Real-time optimization of wind farms using modifier adaptation and machine learning

Leif Erik Andersson[1] and Lars Imsland[1]

[1]Norwegian University of Science and Technology, Department of Engineering Cybernetics, 7491 Trondheim, Norway

Correspondence: Leif Erik Andersson (leif.e.andersson@ntnu.no)

Abstract. Coordinated wind farm control takes the interaction between turbines into account and improves the performance of the overall wind farm. Accurate surrogate models are the key to model-based wind farm control. In this article a modifier adaptation approach is proposed to improve surrogate models. The approach exploits plant measurements to estimate and correct the mismatch between the surrogate model and the actual plant. Gaussian process regression, which is a probabilistic non-parametric modelling technique, is used in the identification of the plant-model mismatch. The efficacy of the approach are illustrated in several numerical case studies. Moreover, challenges in applying the approach to a real wind farm with a truly dynamic environment are discussed.

1 Introduction

Currently wind turbines in a wind farm are operated to maximise their power production and minimise the loads on their structure and power electronics. The impact on the downstream turbines due to wake interactions is ignored. Such a control strategy is called *greedy* since it only focuses on the operation of an individual wind turbine. It is expected that a wind farm control strategy that takes the interaction between turbines into account can improve the overall performance of the wind farm (Steinbuch et al., 1988; Johnson and Thomas, 2009; Barthelmie et al., 2009).

The two main wind farm control strategies are axial induction control and wake steering (Kheirabadi and Nagamune, 2019). The idea behind axial induction control is to deviate the blade pitch and generator torque of the upwind turbine from the greedy control settings. As a consequence, the velocity deficit in the wake behind the turbine decreases. The target net effect is an overall increase of the power production and possibly a decrease of fatigue loads. However, evaluating wind tunnel experiments (Campagnolo et al., 2016; Bartl and Sætran, 2016), high-fidelity simulations (Annoni et al., 2016) and field tests (van der Hoek et al., 2019) it is suggested that axial induction control using steady-state surrogate models to calculate the optimal control settings may be unable to improve the power production of a wind farm.

Currently the more promising wind farm control strategy using stead-state surrogate models is wake steering. The goal of wake steering is to deflect the wake away from the downwind turbine by using the yaw settings of the upwind turbine (Kheirabadi and Nagamune, 2019). Field experiments showing encouraging results were conducted by Fleming et al. (2017, 2019); Howland et al. (2019). In these experiments lookup tables with optimal yaw settings depending on the wind conditions were created using steady-state models. The look-up tables were not updated using plant measurements. Therefore, these approaches can be

seen as open-loop.

The steady-state surrogate models must be simple to allow optimization but also accurate to permit good performance of the model-based controller. The development of surrogate models is an active research field. One of the most popular wake models is the Jensen Park model (Jensen, 1983; Katic et al., 1987). Jiménez et al. (2010) developed one of the first steady-state wake models that described wake deflection due to yaw. A recent wake model, which is also used in this study, was presented by Bastankhah and Porté-Agel (2016). It is based on mass and momentum conservation and assumes a Gaussian distribution of the velocity deficit in the wake. Other extensions to the Jensen Park model were presented by Park and Law (2015), who assumed an inverted Gaussian function of the wake profile, Tian et al. (2015), who used a cosine shape function, and Ge et al. (2019) who analytically derived a Gaussian-shape velocity profile. The steady-state wake models are able to describe the general behaviour of the wake (Barthelmie et al., 2013; Annoni et al., 2014). Nevertheless, they are just vague approximations of a complex phenomena that is, in fact, not well understood (Veers et al., 2019).

Model-free methods using Extremum-seeking (Johnson and Fritsch, 2012; Ciri et al., 2017) or game-theoretic methods (Marden et al., 2013; Gebraad et al., 2013) were proposed to circumvent possible error-prone models in the control of wind farms. However, these methods suffer from slow convergence. Park et al. (2016, 2017) suggested to use a Bayesian Ascent (BA) algorithm fitting a Gaussian Process (GP) regression to input-output data of the plant. A new data-driven surrogate model was created. In Doekemeijer et al. (2019a) the upstream wind velocity and turbulence intensity in the FLORIS model are first estimated from the data. The improved FLORIS model is then used in Bayesian optimization to find a GP surrogate model and optimal yaw angles of the turbines in the wind farm. Another data-driven surrogate model, using polynomial chaos expansion, was presented by Hulsman et al. (2019). Estimating the model parameters of the surrogate model to improve closed-loop control was proposed by Doekemeijer et al. (2019b). However, if the parametric model is structurally incorrect parameter estimation is not able to remove the mismatch between surrogate model and the plant. An example that an improved parameterisation of a surrogate model was not able to remove the mismatch between a low order model and a high fidelity model is given in Fleming et al. (2018). Therefore, a two-step approach iteratively optimizing the plant and updating the model parameter of the surrogate model as plant measurements become available was not pursued here.

Instead, in this article a modifier adaptation (MA) approach (Marchetti et al., 2016) to wind farm control is proposed. The plant-model mismatch is identified exploiting plant measurements, improving the surrogate model. In the identification of the plant-model mismatch GP regression is used. GP is a probabilistic, non-parametric modelling technique well known in the machine learning community (Rasmussen and Williams, 2006). The advantage of using GP regression in MA is that it is not bounded by specific model structures as e.g. parametric models. Consequently, the MA-GP approach is able to correct the surrogate model in a flexible manner (de Avila Ferreira et al., 2018) and improve the performance of the wind farm controller. The article is structured as follows: In Section 2 the optimization problem is formulated. In Section 3 the modifier adaptation using Gaussian process regression is presented and the numerical turbine and wake models are introduced. The approach is tested numerically in Section 4. Section 5 discusses the application of the MA-GP approach to real wind farms. The article ends with a conclusion.

2 Problem formulation

Model-based wind farm optimization usually employs a steady-state surrogate model. Consequently, a plant-model mismatch exists, which can degrade the performance of a controller. In this article, we study the optimization problem of optimizing the power production, noting that the approach in general can handle different objective functions. The optimization problem can be formulated as

$$\mathbf{u}_p^* = \arg\max_{\mathbf{u}} P_p(\mathbf{u}), \quad \mathbf{u} \in \mathcal{U}, \ \ \mathbf{u} = [u_1^T, u_2^T, \ldots, u_N^T]^T, \tag{1}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$ denote the plant input variables, which are the axial induction factors and yaw angles of each turbine; $P_p : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$ is the power production to be maximised; and $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ is the control domain, e.g. box constraints on the control inputs.

The challenge of optimizing the power production of a wind farm is that only an approximate surrogate model of the plant is

available. Consequently, it is not guaranteed that the optimal point of the surrogate model coincide with the optimal point of the plant. MA treats this challenge by directly adapting the optimization problem using plant measurement to allow convergence to the overall plant optimum (Marchetti et al., 2009). The standard MA adds first order modifiers to correct the gradient of the surrogate model. However, the estimation of the plant gradients in each iteration is experimentally expensive and the main bottleneck of the MA implementation in practice (Marchetti et al., 2016). In this article, GPs are used instead to correct the

surrogate model (de Avila Ferreira et al., 2018), and by this alleviating the limitation of MA. The next section gives a brief introduction to GPs, before the new optimization problem of the MA-GP approach is stated.

3 Methodology

In this section the modifier adaptation approach with Gaussian processes for wind farm control is introduced in Sections 3.1 and 3.2. Thereafter, in Section 3.3, the turbine and wake models used in the case study are explained.

