Response to Referees for Wind Energy Science submission "Real-time optimization of wind farms using modifier adaptation and machine learning" by Leif Erik Andersson and Lars Imsland

Dear Referees,

Thank you for the detailed and thorough review of our manuscript "Real-time optimization of wind farms using modifier adaptation and machine learning" by Leif Erik Andersson and Lars Imsland, which we submitted for publication in Wind Energy Science. We appreciate the suggestions and comments to improve the draft. The following changes were made to the manuscript:

- We rewrote the abstract to make it more related to the contribution of the article
- We improved to introduction to shift the focus also on surrogate modelling and adaptation.
- We included a discussion section, which includes some points of our discussion. It gives an outlook on challenges and possible solution to apply the presented method to real wind farms with wake propagation delay, dynamic environment etc.
- We condensed the presentation of the methods, e.g. Gaussian processes and modifier adaptation.
- We rearranged some parts in the presentation of the numerical case study to ease the read and avoid repetition.

These are the main changes in the revised version of our manuscript. Please find below a detailed answer to your comments.

Again, thank you for your time and suggestions to improve our manuscript

On behalf of the authors, yours sincerely,

Leif Erik Andersson

- Attached:
    - Response to Referee #1
    - Response to Referee #2
    - Response to comments in the pdf-file of Referee #1
    - Marked-up manuscript version
    - Marked-up manuscript version (without strikethrough text)

Response to Referee #1:
Interactive comment on “Real-time optimization of wind farms using modifier adaptation and machine learning” by Leif Erik Andersson and Lars Imsland

Dear Bart,

Thank you for your very detailed review on our paper. Including your comments and recommendation improved the paper considerably. Here we will respond to your four main points. In addition, we answered the questions in your attached pdf. This we will attach to this document.

Please see below a more detailed response:

Point 1:

1. The article seems heavily focused on the algorithm. A more in-depth discussion on the practicality of the algorithm is missing. Questions you raised are:
    a. How can the algorithm be applied to wind farm control on real sites?
    b. What would you consider a training set in real life?
    c. Do you need to time average measurements?
    d. How would you deal with time-varying inflow conditions?
    e. Are time delays due to wake propagation an obstacle?

These are extremely interesting and important questions, which were also raised by the other referee. We included a discussion section in the revised version of the article to highlight some of these points. We performed also a LES study, which will be presented at the TORQUE 2020. Based on my own expectations and the LES study I can say:

The training set would be at least the wind velocity, wind direction, and the power outputs of the turbines. Turbulence intensity could be considered as an input. I would recommend doing a sensitivity study to evaluate how much the variance in the turbulence intensity affects the outputs (it should be considerably larger than the effect of the input noise in the wind measurements).

The algorithm needs time averaged measurements. Without an appropriate filter the variance in the data will be large, which will degrade the performance of the learning algorithm. In the LES study we used 5-minutes averaging, which was enough. However, we used quasi-static inflow conditions without wind direction changes. Wind direction changes will degrade the performance. In general, input noise can be to some extend counteracted with more data. I am unsure how the algorithm will react on biased data. If the bias is consistent – e.g. 5 degrees of in all measurements – it should not degrade the algorithm.

For a real-life application, it would be necessary to first collect data – it is possible to create data with high-fidelity models, which would already improve the surrogate model using our adaptation approach. The wind direction can be included as an input variable. Otherwise, for many small circle sectors a new model would have to be identified, which could be impractical. I expect, it is necessary to differentiate between atmospheric conditions and identify separate models for each of these conditions. A multi-model approach could be used. Time delays are difficult to handle. Steady-state

data is wanted so it is necessary to wait until the first downstream turbine is affected by changes in the upwind turbine. It is difficult to evaluate how much time delays will degrade the performance of the algorithm. In simulation we performed it seems most of the energy transfer in the plant can be captured with the upstream and the next two downstream turbines.

Point 2:

2. The article can be significantly reduced especially the sections about the MA and GP algorithms can be condensed.

We condensed the sections about the MA and GP algorithms. However, the additions of the discussion section did not reduce the overall length of the article.

Point 3:

3. The literature survey presented in the introduction seems to focus on general wind farm control. It would be useful to shift to surrogate modelling and adaptation in wind farms.

