ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021-2022), September 13-17, 2022, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Convolution-fed Gaussian Process with active learning for probabilistic power prediction of large-scale wave farm

Min Li[a], Gaofeng Jia[a], Hussam Mahmoud[a], Yi-Hsiang Yu[b], and Nathan Tom[c]

[a]*Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, USA. E-mail: Min.Li@colostate.edu; Gaofeng.Jia@colostate.edu; Hussam.Mahmoud@colostate.edu.*

[b]*National Yang Ming Chiao Tung University, Hsinchu, Taiwan. E-mail: yyu@nycu.edu.tw.*

[c]*National Renewable Energy Laboratory, Golden, CO, USA. E-mail: nathan.tom@nrel.gov.*

ABSTRACT: Predicting the total power generation of a wave farm under actual operational conditions (e.g. random sea) is an indispensable task in the layout design of wave energy converters (WECs). However, such prediction usually entails significant computational cost, especially for large arrays of WECs. To address the challenge, a convolution-fed Gaussian process (CFGP) model with active learning is proposed to efficiently predict the probabilistic power generation of large-scale wave farms. The layout configurations of WECs are first transformed into binary images to automatically consider the permutation invariance property of WECs in the array. Then the images are fed into a convolutional neural network to extract the low-dimensional features. These top-level features are then used to fit a GP model to predict the probabilistic power generation. Furthermore, an active learning scheme is developed to adaptively enrich the training data to improve the prediction accuracy of the CFGP model. The efficiency and accuracy of the proposed CFGP model is demonstrated through an application to arrays with up to 10 WECs.

KEYWORDS: Gaussian Process; convolutional neural network; active learning; wave energy

1 INTRODUCTION

The deployment of wave energy converters (WECs) in large arrays in the form of wave farms offers great prospects for harnessing renewable wave energy. To improve the power generation of wave farms, the layout of WECs needs to be carefully designed so that the hydrodynamic interactions between WECs can be positively exploited. Thus, estimating the power generation of WEC arrays has drawn extensive attention in this field.

However, most of the relevant research on estimating array power generation has focused on applications under deterministic sea states, which assume that waves have a predominant period and do not accurately represent the actual operational conditions for WECs. To establish layouts with robust performance (e.g. expected power production), the uncertainties in the random wave characteristics need to be incorporated (Borgarino et al., 2012; Balitsky et al., 2014; Jia et al., 2015). Meanwhile, in order to calculate the power generation of the array, hydrodynamic interactions between WECs need to be modeled. To understand the hydrodynamic performances of WEC arrays, different models have been proposed by Folley et al. (2012). Among different methods, the multiple-scattering method (Mavrakos, 1991) attempts to converge to exact solutions and strikes a good balance between computa-

tional complexity and accuracy. However, if high accuracy is required, the computational efforts may still be burdensome, especially when the number of WECs is large. Besides multiple-scattering, other numerical models have also been developed. Recently, the open-source WEC-Sim tool (Yu et al., 2014; Tom et al., 2015) was developed to facilitate time domain modeling of a WEC. WEC-Sim has the ability to model devices that comprise rigid bodies, power take-off systems, and mooring systems. Besides individual WECs, WEC-Sim has the capability of modeling arrays (Mankle et al., 2019). Regardless of the models used, additional challenges also arise when dealing with problems, such as uncertainty quantification and optimization, which typically require a large number of model evaluations.

To address the above challenges, a convolution-fed Gaussian process (CFGP) model is proposed to efficiently predict the power generation of WEC arrays considering the uncertainties of the random sea state. The layout configurations of WECs are first transformed into binary images to automatically consider the permutation invariance property of WECs in the array. Then, the images are fed into a convolutional neural network (CNN) to extract the top-level features (corresponding to low-dimensional latent representations), which are then used to fit a GP model. Furthermore, the local variance information provided by the GP is exploited to develop an active learning scheme to adaptively enrich the training data and improve the prediction accuracy. The proposed model is applied to estimate the probabilistic power generation of a wave farm with different numbers of WECs, and its efficiency and effectiveness are investigated.

2 PROBABILISTIC POWER GENERATION OF WEC ARRAYS IN A RANDOM SEA STATE

Among different types of WECs, this paper mainly focuses on cylindrical heave point absorbers, which are heaving floating resonant buoys connected to a power take-off moored to the seafloor. Our ultimate goal is to calculate the power generation from an array of such devices in a random sea (typically characterized by a chosen wave spectrum), and propagate the uncertainties of the random sea characteristics (i.e. the model parameters of the selected spectrum) to the calculation of the power generation.