3.1 Gaussian processes

In this section we give a brief outline of GP regression, for more information consult Rasmussen and Williams (2006). GP regression identifies an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ from data. It is assumed that the noisy observations of $f(\cdot)$ are given by:

$$y_k = f(\mathbf{u}_k) + \nu_k \tag{2}$$

where the value $f(\cdot)$ is perturbed by Gaussian noise ν_k with zero mean and variance σ_ν^2, $\nu_k \sim \mathcal{N}(0, \sigma_\nu^2)$.

In GP regression, $f(\cdot)$ is considered a distribution over functions. In this paper, we assume this distribution has a zero mean function and the squared-exponential (SE) covariance function. The choice of the mean and covariance functions assume certain smoothness and continuity properties of the underlying function (Snelson and Ghahramani, 2006), which seems to be

a good fit for the plant-model mismatch of the surrogate model. The SE covariance function can be expressed as follows:

$$k(\mathbf{u}_i, \mathbf{u}_j) = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{u}_i - \mathbf{u}_j)^T \mathbf{\Lambda}^{-1} (\mathbf{u}_i - \mathbf{u}_j)\right) \tag{3}$$

where σ_f^2 is the covariance magnitude and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_{n_u}^2)$ is a scaling matrix.

Due to the GP assumption the predictive distribution of $f(\mathbf{u})$ at an arbitrary input $\mathbf{u}$ given the training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ has a closed-form solution. The resulting mean $\mu_{\mathrm{GP}}(\mathbf{u}; \mathcal{D}, \hat{\psi})$ can be seen as the GP prediction at $\mathbf{u}$ and the variance $\sigma_{\mathrm{GP}}^2(\mathbf{u}; \mathcal{D}, \hat{\psi})$ as a corresponding measure of uncertainty to this prediction.

The performance of the GP is dependent on hyperparameters $\hat{\psi}$. They are commonly unknown and hence need to be inferred from data. In this article the maximum likelihood estimate is used to calculate the hyperparameters. Finally, we note that the training data are explicitly required to construct the predictive distribution. For this a matrix of size $M \times M$ must be inverted, where M is the number of measurements. Clearly, this makes large data sets challenging.

3.2 Modifier Adaptation with Gaussian processes

In the MA-GP approach the limitation of standard MA are overcome by replacing the modifiers with GPs (de Avila Ferreira et al., 2018). As a result, estimating the plant gradients (modifiers) in each iteration are avoided, at the cost of instead updating the GP. The optimization problem of the MA-GP becomes

$$\hat{\mathbf{u}}_{k+1}^* = \arg\max_{\mathbf{u}} P(\mathbf{u}) + \mu_{\mathrm{GP},k}(\mathbf{u}; \mathcal{D}_k, \hat{\psi}_k), \qquad s.t. \quad \mathbf{u} \in \mathcal{U}, \tag{4}$$

where the plant-model mismatch of the cost function is modelled by μ_{GP}. The training set $\mathcal{D}$ of the GP are the control inputs of the wind farm and the difference in the power production between surrogate model and plant measurements.

The MA-GP approach for wind farm optimization is visualised in Fig. 1. The power output of the surrogate model is subtracted from the noisy power measurements of the plant. The difference in power production and the control inputs create the data set, which is used in the GP training to estimate the hyperparameters. A initial training set is required before initialising the MA-GP approach. In the plant optimization the surrogate model is corrected by the GP regression model, which uses the current data set and hyperparameters. The new optimal control input is applied to the wind farm. The MA-GP is a closed-loop control approach to wind farm optimization.

In Algorithm 1 two additional steps are included in the MA-GP scheme:

– The new optimal control input is filtered with

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{L}(\hat{\mathbf{u}}_{k+1} - \mathbf{u}_k), \quad \mathbf{L} = \mathrm{diag}(l_1, \ldots, l_{n_u}), \quad l_i \in (0, 1]. \tag{5}$$

In the basic MA approach filtering the control input prevents excessive corrections. In the MA-GP approach it permits exploration around the optimal point.

– The hyperparameters are only updated when *HypOpt* is true, which is a user-defined condition. The hyperparameter update is usually the computational bottle-neck of the MA-GP algorithm. We observed that especially for large data

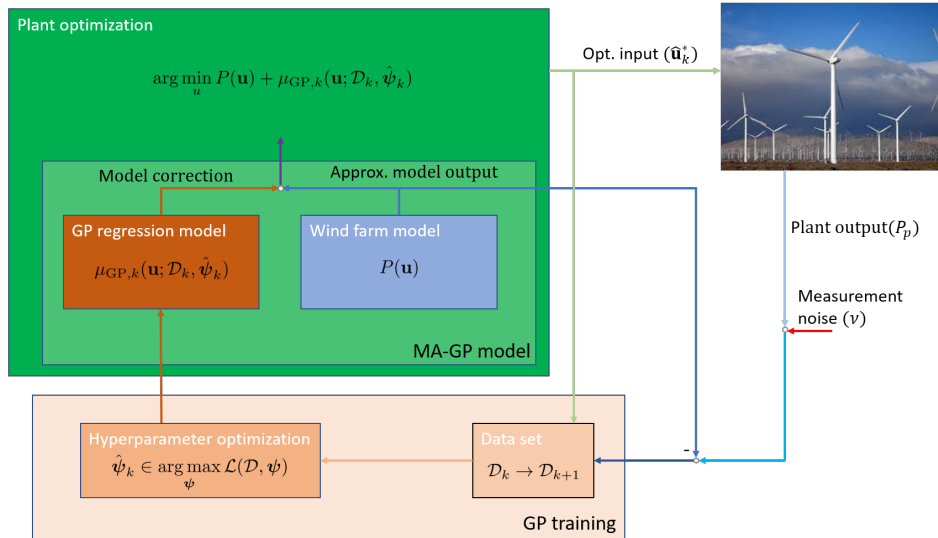



Figure 1. The basic idea of the MA-GP scheme for a wind farm. The GP regression model creates an input-output map of the control inputs to the plant-model mismatch. In the MA-GP model the GP regression model is used to correct the output of the approximate model. This MA-GP model is used in the optimization to compute optimal control inputs for the wind farm. The inputs and the difference between the measured and estimated output of plant and model, respectively, are used to update the data set $\mathcal{D}$ and the hyperparameter ψ. The measured outputs of the plant are corrupted by noise [1]

sets the hyperparameters do not change much from one iteration to the next. Therefore, the hyperparameters can be
120   updated less frequent to decrease computational delay. However, it is recommend to update the hyperparameters as often
as possible.

In the next subsection the turbine and wake models used in the case study are presented.