We refocused the introduction. We included surrogate modelling and adaptation algorithms for wind farms.

Point 4:

4. Shorten the article, clearly state the contribution in the introduction and also start each section with one or two sentences relating the upcoming section to the previous section. In addition, it may be helpful to gather some information in tables.

We followed these recommendations.

Response to Referee #2:
Interactive comment on “Real-time optimization of wind farms using modifier adaptation and machine learning” by Leif Erik Andersson and Lars Imsland

Dear Reviewer,

Thank you for the response to our article. We investigated some of your questions in a LES study. However, as the nature of these studies the data and test tests are limited by computational constraints. The other referee also pointed also out that some parts of the paper can be condensed especially the explanation of the algorithms. We followed the suggestion in the revised version.

Please see below a more detailed response to your questions:

Questions in comment 1

1. Can this approach work in truly dynamic environment?
2. Will the approach work with varying wind directions and wake propagation delay?

> These are extremely interesting and important questions. It was also pointed out by the other referee. If the approach would not be applicable to such environments it would be inapplicable to wind farm control. In our LES study that we plan to present at the TORQUE 2020 we simulated a nine-turbine plant with quasi-static wind conditions. The wind direction did not change, but we applied a turbulence intensity of 5%. We filtered the power output with an averaging filter. The approach was able to improve the power production compared to the Gaussian model (with tuned parameter via parameter estimation) about 2-3%. How a complete dynamic case with uncertainties in wind velocity, direction and yaw angle will affect the approach is difficult to say. The approach will require more data to cope with the variance in the training data. The performance will decrease like in robust approaches, which consider uncertainties explicitly. Nevertheless, we expect the approach will still improve the performance of the wind farm.

Questions in comment 2

1. Could the problem to solve large layouts be addressed by decomposing the large wind farm into manageable subsets according to wake interactions?

> One way is to separate the farm into subsets according to wake interactions. Park and Law (2016) proposed such an approach. The other approach is to include the power measurements of each wind turbine in the model identification. Currently, we use a MISO approach approximating the total power production of the plant. A more efficient use of the available measurements is to identify the power production of each turbine and combine these N models in the optimization to optimize the total power output. It is a distributed learning strategy. In simulation studies we were able to show that this distributed (MIMO) approach scales much better for large wind farms. It needs much less data to achieve the same performance as the MISO approach. The distributed

learning approach can be combined with the subset approach. Even thought, the GP learning can identify these subsets it can be helpful to specify them explicitly. The disadvantage of the distributed approach is the requirement to identify N models (which can be parallelized). The disadvantage of the subset approach is the inflexibility it introduces. Depending on the wind directions and the resulting different subset structures x models would have to be identified for each of these structures.

Questions in comment 3

1. How would the approach handle a non-input-output dependency, like turbulence, which varies on day/night basis?
2. If in the extreme two models for stable and unstable atmospheric conditions are needed, is there a possibility of modeling hidden confounders?

It depends heavily on the influence of the non-measured input to the output. The approach can work without measuring every input. However, if for example turbulence is not explicitly considered in the GP model's inputs its influence will be averaged (over the training data set). In addition, it will increase the variance of the output of the GP.

Conditions like stable and unstable atmospheric conditions where the response of the wind farm can differ drastically have to be approached by separate models. If approximated by one model the model will again average the output of these two conditions. This might decrease performance of the control approach.

I would propose to differentiate in the data collection of the training data between atmospheric conditions and create several models. It would not be necessary to consider the atmospheric condition as an explicit input to the model. During operations it should be possible to estimate which model is most accurate in the current situation and hence estimate the atmospheric condition. The most accurate model would be used in the optimization. Another way would be a multi-model approach in which each model is weighted: $Power = \phi_1 M_1 + \phi_2 M_2 + (1 - \sum \phi) M_3$. The parameter ϕ_i would be estimated using approach proposed in the literature about statistical learning. However, for the multi-model approach I am unsure if an interpolation between models for different atmospheric condition would be appropriate.

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

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. Real-time optimization (RTO) covers a family of optimization methods that incorporate process measurements in the optimization to drive the real process (plant) to optimal performance while guaranteeing constraint satisfaction. Modifier Adaptation (MA) introduces zeroth and first-order correction terms (bias and gradients) for the cost and constraint functions. Instead of updating the plant model, in MA the optimization problem is updated directly from data guaranteeing to meet the necessary condition of optimality upon convergence.

