TotalControl

Advanced integrated supervisory and wind turbine control for optimal operation of large Wind Power Plants

Control algorithms based on surrogate modelling D4.4

Delivery date: 15.03.2022
Lead beneficiary: DTU
Dissemination level: Public

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 725076.

Authors information:

Name	Organisation	Email
Alan Wai Hou Lio	DTU	wali@dtu.dk

Acknowledgements/Contributions:

Name	Name	Name
Jaime Liew (DTU)	Riccardo Riva (DTU)	Nikolay Dimitrov (DTU)

Document information:

Version	Date	Prepared by	Reviewed by	Approved by
1	15.03.2022	Alan Wai Hou Lio	Karl Merz	

Definitions/abbreviations

WPP	Wind power plant
BO	Bayesian optimisation
GPR	Gaussian process regression
GP	Gaussian process
UCB	Upper confidence bound
DEL	Damage equivalent load

Executive summary

This study presents a surrogate model-based control algorithm for wind farm control problems. Specifically, a data-driven method, Bayesian optimisation with Gaussian process, is used to construct a surrogate to model the non-convex function of typical wind farm control objectives. Based on the surrogate, wind farm operation is optimised and coordinated by derating the power output of individual turbines. The proposed method takes into account internal and external information in the wind farm. For example, the surrogate is built using turbine SCADA (supervisory control and data acquisition) data such as the wind turbine power and tower fore-aft fatigue loads. In addition, the control algorithm can consider external signals such as the grid command, electricity price and the cost of maintenance and fatigue structural loads.

The proposed surrogate model-based algorithm was validated for three common wind farm control objectives: (i) wind power plant (WPP) production maximisation; (ii) WPP economical operation to balance between the energy yield, turbine structural fatigue and maintenance frequency; (iii) WPP operation constrained by the grid command (active power control). The surrogate to model the control objective function is built based on Gaussian process (GP), which can inherently handle the uncertainty of the objective function. By leveraging the GP, Bayesian optimisation can determine the next optimal query sampling location by balancing between exploration and exploitation. In exploration, a sampling location is chosen to minimise the uncertainty, whilst in exploitation, the algorithm tries to improve the current maximum value of the surrogate. By selecting the sampling location wisely, the Bayesian optimisation framework can reach the optimum of the control objective by using only a few trial actions. The results from a middle fidelity wind farm software, HAWC2Farm, show that the proposed control algorithm managed to optimise the wind farm operation. Three different control objectives were investigated and for each one of them, the algorithm was able to find the optimal power set-points for each turbine that can optimise the aggregated control objective of the wind farm.

1 Introduction

Placing wind turbines in wind farms closely can reduce the use of lands and costs of installation, operations and maintenance. However, when wind turbines are installed in close proximity, the downstream turbines are likely to be subject to the wake generated by the upstream turbine. These complex interactions cause a loss in wind velocity (wake velocity deficit) and increase the turbulence (added turbulence) for the downstream turbine, leading to decreased power capture and accelerated structural degradation. These challenges motivate the research of wind farm control, coordinating the operation of the turbines to mitigate the wake effect.

Traditionally, each turbine in a wind farm follows a greedy control strategy that attempts to maximise its own power production regardless of the wake influence for downstream turbines. Recent studies suggested that better coordination of turbine operations can improve the performance of the entire wind farm in terms of power output and turbine structural loads. (See [1]). Of the many wind power plant control that has been published in recent years, most can be grouped into two distinct classes, wake steering and induction control. Wake steering is sometimes known as "wake redirection control" [1], whereas induction control is also known as "axial-induction-based control" [2],"derating" [3] or "down-regulation" [4, 5]. In wake steering, the downstream wakes are redirected by creating a misalignment of the upstream rotor from the ambient wind direction using yaw actuation [6, 7]. The upstream wake is deflected so that it will not be overlapped or partially overlapped with the downstream turbines. Therefore, the adverse effect of the wake on the downstream turbines can be mitigated. In induction control, the pitch and/or torque of an upstream turbine is adjusted to reduce its power capture, thus, decreasing the axial induction that allows higher wind velocity to reach a downstream turbine [8, 9]. Higher wind velocity inevitably leads to higher power capture of the downstream turbine. The concept of wake steering is relatively novel and its implication on turbine structural loads is still a on-going research topics. In addition, the time-scale of wake steering actuation is much longer than induction control. This is because the dynamics of yaw actuators is relatively slower than turbine power controllers which use the turbine generator torque or pitch angle. Since building a real-time surrogate requires some trial actions of turbines, a long time-scale might not deteriorate the quality of the surrogate as changes in wind condition. Thus, the focus of this study is on derating control.

The control objective of a wind farm controller is to find the optimal power set-point for each turbine that can maximise the system aggregated cost (or objective) function, which can be (i) power maximisation of the whole wind farm, (ii) turbine structural fatigue load alleviation, (iii) tracking the power command from the transmission system operator (TSO), or their combinations. To achieve these control objective, farm controllers can be classified into model-based and data-driven (model-free).

The chaotic and variable nature of the wake-turbine interactions makes modelling incredibly challenging. In recent years, wind farm wake-turbine interactions have been studied intensively. Different analytical and engineering models were developed to model its complicated dynamics. For example, Jensen[10], Frandsen[11], GCLarsen[12], Bastankhah and Porté-Agel[13], Fuga[14] , Floris[15], etc. The benefit of using analytical and engineering models is that they are fast to provide the quasi-steady wind flow prediction within a wind farm. For example, a study by [16] used an engineering model to establish a fast surrogate to facilitate an efficient optimisation of the power production of a wind farm. The use of analytical and engineering model offers fast and efficient prediction but the drawback is that there might exists mismatch between the model and true system. If the model is largely different to the plant, the optimal power set-points that maximise the surrogate function might be different to the system's true optimum. In [17], a closed-loop adjoint model predictive control strategy was proposed to maximise the power production based on WFSim, where the flow model is updated based on real-time measurements. A similar method [18] is extended to active power control. Another closed-loop control example is by [19], where a dynamic flow predictor [20] is built based on the Frandsen wake model to track the power command from the TSO using model predictive control.

In data-driven approaches, one of the benefits is that modelling of the wake interactions is not necessary to achieve the control objective. For example, an early study by [21] proposed a learning control strategy based on extreme seeking control to maximise the wind farm power under time-varying wind conditions. Another study by [22] demonstrated a learning strategy based on game-theoretic methods to maximise energy production without explicitly knowing the aerodynamic interactions amongst the turbines. A study by [23] proposed another data-driven strategy to maximise power production. However, these methods suffer from slow convergence. Besides power maximisation, a study by [24] proposed a model-free proportional-integral (PI) wind farm control strategy to track the power reference from the TSO (active power control). Later, a study by [25] extended this model-free concept to alleviate aggregated structural loads of turbines.

Another benefit of data-driven is to model the relationship between turbine structural loads, wind condition and control variables. Typical analytical and engineering models can provide a fast and accurate prediction of wind speed, which is useful for wind farm power maximisation and active power control. However, in terms of turbine structural loads, a mapping between the wind condition and fatigue loads needs to be established. Some recent works use surrogate modelling to build this mapping for example by neural network [26], polynomial chaos expansion (PCE) [27], Gaussian process regression[28], fatigue load sensitivity [29], etc. A good control-oriented surrogate enables power and load optimisation of a wind farm where the optimal power set-point can be found for each turbine. One of the problems with surrogate modelling is that to reach the global optimum, the input space needs to be exhaustively explored thus a large number of sampling points are typically required. For example, 10^4 samples were used to train a site-specific load prediction model using PCE without considering control variables [27]. By taking into account control variables (e.g. derating), a study by [26] constructed a control-oriented load surrogate model of two turbines using roughly 10^4 training samples. These methods suffer from the curse of dimensionality. Thus, finding the system optimum using a control-oriented surrogate model for a wind farm is intractable. The trial actions required in the training phase could be enormous. For example, for a wind farm with 32 turbines, 10 samples of the control variable (e.g. different derating levels) of each turbine could require 10^{32} samples per wind case.

As a consequence, this begs two questions: (i) can a surrogate (to model the control objective function) be built by executing only a small number of trials actions where the sampling location is optimised? (ii) can this surrogate be used for optimising the power set-point for achieving the control objectives? In this work, we propose a data-driven wind farm control method using Gaussian process regression and Bayesian optimisation. Specifically, the system cost function (or control objective function) is approximated by the Gaussian process regression using the external (e.g. grid supervisory command or electricity price) and internal information (e.g. turbine SCADA data). The Bayesian optimisation is then used to direct sampling to areas where an improvement over the current best observation. The improvement is typically traded off between exploration and exploitation. Exploitation implies sampling where the surrogate predicts the highest value (cost) of the surrogate whereas exploration means sampling at locations where the prediction uncertainty is high. The proposed control scheme can incorporate cost functions that covers (i) the wind power plant (WPP) production maximisation; (ii) most economic WPP operation (e.g. a balance between energy yield and turbine component fatigue); and (iii) most economic WPP production conditioned on a grid specified WPP production level (active power control). Gaussian process regression and Bayesian optimisation have been recently adopted in wind farm control and most of them was focusing on power maximisation (e.g. [30, 31]). This paper considers the load aspect in the optimisation problem. The proposed controller will be validated in a mid-fidelity wind farm tool Hawc2farm. In addition to WPP production optimisation, different control objectives such as economic WPP operation and economic WPP operation conditioned on a grid command will be studied.

The remainder of this report is organised as follows. Section 2 presents the background and formulation of wind farm control problems, including wind turbine derating design. In Section 3, the proposed wind farm control algorithm is presented and also Bayesian optimisation and Gaussian process regression are discussed. Section 4 validated the proposed algorithm in

HAWC2Farm for cases of two and three turbines in a row. Finally, it is followed by conclusions and discussions of future work in Section 4.4.