3.3   Numerical turbine and wake models

A turbine and wake model are necessary to create a model of a wind farm. The wind turbines are represented using the actuator
125   disc theory, which couples the power and thrust coefficient, C_P and C_T (Burton et al., 2011)

$$C_P = 4a(1-a)^2, \tag{6}$$
$$C_T = 4a(1-a), \tag{7}$$

where a is the axial induction factor. The axial induction factor indicates the ratio of wind velocity reduction at the turbine disk compared to the upstream wind velocity. The steady-state power of each turbine under yaw misalignment is given by (Gebraad

---

[1]The wind farm picture is by Erik Wilde from Berkeley, CA, USA https://www.flickr.com/photos/dret/24110028330/, *Wind turbines in southern California 2016*, https://creativecommons.org/licenses/by-sa/2.0/legalcode

Algorithm 1: Basic MA-GP scheme (del Rio-Chanona et al., 2019)

Initialisation: GP regression model μ_{GP} and hyperparameters $\hat{\psi}_0$ found with initial data set $\mathcal{D}_0$; Optimal operation point of the approximate model $\mathbf{u}_0$; $k = 0$;

while $t < t_{end}$ **do**

    Solve modified optimization problem Eq. (4);

    Filter new operating point $\mathbf{u}_{k+1}$ with Eq. (5);

    Evaluate approximate model at new operating point $\mathbf{u}_{k+1}$;

    Obtain power measurement $P_p(\mathbf{u}_{k+1})$;

    Update the data set $\mathcal{D}_{k+1}$ with input $\mathbf{u}_{k+1}$ and output $y_{k+1} = P_p(\mathbf{u}_{k+1}) - P(\mathbf{u}_{k+1})$. ;

    if *HypOpt* **then**

        Update hyperparameters $\hat{\psi}_{k+1}$ using the updated data set $\mathcal{D}_{k+1}$;

    end

    Update GP regression term μ_{GP} using $\mathcal{D}_{k+1}$ and hyperparameters $\hat{\psi}_{k+1}$;

    k = k+1, $t = t + \Delta t$;

end

---

130 et al., 2016)

$$P = \frac{1}{2}\rho A C_P \cos\gamma^p v^3, \tag{8}$$

where A is the rotor area, ρ the air density, p a correction factor and v is the wind velocity. In actuator disc theory $p = 3$ (Burton et al., 2011). However, based on large-eddy simulations, the turbine power yaw misalignment has been shown to match the output when $p = 1.88$ for the NREL 5MW turbine (Annoni et al., 2018), which we will use in this article. In the numerical

135 study it will be important to implement a "plant" and model, which are different from each other. The actuator disk model will be referred to as the *plant turbine model*. A second adjusted actuator disk turbine model is created, which will be referred to as the *approximate turbine model*. The FLORIS toolbox (NREL, 2019) contains a table with wind velocities and corresponding thrust and power coefficients of the NREL 5MW turbine. These data are fitted to create the approximate turbine model. The equation for the thrust coefficient C_T is given by Eq. (7) while for the power coefficient C_P three new parameter are identified

140 resulting in

$$C_P = 7.037a(0.625 - a)^{1.364}. \tag{9}$$

The approximate turbine model fit is visualised in Fig. 2. Important in the numerical example is the different connection between thrust and power coefficients of plant and approximate turbine model (Fig. 2b). For the turbine dimensions the NREL 5-MW wind turbine is used (Jonkman et al., 2009). Consequently, the rotor diameter is $D = 126.4\,\text{m}$ and the hub height

145 $H_H = 90\,\text{m}$.

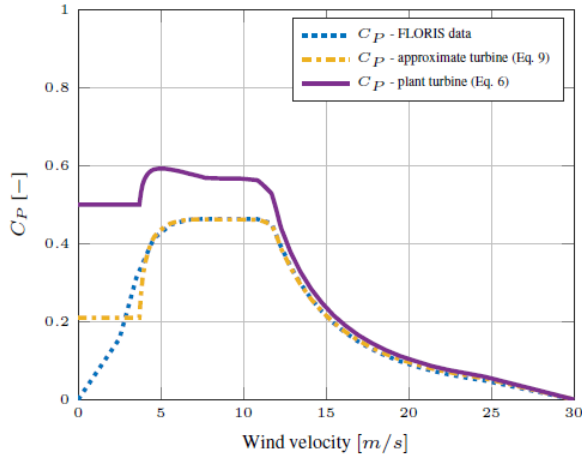
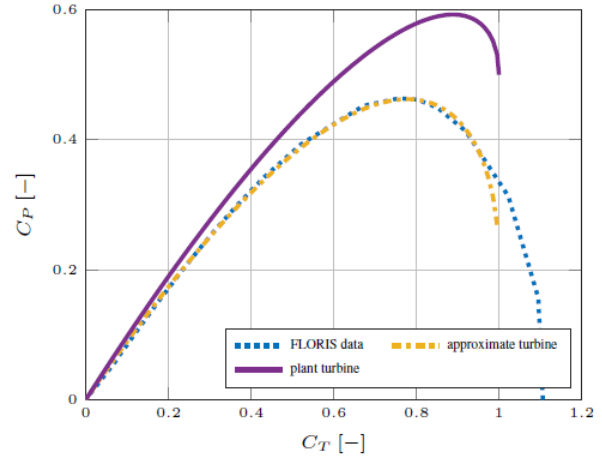
The Gaussian wake model by Bastankhah and Porté-Agel (2014, 2016) is used to model the flow in the wind farm. The

(a) Comparison of C_P values in dependency of the wind velocity.

(b) C_P - C_T curve of the data and both models.

Figure 2. Comparison between data, the plant turbine and the approximate turbine model. The models give different connections between thrust and power coefficients.

three-dimensional steady-state far wake velocity deficit is Gaussian distributed and can be estimated by

$$\frac{\bar{v}(x,y,z)}{\bar{v}_{\infty}} = 1 - Ce^{-0.5((y-\delta)/\sigma_y)^2} e^{-0.5((z-z_h)/\sigma_z)^2}, \tag{10a}$$

$$C = 1 - \sqrt{1 - \frac{C_T \cos\gamma}{8(\sigma_y\sigma_z/d^2)}}, \tag{10b}$$

where z_h is the tower height, δ is the wake deflection, and σ_y and σ_z are the wake widths in lateral and vertical directions. An important variable for the model is the skew angle of the flow past a yawed turbine. The flow skew angle is approximated by

$$\theta \approx \frac{\alpha_1 \gamma}{\cos\gamma} \left(1 - \sqrt{1 - C_T \cos\gamma}\right), \tag{11}$$

where α_1 is a parameter. Bastankhah and Porté-Agel (2016) use $\alpha_1 = 0.3$ and NREL (2019) uses $\alpha_1 = 0.6$ to better fit high-fidelity observations. We will use the Gaussian wake model with $\alpha_1 = 0.3$ as the *approximate wake model* and with $\alpha_1 = 0.6$ as the *plant wake model*.

In the next section the case study using the MA-GP approach and the here presented turbine and wake models is discussed.

4   Numerical case study

In this section numerical results of the MA-GP approach are presented. The control inputs of the wind farms are the yaw angles γ_i and the thrust coefficients $C_{T,i}$ of each turbine. Hence, the wind farm has $2N$ control inputs, where N is the number of wind turbines. The objective of the optimization is to maximize the power production $P_{tot} = \sum_i P_i$ of the wind farm. The relative

Table 1. Overview over the case studies

Case	Control inputs	Size of initial training set	measurement noise	Final error Θ (after x-iterations)
Two turbines	C_{T1}, γ_1	4	no	$0.0009\,\%(10)$
Tow turbines	C_{T1}, γ_1	20	yes	$0.6\,\%(10)$
Tow turbines	C_{T1}, γ_1	30	yes	$0.35\,\%(10)$
n turbines in a row	$\mathbf{C}_T, \boldsymbol{\gamma}$	$20n$	no	Fig. 6
$n \times n$ turbine grid	$\mathbf{C}_T, \boldsymbol{\gamma}$	$20n^2$	no	Fig. 7

error in the power production is given by

$$\Theta = 100 \frac{P_p^* - \hat{P}_p}{P_p^*}, \tag{12}$$

where P_p^* is the optimal power production of the plant and $\hat{P}_p$ is the power production achieved by the MA-GP approach. The control inputs are constrained by box constraints with

$$0 \leq C_{T,i} \leq 0.95, \quad \text{and} \ \ 0° \leq \gamma_i \leq 40°. \tag{13}$$

The yaw angles γ_i are constrained to positive yaw angles since the Gaussian wake model is symmetric. Asymmetry as in a real wind farm is not represented in the models used in this article. If the MA-GP approach is applied to a real wind farm it would be unnecessary to constrain the yaw angle to positive angles since the MA-GP approach would automatically converge to the superior yaw rotation.