2.1 *Power generation of WEC arrays in a random sea state*

The calculation of the power generation in a random sea state for an array of such devices is based on the modified method from Scruggs et al. (2013), which is extended here to consider multiple converters oscillating only in the heave direction. The random sea state is first characterized by a stochastic stationary excitation, which will be discussed in detail in Section 2.1.2. For a given WEC array, the (average) power generation in the random sea state is then estimated by a frequency domain numerical model

$$P_{gen}^{st} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_p(\omega)\mathrm{d}\omega \tag{1}$$

where ω is the frequency and $S_p(\omega)$ is the spectral density of generated power in the frequency domain, which is given by

$$\begin{aligned} S_p(\omega) =& \mathbf{G}_a^H[\mathbf{I}+\mathbf{G}_i\mathbf{Y}]^{-H}[Re\{\mathbf{Y}\}-R_c\mathbf{Y}^H\mathbf{Y}] \\ & [\mathbf{I}+\mathbf{G}_i\mathbf{Y}]\mathbf{G}_a S_a \end{aligned} \tag{2}$$

where $(.)^H$ denotes the Hermittian adjoint (i.e. complex-conjugate transpose), $Re\{.\}$ denotes the real component, and R_c is the stator resistance of the generators. The functional relationships on ω in the right-hand side of Eq. (2) were suppressed for simplicity. $\mathbf{G}_a = \mathbf{G}_a(j\omega)$ and $\mathbf{G}_i = \mathbf{G}_i(j\omega)$ are the transfer functions for wave (input) to energy (output) and need to be calculated based on the hydrodynamic coefficients of the WECs. $\mathbf{Y} = \mathbf{Y}(j\omega)$ is the frequency domain representation for the controller, and in order to maximize $S_p(\omega)$ (i.e. maximize P_{gen}^{st} under given S_a), the optimal controller is given by (Scruggs et al., 2013)

$$\mathbf{Y}(j\omega) = \left[\mathbf{G}_i^T(-j\omega)+2R_c\mathbf{I}\right]^{-1} \tag{3}$$

Plugging Eq. (3) into Eq. (2) yields $S_p(\omega)$. $S_a = S_a(\omega)$ is the spectral density of the stationary stochastic process used to characterize the random sea state.

In the end, P_{gen}^{st} can be calculated through standard numerical integration of the scalar integral in Eq. (1). This first requires partitioning the frequency range of interest (where $S_a(\omega)$ has nonzero contribution) into appropriate frequencies $\{\omega_i; i=1,\ldots,N_\omega\}$. Then, the hydrodynamic coefficients for each frequency ω_i need to be calculated to obtain transfer functions $\mathbf{G}_a$ and $\mathbf{G}_i$. For given $S_a(\omega_i)$, $S_p(\omega_i)$ can then be calculated. Finally, P_{gen}^{st} can be evaluated through the contribution from each frequency.

Therefore, in order to calculate the power production of WEC arrays in a random sea, hydrodynamic modeling of WEC arrays and characterization of the random sea state are the key steps. These will be discussed in Sections 2.1.1 and 2.1.2, respectively.

2.1.1 *Hydrodynamic modeling*

Assume the considered WEC array has N identical buoys floating in the water, and a two-dimensional Cartesian coordinate system is defined to express the locations of the WECs. The WECs in the array are under the incident wave propagating along the positive x-axis, and the corresponding incident angle is an arbitrary number β. Figure 1 shows an example layout of the array with N WECs. Without loss of generality, the leftmost buoy of the layout is assumed to be located at the origin of the coordinate system. The layout can then be characterized by the locations of the remaining $N-1$ buoys, i.e. $\mathbf{x} = [\bar{x}_1\ \bar{x}_2\ \ldots\ \bar{x}_{N-1}\ \bar{y}_1\ \bar{y}_2\ \ldots\ \bar{y}_{N-1}] \in X \subset \mathbb{R}^{2(N-1)}$, where the pair $(\bar{x}_i, \bar{y}_i)$ represents the coordinates for the ith WEC and X denotes the admissible layout space. For convenience, the layouts of the WEC arrays are transformed into equivalent layouts with all buoys on the right half plane with $\beta = 0$ beforehand, which can be established by rotating the coordinate system and adjusting the origin.

The multiple-scattering method is adopted in this paper to calculate the hydrodynamic interaction within the array of WECs because of its versatility in achieving enhanced accuracy. This method relies on estimating single-body hydrodynamic characteristics, and describes the interaction between the different bodies by superimposing the incident wave potential and various orders of propagating and evanescent modes that are scattered and radiated by all the devices in the array, so that an accurate representation of the total wave field around each device can be obtained. Ultimately, for each considered frequency ω_i, the multiple-scattering method provides the forces exerted by the incident wave for each body. It also provides the added mass and damping coefficients exerted on each degree of freedom of body p by unit-displacement oscillation of each degree of freedom of body q. These hydrodynamic characteristics are sufficient for deriving the transfer function between wave amplitude and WEC velocities for the entire array, which can be further used to calculate the power generation of the array.