2 Background and Problem formulation

Considering a wind farm with the number of turbines denoted by $N_{\mathrm{T}} \in \mathbb{Z}$, a wind farm control problem can typically be generalised as an optimisation problem with some control variables and constraints:

$$\delta_i = \arg\max_{\delta_i} \; f(\delta_i, \phi) \tag{1a}$$

$$\text{s.t.} \;\; h_{\mathrm{eq}}(\delta_i, \phi) = H_{\mathrm{eq}} \tag{1b}$$

$$h_{\mathrm{ineq}}(\delta_i, \phi) \leq H_{\mathrm{ineq}} \tag{1c}$$

where δ_i is the control variables and in this case, it is the derating set-point of turbine $i \in \{1, \dots N_{\mathrm{T}}\}$). The notation ϕ denotes the exogenous conditions such as atmospheric conditions (the wind speed, turbulence intensity, wind shear and wind direction) and electricity price. The cost function (or control objective function) f to be maximised (or minimised with a negative sign) can be defined as either (i) wind power plant power production (ii) wind power plant structural loads. The constraints (1b) and (1c) denote the equality and inequality constraints, which can be power curtailment requirements from the grid or limitations on the turbine structural load.

As mentioned in Section 1, this work will focus on three control objectives, namely (i) wind power plant power maximisation; (ii) wind power plant economical operation; and (iii) economical operation under grid power requirement. First, the wind power plant power maximisation can be described as follows:

$$\max_{\delta_i} \;\; \sum_{i=1}^{N_{\mathrm{T}}} P_i(\delta_i, \phi) \tag{2}$$

where P_i denotes the power output of turbine $i \in \{1, \dots N_{\mathrm{T}}\}$.

In wind power plant economical operation, a balance between energy yield, cost of turbine component degradation, maintenance cost needs to be optimised. Turbine component degradation known as fatigue damage is typically quantified by the stress amplitudes and their number of cycles on the material characterised by its material exponent. A quantity known as 1-Hz Damage Equivalent loads (DELs) is often used to quantify the fatigue damage. Associating a monetary cost with the DEL of a turbine component is non-trivial as all turbine components are interconnected. But this is beyond the scope of this study. Besides turbine component degradation, turbine components need to be serviced from time to time. In a wind farm, downstream turbines experience higher structural load degradation due to the wake effect. Minimising the deviation of structural loads between turbines could be beneficial. Therefore, the optimisation problem for wind power plant economical operation can be formulated as follows:

$$\max_{\delta_i} \;\; \sum_{i=1}^{N_{\mathrm{T}}} P_i(\delta_i, \phi) - \sum_{i=1}^{N_{\mathrm{T}}} \sum_{j=1}^{N_{\mathrm{comp}}} \lambda_{\mathrm{d},j} D_{i,j}(\delta_i, \phi) - \sum_{j=1}^{N_{\mathrm{comp}}} \lambda_{\mathrm{m},j} \sigma \left(D_{i,j}(\delta_i, \phi) \right) \tag{3}$$

where $D_{i,j}$ denotes the DEL of the component $j \in \{1, \dots, N_{\mathrm{comp}}\}$ of turbine i. The weights $\lambda_{\mathrm{d},j}, \lambda_{\mathrm{m},j}$ are used to balance between the importance of turbine component degradation and maintenance cost over the energy yield. The notation σ denotes standard deviation.

In active power control, wind power plants participate in ancillary services that require the wind farm to maintain the proper power flow to the grid and provide the additional reserved power if needed. One of the main objectives of active power control is to stabilise the power grid frequency. The wind farm power follows the power reference requirement from the grid by

derating the turbines. Typically, the power reference from the TSO is lower than the maximum available power in the wind farm. Thus, the turbines can be coordinated to output the required power while alleviating the structural loads. The optimisation problem can be formulated as follows.

$$\min_{\delta_i} \quad -\sum_{i=1}^{N_\mathrm{T}} \sum_{j=1}^{N_\mathrm{comp}} \lambda_{\mathrm{d},j} D_{i,j}(\delta_i, \phi) - \sum_{j=1}^{N_\mathrm{comp}} \lambda_{\mathrm{m},j} \sigma\left(D_{i,j}(\delta_i, \phi)\right) \tag{4a}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_\mathrm{T}} P_i(\delta_i, \phi) = P_\mathrm{ref} \tag{4b}$$

where P_ref is the power reference requirement of the entire wind farm from the grid.

In order to solve the optimisation problems Eq. (2), (3) and (4), a model that can accurately predict the power output and turbine component DEL of all turbines is required. The model needs to take into account the wind farm atmospheric conditions including the wake effect as well as the effect of turbine derating. Some studies solved these optimisation problems based on analytical, engineering or surrogate models. For example, Fuga can take into account the wind farm geometry and turbine derating in terms of thrust coefficient. The wake deficit can be fast and accurately predicted, which can then be used to calculate the power output of turbines and solve the optimisation problem (2). Another example [32] is to train the surrogate model with a large amount of measurement and solve the optimisations based on the surrogate model predictions. However, due to plant-model mismatch (e.g. wind farm atmospheric conditions, the effect of de-rating on the wake), purely open-loop model-based optimisation is often unable to reach the plant optimality. Therefore, this work seeks to use a surrogate model that is trained online with minimal real-time measurements and such a model needs to be used to solve the above-mentioned optimisation problems.

2.1 Turbine controller and down-regulation strategy

The turbine derating strategy has a large impact on the wake effect in a wind farm, thus resulting in different power and structural cost function. In this section, the turbine controller and derating operation are discussed.

2.1.1 Basic turbine controller

Basic wind turbine controllers are typically consisted of the generator torque and blade pitch controllers. The generator torque control law is defined as follows:

$$\tau_\mathrm{g}(t) = \begin{cases} K_\mathrm{opt}\omega(t)^2, & \text{if } \theta(t) \leq \theta_\mathrm{s}, \quad (5a) \\ \dfrac{P_\mathrm{rated}}{\omega(t)}, & \text{if } \theta(t) \geq \theta_\mathrm{s}, \quad (5b) \end{cases}$$

$$\tau_\mathrm{g}(t) \in [\underline{\tau}_\mathrm{g}, \bar{\tau}_\mathrm{g}], \tag{5c}$$

The operating conditions are dependent on the switching parameter of the pitch angle θ_s. In below-rated wind condition (5a), the generator torque controller maximises the turbine power by tracking the optimal tip-speed ratio with the optimal gain $K_\mathrm{opt} \in \mathbb{R}$, whilst in the above-rated wind condition (5b), the controller maintains the power at the rated value $P_\mathrm{rated} \in \mathbb{R}$. The generator torque is constrained by the minimum and maximum limits denoted as $\underline{\tau}_\mathrm{g}, \bar{\tau}_g \in \mathbb{R}$. For brevity, the aspects of how the controller (5) handles transitions around the start-up and the rated rotor speed are omitted from this paper and more details can be found in [33].

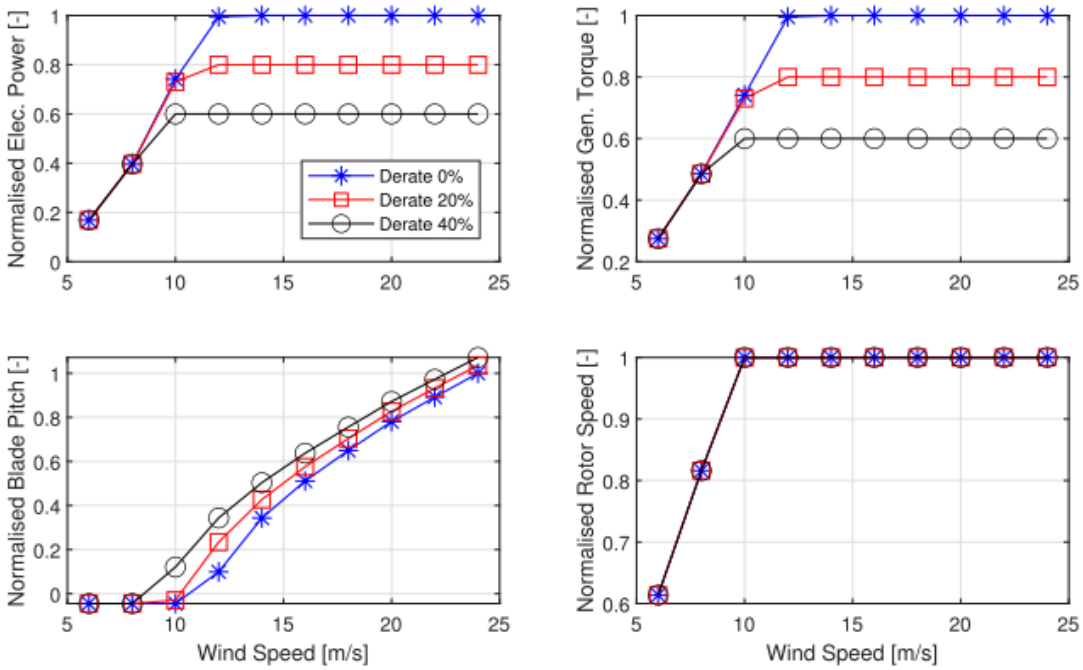

Figure 1: Normalised wind turbine mean electrical power, torque, blade pitch and rotor speed curves as a function of wind speed for different derating percentages [26].

The blade pitch controller is typically designed as a gain-scheduled proportional-integral (PI) controller, defined as follows:

$$\theta(t) = f_{\text{PI}}(\omega(t) - \omega_{\text{rated}}), \qquad \theta \in [\theta_{\min}, \theta_{\max}], \tag{6}$$

The PI control law $f_{\text{PI}} : \mathbb{R} \to \mathbb{R}$ drives the rotor speed to the rated value $\omega_{\text{rated}} \in \mathbb{R}$ and typically, it is gain-scheduled by the pitch angle (e.g. [33]), whilst the pitch angle is limited by $\theta_{\min}, \theta_{\max} \in \mathbb{R}$.

2.1.2 Down-regulation strategies

The power produced by a turbine is a product of the generator torque and the generator speed. Down-regulation can be achieved by either manipulating the generator torque or rotor speed set-point [34]. Therefore, a number of derating strategies exist in the literature [9]. This work considers torque-based down-regulation strategies, which is also known as Max-Omega strategy.