The approximate turbine and wake models are used as the *approximate model* while the plant turbine and wake models are used as the *plant model*. In the MA-GP approach only measurements of the total power output of the wind farm are used. The hyperparameter optimization is performed using the MATLAB optimization toolbox and the nonlinear programming solver *fmincon*. For the optimization of the control inputs of the wind farm the open source software tool *CasADi* (Andersson et al., 2019) is used. *CasADi* is a symbolic framework that provides gradients using Algorithmic Differentiation. The software package *Ipopt* is used as a solver for the nonlinear program (Wächter and Biegler, 2006).

In the following three different wind farms are discussed:

- Two turbines, in which only the upstream turbine is controlled (Sec. 4.1)

- A row of turbines, in which all turbines are controlled (Sec. 4.2).

- A grid of turbines, in which all turbines are controlled (Sec. 4.3).

An overview over the case studies discussed in the following sections is given in Tab. 1.

4.1 Two turbine case

The operating points of two turbines in a row are optimized. The thrust and yaw angle of the downwind turbine are fixed resulting in only two optimization variables in the MA-GP approach. The downwind turbine is operated at its greedy operation

(a) Surface of objective function of plant. (b) Surface of objective function of approximate model.

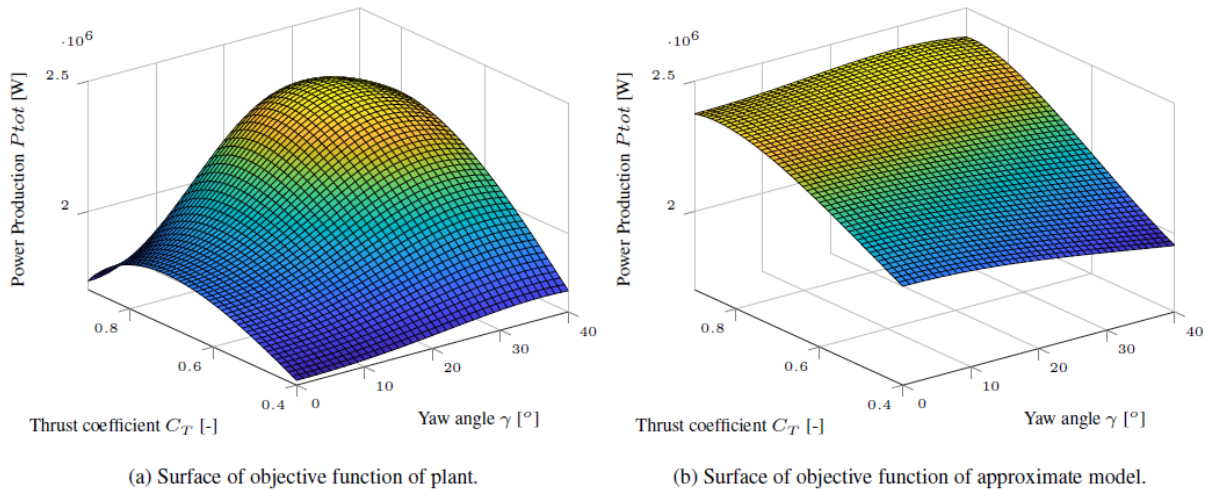



Figure 3. The power production of plant and approximate model in dependency of the control inputs of the upwind turbine.

point. The turbine row is facing the wind and the spacing between turbines is $5\,\mathrm{D}$. The power production of the wind farm
in dependency of the control inputs of the upwind turbine in shown in Fig. 3. The optimal operation point of the plant is
$C_{T,p} = 0.82$ and $\gamma_p = 31°$ and of the approximate model $C_{T,p} = 0.89$ and $\gamma_p = 29°$. Indeed, the relative optimization error of
the model is only $\Theta = 1.67\%$. Still, the model assumes that the power production is much less sensitive to changes in the yaw
angle, which should be corrected by the MA-GP approach.

Four training points at $C_T = [0.4, 0.8]^T$ and $\gamma = [0°, 25°]^T$ are used to create the initial training set of the GP regression model.
The power production of the corrected model in dependency of the control inputs is shown in Fig. 4a. The contour plot of the
objective function of the plant, approximate model and MA-GP model after the initial training is shown in Fig. 4b. Clearly
four operating points are not sufficient to correct the approximate model correctly. In fact, the optimal operating point of the
MA-GP model has an error of 2.87%, which is larger than the original error of the approximate model.

The MA-GP approach is initialised at the optimal operating point of the approximate model. In each iteration the hyper-
parameters and the data set of the GP regression model are updated. The new operating point is filtered with Eq. (5) and
$\mathbf{L} = \mathrm{diag}(0.4, 0.4)$. The MA-GP approach is able to correct the approximate model and drive the process to its optimal oper-
ating point (Fig 5). After four iterations the relative error Θ is about 0.2% and after ten iterations it is 0.0009%. In addition,
the contour lines of the objective function are well approximated (Fig. 5). A larger difference between MA-GP model and the
plant can be observed at the edges away from the current operating points. Data points at the edges are necessary to improve
the identification there. However, to drive the process to its optimal operating points a correct identification of the objective
function far away from the maximum is unnecessary. Clearly the initial training set with only four operating points could be
increased to improve the identification of the initial model of the MA-GP approach.

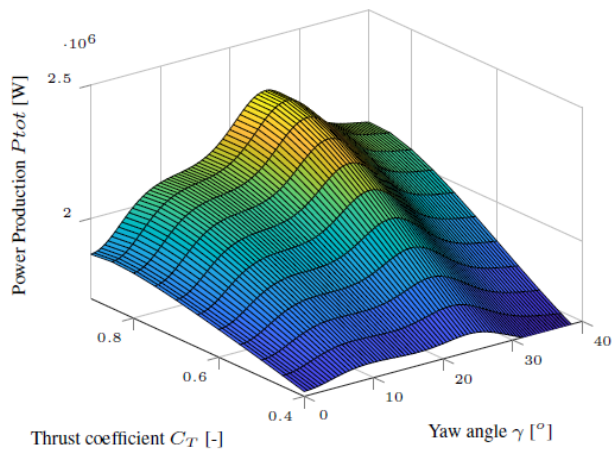
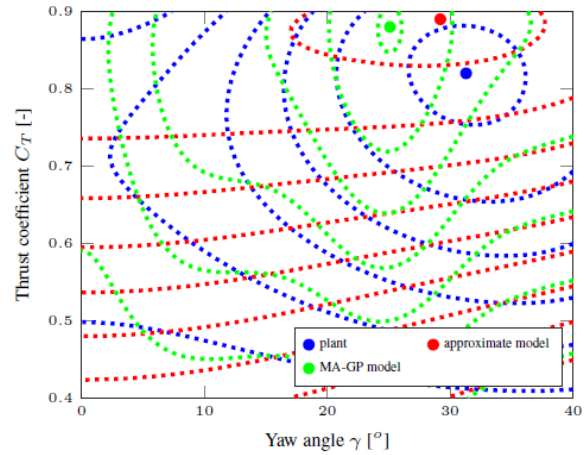
In the current example it was assumed that the measurements are noise-free. If noise is added to the power measurements the

(a) Surface of objective function of MA-GP corrected model after initial training.

