no effects in the spanwise direction. This assumption was also adopted for numerical computations of free surface flows over a submerged body in many studies [9]. The boundary conditions are applied as follows: (i) the fixed uniform velocity is specified at the inlet condition, (ii) at the outlet condition, the extrapolation values are applied for the pressure and velocity fields, (iii) at the top boundary condition, open conditions are applied and (iv) at the bottom line and cylinder, non-slip wall conditions are used.

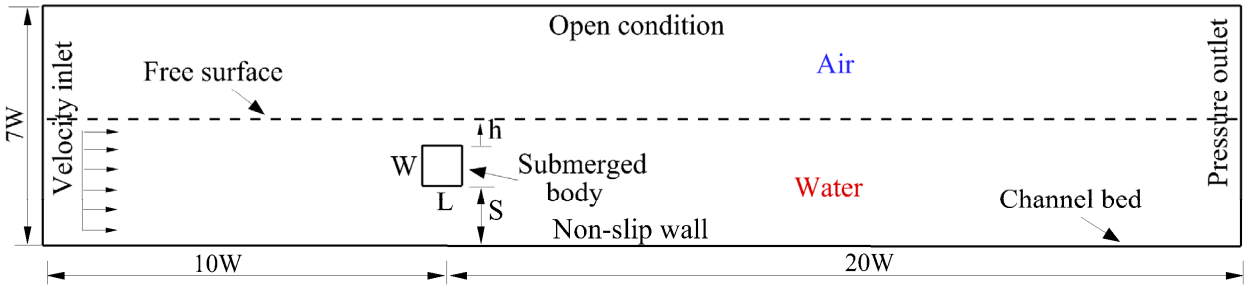


Figure 1. Computational model of the free surface flow around a submerged body.

The mesh distribution for the whole computational domain and zoomed regions near the submerged body are presented in Figure 2. Here, the meshing strategy with a high-resolution value close to the body and free surface is used to obtain high accuracy predictions for the pressure and velocity fields, particularly for the free surface shape motion. The grid and time step sensitivity tests were performed through convergency analyses in the previous work [9]; thus the fine grid with a total number of nodes of 112,649, the $y+$ value at the body surface of 1.1, and the time step of 0.002 (s) are used in the present simulations.

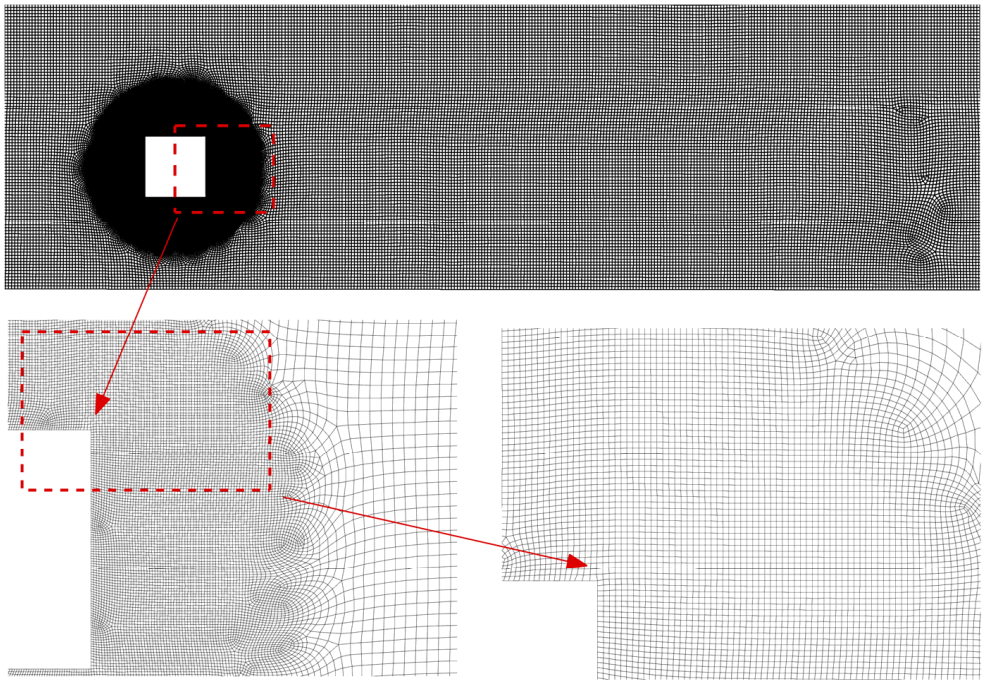


Figure 2. The mesh domain and mesh distribution around the submerged body.

The complex simulation of the above-mentioned CFD model reveals the need for developing a more efficient surrogate model to possibly and rapidly identify nonlinear hydrodynamic pressure coefficients on a submerged body under free surface flows. The development of a surrogate model or metamodel first requires the generated samples that involve modeling parameters. In this study, Latin Hypercube Sampling (LHS) [24] is utilized to generate training samples; this technique is widely used and has demonstrated its efficiency in the construction of surrogate models, especially ones based on the Gaussian process [25,26].

Many studies have demonstrated that the shape ratio (AR) (i.e., the length and depth ratios of the rectangular body, $AR = L/W$) and the Re number are the most significant parameters that affect the flow characteristic and impact hydrodynamic pressure [8,10]. Therefore, these two parameters are chosen as input random variables in the study with the ranges selected and presented in Table 1, which are assumed to be a uniform distribution. The other geometry parameters, such as the depth h and the clearance distance S, are deterministic.

Table 1. Modeling parameters for the design of experiments.

Parameter	Distribution Function	Lower Value	Upper Value
Shape ratio AR	Uniform	0.2	4
Re number	Uniform	8000	16,000

By using the LHS on two modeling parameters, a total of 40 samples are generated, as distributed in Figure 3. It should be noticed that there is no specific standard for the number of initial training samples, depending on the number of input variables, particular problem and training method. Since the Gaussian process has not needed the pre-assuming of a specified model and just requires a small number of initial training samples, an optimized design of 40 samples is chosen, as proposed by [26].