The torque-based strategy performs turbine down-regulation by changing the generator torque input solely. To implement the torque-based strategy, a new maximum torque limit $\bar{\tau}_{\text{g,derated}} \in \mathbb{R}$ is imposed on the generator torque in (5c), defined as follows:

$$\bar{\tau}_{\text{g,derated}} = \frac{P_{\text{derated}}}{\omega_{\text{rated}}} = \frac{\delta P_{\text{rated}}}{\omega_{\text{rated}}}, \tag{7}$$

where $P_{\text{derated}} \in \mathbb{R}$ denotes the derated power and the derating set-point δ is defined as a percentage of the rated turbine power. One of the benefits of such a strategy is that during power curtailments, the rotor speed is operating at rated and thus reserving the maximum amount of spinning energy for providing fast frequency response support to the grid [35]. As soon as the nominal rotor speed is reached, the blades are pitched towards feathering to reduce the power to the desired level. Figure 1 shows the steady-state turbine operational pitch angle and rotor speed, as well as the electrical power output and generator torque at different derating set-point.

3 Wind farm control using Bayesian Optimisation and Guassian Process

Bayesian optimisation (BO) is an iterative algorithm to search the global maximum of an unknown function f. Namely, it is equivalent to find $x^* = \arg\max_{x\in\mathcal{X}} f(x)$, where $\mathcal{X}$ is the design space. A BO framework contains two elements: (i) a surrogate model for approximating the unknown function f, and (ii) an acquisition function for determining which point to query for the next evaluation of f.

3.1 Surrogate modelling using Gaussian process regression

This section presents the use of Gaussian process regression (GPR) to model the cost function f and constraints g, h in the wind farm optimisation problem in Eq.(1).

3.1.1 Motivation of the choice of GPR

In Figure 1, the relationships between derating set-point, power and structural loads are nonlinear. Thus, a nonlinear regression method needs to be used to approximate these relations. Unlike many other nonlinear regressions methods that learn the exact parameters in a function, GPR is a non-parametric and Bayesian framework for nonlinear function approximation. Most importantly, GPR can provide the mean and its uncertainty (standard deviation) of the query sample, which is crucial for Bayesian optimisation is to find the global maximum of the unknown function f in Eq. (1). The GPR method requires little data-set and prior process knowledge, thus, it has been widely adopted in many applications in wind energy. For example, wind turbine power curve fitting [36], model parameter estimation [37], condition monitoring [38], wind speed prediction [39], rotor effective speed estimation [40].

3.1.2 Basics of Gaussian process regression

Given a set of N pair of training data or observations $\mathcal{D} \equiv \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N = \{\mathbf{X}, \mathbf{y}\}$, the main task of the regression problem is to learn the unknown model or function $f : \mathbb{R}^d \to \mathbb{R}$ that maps the input vector $\mathbf{x}_i \in \mathbb{R}^d$ to the observation $\mathbf{y}_i \in \mathbb{R}$. Let the noisy observation be given as follows:

$$\mathbf{y}_i = f(\mathbf{x}_i) + \epsilon_i,\ \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2), \tag{8}$$

where the difference between the observed value $\mathbf{y}_i$ and function value $f(\mathbf{x}_i)$ is assumed to be Gaussian noise $\epsilon \in \mathbb{R}$ with zero mean and variance σ_ϵ^2. The key idea of GPR is to model the unknown function $f(\mathbf{x})$ as a Gaussian process (GP), where each function value $f(\mathbf{x}_i)$ evaluated at different $\mathbf{x}_i$ is jointly Gaussian distributed with mean function $m(\mathbf{x}_i)$ and covaraince function $k(\mathbf{x}_i, \mathbf{x}_j)$, defined as follows:

$$f(\mathbf{X}) \sim \mathcal{GP}\left(m(\mathbf{X}), K(\mathbf{X}, \mathbf{X})\right), \tag{9a}$$

$$m(\mathbf{X}) = \mathbb{E}\left(f(\mathbf{X})\right), \tag{9b}$$

$$K(\mathbf{X}, \mathbf{X}) = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_1, \mathbf{x}_N)^T & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}. \tag{9c}$$

The major advantage of this Gaussian process assumption (9) is that the target value $\mathbf{y}$ in (8) is a linear combination of Gaussian distributions, thus, it also possesses the following Gaussian distribution property:

$$\mathbf{y} \sim \mathcal{N}(m(\mathbf{X}), K + I\sigma_\epsilon^2), \tag{10}$$

where $K = K(\mathbf{X}, \mathbf{X})$. Notice that the Gaussian process in (10) is solely described by the covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$ and the noise variance σ_ϵ^2. The covariance function and its parameters describe the Gaussian distribution of f and are typically chosen based on prior process knowledge.

This work considers the squared exponential kernel function as the covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$, defined as follows:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)}{2l^2}\right), \tag{11}$$

where $l \in \mathbb{R}_{>0}$ denotes the length-scale of the kernel function, which specifies the smoothness of the model $f(\mathbf{x})$. The squared exponential kernel function is a widely-used Gaussian kernel because of its smoothness and continuity properties, that indicate the function is infinitely differentiable. The kernel function decays rapidly if an increasingly distant pair of input $\mathbf{x}_i$ and $\mathbf{x}_j$ is being evaluated, showing weak correlations between $f(\mathbf{x}_i)$ and $f(\mathbf{x}_j)$. Any other differentiable kernel functions (e.g. [41]) can be used without loss of generality. The hyperparameters of the Gaussian process (10) are given by $\theta_{\text{hp}} = [l, \sigma_\epsilon]$. The optimal values for these hyperparameters can be found by maximizing the log marginal likelihood $p(\mathbf{y}|\theta_{\text{hp}})$ [41]:

$$\log p(\mathbf{y}|\theta_{\text{hp}}) = -\frac{1}{2}\log|K| - \frac{1}{2}\mathbf{y}^T K^{-1}\mathbf{y} - \frac{N}{2}\log(2\pi). \tag{12}$$

Based on the Gaussian process (10), the joint distribution between the training data set $\mathcal{D} = \{\mathbf{X}, \mathbf{y}\}$ and an arbitrary set of test data $\mathbf{x}_*$ is as follows:

$$\begin{bmatrix} \mathbf{y} \\ f(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(\mathbf{X}) \\ m(\mathbf{x}_*) \end{bmatrix}, \begin{bmatrix} K + I\sigma_\epsilon^2 & k(\mathbf{X}, \mathbf{x}_*) \\ k(\mathbf{x}_*, \mathbf{X}) & k(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right), \tag{13}$$

where $k_* = k_*^T = k(\mathbf{X}, \mathbf{x}_*) = [k(\mathbf{x}_1, \mathbf{x}_*), \cdots, k(\mathbf{x}_N, \mathbf{x}_*)]^T$, and k_* and $k(\mathbf{x}_*, \mathbf{x}_*)$ can be calculated using (11). Subsequently, conditioning the joint distribution yields the predicted function values $f(\mathbf{x}_*)$ with a posterior Gaussian distribution, $p\left(f(\mathbf{x}_*)|\mathbf{y}, \mathbf{X}, \theta_{\text{hp}}\right)$, where its mean $\mu(\mathbf{x}_*)$ and corresponding variance $\sigma^2(\mathbf{x}_*)$ are as follows:

$$\mu(\mathbf{x}_*) = k_*^T(K + I\sigma_\epsilon^2)^{-1}\mathbf{y}, \tag{14a}$$

$$\sigma^2(\mathbf{x}_*) = k(\mathbf{x}_*, \mathbf{x}_*) - k_*^T\left(K + I\sigma_\epsilon^2\right)^{-1}k_*. \tag{14b}$$

The mean $\mu(\mathbf{x}_*)$ and variance $\sigma^2(\mathbf{x}_*)$ can be seen as the predictions of the hidden function f output for a given input $\mathbf{x}_*$.

Gaussian process regression is a powerful tool to estimate the cost function in a wind farm optimisation. However, similar to any other machine learning methods, it might require a large number of samples to produce an accurate representation of the unknown function f (or the surrogate model). A large number of trial control actions in real turbines is not feasible. Thus, it is crucial to wisely select the trial actions to build a high-quality surrogate model that can be used in the wind farm controller to achieve the control objective. Thus, Bayesian optimisation can compute the best sampling location by evaluating the acquisition function.

3.2 Acquisition function

The acquisition function is typically a computationally inexpensive function that can be used to evaluate how optimal a query sample x is for the optimisation problem (1). In other words, instead of evaluating the unknown, non-convex and expensive function f real-time, the acquisition function is a function that can be evaluated in a computationally efficient manner. The acquisition function values can be computed within a computer with a large number of samples. A GP-based surrogate is ideal for this fast function evaluation.

Since a GP-based surrogate is an approximate of the unknown function f. A simple strategy to select the next query point is to choose where the surrogate has the highest mean $\bar{f}(x)$, which

can help to search for a superior value than the current one. This is known as exploitation or passive learning. Nonetheless, due to potential mismatch between the surrogate and unknown function f, it is beneficial sometimes to take account the uncertainty of the surrogate. Another strategy is to select the next query point where the surrogate has some degrees of uncertainty, namely high surrogate variance, which can help to reduce the uncertainty around the selected query point. This is known as exploration or active learning, which can reduce the uncertainty of the surrogate model to the unknown function f. Clearly, a balance between exploration and exploitation is crucial to operate at the maximum of the unknown function f.

The next query point is determined by maximising the acquisition function that takes into account the exploration and exploitation. The widely used acquisition function is upper bound confidence (UBC), which has a explicitly clear structure showing the importance between the surrogate mean and variance. The maximisation problem is defined as follows.

$$\mathbf{x}^{n+1} = \arg\max_{\mathbf{x}} \left(\mu(\mathbf{x}) + \kappa\sigma(\mathbf{x})\right), \tag{15}$$

where κ is related to the trade-off between exploration and exploitation. While high values of κ lead to an explorative behaviour of the algorithm, lower values of *kappa* favour exploitation near known sampled locations. The optimisation (15) can be solved by methods such as sequential least squares programming algorithm. The problem (15) sometimes includes constraints on the input space $\mathbf{x}$. For example, in the active control problem, the sum of turbine power set-points needs to fulfil the grid requirement.