(b) Contour plot.

Figure 4. The power production of MA-GP model in dependency of the control inputs of the upwind turbine and the contour plot of plant, approximate model and MA-GP model after the initial training.

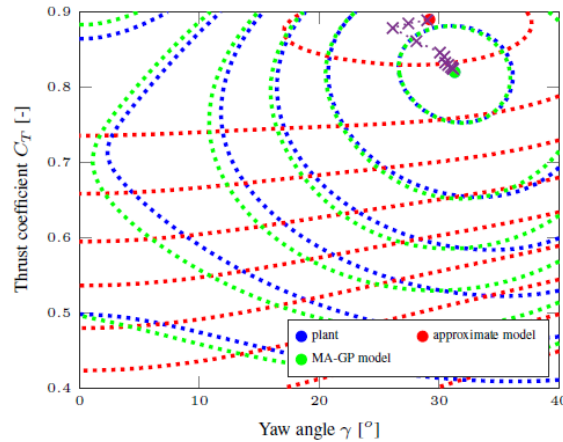



Figure 5. The contour plot of plant, approximate model and MA-GP model after ten iterations. The operating points of each iteration are marked with a cross.

correct identification becomes more challenging and a larger training data set is necessary. A noise with a standard deviation of
205 $50\,\text{kW}$ is added to the measurement, which in the current set-up translates to a turbulence intensity of about $3\,\%$. The standard deviation is of the same size as the error in the power production of plant and approximate model at the optimal operating point of the plant. A training data set of 20 points is created. After ten iterations the relative error Θ is about $0.6\,\%$. The algorithm

(a) Boxplot.
(b) Initial model error.

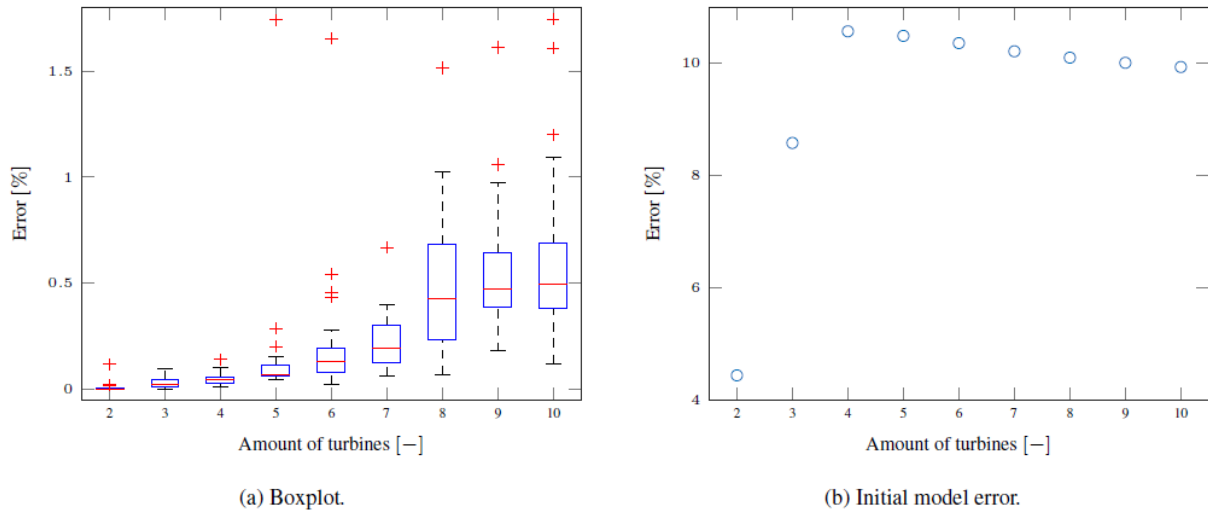



Figure 6. The boxplot of the optimization results for the differently long wind turbine rows on the left. The red line indicates the median. The bottom and top edges of the blue box indicate the 25th and 75th percentiles, respectively. The red marker indicate outliers and the whiskers extend to the most extreme data points not considered as outliers. The error of the MA-GP approach and the initial error dependent of the amount of turbines in the row. The initial error in the model depending on the amount of turbines in the row on the right.

is able to converge. However, due to the measurement noise a small error remains after ten iterations. The error can be easily decreased with a larger initial data set, e.g. with a training set of 30 points the error after ten iterations is about $0.35\,\%$.

210   4.2 n turbine row case

In this subsection the optimization of n turbines aligned in a row with a spacing of $5\,\mathrm{D}$ is discussed. It is difficult to know the required size of the training set for a satisfying performance of the MA-GP approach a priori. It depends on the sensitivity of the output to the input variables. It is, however, recommended to have about ten training points for each input (Loeppky et al., 2009). Therefore, the size of the initial training set is chosen to be $n_d = 10n_u$, where n_u is the amount of control inputs.
215   The operating points of the training set are chosen randomly using Latin hypercube sampling. The convergence of the MA-GP algorithm is tested on 25 Monte Carlo simulations. The difference between each run is the initial training set.

  The error increases with the amount of turbines while it is almost zero for 2 to 4 turbines (Fig. 6). A reason for the increase in the error with more turbines is the similar sensitivity of the control inputs of each turbine to the power output of the plant. It makes it challenging to correctly identity the input-output map. The error can be decreased with more data in the training set.
220   Currently, the optimization of the process and the optimization of the hyperparameters takes less than a second even for the ten turbine case. Consequently, it is possible to increase if available the data set.

Assuming a sufficient large initial training set the MA-GP approach is able to find the near optimal point in one iteration since the approach basically just improves the surrogate model. This stands in contrast to purely model-free approaches, e.g.

(a) Boxplot - initial data $n_d = 10n_u$. (b) Boxplot - initial data $n_d = 30n_u$.

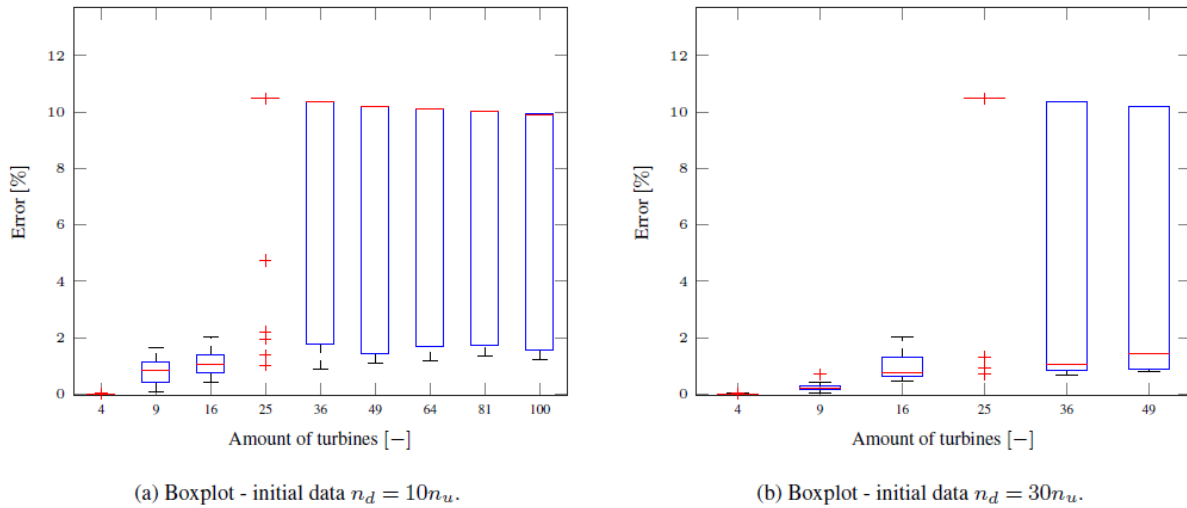



Figure 7. The boxplot of the optimization results for the differently large wind turbine grids. The red line indicates the median. The bottom and top edges of the blue box indicate the 25th and 75th percentiles, respectively. The red marker indicate outliers and the whiskers extend to the most extreme data points not considered as outliers. The error of the MA-GP approach and the initial error dependent of the amount of turbines in the row. The difference between both runs is the size of the initial training set.