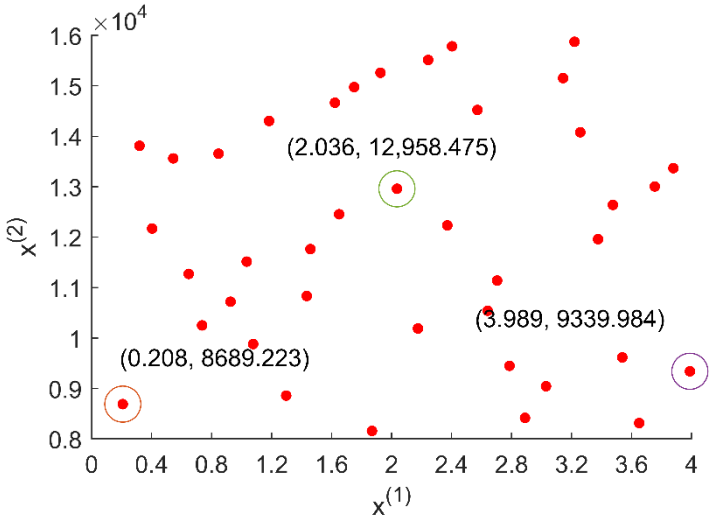


Figure 3. Design of experiments using LHS.

As examples of the flow characteristic, Figure 4 shows the velocity field and pressure contour for three cases of AR and Re values, marked by large circles in Figure 3. The free surface shape is plotted by the red solid line using the air volume fraction value of 0.5. Under the presence of the submerged body, significantly increasing free surface flow and reduction in the depth at the upward and the downward regions is observed. Therefore, a high-velocity water flow in the downstream region is formed. In addition, submerged wake vortices behind the body are generated under the effects of the inclination of the free surface, as shown by streamline fields on the left side of Figure 4. This nonlinear evolution behavior is found to be a major mechanism that increases the hydrodynamic force coefficients in comparison with the unbounded free surface flow. The pressure distributions around the submerged body are also shown on the right side of Figure 4. In the presence of the free surface, asymmetric low-pressure regions at the top and bottom of the submerged body are observed.

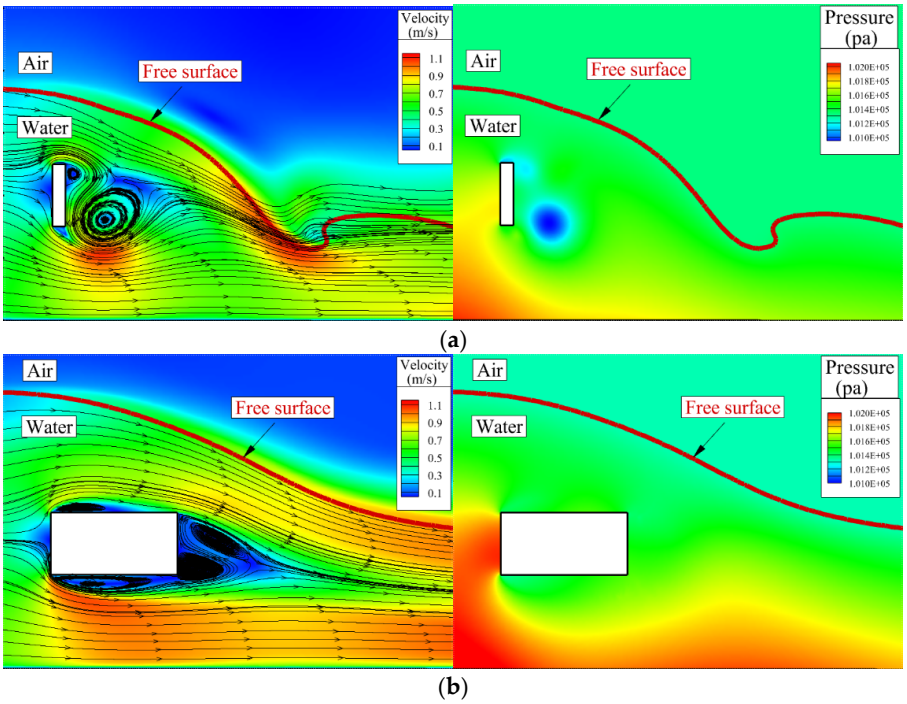
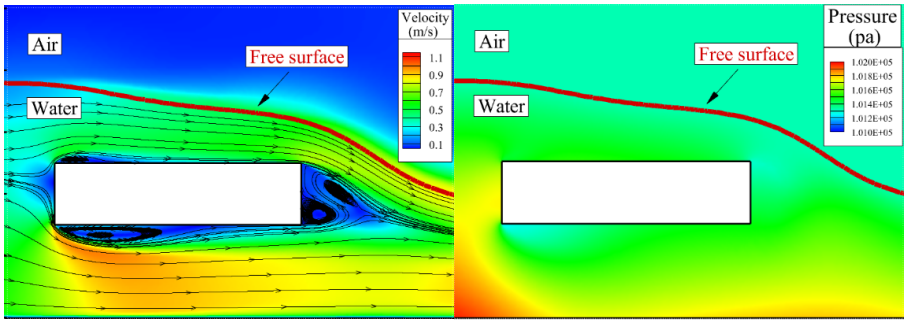


Figure 4. *Cont.*



(**c**)

Figure 4. Velocity and pressure fields around the submerged body with the free surface interactions: (**a**) AR = 0.208, Re = 8689.223, (**b**) AR = 2.036, Re = 12,958.475, (**c**) AR = 3.989, Re = 9339.984.