After the new query point is obtained, it will be applied to the true system (or cost function). Then new measurement samples are gathered to update the surrogate model. The process is repeated until the maximum iteration is reached or the termination condition is met.

3.3 Phases of the BO

Bayesian Optimisation operates by continuously exploiting and exploring the cost function to search for the optimum. In exploration, some choices of query points might potentially jeopardise the performance of the algorithm. Thus, it is crucial to balance the exploration and exploitation in real-time. If the cost function is learnt good enough to facilitate good prediction of the optimum, a pure exploitation strategy should be used to maximise the performance (e.g. wind farm power output). Similarly, if the surrogate model fails to predict the true optimum due to dynamic changes in atmospheric conditions, the algorithm should switch to an exploration/exploitation strategy for re-learning the cost function by using the new measurement samples.

One simple and heuristic strategy is for pure exploitation, the parameter κ in the acquisition function can be set to zero while for exploring/exploiting, the parameter κ is set to a non-zero value to ensure the acquisition function considering the uncertainty of the surrogate. To make these switchings, the termination and re-learning conditions are needed to be defined. For example, the termination condition that stops exploration can be triggered when the surrogate improvement is lower than a threshold:

$$|\mu(\mathbf{x}_{k+1}) - \mu(\mathbf{x}_k)| < \epsilon \tag{16}$$

For the re-learning condition, one heuristic condition that re-initialise exploration/exploitation can be that if the true optimal $f(\mathbf{x})$ lies outside three standard deviations of the surrogate mean function:

$$|f(\mathbf{x}) - \mu(\mathbf{x})| > 3\sigma(\mathbf{x}) \tag{17}$$

If the surrogate model fails to predict accurately, the algorithm will discard the current set of training data and collect new samples from scratch. One benefit of re-learning is to relieve the computation burden of the algorithm. To make a GP prediction, each time requires to solve the inversion of the covariance matrix and the size of the matrix grows as more samples are collected. Discarding the samples that fail to make accurate predictions can speed up the algorithm.

The Bayesian optimisation is described in Algorithm 1.

Algorithm 1 Bayesian Optimisation

Initialisation. Set $k = 0$. Choose the initial $\mathbf{x}_0$ and observe output $\mathbf{y}_0$;
while $k < n_{\max}$ **or** *Condition* = True **do**
  Update the training data set $\mathcal{D}^k = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=0}^k = \{\mathbf{X}^k, \mathbf{y}^k\}$;
  Train the GP model using $\mathcal{D}^k$;
  Optimise the hyperparameter θ_{hp};
  Compute the next query sample point $\mathbf{x}_{k+1} = \arg\max_{\mathbf{x}} \big(\mu(\mathbf{x}|\mathcal{D}^k) + \kappa\sigma(\mathbf{x}|\mathcal{D}^k)\big)$;
  Evaluate the unknown function f with $\mathbf{x}_{k+1}$ to obtain $\mathbf{y}_{k+1}$;

  if $|\mu(\mathbf{x}_{k+1}) - \mu(\mathbf{x}_k)| < \epsilon$ **then** ▷ Termination condition
    Set *Condition* = False; ▷ Terminate the optimisation
    Deploy the optimal $\mathbf{x}^* = \mathbf{x}_{k+1}$ for every time step;
  end if
  $k = k + 1$
end while
if $|f(\mathbf{x}) - \mu(\mathbf{x})| > 3\sigma(\mathbf{x})$ **then** ▷ re-learning condition
  Discard the training dataset $\mathcal{D}^k$;
  Set *Condition* = True and go to **Initialisation**;
end if

4 Results and Discussion

4.1 Simulation Set-up

The wind farm is simulated using a middle-fidelity wind farm framework, HAWC2Farm, which can accurately and efficiently compute the wind turbine production and structural loads within a wind farm. In HAWC2Farm, the turbine dynamics is modelled by HAWC2, which is an aeroelastic turbine simulation code [42] and the wake dynamics and interactions between turbines are based on the dynamic wake meandering model (DWM) [43]. The turbulent wind flow is generated by the Mann turbulence box. The turbulence intensity in this study is around 5%. In HAWC2, DTU Wind Energy Controller (DTUWEC) is used to derate the turbine [44] and the derating level is bounded between 50% and 100% of the rated power. The wind turbine model used in this study is the DTU 10MW reference wind turbine with a rotor diameter of 178.3 m, rated power of 10 MW and rated wind speed of 11.4 m/s [45]. The training data set consists of observations of 10-min averages. This is because, in order to capture the turbine structural loads, DELs of 10-min time-series are recommended according to the IEC standard [46].

4.2 Two turbines in a row

To illustrate the BO algorithm, we consider two turbines in a row as depicted in Figure 2. The streamwise spacing is 5 rotor diameters. In this study, we assume the wind condition remain unchanged during the optimisation. Namely, the 10-min mean wind speed and turbulence intensity remain the same.

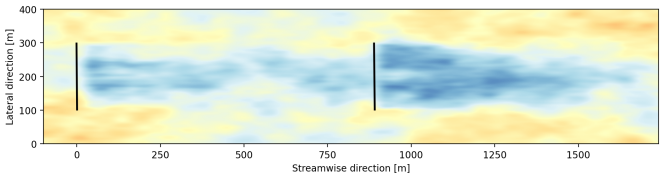

Figure 2: Two turbines in a row. The streamwise spacing is 5 rotor diameters.

4.2.1 Wind Power Plant Power Maximisation

Figure 3 shows that the BO algorithm is searching for the maximum wind farm power output over ten (10) iterations. The mean wind speed $\bar{U}$ is 10 m/s, which is below-rated. By derating the upstream turbine, more power becomes available to the downstream one. The example investigates what is the optimal power set-point for the upstream turbine in regard to maximising the aggregated power output of the wind farm. Notice that the downstream turbine is operating at a greedy strategy without derating. To obtain the true cost function, a set of simulations is performed in HAWC2Farm. The upstream turbine is derated between 50% to 100% by increments of 5%, which results in 11 samples of 10-min mean of total power generated by the wind farm. The 10-min mean total power is then interpolated across the input space. By inspecting the true cost function, the maximum 10-min mean power output of the wind farm is 9.62 MW and the optimal power set-point for turbine 1 is 77%.

To obtain a rough trend of the surrogate, the initial sampling points are selected at the bounded values, which are 100% (greedy strategy) and 50% derating of turbine 1. By evaluating the surrogate across the input space, the surrogate mean $\mu(\mathbf{x})$ and standard deviation $\sigma(\mathbf{x})$ were calculated, which are shown in the blue line and light blue shading area respectively in Figure 3. The next query point is determined by maximising the acquisition function. The tuning parameter in acquisition function is set to be 1 ($\kappa = 1$) to balance exploitation and exploration. For example, in iteration 2, the query point was not chosen where the surrogate

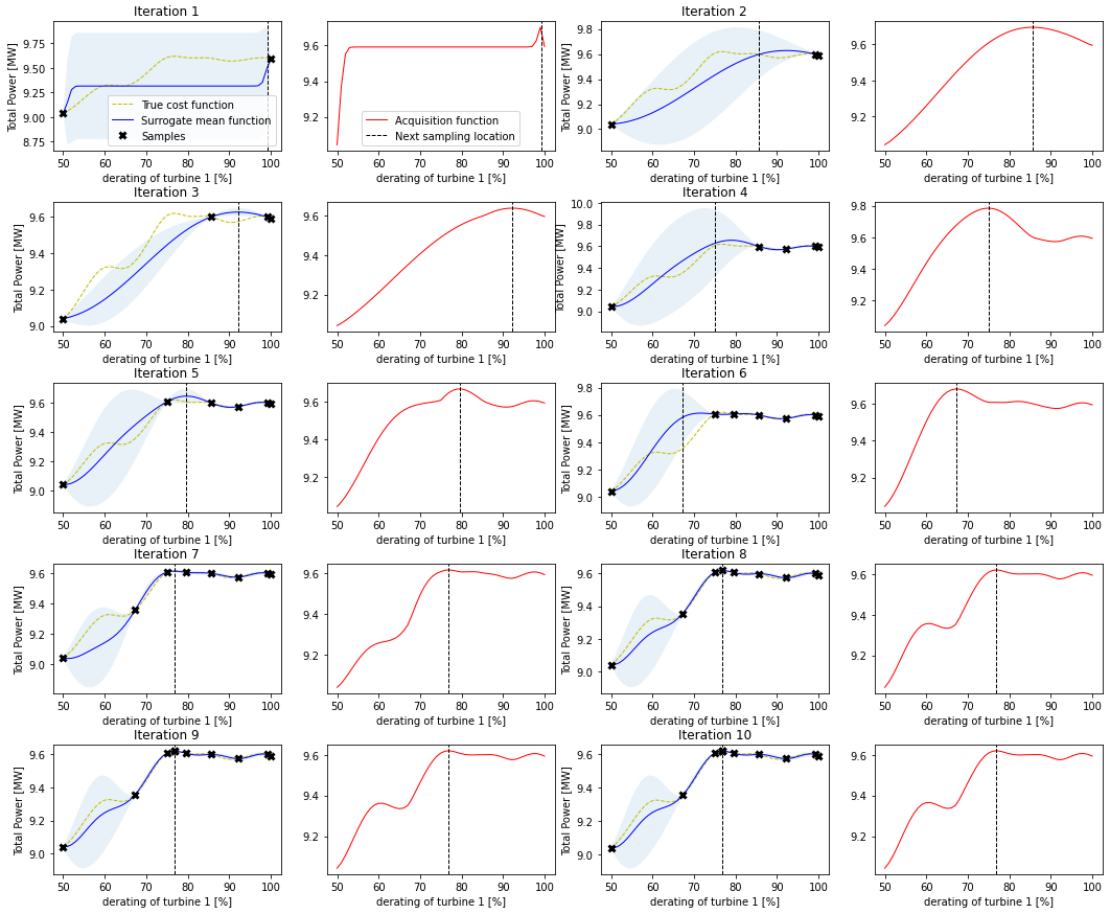

Figure 3: An example of WPP power maximisation of two turbines by derating the upstream one. For each iteration, Left: The true cost function (green) is approximated by the surrogate mean (blue) and 95% confidence interval (light blue) evaluating the system with samples (cross). Right: The next query sampling point (dashed line) is determined by maximising the acquisition function (red).

mean (blue line) is at maximum, instead, the acquisition function takes into account the surrogate uncertainty (standard deviation) and resulted in an exploring decision. After 7 iterations, the sampling location is converging the optimum, which is 77%. Beyond iteration 7, it can be seen that the acquisition function and optimal power set-point were not changing.