extremum seeking (Johnson and Fritsch, 2012) or MPPT (Gebraad et al., 2013), which usually need several iterations to find
225 an optimum. Moreover, after the initial training the MA-GP model usually represents of the plant better than the approximate model. Nonetheless, measurements close to the optimum of the MA-GP model can help to improve the model further.

4.3 $n \times n$ turbine grid case

In this subsection the optimization of a wind farm with turbines arranged in a $n \times n$ grid with a spacing of $5\,\mathrm{D}$ is presented. Consequently, the wind farm consist of n turbine rows each containing n turbines. The wind direction is aligned with the rows
230 of the grid. Interaction between parallel rows is neglectable, which is, however, not known to the MA-GP approach. Again the size of the initial training set is chosen to depend linearly on the size of the amount of control inputs with $n_d = 10n_u$ and the MA-GP approach is tested on 25 Monte Carlo simulations.

Again the algorithm converges for a small amount of turbines (Fig. 7a). However, the error in the optimization increases as the amount of turbine increase. Moreover, for grids with 25 and more turbines the majority of the optimizations get stuck at the
235 initial conditions, which is defined by the optimal operation point of the model (Fig. 6b)[2]. This behavior might be caused by overfitting causing multiple local optima in the MA-GP model. Moreover, even in the cases where the MA-GP improves the performance of the wind farm the algorithm converges to errors in the range of $1\,\%$ to $2\,\%$ after 25 iterations. These are much larger than observed in the turbine row case.

---

[2]The percentage in initial error of the turbine row (Fig. 6b) is equal to the percentage in initial error of the grid.

If the MA-GP algorithm for larger wind farms converges to an optimum it usually takes first a few iterations, where the wind
farm is operated at the optimal point of the approximate model, before the error reduction begins. Obviously the algorithm
needs the additional information around the operating point. Interestingly, once the algorithm actually left the initial operating
point it converges relatively quickly to an operating point close to the plant optimum. This is a strong indication that exploration
or even just small excitation around an operating point should be activated if the operating point does not change for some time.
A reason for the increase of the error of larger wind farms is the decrease of the sample density. The size of the initial training
set is increased linearly while it would have to increase exponentially to preserve the same sampling density. For the wind farm
with 100 wind turbines and the current setup the hyperparameter optimization takes usually about 15 s. In some rare cases it
took about 5 min. In these cases the optimizer was not able to converge to an optimum and the maximum amount of allowed
iterations were used. The plant optimization takes less than 10 s. Consequently, the size of the data set is not a limiting factor
to improve the performance of the larger wind farms.

The increase of the initial training set improves the convergence of the method (Fig. 7b). Nevertheless, even with the larger size
of the initial training set it is challenging to converge to the correct optimum point for cases with a large input space. A larger
training set would be necessary for these cases. On the other hand, the training of the hyperparameters in the GP regression
scales cubic with the amount of data. Obviously this ultimately limits the size of the training set since the approach can become
computational infeasible.

Nonetheless, the results show clearly that the MA-GP is able to improve the performance of the model-based optimization
for some of the cases. It is not clear how the initial data sets differ for these successful cases. However, it is expected that a
large amount of operation points can be excluded from the initial training set of the GP regression since it is known from the
model that they are far away from the optimum operating point. Currently, the initial training set is chosen randomly by Latin
hypercube sampling. A smarter selection with a larger density of points around the optimal operating point of the model may
improve the MA-GP approach without increasing the initial data set.

In the next section the practical implications of the MA-GP approach are discussed.

5 Discussion

In this section an outlook on how to apply the MA-GP approach to a real wind farm is given. It is beyond this article to solve
all the challenges.

A major challenge is the dynamic environment a wind farm operates in. Averaging and filtering is required to approximate
steady-state conditions. In a nine-turbine LES study presented in Andersson et al. (2020c)[3] five-minutes averaging is used. A
longer averaging horizon will make the MA-GP more robust since the variance in the data decreases. A too long averaging
horizon will reduce the performance since the plant response is delayed and averaged. Moreover, measurement and input noise
can degrade the performance of the adaptation. The negative influence of input and measurement noise can be reduced by a
larger training data set.

---

[3]The article is submitted to TORQUE 2020

Another challenge is the wake propagation delay. In the LES study the first five minutes after a change in the control inputs is discarded to remove the transients. A similar approach might be necessary in a real wind farm. A wake propagation through the entire farm is not necessary. Depending on the measurements noise level it suffices to include the interaction of about two to three turbines (Andersson et al., 2020a).

275 The sensitivity of the input-output map can be increased by including the power measurements of each turbine and identifying a multiple-input multiple-output model. It is shown in (Andersson et al., 2020a) that this can help to decrease the necessary size of the training data set and improve the performance of the MA-GP approach for large wind farms. In addition, the wind farm could be separated into subsets. The separation would depend on the turbines' interaction considering a range of wind directions, e.g. a wind farm as presented in Sec. 4.3 could be separated into three subsets for each of the wind directions around

280 $0°$, $45°$, $90°$, $135°$, $180°$, $225°$, $270°$, and $315°$.

For a real wind farm the minimum training set should contain wind velocity, wind direction, the control inputs and the plant-model error of the power outputs of each turbine. The inclusion of other variables, e.g. the turbulence intensity, depends highly on the sensitivity of the variable to the plant-model mismatch of the power productions. Their effect should be larger than the effect of the input noise of the wind. Otherwise, it is not recommended to include them in the MA-GP approach.

285 Atmospheric conditions that considerably change the response of the wind farm could be handled by a multi-model approach. The model error for each atmospheric conditions is identified using a separate model. The multi-model approach can also be used to estimate the current atmospheric condition. If the atmospheric conditions are not considered explicitly in the MA-GP approach the response of the wind farm will be averaged over the atmospheric conditions. In fact, this happens to every variable that is not explicitly considered. On the other hand, the MA-GP approach automatically adapts to constant effects, e.g. terrain

290 effects.

It is important to point out that the MA-GP approach supplements model-based wind farm control. It is still beneficial to have a good surrogate model even though theoretically the MA-GP can work with a bad surrogate model. Moreover, the initial training set of the MA-GP approach can be generated by a high-fidelity model. In that case the MA-GP approach would initially reduce the error between surrogate and high-fidelity model, which should improve the performance of the wind farm controller. During

295 operation the initial data set can be gradually replaced by real measurements. The GP allows to weight different training sets, which should be used when working with two different training sets. Moreover, during operation the data set should be updated continuously replacing old data points with new ones.