For each sample, hydrodynamic pressure coefficients (i.e., drag, lift and moment coefficients) acting on submerged bodies are obtained from CFD analyses. Figure 5 shows an example of the time evolution of hydrodynamic pressure coefficients for AR = 0.208, Re = 8689.223 (Figure 5a), AR = 2.036, Re = 12,958.475 (Figure 5b), and AR = 3.989, Re = 9339.984 (Figure 5c). The numerical results show that the hydrodynamic force coefficients are significantly varied under the nonlinear interaction with the free surface. In the cases of lower AR values (AR = 0.208), the force coefficients are in periodic evolutions with time and are characterized by a mean value and oscillated magnitude. These predicted values are determined by an averaged method over five oscillation cycles. In the case of higher values (AR = 2.036 and 3.989), stable behaviors of the force coefficients without oscillation features are observed.

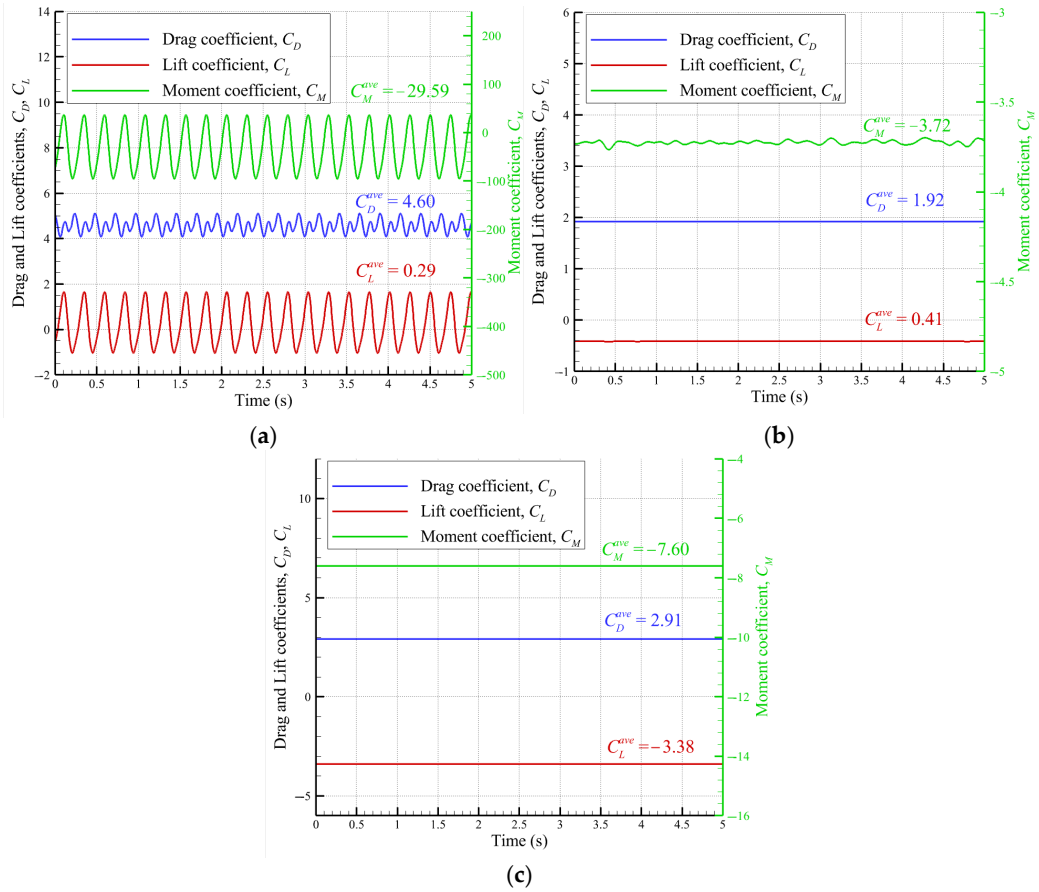


Figure 5. The time evolution of hydrodynamic force coefficients: (**a**) AR = 0.208, Re = 8689.223, (**b**) AR = 2.036, Re = 12,958.475, (**c**) AR = 3.989, Re = 9339.984.

Similarly, the outcomes of interest including six above-mentioned quantities (the mean values of C_L, C_D, C_M and their oscillations) are obtained and summarized in Table A1 (Appendix A), resulting in a total of 40 examples of training data in the dataset. The observed responses from the dataset are later used to train surrogate models of the outcomes.

4.2. Surrogate Model Development

Based on the above model output, the surrogate model is then developed using the above-mentioned Gaussian process modeling. The major advantage of this modeling approach is that it requires less observed data for the regression as compared to other data-driven techniques, such as support vector machine or artificial neural network. To optimize the surrogate model for the prediction, several trend and correlation functions are tested in this study. Due to the nonlinearity of the hydrodynamic pressure coefficients, nonlinear regression and correlation models are selected. In particular, for the trend, polynomial (one, two, and three degrees) functions are employed. On the other term, Matérn 3/2, exponential and Gaussian (in Equations (8)–(10)) correlation functions are examined.

To estimate the accuracy of each tested surrogate model, the leave-one-out (LOO) cross-validation is adopted, where one point is randomly ignored for the cross-validation and the other points are for training the surrogate model. This procedure is repeated until all the points are used. Therefore, to perform the LOO, one point $\boldsymbol{x}^{(\mathrm{i})}$ from the initial DOE is subsequently removed and the surrogate model $\hat{\boldsymbol{Y}}_{0,(-i)}\left(\boldsymbol{x}^{(\mathrm{i})}\right)$ is built from the remaining points of the design. The LOO cross-validation error is calculated on the true response design and its corresponding predicted responses as

$$\epsilon_{LOO} = \frac{1}{n}\frac{\sum_{i=1}^{n}\left[Y(x_i) - \hat{\boldsymbol{Y}}_{0,(-i)}(x_i)\right]^2}{\mathrm{Var}\,(\boldsymbol{Y})}, \tag{19}$$

where $\mathrm{Var}(\boldsymbol{Y})$ defines the estimated variance of the output variable.