To summarise, the BO could find the maximum 10-min mean wind farm power using only 7 trial actions, compared to the traditional surrogate technique of 11 trials actions. In terms of energy production, if a traditional greedy control strategy (100% of derating of turbine 1) is used, the energy yield was 157.2 MWh over 10 iterations (one iteration is 10 min). For the BO approach, the energy yield was 158.2 MWh, which is a significant increase of 0.63% in energy yield. Moreover, if the wind condition remains unchanged for longer, the BO solution would definitely lead to a much higher percentage of the energy yield.

4.2.2 Wind Power Plant Economical Operation

In addition to power maximisation, the BO algorithm can also alleviate turbine structural loads while maximising the energy yield. In this study, for simplicity, only the tower fore-aft DEL is considered. The BO framework can easily extend to other structural loads.

Similar to Section 4.2.1, HAWC2Farm simulations were performed to obtain the 10-min mean

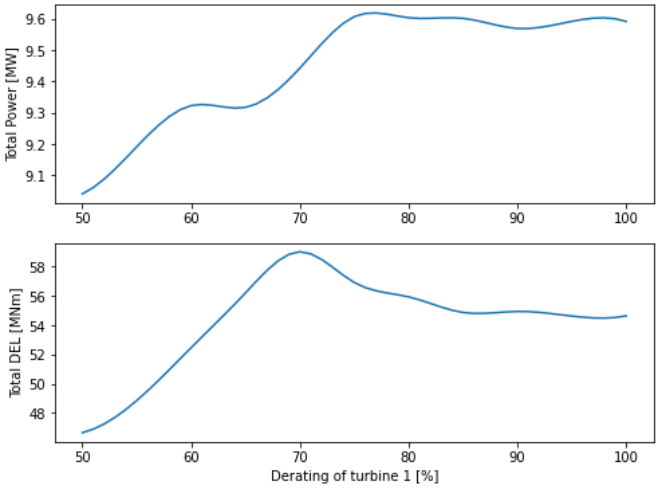

Figure 4: WPP economical operation. True cost function of 10-minute mean of the total wind farm power output (top) and the sum of tower DELs of two turbines (bottom).

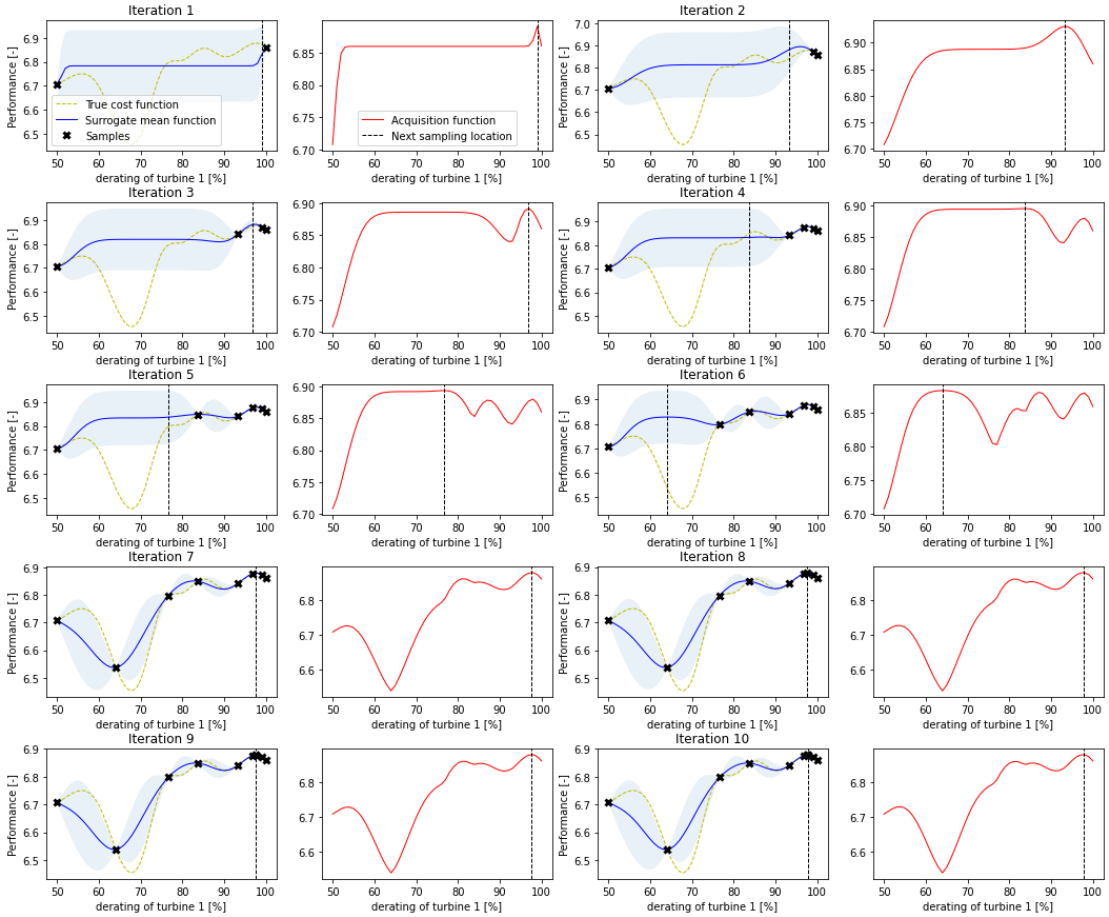

Figure 5: An example of WPP economical operation. For each iteration, Left: The true cost function (green) is approximated by the surrogate mean (blue) and 95% confidence interval (light blue) evaluating the system with samples (cross). Right: The next query sampling point (dashed line) is determined by maximising the acquisition function (red).

of total wind farm power and the sum of turbine tower DELs within the wind farm. To obtain the true cost function, 11 simulations were performed by derating the upstream turbine from 100% to 50% by increments of 5%. Figure 4 shows the true function of 10-min mean of the wind farm power and the sum of tower DELs of the wind farm in respect to the derating levels of turbine 1. Notice that the magnitude of DELs is higher than the 10-min mean of the wind farm power. To form a cost function to maximise i.e. Eq. (3), a weight λ_d needs to be carefully selected to reflect the momentary cost of the energy and fatigue loads. For example, the electricity price on the spot market and the cost of materials. For demonstration purpose, in this study, we set the weight λ_d as 0.05. Based on Eq. (3), the new cost function to maximise is depicted in Figure 5.

In Figure 5, the BO algorithm optimises the combination of wind farm mean power and total tower DEL, thus, the combined cost is denoted as *performance*. The maximum performance is 6.88 and the optimal power set-point for turbine 1 is 98%. Figure 5 shows the BO explored the input space at the first few iterations. This is due to the uncertainty remaining high across the input space, for example, between iteration 4 and 6. The surrogate uncertainty (standard deviation) decreased as more input are sampled across the input space. Thus, after iteration 7, the maximum power set-point for turbine 1 was reached. The cumulative performances over 10 iterations were 81.8 for the greedy strategy and 82.3 for the BO approach, which is a 0.6% increase.

4.3 Three turbines in a row

To reveal the true power of the BO algorithm, a higher dimension problem is considered in this section. The example in this section contains three turbines in a row. Three DTU 10MW reference wind turbines are 5 rotor diameters apart from each other, as depicted in Figure 6. Two control objectives are investigated: (i) WPP economical operation; (ii) WPP economical operation conditioned on the grid specified WPP production level (active power control). The control variables in these problems are the power set-points for the first two upstream turbines, which are bounded between 50% and 100% of the rated power. The last turbine in the wind farm operates based on greedy strategy.

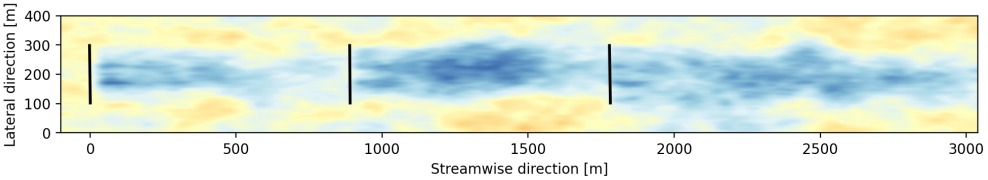

Figure 6: Three turbines in a row with a spacing of 5 rotor diameters in the streamwise direction.

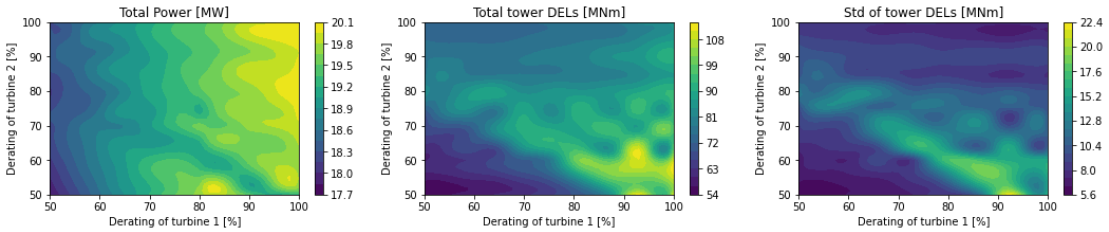

Figure 7: WPP Economical operation. True cost functions of the 10-minute mean of the wind farm power output (left), the sum (middle) and standard deviation (right) of the tower DELs of three turbines.