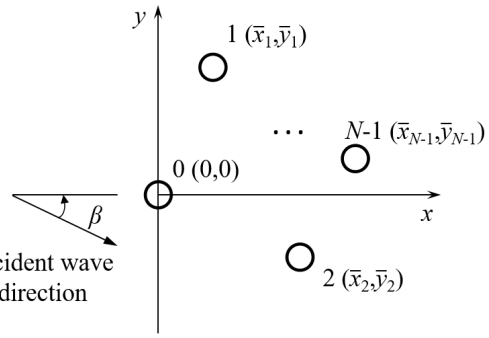

Figure 1. Layout of an array of N WECs in the Cartesian coordinate.

2.1.2 *Random sea state characterization*

The random sea state is characterized in this paper as a stationary stochastic process with spectral density $S_a(\omega)$, described through the standard JONSWAP power spectrum (Faltinsen, 1993). $S_a(\omega)$ is parameterized by its mean wave period T_p, significant wave height H_s, and sharpness factor γ, as

$$S_a(\omega) = 310\pi \frac{H_s^2}{T_p^4 \omega^5} \exp\left(-\frac{944}{T_p^4 \omega^4}\right) \gamma^\eta \quad (4)$$

where

$$\eta = \exp\left[-\left(\frac{0.191\omega T_p - 1}{\sqrt{2}\phi}\right)^2\right];$$

$$\phi = \begin{cases} 0.07; & \omega T_p \le 5.24 \\ 0.09; & \omega T_p > 5.24 \end{cases}$$

The sharpness factor is typically constrained to be between 1 and 7 with γ=1 describing a fully developed sea.

2.2 *Expected power generation of WEC arrays in a random sea state*

The (average) power generation P_{gen}^{st} under a stochastic stationary excitation calculated above is under known characteristics for the sea spectrum (i.e. given T_p, H_s, and γ for the JONSWAP spectrum). However, previous works show that there is significant variability for the parameters of the sea spectrum (Balitsky et al., 2014), which might have an important impact on the array power generation, and potentially on the layout design. Uncertainty propagation can be used to account for the variability. Let $\theta = [T_p \; H_s \; \gamma]$ lying in $\Theta \subset \mathbb{R}^{n_\theta}$ be the vector of uncertain model parameters, where Θ denotes the set of their possible values. A probability density function $p(\theta)$, which incorporates our available knowledge about the excitation, is assigned to characterize the uncertainty in these parameters. The expected power generation considering all the uncertainties in the stationary sea excitation is then expressed as

$$\begin{aligned} P_{gen}(\mathbf{x}) &= E_{p(\theta)}[P_{gen}^{st}(\mathbf{x},\theta)] \\ &= \int_{\Theta} P_{gen}^{st}(\mathbf{x},\theta)p(\theta)\mathrm{d}\theta \end{aligned} \tag{5}$$

where $E_{p(\theta)}[\cdot]$ denotes expectation under probability model $p(\theta)$.

The above integral can be calculated through general stochastic simulation (e.g. Monte Carlo simulation), especially when the dimension is high, or through numerical integration when the dimension is low (e.g. 2 or 3). To reduce the number of dimensions, the proportionality of P_{gen}^{st} to H_s^2 is considered based on Eqs. (1)−(4). The power generation P_{gen}^{st} is then given by $P_{gen}^{st}(\mathbf{x},\theta) = H_s^2 P_{gen}^{st}(\mathbf{x},\underline{\theta}|H_s=1)$, where $\underline{\theta} = [T_p \; \gamma]$. However, even with the computational tricks mentioned above, the computational cost can be expensive. This is mainly because the hydrodynamic modeling (i.e. modeling hydrodynamic interactions between WECs) takes a lot of computational time, especially when the orders of interaction for the multiple-scattering approach are set to be high (corresponding to more accurate modeling) and when the number of WECs is large. The expensive model for calculating the probabilistic power generation mentioned above is referred to as the high-fidelity model hereafter, which calculates $P_{gen}(\mathbf{x})$ for given $\mathbf{x}$.

3 CONVOLUTION-FED GAUSSIAN PROCESS WITH ACTIVE LEARNING FOR POWER PREDICTION

3.1 *Overall idea*

To alleviate the computational burden in calculating the probabilistic power production of WEC arrays, surrogate modeling (such as GP) can be used. However, direct application of GP in this context may face the challenge of high-dimensional input (e.g. corresponding to the coordinates of all WECs). Moreover, a GP model that takes the coordinates of WECs as inputs ignores the physical characteristics of the input/output relationship, i.e. the power generation of the array is permutation invariant with respect to the ordering of the WECs. Directly taking the coordinates as inputs will significantly enlarge the design space. This may significantly impact the prediction accuracy of the GP model (under a given number of training data).