Table 2 shows the cross-validation error of the tested surrogate models as combinations of different trend and observation functions. It can be observed that the estimated LOO errors vary with different combinations and outcomes of interest. In most of the cases, the combination of 3rd degree polynomial function and Matérn 3/2 correlation function results in the best performance with a minimum mean error of the prediction for the outcomes (highlighted in bold in Table 2). Also of note is that combinations of 2nd degree polynomial-Matérn 3/2 and 3rd degree polynomial-Gaussian also exhibit a good prediction.

Table 2. Cross-validation error estimation of the tested surrogate models.

Trend Function	Correlation Function	LOO Error						Mean Error
		C_D (Mean)	C_D (Oscil.)	C_L (Mean)	C_L (Oscil.)	C_M (Mean)	C_M (Oscil.)	
1st degree polynomial	**Matérn 3/2**	0.0222	0.0244	0.0010	0.0631	0.0956	0.0350	0.0402
2nd degree polynomial		0.0102	0.0250	0.0023	0.0620	0.1004	0.0371	0.0395
3rd degree polynomial		**0.0074**	**0.0270**	**0.0045**	**0.0610**	**0.0745**	**0.0259**	**0.0334**
1st degree polynomial	Exponential	0.0264	0.0449	0.0100	0.0728	0.1306	0.0672	0.0586
2nd degree polynomial		0.0079	0.0452	0.0068	0.0787	0.1131	0.0486	0.0500
3rd degree polynomial		0.0072	0.0616	0.0060	0.1173	0.0906	0.0367	0.0532
1st degree polynomial	Gaussian	0.0314	0.2446	0.0026	0.0590	0.1250	0.0441	0.0845
2nd degree polynomial		0.0155	0.2637	0.0042	0.3129	0.1111	0.0327	0.1234
3rd degree polynomial		0.0105	0.0316	0.0046	0.0677	0.0841	0.0216	0.0367

Examples of optimal surrogate models for the drag coefficient outcomes, i.e., mean and oscillation amplitude quantities, are shown in Figure 6, where the red dots represent the DOE. The estimated parameters of all six models for six quantities of interest are summarized in Table A2 (Appendix B). Once a surrogate model is built with its estimated

parameters, the hydrodynamic pressure coefficients of an arbitrary design of AR and Re can be rapidly predicted without the need for an attempt at re-simulation.

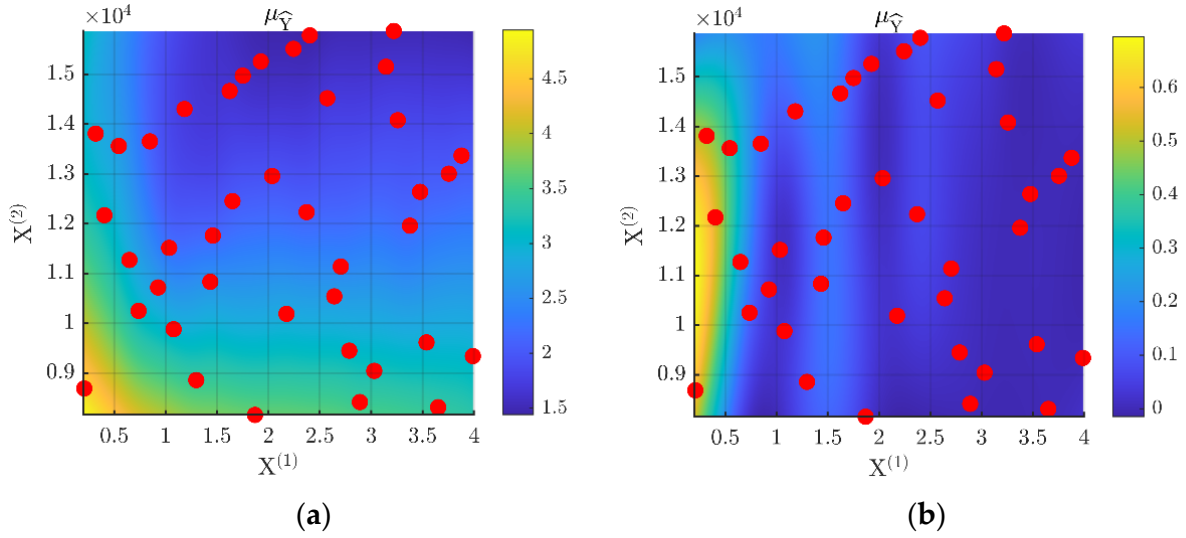


Figure 6. Examples of surrogate models of the drag coefficient in terms of $\mu_{\hat{Y}}$: (**a**) Mean and (**b**) Oscillation amplitude.

4.3. Validation of Surrogate Models with a Test Set

A test set of the hydrodynamic pressure coefficients is obtained from the previous work to validate the developed surrogate models. In particular, five different designs of the submerged body under free surface flows were numerically performed, which were uniformly composed by Re = 11,850 and five body shape ratios, AR = 0.25, 0.5, 1.0, 2.0, and 4.0. As a comparison, the plots of observed mean and oscillation magnitude together with those predicted for three hydrodynamic pressure coefficients are shown in Figure 7. Careful readers can see a good fit between the observed and predicted values both for the mean and oscillation quantities. To quantify the goodness of fit between the observed and predicted data, the mean square error (MSE) and coefficient of determination (R^2) are calculated and presented in Table 3. It can be re-confirmed that the surrogate models predict the mean and oscillation amplitude of the hydrodynamic pressure coefficients with a high degree of accuracy. Hence, the developed models are reliable in prediction and efficient in terms of computational effort.