The MA-GP initial synthesis can be similar to the approach presented in Doekemeijer et al. (2020):

1. Create training data set using high-fidelity simulations.

300 2. Estimate the model parameters of the approximate model using high-fidelity data.

3. Identify a model of the plant-model mismatch of approximate and high-fidelity model using GPs.

If during operation also the free-stream wind velocity or the turbulence intensity are estimated, only the approximate model without the MA-GP correction should be used to avoid a feedback of the identified model to the training set.

6 Conclusions

305 The modifier-adaptation approach with Gaussian processes applied to wind farm control is presented. It is a real-time optimization strategy, which corrects the approximate model used in the optimization by using plant measurements. In the wind farm case the total power production is assumed to be measured and used in the MA-GP approach. The approach works well for small input spaces. Here the GP regression is able to correct the model almost perfectly. Consequently, operating points very close to the real optimum are found in the optimization. For larger input spaces, on the other hand, the error increases.

310 Moreover, for the grid-type wind farm layout with more than 25 turbines convergence with the relatively small initial training sets used in this work could not be achieved at all times.

The MA-GP approach has similarities with Bayesian optimization (BO). Park et al. (2016, 2017) applied BO successfully in wind tunnel tests, and we expect the MA-GP approach to behave similarly. In Section 5 several possible future investigations to make the MA-GP applicable to real wind farms were pointed out. The performance of large wind farms can be improved by

315 the multiple inputs and multiple output approach and subset separation. In addition the following ideas can be tested:

- Increase the training set until it becomes computational unfeasible to increase the training set further.

- Choose the training data points in a smarter way such that they provide enough information about the regions around the expected optimum. Operating points far away from the expected optimum are excluded.

- Extend the algorithm with an exploration part. This can be achieved, for example, by including the variance of the GP

320 regression model in the optimization.

An important investigation is the sensitivity of the approach to measurements, input noise and time delays. In Andersson et al. (2020b) a simple way how to include input noise explicitly in the MA-GP approach is presented. Finally, the model identification should be tested on high fidelity and real data. A preliminary study on a nine turbine wind farm case using data from the high-fidelity simulator SOWFA (Churchfield et al., 2012) will be presented in Andersson et al. (2020c).

325 *Author contributions.* LEA compiled the literature review, performed numerical simulations, post-processed the data, and wrote the article. LI helped formulate the methodology used in the article and participated in structuring and reviewing of the article.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge the financial support from the Norwegian Research Council and the industrial partners of OPWIND: Operational Control for Wind Power Plants (Grant No.: 268044/E20)

References

Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., and Diehl, M.: CasADi – A software framework for nonlinear optimization and optimal control, Mathematical Programming Computation, 11, 1–36, https://doi.org/10.1007/s12532-018-0139-4, 2019.

Andersson, L. E., Bradford, E. C., and Imsland, L.: Distributed learning and wind farm optimization with Gaussian processes, in: American Control Conference (ACC), 2020 - [accepted], 2020a.

Andersson, L. E., Bradford, E. C., and Imsland, L.: Gaussian processes modifier adaptation with uncertain inputs using distributed learning and optimization on a wind farm, in: IFAC World congress 2020 - [accepted], 2020b.

Andersson, L. E., Doekemeijer, B., van der Hoek, D., van Wingerden, J.-W., and Imsland, L.: Adaptation of Engineering Wake Models using Gaussian Process Regression and High-Fidelity Simulation Data, arXiv preprint arXiv:2003.13323, 2020c.

Annoni, J., Seiler, P., Johnson, K., Fleming, P., and Gebraad, P.: Evaluating wake models for wind farm control, in: 2014 American Control Conference, pp. 2517–2523, https://doi.org/10.1109/ACC.2014.6858970, 2014.

Annoni, J., Gebraad, P. M., Scholbrock, A. K., Fleming, P. A., and Wingerden, J.-W. v.: Analysis of axial-induction-based wind plant control using an engineering and a high-order wind plant model, Wind Energy, 19, 1135–1150, 2016.

Annoni, J., Fleming, P., Scholbrock, A., Roadman, J., Dana, S., Adcock, C., Porte-Agel, F., Raach, S., Haizmann, F., and Schlipf, D.: Analysis of control-oriented wake modeling tools using lidar field results, Wind Energy Science, 3, 819–831, https://doi.org/10.5194/wes-3-819-2018, https://www.wind-energ-sci.net/3/819/2018/, 2018.

Barthelmie, R. J., Hansen, K., Frandsen, S. T., Rathmann, O., Schepers, J., Schlez, W., Phillips, J., Rados, K., Zervos, A., Politis, E., et al.: Modelling and measuring flow and wind turbine wakes in large wind farms offshore, Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology, 12, 431–444, 2009.

Barthelmie, R. J., Hansen, K. S., and Pryor, S. C.: Meteorological controls on wind turbine wakes, Proceedings of the IEEE, 101, 1010–1019, 2013.

Bartl, J. and Sætran, L.: Experimental testing of axial induction based control strategies for wake control and wind farm optimization, in: Journal of Physics: Conference Series, vol. 753, p. 032035, IOP Publishing, 2016.

Bastankhah, M. and Porté-Agel, F.: A new analytical model for wind-turbine wakes, Renewable Energy, 70, 116–123, 2014.

Bastankhah, M. and Porté-Agel, F.: Experimental and theoretical study of wind turbine wakes in yawed conditions, Journal of Fluid Mechanics, 806, 506–541, 2016.

Burton, T., Jenkins, N., Sharpe, D., and Bossanyi, E.: Wind energy handbook, John Wiley & Sons, 2011.

Campagnolo, F., Petrović, V., Bottasso, C. L., and Croce, A.: Wind tunnel testing of wake control strategies, in: American Control Conference (ACC), 2016, pp. 513–518, IEEE, 2016.

Churchfield, M. J., Lee, S., Michalakes, J., and Moriarty, P. J.: A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics, Journal of turbulence, p. N14, 2012.

Ciri, U., Rotea, M. A., and Leonardi, S.: Model-free control of wind farms: A comparative study between individual and coordinated extremum seeking, Renewable Energy, 113, 1033–1045, 2017.

de Avila Ferreira, T., Shukla, H. A., Faulwasser, T., Jones, C. N., and Bonvin, D.: Real-Time optimization of Uncertain Process Systems via Modifier Adaptation and Gaussian Processes, in: 2018 European Control Conference (ECC), pp. 465–470, IEEE, 2018.

365   del Rio-Chanona, E. A., Graciano, J. E. A., Bradford, E., and Chachuat, B.: Modifier-Adaptation Schemes Employing Gaussian Processes and Trust Regions for Real-Time Optimization, IFAC-PapersOnLine, 52, 52–57, https://doi.org/10.1016/j.ifacol.2019.06.036, https://linkinghub.elsevier.com/retrieve/pii/S2405896319301211, 2019.

Doekemeijer, B. M., van der Hoek, D. C., and van Wingerden, J.-W.: Model-based closed-loop wind farm control for power maximization using Bayesian optimization: a large eddy simulation study, in: 2019 IEEE Conference on Control Technology and Applications (CCTA),

370   pp. 284–289, IEEE, 2019a.

Doekemeijer, B. M., Van Wingerden, J.-W., and Fleming, P. A.: A tutorial on the synthesis and validation of a closed-loop wind farm controller using a steady-state surrogate model, in: 2019 American Control Conference (ACC), pp. 2825–2836, IEEE, 2019b.