To address this challenge, a CFGP model is proposed to efficiently predict the probabilistic power production of WEC arrays. The wave farm domain is first discretized, and different layouts of WECs are transformed into images. The transformed images can automatically incorporate the physical knowledge (i.e. permutation invariance property) of the problem, because representing WECs as pixels in the images does not specify the ordering of the WECs. Then, the images are fed into a CNN to extract the top-level features, which correspond to the low-dimensional latent representations. The extracted low-dimensional features are then used to fit a GP model to predict the response. Further, the local variance information provided by the GP model is used to develop infill criteria to facilitate the

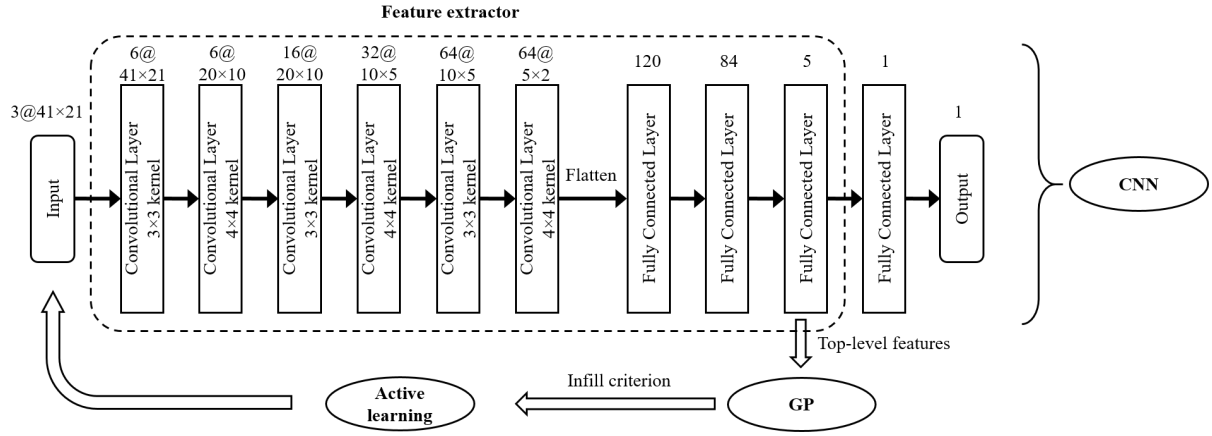

Figure 2. Illustration of the CFGP model with active learning and network architectures.

active learning of the CFGP model. This active learning scheme adaptively enriches the training data to improve the prediction accuracy. The proposed CFGP model with active learning is illustrated in Figure 2. It ultimately will provide an efficient approximation for the output of the high-fidelity model.

3.2 *Convolution-fed Gaussian process*

The original input of the high-fidelity model is a location vector defined by the coordinates of the WECs, i.e. $\mathbf{x} = [\bar{x}_1\ \bar{x}_2\ \ldots\ \bar{x}_{N-1}\ \bar{y}_1\ \bar{y}_2\ \ldots\ \bar{y}_{N-1}]$. To take advantage of the image processing capability of CNN, the input $\mathbf{x}$ is first transformed into an image. More specifically, the wave farm domain is first discretized into a grid, and the numbers of pixels along the x-axis and y-axis are N_x and N_y, respectively. Each node of the grid is assigned with a value to represent the deployment of the WEC (e.g. 1 represents that a WEC is deployed whereas 0 represents that a WEC is not deployed). With such discretization, different layouts of WECs can be transformed into images with the pixel values corresponding to the node values of the grids. Figure 3 demonstrates the transformation of a layout of 6 WECs into a binary image.

It is worth noting that discretizing the wave farm domain into a grid and deploying the devices at the nodes will reduce the design domain due to the discretization. For layout design or optimization problems, the grid size is suggested to be small so that potential optimal layouts can be covered. For power prediction problems, if the WECs of a real layout case are not at the nodes (even if the grid size is small), the WECs can be assigned to the nearest nodes to approximate the real layout. The transformed input is then denoted $\mathbf{x}_{img}$, which is a binary matrix with elements equal to 0 or 1. For the output y, it is defined as the expected power generation (i.e. P_{gen}).

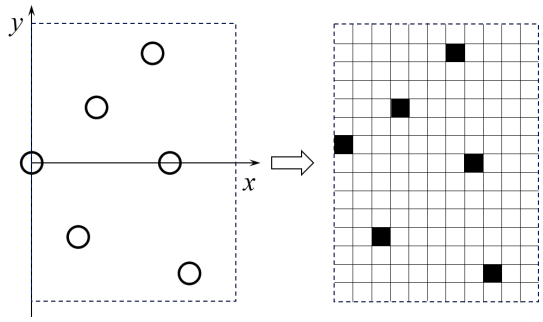

Figure 3. Transformation of a layout of 6 WECs into the binary image (Note: dashed rectangle represents the wave farm region).