Table 3. Error estimations of the observed and predicted hydrodynamic pressure coefficients in the case of Re = 11,850.

Quantity of Interest	MSE	R^2
C_D (Mean)	0.0033	0.9776
C_D (Oscil.)	0.0009	0.9868
C_L (Mean)	0.0060	0.9866
C_L (Oscil.)	0.0150	0.9955
C_M (Mean)	4.4359	0.9727
C_M (Oscil.)	8.8146	0.9030

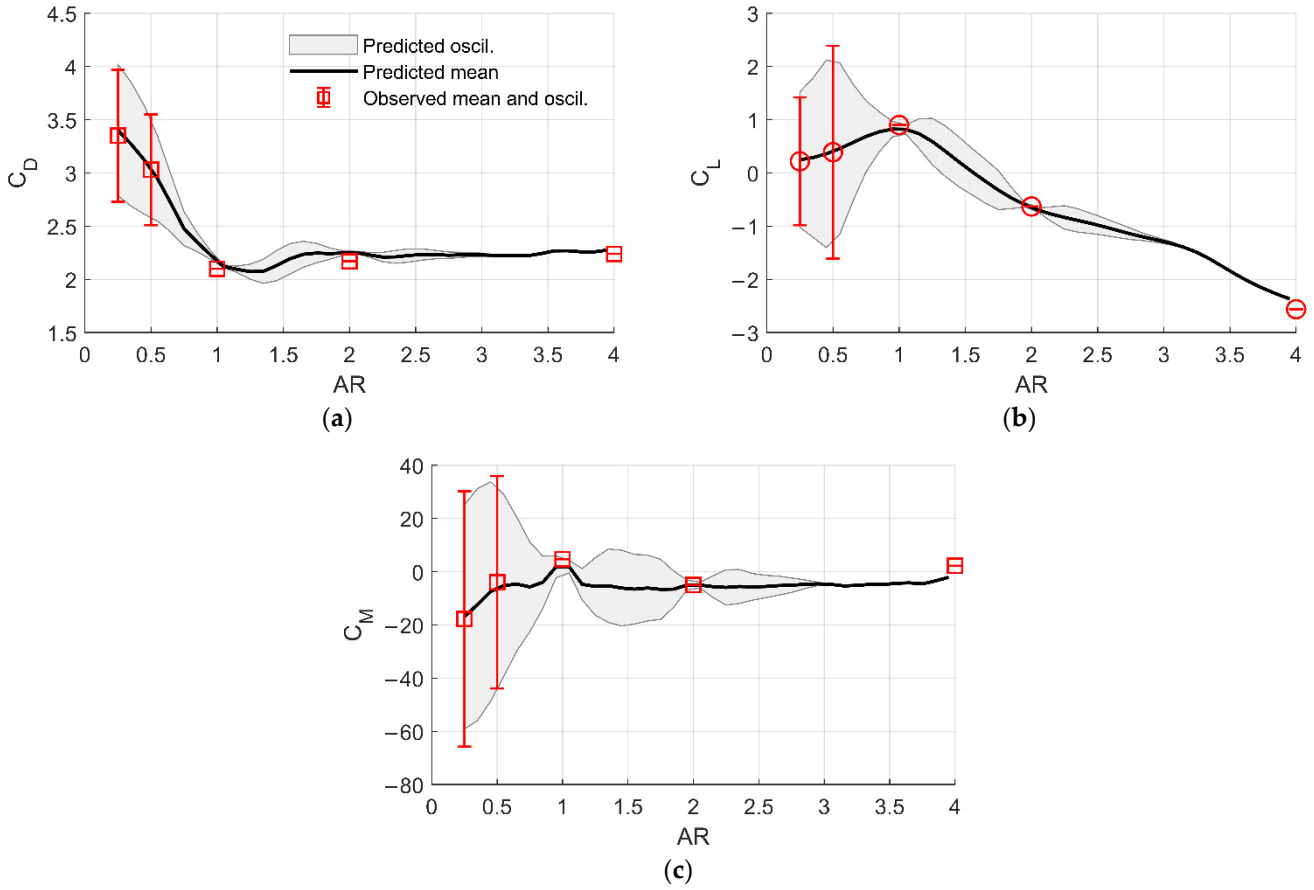


Figure 7. Comparisons of predicted hydrodynamic pressure coefficients with observed values from Nguyen et al. [9]: (**a**) Drag, (**b**) Lift, and (**c**) Moment coefficients.

In the practice design of potentially submerged civil structures, such as bridge decks, offshore platforms and hydrokinetic turbine blades, it is important to avoid unstable regions caused by the oscillation of the hydrodynamic forces. By considering a wide range of *Re* number and shape aspect ratio (Re = 6000–20,000, AR = 0.1–10), the unstable or oscillation regions are plotted based on the developed surrogate models for the three coefficients, as shown in Figure 8, in which the bar color represents the oscillation amplitude of the examined coefficients. It can be observed that the most unstable region appears with small aspect ratios of the submerged body. The amplitude of the oscillation mostly decreases with the increase of both Re and AR. In the cases of $AR > 2$, the oscillation amplitude significantly drops and almost equals zero. These observations are criteria for the dynamic instability assessment of existing submerged civil structures or for practice design of new ones under the free surface flow to avoid adverse effects of the dynamic impact.