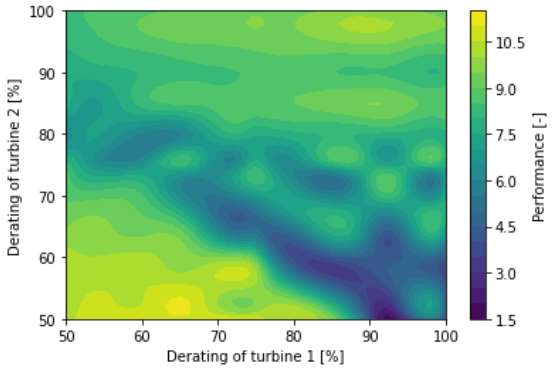

Figure 8: True cost functions to be maximised for WPP economical operation.

4.3.1 WPP economical operation

In this example of WPP economical operation, additional maintenance cost is considered, where the standard deviation of the tower DELs is taken into account when forming the cost function. The mean wind speed is 12 m/s and turbulence intensity is around 5%. The true functions are computed by simulating 121 combinations of cases where turbine 1 and turbine 2 are derated between 50% and 100% of the rated power by increments of 5%. Then, the 10-min mean power of the wind farm, sum and standard deviation of the tower DELs are collected. By interpolating the data, the colour maps of the true functions are shown in Figure 7. These maps are highly nonlinear and have many local maxima and minima. To form a cost function to maximise for WPP economical operation, based on the Eq. (3), the weights λ_{d} for fatigue and λ_{m} for maintenance are needed. These weights should be selected based on the momentary cost of the loads and the spot price of electricity. In this study, they are selected as 0.1 and 0.3, respectively. Thus, the function to be optimised is depicted in Figure 8. The optimal performance is 11.15 when turbine 1 is operating at 65% derating and turbine 2 is operating at 52% derating.

Figure 9 shows the BO iteration for the WPP economical operation problem. Ten (10) iterations are illustrated in Figure 9. To capture a rough trend of the surrogate, the initial derating set-points were selected as {100%, 100%} and {50%,50%}. The parameter κ for the acquisition function is set to 1. By maximising the acquisition function, the next query point was selected. After the six iterations, the algorithm stopped exploring as the uncertainty (or standard deviation) decreased. The acquisition function was roughly not changing and the query sampling location converged into the optimal point. Although the colour map of the surrogate mean does not exactly match that of the true cost function. But what is important is that the optimal power set-points were found by the BO algorithm within 6 iterations. Compared to the traditional surrogate modelling approach, 121 simulations were conducted to build a surrogate model finding the optimum, where turbine 1 and 2 were derated between 50% and 100% of the rated power by increments of 5%.

By inspecting Figure 9, one might argue that after iteration 10, some local optima were not captured or predicted correctly by the BO/GP surrogate model. This is reasonable. First, the most important task for the BO is to find to the global maximum of the function. Once the global maximum is found, it is more beneficial to have exploitation behaviour such that the next query samples are chosen close to the points that provide global maximum. Second, if the accuracy of the surrogate across the whole input space is important to the designer, the parameter in the acquisition function κ can be increased to promote the exploring behaviour of the algorithm such that the uncertainty or standard deviation of the surrogate would reduce across the entire input space.

If the BO algorithm ran longer, the control variables (power set-points) would stay at the peak of the surrogate mean. The cumulative performance is compared with the greedy control (no derating) and the BO algorithm in Figure 10. Notice that after iteration 6, the BO algorithm outperformed the greedy control.

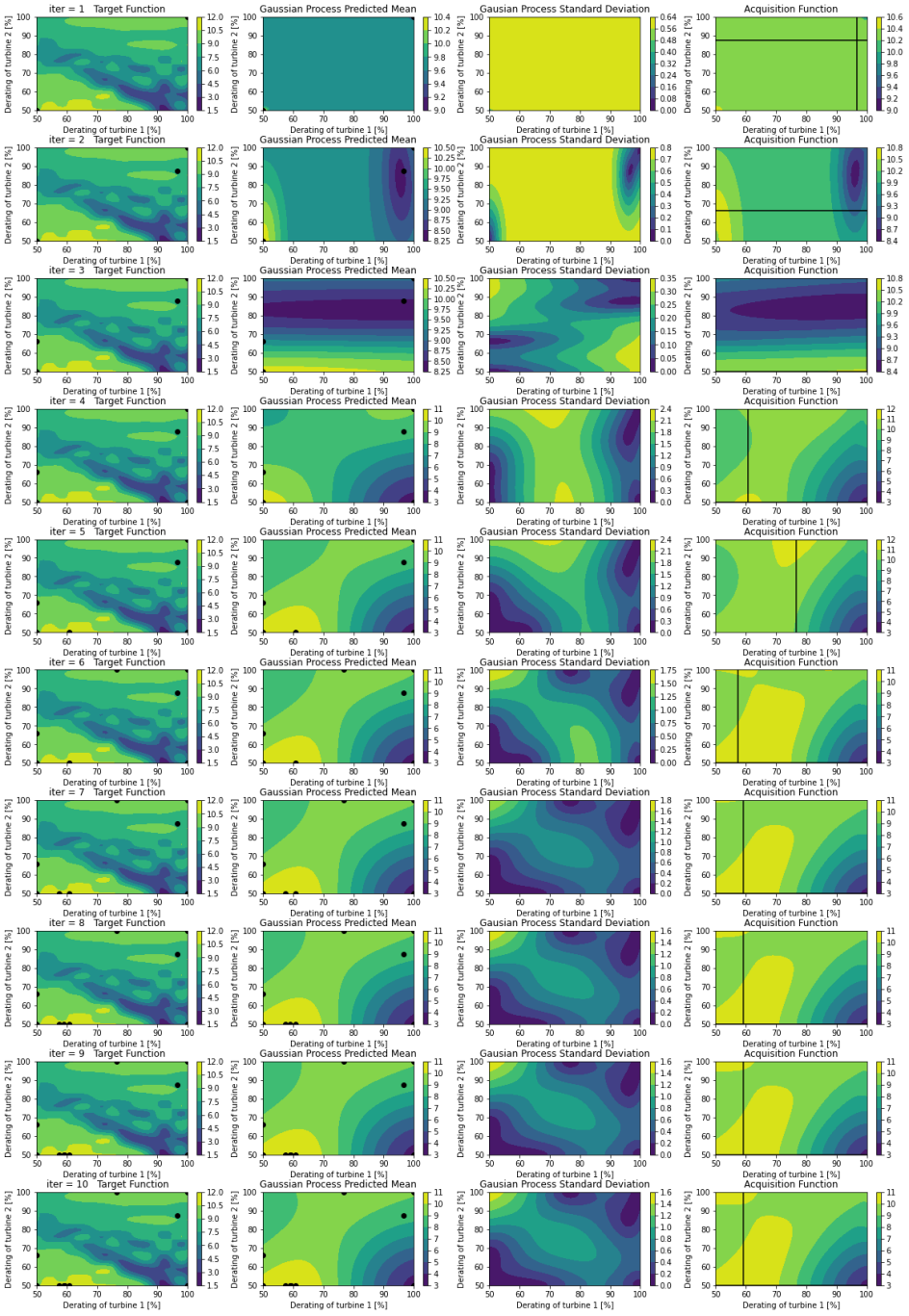

Figure 9: An example of WPP economical operation of three turbines by derating the first two upstream turbines. For each iteration (*iter*), Left: the true cost function; Middle-left: the mean of surrogate prediction fitted by the sample measurements (dots); Middle-right: the surrogate standard deviation Right: the next query sampling point (solid line) is determined by maximising the acquisition function.

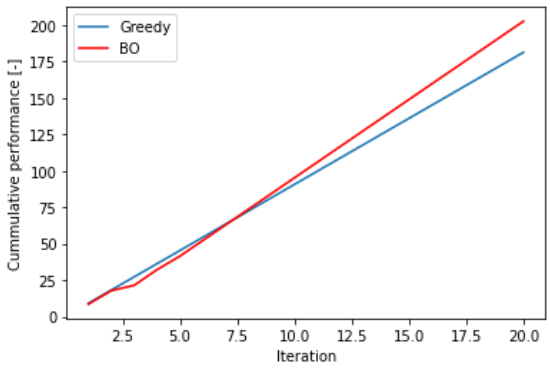

Figure 10: WPP economical operation of three turbines in a row. Cumulative performance of the greedy and BO strategies.

4.3.2 Active power control

This section demonstrates the application of BO to active power control problem. In active power control, the wind farm is commanded to provide a fixed level of power by the grid. For brevity, this study assumes that the power requirement by the grid to the wind farm was not changed during the optimisation. It also implies the grid command signal was not changing over 10-min. This assumption is imposed due to the fact that the fatigue load calculation is based on 10 minutes average. The assumption can be eased if the fatigue damage can be computed in real-time using online estimation techniques e.g. [47].

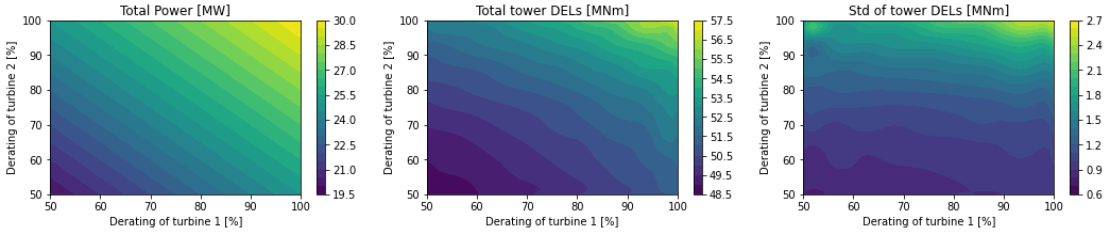

Figure 11: True functions of the 10-minute mean of wind farm power output (left), the sum (middle) and standard deviation (right) of the tower DELs of three turbines under a wind flow with a mean wind speed of 15 m/s.