In the framework of the CFGP model, a CNN is first trained as a feature extractor. The architecture of the CNN used in this paper is shown in Figure 2 which will be explained in detail later. It takes $\mathbf{x}_{img}$ (i.e. images) as the input and passes it through a sequence of layers (including convolutional layers and fully connected layers), and in the end outputs the prediction $\hat{y}_{CNN}(\mathbf{x}_{img})$ of the response $y(\mathbf{x}_{img})$. In this problem, the CNN

can be divided into two parts: (i) a feature extractor that consists of a sequence of convolutional layers and fully connected layers, (ii) a regressor defined by the last fully connected layer. The loss function of the CNN is selected as the Log-Cosh Loss

$$L = \sum_{h=1}^{n} \log[\cosh(\hat{y}_{CNN}(\mathbf{x}_{img}^h) - y(\mathbf{x}_{img}^h))] \quad (6)$$

where n is the total number of the training data, $\hat{y}_{CNN}$ is the prediction of the response y from the CNN, and $\mathbf{x}_{img}^h$ is the hth input. To improve the accuracy of the trained CNN, the training set is augmented by flipping the images with respect to the x-axis because the output (i.e. expected power of the array) is symmetric with respect to the x-axis in this case.

After the calculation of the feature extractor, the top-level features (i.e. low-dimensional latent representations) of the input $\mathbf{x}_{img}$ can be automatically extracted. Then, the low-dimensional features, denoted as $\mathbf{z}$, are used to train a GP model to predict the final response y. GP model is selected for this implementation because it provides not only the mean prediction but also the local uncertainty information, which can be useful for guiding the selection of training data. For a system model with input $\mathbf{z}$, a GP model $\hat{y}(\mathbf{z})$ can be trained to approximate the deterministic system response $y(\mathbf{z}) \in \mathbb{R}$. The predictive distribution conditioned on the training data $\mathscr{D}$ is given by

$$\hat{y}(\mathbf{z}^*)|\mathscr{D} \sim \mathscr{N}(\mathbf{K}^*[\mathbf{K}+\sigma^2\mathbf{I}]^{-1}\mathbf{y}, \mathbf{K}^{**} - \mathbf{K}^*[\mathbf{K}+\sigma^2\mathbf{I}]^{-1}\mathbf{K}^* + \sigma^2) \quad (7)$$

where $\mathbf{z}^*$ is a new input point, $\mathscr{N}(\cdot)$ represents the normal distribution, $\mathbf{y}$ is the observation vector, σ^2 is the variance of the measurement error and I is a identity matrix. $\mathbf{K}$, $\mathbf{K}^*$, and $\mathbf{K}^{**}$ represent the covariance functions (i) between the training data, (ii) between $\mathbf{z}^*$ and the training data, and (iii) between $\mathbf{z}^*$, respectively. Through the proper tuning of the covariance functions and the corresponding hyperparameters (e.g. lengthscale and variance), GP model can approximate very complex functions. The optimal selection of the hyperparameters can be based on the Maximum Likelihood Estimation (MLE) principle, where the likelihood is defined as the probability of the observations.

The development of the CFGP model requires, the creation of a database of n observations $\mathbf{y}$ based on the high-fidelity model, corresponding to evaluations of the response vector $\{y^h; h = 1,\ldots,n\}$ for different inputs $\{\mathbf{x}_{img}^h; h = 1,\ldots,n\}$ in the training set. These inputs can be selected using some space filling technique such as Latin Hypercube Sampling (LHS), perhaps augmented through an adaptive refinement to improve the accuracy in target regions (Zhang et al., 2017; Kyprioti et al., 2020). For the latter, as mentioned earlier, the variance of GP prediction can guide the design of experiments (i.e. adaptive sampling of the training data) and one example is to use expected improvement (EI) as infill criterion to explore design regions with high predictive uncertainty. Note that the CNN and the GP share the same training set in this implementation.

3.3 *Active learning*

In order to build a CFGP model accurately with as few samples as possible, active learning schemes can be used. More specifically, new samples are selected in an iterative manner based on the information extracted from the previously trained model. In this manner, the new samples typically are drawn from locations of interest in the design space (e.g. locations with high uncertainty).

A key step of active learning is to select an appropriate infill criterion to balance the local exploitation as well as the global exploration. Expected Improvement for Global Fit (EIGF) (Lam, 2008) is adopted in this paper. The reason is that (i) it aims to accurately estimate the whole design region, (ii) it can make the most of the variance information extracted from the model trained in previous iterations, and (iii) it is more computationally efficient than most

of the other infill criteria. In this paper, EIGF is calculated in the feature space, and at the new point $\mathbf{z}^*$ it is given by

$$EIGF(\mathbf{z}^*) = (\hat{y}(\mathbf{z}^*) - y(\mathbf{z}^{closest}))^2 + \sigma^2_{\hat{y}}(\mathbf{z}^*) \tag{8}$$

where $\mathbf{z}^{closest}$ is the point in the training set which is the closest neighbour of $\mathbf{z}^*$, and $\sigma^2_{\hat{y}}$ is the predictive variance of the GP model. By solving the optimization problem (i.e. maximize EIGF), one or multiple new sample point can be selected and added to the training set. It should be noted that the feature space is not continuous in this problem since not every point in the feature space corresponds to a valid layout configuration. Therefore, instead of directly solving the optimization problem, an alternative option is used. It first generates a large reference set which covers most of the possible layout configurations and then optimizes EIGF over the reference set.