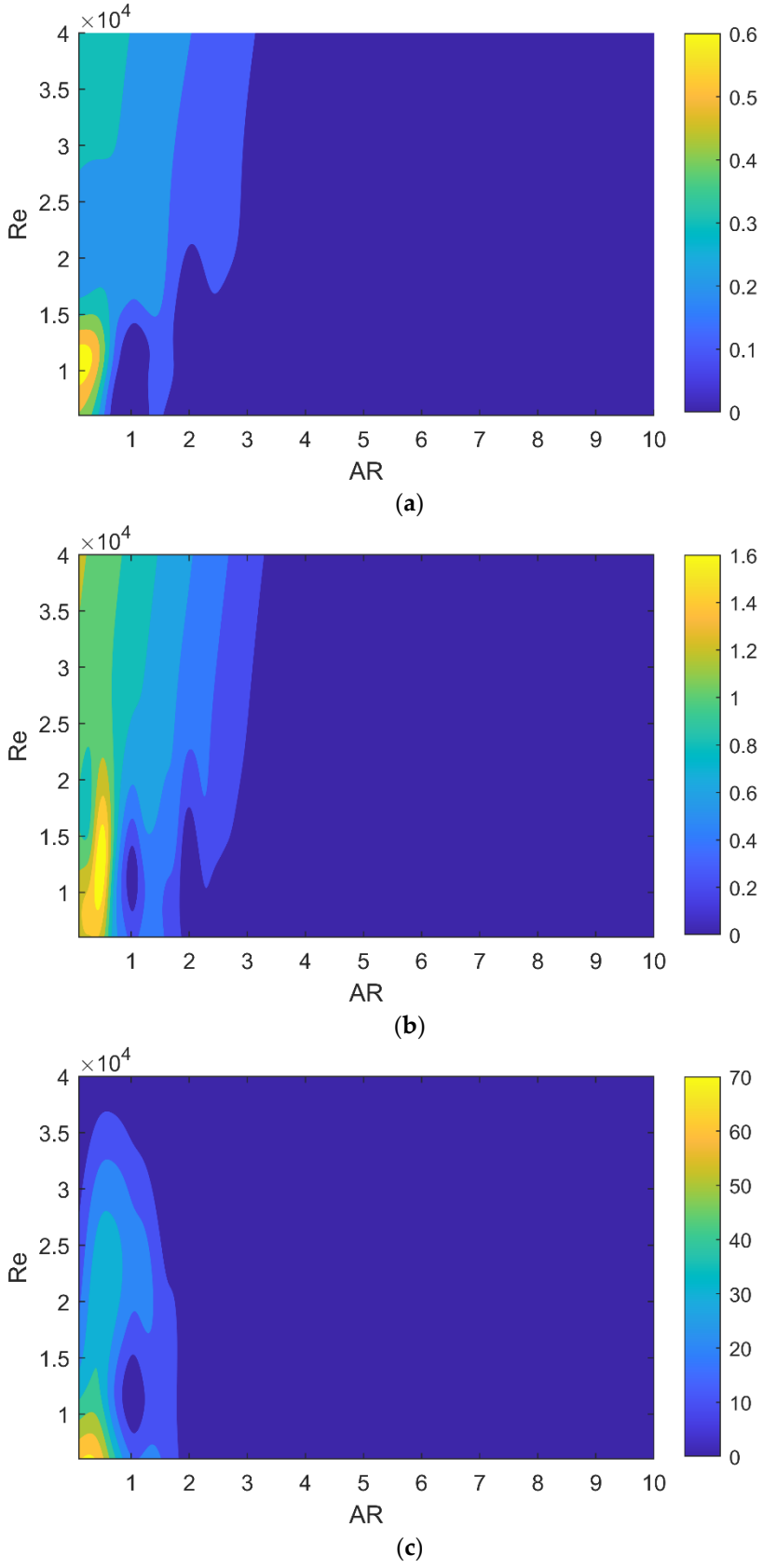


Figure 8. Predictions of unstable zones of hydrodynamic pressure coefficients: (**a**) Drag, (**b**) Lift, and (**c**) Moment coefficients.

5. Conclusions

This study aimed to develop a computationally efficient and accurate surrogate model to estimate hydrodynamic pressure coefficients on submerged bodies beneath the water surface. Using the LHS sampling method, several computational fluid dynamics analyses, based on a Navier-Stokes solver implemented with the shear stress transport turbulence model and the volume of fluid method, were performed to extract hydrodynamic pressure coefficients and their instability.

From the outcomes of the CFD analyses, a Gaussian process modeling-based surrogate model was trained to predict the hydrodynamic pressure coefficients around submerged bodies with a rectangular shape considering a range of shape ratio and Re number values.

As cross-validation for several testing surrogate models, the optimized model was found to be a combination of the 3rd degree polynomial and Matérn 3/2 correlation functions.

Since the surrogate models were developed, the hydrodynamic pressure coefficients were then predicted for a wide range of input parameters. The finding from the study highlighted the efficiency of the surrogate model in rapidly estimating the hydrodynamic pressure coefficients in place of complex and expensive CFD analyses.

By plotting unstable regions of the hydrodynamic pressure coefficients within the ranges of the shape ratio and *Re* number, it is concluded that the most unstable region appeared at small aspect ratios of the submerged body. In most of the cases, the oscillation amplitude significantly dropped with the increase of both AR and *Re* and reached almost zero with $AR > 2$.

The surrogate model in this study can be practically applied in rapidly assessing the hydrodynamic force and its instability effect on existing submerged civil structures, or in designing new structures, where a suitable shape ratio should be adopted to avoid flow-induced instability of hydrodynamic forces.

The present study can also be enabled and facilitate future sensitivity, fragility and reliability studies across a broad range of submerged bodies and flow conditions that are involved in civil structures under flood and wave flows.

Author Contributions: Conceptualization, V.M.N. and H.N.P.; methodology, V.M.N. and H.N.P.; software, H.N.P. and T.H.P.; validation, V.M.N., H.N.P. and T.H.P.; formal analysis, H.N.P. and T.H.P.; investigation, H.N.P. and T.H.P.; writing—original draft preparation, H.N.P. and T.H.P.; writing—review and editing, V.M.N., H.N.P. and T.H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by Funds for Science and Technology Development of the University of Danang under project number B2020-DN02-80.