Doekemeijer, B. M., van der Hoek, D., and van Wingerden, J.-W.: Closed-loop model-based wind farm control using FLORIS under time-varying inflow conditions, Renewable Energy, 2020.

375   Fleming, P., Annoni, J., Shah, J. J., Wang, L., Ananthan, S., Zhang, Z., Hutchings, K., Wang, P., Chen, W., and Chen, L.: Field test of wake steering at an offshore wind farm, Wind Energy Science, 2, 229–239, 2017.

Fleming, P., Annoni, J., Churchfield, M., Martinez-Tossas, L. A., Gruchalla, K., Lawson, M., and Moriarty, P.: A simulation study demonstrating the importance of large-scale trailing vortices in wake steering, Wind Energy Science, 3, 243–255, https://doi.org/10.5194/wes-3-243-2018, https://www.wind-energ-sci.net/3/243/2018/, 2018.

380   Fleming, P., King, J., Dykes, K., Simley, E., Roadman, J., Scholbrock, A., Murphy, P., Lundquist, J. K., Moriarty, P., Fleming, K., van Dam, J., Bay, C., Mudafort, R., Lopez, H., Skopek, J., Scott, M., Ryan, B., Guernsey, C., and Brake, D.: Initial results from a field campaign of wake steering applied at a commercial wind farm – Part 1, Wind Energy Science, 4, 273–285, https://doi.org/10.5194/wes-4-273-2019, https://www.wind-energ-sci.net/4/273/2019/, 2019.

Ge, M., Wu, Y., Liu, Y., and Yang, X. I.: A two-dimensional Jensen model with a Gaussian-shaped velocity deficit, Renewable Energy, 141,

385   46–56, 2019.

Gebraad, P., Teeuwisse, F., Van Wingerden, J., Fleming, P. A., Ruben, S., Marden, J., and Pao, L.: Wind plant power optimization through yaw control using a parametric model for wake effects – a CFD simulation study, Wind Energy, 19, 95–114, 2016.

Gebraad, P. M., van Dam, F. C., and van Wingerden, J.-W.: A model-free distributed approach for wind plant control, in: American Control Conference (ACC), 2013, pp. 628–633, IEEE, 2013.

390   Howland, M. F., Lele, S. K., and Dabiri, J. O.: Wind farm power optimization through wake steering, Proceedings of the National Academy of Sciences, 116, 14 495–14 500, 2019.

Hulsman, P., Andersen, S. J., and Göçmen, T.: Optimizing Wind Farm Control through Wake Steering using Surrogate Models based on High Fidelity Simulations, Wind Energy Science Discussions, pp. 1–34, 2019.

Jensen, N. O.: A note on wind generator interaction, 1983.

395   Jiménez, Á., Crespo, A., and Migoya, E.: Application of a LES technique to characterize the wake deflection of a wind turbine in yaw, Wind energy, 13, 559–572, 2010.

Johnson, K. E. and Fritsch, G.: Assessment of extremum seeking control for wind farm energy production, Wind Engineering, 36, 701–715, 2012.

Johnson, K. E. and Thomas, N.: Wind farm control: Addressing the aerodynamic interaction among wind turbines, in: American Control

400   Conference, 2009. ACC'09., pp. 2104–2109, IEEE, 2009.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G.: Definition of a 5-MW reference wind turbine for offshore system development, National Renewable Energy Laboratory, Golden, CO, Technical Report No. NREL/TP-500-38060, 2009.

Katic, I., Højstrup, J., and Jensen, N. O.: A simple model for cluster efficiency, in: European wind energy association conference and exhibition, 1987.

405 Kheirabadi, A. C. and Nagamune, R.: A quantitative review of wind farm control with the objective of wind farm power maximization, Journal of Wind Engineering and Industrial Aerodynamics, 192, 45–73, 2019.

Loeppky, J. L., Sacks, J., and Welch, W. J.: Choosing the sample size of a computer experiment: A practical guide, Technometrics, 51, 366–376, 2009.

Marchetti, A., Chachuat, B., and Bonvin, D.: Modifier-adaptation methodology for real-time optimization, Industrial & engineering chemistry

410 research, 48, 6022–6033, 2009.

Marchetti, A. G., François, G., Faulwasser, T., and Bonvin, D.: Modifier Adaptation for Real-Time Optimization – Methods and Applications, Processes, 4, 55, 2016.

Marden, J. R., Ruben, S. D., and Pao, L. Y.: A Model-Free Approach to Wind Farm Control Using Game Theoretic Methods, IEEE Transactions on Control Systems Technology, 21, 1207–1214, https://doi.org/10.1109/TCST.2013.2257780, 2013.

415 NREL: FLORIS. Version 1.0.0, https://github.com/NREL/floris, 2019.

Park, J. and Law, K. H.: Layout optimization for maximizing wind farm power production using sequential convex programming, Applied Energy, 151, 320 – 334, https://doi.org/https://doi.org/10.1016/j.apenergy.2015.03.139, http://www.sciencedirect.com/science/article/pii/S0306261915004560, 2015.

Park, J., Kwon, S.-D., and Law, K. H.: A data-driven approach for cooperative wind farm control, in: American Control Conference (ACC),

420 2016, pp. 525–530, IEEE, 2016.

Park, J., Kwon, S.-D., and Law, K.: A data-driven, cooperative approach for wind farm control: a wind tunnel experimentation, Energies, 10, 852, 2017.

Rasmussen, C. E. and Williams, C. K.: Gaussian processes for machine learning, the MIT Press, 2, 4, 2006.

Snelson, E. and Ghahramani, Z.: Sparse Gaussian processes using pseudo-inputs, in: Advances in neural information processing systems,

425 pp. 1257–1264, 2006.

Steinbuch, M., De Boer, W., Bosgra, O., Peeters, S., and Ploeg, J.: Optimal control of wind power plants, Journal of Wind Engineering and Industrial Aerodynamics, 27, 237–246, 1988.

Tian, L., Zhu, W., Shen, W., Zhao, N., and Shen, Z.: Development and validation of a new two-dimensional wake model for wind turbine wakes, Journal of Wind Engineering and Industrial Aerodynamics, 137, 90–99, 2015.

430 van der Hoek, D., Kanev, S., Allin, J., Bieniek, D., and Mittelmeier, N.: Effects of axial induction control on wind farm energy production - A field test, Renewable Energy, 140, 994 – 1003, https://doi.org/https://doi.org/10.1016/j.renene.2019.03.117, http://www.sciencedirect.com/science/article/pii/S096014811930429X, 2019.

Veers, P., Dykes, K., Lantz, E., Barth, S., Bottasso, C. L., Carlson, O., Clifton, A., Green, J., Green, P., Holttinen, H., Laird, D., Lehtomäki, V., Lundquist, J. K., Manwell, J., Marquis, M., Meneveau, C., Moriarty, P., Munduate, X., Muskulus, M., Naughton, J., Pao, L., Paquette,

435 J., Peinke, J., Robertson, A., Sanz Rodrigo, J., Sempreviva, A. M., Smith, J. C., Tuohy, A., and Wiser, R.: Grand challenges in the science of wind energy, Science, 366, https://doi.org/10.1126/science.aau2027, https://science.sciencemag.org/content/366/6464/eaau2027, 2019.

Wächter, A. and Biegler, L. T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Mathematical Programming, 106, 25–57, https://doi.org/10.1007/s10107-004-0559-y, http://dx.doi.org/10.1007/s10107-004-0559-y, 2006.