To implement the active learning, we first prepare a reference set which includes a large amount of inputs (i.e. layout configurations), denoted $\mathbf{X}_{ref}$. At the same time, we generate a small number of training data $\mathbf{X}_0$, and evaluate the high-fidelity model to obtain corresponding response $\mathbf{y}_0$. The sizes of the reference set and the training set are denoted n_{ref} and n_0, respectively. The CFGP model is initially trained based on the the pair $\{\mathbf{X}_0, \mathbf{y}_0\}$, and the model is then used to predict the response for the candidate samples in the reference set (denoted as $\hat{\mathbf{y}}_{ref}$). Based on Eq. (8), EIGFs are calculated for the candidate samples, and the ones with top n_{add} highest EIGFs are selected. The selected samples (denoted $\mathbf{X}_{add}$) are removed from the reference set and added to the training set, and the corresponding response of the selected samples (denoted $\mathbf{y}_{add}$) is then calculated with the high-fidelity model. Finally, the CFGP is updated with the new training set. This adaptive addition of new sample points is carried out iteratively until some convergence criterion is reached, e.g. setting an upper limit on the number of iterations, or total number of model evaluations.

4 EXAMPLE

The performance of the proposed CFGP with active learning method is demonstrated by applying it to predict the expected power generation of a wave farm with different numbers of WECs. The wave farm is in a rectangular domain with 200 m along the x-axis and 400 m along the y-axis, and the water depth is 60 m. The WECs are identical cylinders oscillating in the heave direction only and are the same as the one considered in (Jia et al., 2015). The radius of each buoy is r_b=3 m and its mass is m_b=1.8e5 kg, corresponding to a draft of D_r=6.37 m. The transducer damping is c_g=2 kN/m/s, the stator resistance is R_c=1 Ohms, and the short circuit damping is C_e=10 kN/m/s. For the high-fidelity numerical model, when calculating the hydrodynamic coefficients, the order of interaction is set as 5 for the multiple-scattering approach, whereas the eigenfunction series are truncated at 5 and 40 for the main fluid and the fluid below the cylinder, respectively. These values are selected through a convergence study of the numerical model. The wave climate statistics are based on the recommendations in the literature. The significant wave height H_s is assumed to follow a Weibull distribution with a shape parameter of 1.53 and a scale parameter of 3.1. The mean wave period T_p follows a conditional on H_s lognormal distribution with median and coefficient of variation given by

$$\begin{aligned} T_{p,med} &= \exp(1.78 + 0.288 {H_s}^{0.474}) \\ T_{p,cov} &= \sqrt{0.001 + 0.097 \exp(-0.255 H_s)} \end{aligned} \tag{9}$$

The sharpness factor γ is assumed to follow a beta distribution with shape parameters 1.5 and 2 constrained between 1 and 7. The uncertain model parameter θ corresponds to $\theta = [H_s, T_p, \gamma]$, which leads to $\underline{\theta} = [T_p, \gamma]$. Because $\underline{\theta}$ has only two components in this implementation, standard numerical quadrature is selected for estimating the probabilistic performance with respect to them.

4.1 *Implementation details*

For the discretization of the design domain, $N_x = 20$ and $N_y = 40$ are selected, which leads to a grid with 21×41 nodes. The nodes

at which WECs are deployed are assigned a value of 1; other nodes are assigned a value of 0. The grid is then transformed into a binary image where the pixels values correspond to the original node values. For the initial number of the training data for the CFGP, $n_0 = 200$ is selected, and the samples are generated using LHS. To establish the reference set, $n_{ref} = 10,000$ candidate samples are generated by LHS. For the active learning, in each iteration, a total number of 200 samples is selected from the reference set (i.e. $n_{add} = 200$). The stopping criterion of active learning is selected as when the maximum number of iterations reaches 10.

The architecture of the CNN is illustrated in Figure 2. The input fed into the network is the image with 3 channels and 21×41 pixels, and the output dimension after each layer is shown on top of each layer representation. The dimension of the top-level features is selected as 5 for all cases. As to the augmentation of the training data, augmentation on the fly is adopted, and a random flip across the x-axis is applied. More specifically, in each epoch, each training data has the probability of 0.5 to be flipped and the probability of 0.5 to remain unchanged. The CNN is trained using the Adam version of minibatch stochastic gradient descent with a minibatch size of 32, and the learning rate is 0.001. The loss function, as mentioned earlier, is selected as the Log-Cosh Loss. Pytorch is used to implement the networks, and the number of epochs is set as 500. For the GP part, a radial basis function kernel is selected and maximum likelihood estimate (MLE) is used to optimize the hyperparameters.