Data Availability Statement: Not applicable.

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

Appendix A

Table A1. Observed response from CFD simulations for the initial training samples.

Input Variable		Output Hydrodynamic Pressure Coefficients					
AR	*Re*	C_D		C_L		C_M	
		Mean	**Oscillation Amplitude**	**Mean**	**Oscillation Amplitude**	**Mean**	**Oscillation Amplitude**
1.623	14,665.087	1.653	0.146	0.405	0.492	−3.511	12.653
3.476	12,640.045	2.087	0	−1.465	0	−3.957	0
2.403	15,783.894	1.451	0.087	0.260	0.305	−1.895	5.741

Table A1. *Cont.*

Input Variable		Output Hydrodynamic Pressure Coefficients					
AR	***Re***	C_D		C_L		C_M	
		Mean	**Oscillation Amplitude**	**Mean**	**Oscillation Amplitude**	**Mean**	**Oscillation Amplitude**
3.144	15,153.660	1.632	0.018	−0.273	0.084	−1.909	0
2.036	12,958.475	1.920	0.007	0.410	0.042	−3.720	1.029
2.890	8416.533	3.441	0	−2.661	0	−8.266	1.504
1.297	8857.737	3.429	0.102	−0.361	0.433	−9.072	15.586
0.208	8689.223	4.600	0.578	0.290	1.433	−29.590	57.009
2.704	11,139.096	2.458	0.030	−1.416	0.103	−5.916	3.124
1.435	10,833.029	2.569	0.139	−0.065	0.446	−7.160	14.355
2.244	15,512.939	1.487	0.055	0.283	0.333	−2.195	6.928
3.538	9615.309	2.959	0	−2.973	0	−6.677	0
1.926	15,258.790	1.521	0.029	0.382	0.145	−0.823	3.406
1.183	14,304.394	1.737	0.128	0.600	0.465	−1.907	11.747
1.751	14,974.758	1.589	0.098	0.396	0.464	−3.120	11.481
3.755	13,004.791	2.040	0	−1.717	0	−3.330	0
3.879	13,368.033	1.987	0	−1.747	0	−1.323	0
3.377	11,959.480	2.199	0	−1.608	0	−4.654	0
3.651	8313.764	3.372	0	−3.584	0	−7.782	0
0.927	10,719.160	2.701	0.040	0.650	0.190	−1.448	6.000
0.404	12,170.650	3.099	0.541	0.303	1.663	−9.008	42.218
0.846	13,654.136	2.122	0.142	0.597	0.472	−1.460	12.568
1.459	11,762.251	2.170	0.143	0.171	0.463	−6.291	14.129
0.647	11,271.221	2.905	0.262	0.575	1.080	−5.339	26.001
2.786	9449.139	3.077	0.012	−2.200	0.037	−7.400	2.408
1.869	8157.854	3.678	0.036	−1.774	0.158	−11.104	6.056
3.258	14,079.809	1.828	0.001	−0.660	0.025	−2.805	0
0.736	10,250.214	3.081	0.128	0.617	0.650	−7.770	20.596
3.220	15,872.356	1.509	0.016	−0.151	0.081	−1.194	0
1.077	9880.575	2.986	0.028	0.336	0.153	−3.508	4.934
2.641	10,540.904	2.681	0.031	−1.625	0.105	−6.511	3.687
1.650	12,453.679	2.040	0.122	0.010	0.427	−5.513	12.430
0.319	13,810.938	2.893	0.483	0.162	1.227	−11.899	39.283
1.034	11,514.888	2.253	0.013	0.775	0.071	2.268	2.172
3.030	9043.564	3.200	0	−2.554	0	−7.586	0.260
2.176	10,188.772	2.835	0.024	−1.431	0.146	−7.570	5.035
2.371	12,232.946	2.111	0.064	−0.759	0.223	−5.168	6.053
0.544	13,561.550	2.693	0.382	0.291	1.676	−3.352	34.293
2.572	14,520.981	1.699	0.063	−0.116	0.225	−2.922	4.254
3.989	9339.984	2.910	0	−3.380	0	−7.600	0

Appendix B

Table A2. Estimated parameters of the Gaussian process-based surrogate model.

Surrogate Model	β	σ^2	ℓ
C_D (Mean)	(2.120; −0.665; 0.150; 0.235; −0.003; 0.002; −0.138; −0.0690; 0.037)	0.009	(0.220; 0.9696)
C_D (Oscil.)	(0.445; 0.0807; −0.402; −0.027; −0.067; 0.0667; 0.003; 0.052; 0.034; −0.063)	3.661	(1.886; 8.554)
C_L (Mean)	(−1.531; 0.743; −0.521; −0.037; 0.152; 0.193; −0.061; −0.060; 0.031; −0.094)	58.317	(2.010; 9.945)
C_L (Oscil.)	(0.140; 0.070; −0.181; 0.014; 0.198; 0.010; −0.008; −0.101; 0.004; −0.035)	0.036	(0.137; 9.991)
C_M (Mean)	(−3.577; 2.452; −3.368; −0.934; −1.127; 0.049; 0.03; 2.114; 1.166; 0.033)	2.186	(0.148; 1.225)
C_M (Oscil.)	(3.223; 0.982; −3.512; 0.696; 6.151; 0.755; −0.183; −3.423; −0.919; −1.282)	9.040	(0.178; 1.785)