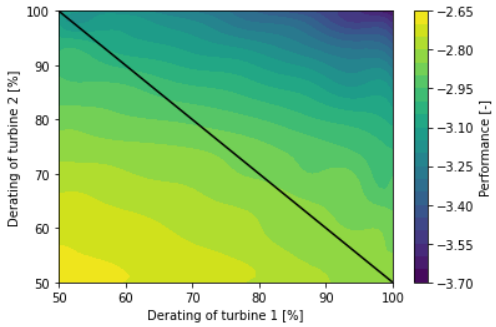

Figure 12: True cost functions to be maximised for WPP active power control. The solid line denotes the power constraint by the grid where the degrees-of-freedom can operate.

For illustration purposes and brevity, only two degrees of freedom are considered in the wind farm. Turbine 3 operates at greed control strategy. In addition, the mean wind speed in this

study is 15 m/s, which ensures all wind turbines can operate at the rated power. The command power in this study is 25 MW.

Notice that for cases where the wind speed is not sufficient for turbines to produce the rated power, turbine wind speed measurement is needed to calculate available power from the wind. This available power is then used as the bound in the optimisation in order to ensure the optimised power set-point is within the available power limit.

Figure 11 shows the true functions of 10-min mean wind farm power, total and standard deviation of the tower DELs. The function of the mean power output is linear with respect to the degrees of freedom of control variables. It is expected as the wind condition is sufficient for turbines to produce the rated power. The 10-min mean power of the wind farm decreases linearly as the derating happens in turbine 1 or 2. In contrast, the fatigue loads (DEL and standard deviation) are nonlinear with respect to the degrees of freedom. This is because the fatigue damage is a result of the complicated wake interactions between turbines. In active power control, the control objective is to follow the power requirement from the grid while alleviating the turbine structural loads and maintenance costs. In this example, the weights $\lambda_{\mathrm{d}}, \lambda_{\mathrm{m}}$ for total DELs and standard deviation of DELs of turbine towers are 0.05 and 0.3, respectively. The true cost function to maximise is depicted in 12. Notice that the command power is 25 MW and turbine 3 is operating at 10 MW. Thus, the combination of 10-min mean power output of turbine 1 and 2 must be 15 MW, which is represented by the solid line in Figure 12.

Figure 13 shows the BO optimisation process. To capture a general trend of the surrogate, two initial samples were selected and constrained by the grid requirement, which are {100%,50%} and {50%,100%} of the rated power for turbine 1 and 2. The maximum point of the cost function is -2.82 at power set-points {94%,56%}. In Figure 13, due to the constraint, the global maximum of acquisition function cannot be reached as illustrated in iteration 2. Thus, the next query point is determined by maximising the maximum along with the constraint instead. Before iteration 8, the algorithm shows exploring behaviour as some areas are highly uncertain along the power constraint. After iteration 8, the optimum is reached and the optimal power set-point is found. By inspecting the colour map of the surrogate mean, it looks relatively different to the true function across the input space except for the sampling location along with the constraint. This is due to the power constraint from the grid. If the grid power constraint changes over time, this allows sampling happening in more 'exotic' locations, that can help improving the surrogate predictions across the input space.

However, the function along the constraint is well captured by the surrogate, which is shown in Figure 14. The maximum value of the surrogate mean function matches well with that of the true cost function. Most importantly, the last sampling point is around the true optimal set-point that maximises the cost function.

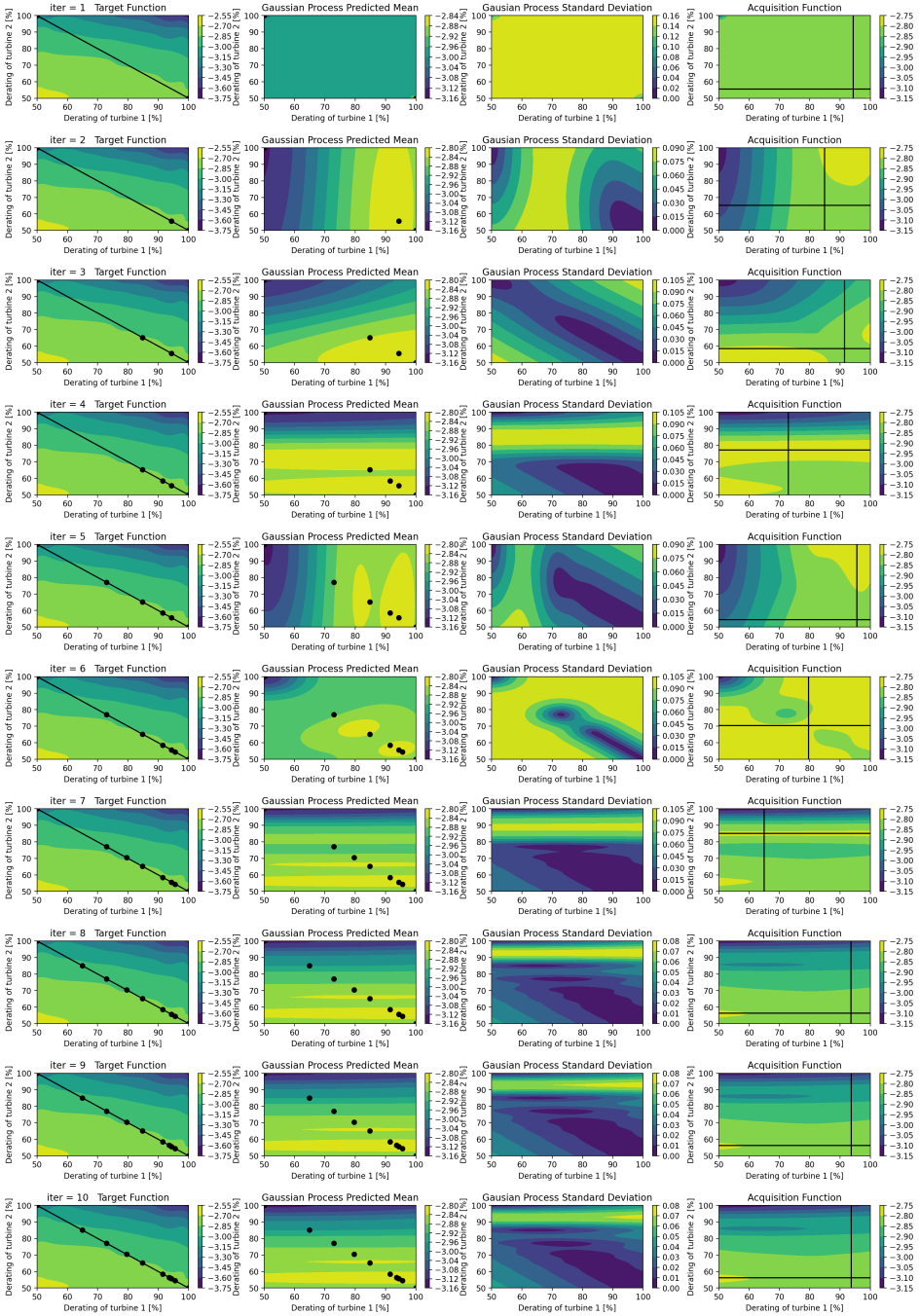

Figure 13: An example of WPP active power control of three turbines by derating the first two upstream turbines. For each iteration (*iter*), Left: the true cost function and power constraints from the grid (solid line); Middle-left: the prediction mean of the surrogate fitted by the sample measurements (dots); Middle-right: the surrogate standard deviation Right: the next query sampling point (solid line) is determined by maximising the acquisition function.

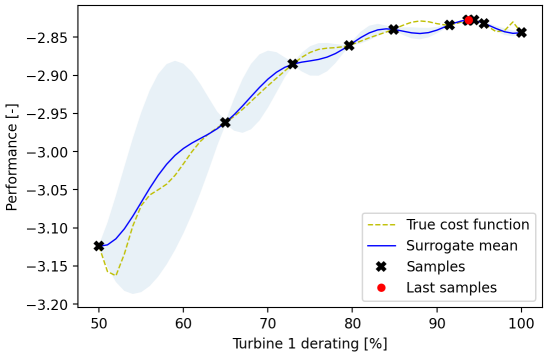

Figure 14: Active power control. True function and surrogate mean function along the constraint. The light blue area highlights the 95% confidence interval of the surrogate mean. The sampling locations are denoted by the crosses. The last sample location is highlighted by the red dot.

4.4 Conclusions and future work

This study presented a Bayesian optimisation with Gaussian process framework to a surrogate model-based control problem. The framework was applied to examples of two and three turbines in a row. The proposed framework was able to handle three common wind farm control objectives, namely (i) WPP power maximisation; (ii) WPP economical operation; (iii) active power control. To demonstrate the power of the BO framework, data from a middle fidelity wind farm simulation software, HAWC2Farm, were used. By optimising the sampling strategy and balancing between exploitation and exploration, the BO framework managed to construct a reliable surrogate model that can be used to achieve the control objective using only a few trials. The efficiency and benefit of the BO become noticeably different to traditional surrogate modelling approaches for problems with higher input dimensions.

In this study, the wind condition is assumed to be unchanged during the execution of the BO algorithm. In a wind farm, the wind condition changes gradually, for example, the wind conditions might change every hour. Thus, 10 iterations for the BO to find the optimum and each iteration that takes 10 minutes might not be infeasible in real-life. Future work will investigate the possibility to shorten the time of each iteration by using online estimation of fatigue damage techniques [47]. Another direction could be to use adaptive BO and treat the wind farm problem as a dynamic optimisation problem [48]. Moreover, the BO framework can be extended to a larger wind farm such as the TotalControl reference wind power plant in the future.

References

[1] S Boersma, B.M. Doekemeijer, P.M.O. Gebraad, P.A. Fleming, J Annoni, A.K. Scholbrock, J.A. Frederik, and J-W. van Wingerden. A tutorial on control-oriented modeling and control of wind farms. In *2017 American Control Conference (ACC)*, pages 1–18. IEEE, 2017.