To validate the accuracy of the established CFGP with active learning, validation metrics are calculated over a testing set in each iteration. The test size is set as $n_{test} = 1000$, and the testing data are also generated by LHS. Here, we use the coefficient of determination R^2 and the mean absolute percentage error $MAPE$ as the validation metric. Large R^2 values or small $MAPE$ values indicate that the trained model has good accuracy.

4.2 *Results and discussions*

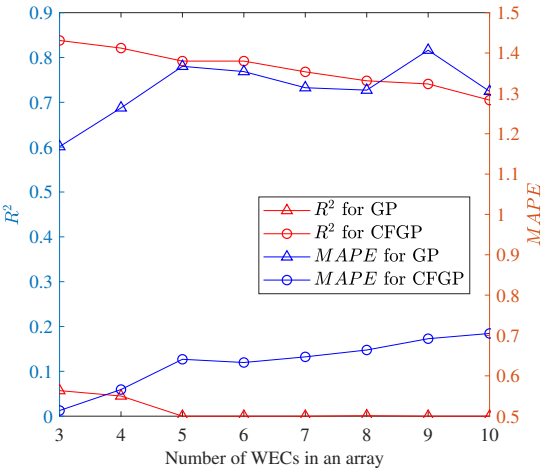

Figure 4. Comparison of the model accuracy between GP and CFGP.

Figure 4 reports the prediction accuracy metrics of the CFGP model with active learning over the 1000 testing layout configurations. For all the cases, R^2 is over 0.70 and $MAPE$ are no more than 0.7%, which indicates good prediction accuracy. As the number of WECs increases, the prediction accuracy decreases. This is expected because the transformed images corresponding to wave farms with more WECs have more complex features than those with fewer WECs, and therefore it is more difficult for the surrogate model to effectively capture the features with the same amount of training data.

In order to compare the performance of the proposed CFGP model with the basic GP model, we also build a GP with the same total number of training data and the same type of kernel function, and predict the output over the same set of testing data. The input of the GP model is selected as the coordinates of the WECs, i.e. $\mathbf{x} = [\bar{x}_1\ \bar{x}_2\ \ldots\ \bar{x}_{N-1}\ \bar{y}_1\ \bar{y}_2\ \ldots\ \bar{y}_{N-1}]$. The accuracy metrics are also shown in Figure 4. It can be observed that with only 2000 training data, the established GP model shows significantly worse performance than the CFGP model in terms of both R^2 and $MAPE$. Note that R^2 values for the GP model are nearly 0 as observed from Figure 4. By examining the results, we find that the predictions significantly deviate from the ground truth values in the domain with a

small amount of training data. This indicates bad generality of the GP model. The reasons for low prediction accuracy of the GP model are twofold: (i) the GP model takes the coordinates as the inputs which inevitably specifies the ordering of the WECs (e.g. which WEC is the first, which one is the second, etc.); however, for a given layout the response is permutation-invariant with respect to the ordering of the WECs. By contrast, the CFGP model can take into account the permutation invariance automatically and also consider x-axis symmetry through augmenting the training data; (ii) the training data of the GP model are generated one time by LHS, whereas the training data of the CFGP model are adaptively enriched through active learning.

Figure 5 shows the variation of validation metric R^2 with the iteration for arrays of 3 to 10 WECs. As the iteration evolves, more samples are adaptively added to the training set and the prediction accuracy thus increases gradually. This increasing trend converges quickly, and when 1200 additional samples are added, the accuracy for almost all cases can be up to 70%. The figure demonstrates the effectiveness of the active learning on improving the prediction accuracy. It should be noted that R^2 does not always increase when more samples are added, partly because the selected candidate samples might not always be helpful in improving the prediction accuracy of the CNN part because the uncertainty of the CNN prediction is not included in developing the infill criterion. In addition, the candidate samples are added in batches, which might make some of the updated training data clustered in the design space. One remedy for this issue is to introduce a distance metric between the candidate samples to encourage adding more diversified samples.

Finally, we discuss the efficiency gain provided by the CFGP model. On average, one evaluation of the high-fidelity model for calculating the probabilistic power generation of arrays of 3 to 10 WECs takes 12.4 s, 12.6 s, 15.6 s, 20.7 s, 27.6 s, 41.3 s, 48.5 s, 71.9 s, respectively. Overall, the computational time increases exponentially with the number of WECs in the high-fidelity model, further highlighting the computational challenges in modeling large arrays. On the other hand, the CFGP model takes an approximately similar amount of time for different number of WECs, i.e. only around 0.0007s. Therefore, several orders-of-magnitude speedup can be obtained. Note that when more expensive numerical models (such as time domain models) are used, the computational gain by using CFGP will be even greater. This speedup is of great importance for solving problems, such as layout optimization of large-scale wave farms where typically a large number of model evaluations are required and using a high-fidelity model may be prohibitive.