References

1. Palermo, D.; Nistor, I.; Saatcioglu, M.; Ghobarah, A. Impact and damage to structures during the 27 February 2010 Chile tsunami Canadian. *J. Civ. Eng.* **2013**, *40*, 750–758. [CrossRef]
2. Kim, H.; Sim, S.H.; Lee, J.; Lee, Y.J.; Kim, J.M. Flood fragility analysis for bridges with multiple failure modes. *Adv. Mech. Eng.* **2017**, *9*. [CrossRef]
3. Balomenos, G.P.; Padgett, J.E. Fragility analysis of pile-supported wharves and piers exposed to storm surge and waves. *J. Waterw. Port Coast. Ocean Eng.* **2018**, *144*, 04017046. [CrossRef]
4. Liu, T.L.; Guo, Z.M. Analysis of wave spectrum for submerged bodies moving near the free surface. *Ocean Eng.* **2013**, *58*, 239–251. [CrossRef]
5. Ren, H.; Fu, S.; Liu, C.; Zhang, M.; Xu, Y.; Deng, S. Hydrodynamic Forces of a Semi-Submerged Cylinder in an Oscillatory Flow. *Appl. Sci.* **2020**, *10*, 6404. [CrossRef]
6. Olaya, S.; Bourgeot, J.; Benbouzid, M.E.H. Hydrodynamic Coefficient Computation for a Partially Submerged Wave Energy Converter. *IEEE J. Ocean. Eng.* **2015**, *40*, 522–535. [CrossRef]
7. Lagrange, R.; Delaune, X.; Piteau, P.; Borsoi, L.; Antunes, J. A new analytical approach for modeling the added mass and hydrodynamic interaction of two cylinders subjected to large motions in a potential stagnant fluid. *J. Fluids Struct.* **2018**, *77*, 102–114. [CrossRef]
8. Chen, X.; Liang, H. Wavy properties and analytical modeling of free-surface flows in the development of the multi-domain method. *J. Hydrodyn. Ser. B* **2018**, *28*, 971–976. [CrossRef]
9. Nguyen, V.M.; Le, A.T.; Phan, H.N.; Nguyen, Q.K.; Chau, V.T.; Phan, T.H. On the behavior of nonlinear hydrodynamic pressure coefficients of a submerged cylinder beneath the water surface. *Vietnam. J. Mech.* **2021**, *43*, 371–387. [CrossRef]
10. Hoang, P.H.; Phan, H.N.; Nguyen, D.T.; Paolacci, F. Kriging Metamodel-Based Seismic Fragility Analysis of Single-Bent Reinforced Concrete Highway Bridges. *Buildings* **2021**, *11*, 238. [CrossRef]
11. Phan, H.N.; Paolacci, F.; Di Filippo, R.; Bursi, O.S. Seismic vulnerability of above-ground storage tanks with unanchored support conditions for Na-tech risks based on Gaussian process regression. *Bull. Earthq. Eng.* **2020**, *18*, 6883–6906. [CrossRef]
12. Bhatti, M.M.; Marin, M.; Zeeshan, A.; Abdelsalam, S.I. Recent trends in computational fluid dynamics. *Front. Phys* **2020**, *8*, 593111. [CrossRef]
13. Ahmadi, M.A.; Chen, Z. Machine learning models to predict bottom hole pressure in multi-phase flow in vertical oil production wells. *Can. J. Chem. Eng.* **2019**, *97*, 2928–2940. [CrossRef]
14. Guillén-Rondon, P.; Robinson, M.D.; Torres, C.; Pereya, E. Support vector machine application for multiphase flow pattern prediction. In Proceedings of the Workshop on Data Mining for Geophysics and Geology (DMG2), San Diego, CA, USA, 5 May 2018.
15. Ganti, H.; Khare, P. Data-driven surrogate modeling of multiphase flows using machine learning techniques. *Comput. Fluids* **2020**, *211*, 104626. [CrossRef]
16. Quinonero-Candela, J.; Rasmussen, C.E.; Williams, C.K. Approximation Methods for Gaussian Process Regression. In *Large-Scale Kernel Machines*; MIT Press: Cambridge, MA, USA, 2007.
17. Duy, T.N.; Nguyen, V.T.; Phan, T.H.; Park, W.G. An enhancement of coupling method for interface computations in incompressible two-phase flows. *Comput. Fluids* **2021**, *214*, 104763. [CrossRef]
18. Phan, H.N.; Lee, J.H. Flood Impact Pressure Analysis of Vertical Wall Structures using PLICVOF Method with Lagrangian Advection Algorithm. *J. Comput. Struct. Eng. Inst. Korea* **2010**, *23*, 675–682.
19. Nguyen, V.T.; Park, W.G. A free surface flow solver for complex three-dimensional water impact problems based on the VOF method International. *J. Numer. Methods Fluids* **2016**, *82*, 3–34. [CrossRef]
20. ANSYS Inc. *ANSYS Fluent User's Guide 2019R1*; ANSYS Inc.: Canonsburg, PA, USA, 2019.
21. Santner, T.J.; Williams, B.J.; Notz, W.I. *The Design and Analysis of Computer Experiments*; Springer: New York, NY, USA, 2003.

22. Stefano, M.; Bruno, S. UQLab: A Framework for Uncertainty Quantification in MATLAB. In Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM 2014), Liverpool, UK, 13–16 July 2014. [CrossRef]
23. Chu, C.R.; Lin, Y.A.; Wu, T.R.; Wang, C.Y. Hydrodynamic force of a circular cylinder close to the water surface. *Comput. Fluids* **2018**, *171*, 154–165. [CrossRef]
24. Mckay, M.D.; Beckman, R.J.; Conover, W.J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **1979**, *21*, 239–245. [CrossRef]
25. Pebesma, E.J.; Heuvelink, G.B. Latin hypercube sampling of Gaussian random fields. *Technometrics* **1999**, *41*, 303–312. [CrossRef]
26. Iooss, B.; Boussouf, L.; Marrel, A.; Feuillard, V. Numerical study of algorithms for metamodel construction and validation. In *Safety, Reliability and Risk Analysis: Theory, Methods and Applications*; Martorell, S., Soares, C.G., Barnett, J., Eds.; Taylor & Francis Group: London, UK, 2009.