[2] Jennifer Annoni, Pieter M. O. Gebraad, Andrew K. Scholbrock, Paul A. Fleming, and Jan-Willem van Wingerden. Analysis of axial-induction-based wind plant control using an engineering and a high-order wind plant model. *Wind Energy*, 19(6):1135–1150, 2016.

[3] D. Astrain Juangarcia, I. Eguinoa, and T. Knudsen. Derating a single wind farm turbine for reducing its wake and fatigue. *Journal of Physics: Conference Series*, 1037(3), 2018.

[4] Daan Van Der Hoek, Stoyan Kanev, and Wouter Engels. Comparison of Down-Regulation Strategies for Wind Farm Control and their Effects on Fatigue Loads. *Proceedings of the American Control Conference*, 2018-June:3116–3121, 2018.

[5] Wai Hou Lio, Christos Galinos, and Albert Meseguer Urban. Analysis and design of gain-scheduling blade-pitch controllers for wind turbine down-regulation *. In *2019 IEEE 15th International Conference on Control and Automation (ICCA)*, pages 708–712. IEEE, 2019.

[6] Steffen Raach, Jan-Willem van Wingerden, Sjoerd Boersma, David Schlipf, and Po Wen Cheng. Hinf controller design for closed-loop wake redirection. In *2017 American Control Conference (ACC)*, pages 703–708. IEEE, 2017.

[7] Paul Fleming, Jennifer Annoni, Andrew Scholbrock, Eliot Quon, Scott Dana, Scott Schreck, Steffen Raach, Florian Haizmann, and David Schlipf. Full-Scale Field Test of Wake Steering. *Journal of Physics: Conference Series*, 854:012013, 2017.

[8] Wim Munters and Johan Meyers. Effect of wind turbine response time on optimal dynamic induction control of wind farms. *Journal of Physics: Conference Series*, 753:052007, 2016.

[9] Wai Hou Lio, Mahmood Mirzaei, and Gunner Chr. Larsen. On wind turbine down-regulation control strategies and rotor speed set-point. *Journal of Physics: Conference Series*, 1037:032040, 2018.

[10] N O Jensen. A note on wind generator interaction. Technical report, Risø National Laboratory, Roskilde, Denmark, 1983.

[11] Sten Frandsen, Rebecca Barthelmie, Sara Pryor, Ole Rathmann, Søren Larsen, Jørgen Højstrup, and Morten Thøgersen. Analytical modelling of wind speed deficit in large offshore wind farms. *Wind Energy*, 9(1-2):39–53, 2006.

[12] Gunner C Larsen. A simple stationary semi-analytical wake model. Technical Report August, Risø National Laboratory for Sustainable Energy, Technical University of Denmark. (Denmark., 2009.

[13] Majid Bastankhah and Fernando Porté-Agel. A new analytical model for wind-turbine wakes. *Renewable Energy*, 70:116–123, 2014.

[14] Søren Ott, Jacob Berg, and Morten Nielsen. Linearised CFD Models for Wakes. 2011.

[15] Pieter M.O. Gebraad and J. W. Van Wingerden. A control-oriented dynamic model for wakes in wind plants. *Journal of Physics: Conference Series*, 524(1), 2014.

[16] Mads M. Pedersen and Gunner C. Larsen. Integrated wind farm layout and control optimization. *Wind Energy Science*, 5(4):1551–1566, 2020.

[17] Mehdi Vali, Vlaho Petrović, Sjoerd Boersma, Jan Willem van Wingerden, Lucy Y. Pao, and Martin Kühn. Adjoint-based model predictive control for optimal energy extraction in waked wind farms. *Control Engineering Practice*, 84(September 2018):48–62, 2019.

[18] Mehdi Vali, Vlaho Petrovi, Lucy Y Pao, and K Martin. Model Predictive Active Power Control for Optimal Structural Load Equalization in Waked. pages 1–16, 2021.

[19] Jonas Kazda and Nicolaos A. Cutululis. Model-optimized dispatch for closed-loop power control of waked wind farms. *IEEE Transactions on Control Systems Technology*, 28(5):2029–2036, 2020.

[20] Jonas Kazda and Nicolaos Antonio Cutululis. Fast control-oriented dynamic linear model of wind farm flow and operation. *Energies*, 11(12), 2018.

[21] Kathryn Johnson and Geraldine Fritsch. Assessment of extremum seeking control for wind farm energy production. *Wind Engineering*, 36(6):701–716, 2012.

[22] Jason R. Marden, Shalom D. Ruben, and Lucy Y. Pao. A Model-Free Approach to Wind Farm Control Using Game Theoretic Methods. *IEEE Transactions on Control Systems Technology*, 21(4):1207–1214, 2013.

[23] P. M. O. Gebraad and J. W. van Wingerden. Maximum power-point tracking control for wind farms. *Wind Energy*, 18(3):429–447, 2015.

[24] Jan-Willem van Wingerden, Lucy Pao, Jacob Aho, and Paul Fleming. Active Power Control of Waked Wind Farms. *IFAC-PapersOnLine*, 50(1):4484–4491, 2017.

[25] Mehdi Vali, Vlaho Petrović, Gerald Steinfeld, Lucy Y. Pao, and Martin Kühn. An active power control approach for wake-induced load alleviation in a fully developed wind farm boundary layer. *Wind Energy Science*, 4(1):139–161, 2019.

[26] Christos Galinos, Jonas Kazda, Wai Hou Lio, and Gregor Giebel. T2FL: An efficient model for wind turbine fatigue damage prediction for the two-turbine case. *Energies*, 13(6), 2020.

[27] Nikolay Dimitrov. Surrogate models for parameterized representation of wake-induced loads in wind farms. *Wind Energy*, 22(10):1371–1389, 2019.

[28] Georgios Gasparis, Wai Hou Lio, and Fanzhong Meng. Surrogate Models for Wind Turbine Electrical Power and Fatigue Loads in Wind Farm. *Energies*, 13(23):6360, 2020.

[29] Yingming Liu, Yingwei Wang, Xiaodong Wang, Jiangsheng Zhu, and Wai Hou Lio. Active power dispatch for supporting grid frequency regulation in wind farms considering fatigue load. *Energies*, 12(8), 2019.

[30] Jinkyoo Park and Kincho H. Law. Bayesian Ascent: A Data-Driven Optimization Scheme for Real-Time Control with Application to Wind Farm Power Maximization. *IEEE Transactions on Control Systems Technology*, 24(5):1655–1668, 2016.

[31] Bart M. Doekemeijer, D'Aan C.Van Der Hoek, and Jan Willem Van Wingerden. Model-based closed-loop wind farm control for power maximization using Bayesian optimization: A large eddy simulation study. *CCTA 2019 - 3rd IEEE Conference on Control Technology and Applications*, pages 284–289, 2019.

[32] Nikolay Dimitrov and Anand Natarajan. Wind farm set point optimization with surrogate models for load and power output targets. *Journal of Physics: Conference Series*, 2018(1), 2021.

[33] Morten Hartvig Hansen and Lars Christian Henriksen. Basic DTU Wind Energy controller. Technical Report January, 2013.

[34] Mahmood Mirzaei, Mohsen Soltani, Niels K. Poulsen, and Hans H. Niemann. Model based active power control of a wind turbine. In *2014 American Control Conference*, number July 2016, pages 5037–5042. IEEE, 2014.

[35] Jacob Aho, Andrew Buckspan, and Jason H. Laks. A tutorial of wind turbine control for supporting grid frequency through active power control. *Proc. of the American Control Conference*, pages 3120–3131, 2012.

[36] Bartolomé Manobel, Frank Sehnke, Juan A. Lazzús, Ignacio Salfate, Martin Felder, and Sonia Montecinos. Wind turbine power curve modeling based on Gaussian Processes and Artificial Neural Networks. *Renewable Energy*, 125:1015–1020, 2018.

[37] E Hart, W E Leithead, and J Feuchtwang. Wind turbine C p,max and drivetrain-losses estimation using Gaussian process machine learning. *Journal of Physics: Conference Series*, 1037:032024, 2018.

[38] Ravi Kumar Pandit and David Infield. SCADA-based wind turbine anomaly detection using Gaussian process models for wind turbine condition monitoring purposes. *IET Renewable Power Generation*, 12(11):1249–1255, 2018.

[39] Yanhua Liu, Ron J. Patton, and Shuo Shi. Wind Turbine Load Mitigation Using MPC with Gaussian Wind Speed Prediction. *2018 UKACC 12th International Conference on Control, CONTROL 2018*, pages 32–37, 2018.

[40] Wai Hou Lio, Ang Li, and Fanzhong Meng. Real-time rotor effective wind speed estimation using Gaussian process regression and Kalman filtering. *Renewable Energy*, 169:670–686, 2021.

[41] C.E. Rasmussen and C.K. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006.

[42] Torben Juul Larsen and Anders Melchior Hansen. How 2 HAWC2, the user's manual. Technical report, 2019.

[43] Gunner C Larsen, Helge Aa Madsen, J Torben, and Niels Troldborg. Wake modeling and simulation. Technical Report July, Risø National Laboratory for Sustainable Energy, Risø-R-1653(EN), 2008.

[44] Fanzhong Meng, Wai Hou Lio, and Thanasis Barlas. DTUWEC: an open-source DTU Wind Energy Controller with advanced industrial features. *Journal of Physics: Conference Series*, 1618:022009, 2020.

[45] C. Bak, F. Zahle, R. Bitsche, A. Yde, L. C. Henriksen, A. Natarajan, and M. H. Hansen. Description of the DTU 10 MW Reference Wind Turbine. *DTU Wind Energy Report-I-0092*, 2013.

[46] IEC. IEC 61400-1 Ed.3: Wind turbines - Part 1: Design requirements. 2005, 2005.

[47] A. Cetrini, F. Cianetti, M.L. Corradini, G. Ippoliti, and G. Orlando. On-line fatigue alleviation for wind turbines by a robust control approach. *International Journal of Electrical Power Energy Systems*, 109(October 2018):384–394, 2019.

[48] Favour M Nyikosa, Michael A Osborne, and Stephen J Roberts. Bayesian Optimization for Dynamic Problems. 2018.