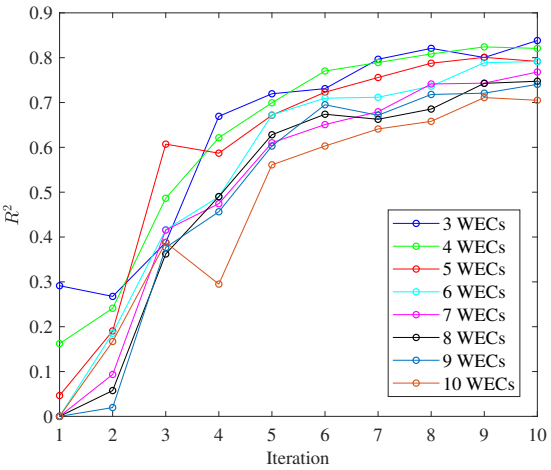

Figure 5. Variation of the R^2 with the iteration for arrays of 3 to 10 WECs.

5 CONCLUSION

This paper proposed an efficient convolution-fed Gaussian process (CFGP) model with active learning algorithm for predicting the probabilistic power generation of WEC arrays in a random sea. The illustrative example verified the effectiveness and efficiency of the proposed algorithm. The proposed algorithm was able to predict the expected power generation of arrays with 3 to 10 WECs in a random sea state with relatively high accuracy. The active learning helps enrich the training set and adaptively improves the prediction accuracy. A future research topic of great interest is how to incorporate the uncertainty of CNN

prediction into the development of the infill criterion for active learning. Also, applications to even larger-scale wave farms will be investigated. Considering the high efficiency in evaluating the CFGP model, the model has great potential to solve optimization problems (e.g. the layout design of wave farms), which will be investigated in future work.

6 ACKNOWLEDGMENTS

The authors thank the Colorado Energy Research Collaboratory Seed Grant Program for support. This material is based upon work supported in part by the National Science Foundation Engineering Design and System Engineering Program under grant number 2034040. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

Balitsky, P., Bacelli, G., and Ringwood, J. V. (2014). "Control-influenced layout optimization of arrays of wave energy converters." *International Conference on Offshore Mechanics and Arctic Engineering*, Vol. 45547, American Society of Mechanical Engineers.

Borgarino, B., Babarit, A., and Ferrant, P. (2012). "Impact of wave interactions effects on energy absorption in large arrays of wave energy converters." *Ocean Engineering*, 41, 79–88.

Faltinsen, O. (1993). *Sea loads on ships and offshore structures*, Vol. 1. Cambridge university press.

Folley, M., Babarit, A., Child, B., Forehand, D., O'Boyle, L., Silverthorne, K., Spinneken, J., Stratigaki, V., and Troch, P. (2012). "A Review of Numerical Modelling of Wave Energy Converter Arrays." *Volume 7: Ocean Space Utilization; Ocean Renewable Energy*, 535–545.

Jia, G., Taflanidis, A., and Scruggs, J. (2015). "Layout optimization of wave energy converters in a random sea." *The Twenty-fifth International Ocean and Polar Engineering Conference*, International Society of Offshore and Polar Engineers.

Kyprioti, A. P., Zhang, J., and Taflanidis, A. A. (2020). "Adaptive design of experiments for global kriging metamodeling through cross-validation information." *Structural and Multidisciplinary Optimization*, 1–23.

Lam, C. Q. (2008). "Sequential adaptive designs in computer experiments for response surface model fit." Ph.D. thesis, The Ohio State University, The Ohio State University.

Mankle, H., Yu, Y.-H., and Dupont, B. (2019). "WEC-Sim Array Development and Experimental Comparison Study." *13th European Wave and Tidal Energy Conference*, (October).

Mavrakos, S. (1991). "Hydrodynamic coefficients for groups of interacting vertical axisymmetric bodies." *Ocean Engineering*, 18(5), 485–515.

Scruggs, J., Lattanzio, S., Taflanidis, A., and Cassidy, I. (2013). "Optimal causal control of a wave energy converter in a random sea." *Applied Ocean Research*, 42, 1–15.

Tom, N., Lawson, M., and Yu, Y.-H. (2015). "Recent Additions in the Modeling Capabilities of an Open-Source Wave Energy Converter Design Tool." *25th International Ocean and Polar Engineering Conference*, (April).

Yu, Y.-H., Lawson, M., Ruehl, K., and Michelen, C. (2014). "Development and Demonstration of the Wec-Sim Wave Energy Converter Simulation Tool." *Proceedings of the 2nd Marine Energy Technology Symposium*, Seattle, WA, 1–8.

Zhang, J., Taflanidis, A., and Medina, J. (2017). "Sequential approximate optimization for design under uncertainty problems utilizing kriging metamodeling in augmented input space." *Computer Methods in Applied Mechanics and Engineering*, 315, 369–395.