The main burden of the MA approach is the estimation of the first-order modifiers of the cost and constraint functions at each RTO iteration. Finite-difference approximation is the most common approach that requires at least nu + 1 steady-state operation points to estimate the gradients, where nu is the number of control inputs. Obtaining these can require a long convergence time. For this reason, this work considers the use of Gaussian process (GP) regression to estimate the plant-model mismatch based on plant measurements, and replace the usual modifiers by these high order regression functions. GP is a probabilistic, non-parametric modelling technique well known in the machine learning community. The approach is tested on several numerical test cases simulating wind farms. It is shown that the approach is able to correct the model and converges to the plant optimal point. Several improvements for large inputs spaces, which is a challenging problem for the approach presented in the article, are discussed.

1 Introduction

Currently the wind turbines in a wind farm are operated at their individual optimal operating point. This control strategy is called *greedy* wind farm control since the interactions between turbines are not taken into account. However, it is expected that the greedy control strategy leads to sub-optimal performance of the wind farm (Steinbuch et al., 1988; Johnson and Thomas, 2009; Barthelmie et al., 2009). A coordinated wind farm controller, which takes the wake interactions between turbines in a wind farm into account, may result in a superior performance compared to the greedy wind farm controller. The two main wind farm control strategies are axial induction control, e.g. Steinbuch et al. (1988); Corten and Schaak (2003); Horvat et al. (2012); Rotea (2014); Munters and Meyers (2016) and wake steering control, e.g. Medici (2005); Adaramola and Krogstad (2011); Wagenaar et al. (2012); Park et al. (2013); Gebraad and Van Wingerden (2014). The idea behind the former 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 and the power production of the downwind turbine changes. The target net effect is an overall

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

increase of the power production and possibly an decrease of fatigue loads. However, recent studies suggest that axial induction control using steady-state models to calculate the optimal control settings may be unable to improve the power production of a wind farm (Schepers and Van der Pijl, 2007; Campagnolo et al., 2016; Bartl and Sætran, 2016; Annoni et al., 2016). The currently more promising wind farm control strategy using steady-state 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. Field experiments showing promising results were conducted by Fleming et al. (2017, 2019); Howland et al. (2019). In these experiments lookup tables with optimal yaw settings of each turbine are created with help of an steady-state model. Hence the wind farm is operated in an open-loop control setting.

The steady-state wake models used in model-based control are usually relatively simple. They estimate the velocity deficit in wakes. For a long time one of the most popular wake models was 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. 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 value approximations of a complex phenomena that is, in fact, not well understood (Veers et al., 2019). Hence, real time optimization (RTO), which incorporates plant measurements to improve the performance of the wind farm controller, is extremely useful for this process.

~~Probably~~ one of the most intuitive RTO strategies is the "two-step" approach. Here, first the model parameters are updated, and then new control inputs are computed based on the updated model. The two steps refer to the parameter optimization and control input optimization, which are performed sequentially (Marchetti et al., 2016). However, the two-step approach cannot guarantee plant optimality upon convergence if the model is structurally incorrect (Marchetti et al., 2016). An example that an improved parameterisation of the steady-state wake model was not able to remove the mismatch between a low order model and a high fidelity model of wake is given in Fleming et al. (2018).

In contrast, modifier adaptation (MA) corrects the cost and constraint functions of the optimization problem directly, and reaches, under suitable assumptions, true plant optimality upon convergence (Marchetti et al., 2009). The bottleneck of the MA approach is the estimation of the gradients of the objective and constraint functions at each RTO iteration. Finite difference approximation is one of the most common approaches that requires $n_u + 1$ steady-state operation points to estimate the gradients, where n_u is the amount of control inputs. These can lead to a long convergence time, especially for processes with high dimensional input spaces. Therefore, in this work Gaussian process (GP) regression is combined with MA (de Avila Ferreira et al., 2018; del Rio Chanona et al., 2019). GP is a probabilistic, non-parametric modelling technique well known in the machine learning community (Rasmussen and Williams, 2006). The GP regression model estimates the plant-model mismatch using plant measurements. Then the GP model is used to correct the original optimization problem and by this improve the optimization of the plant inputs.

The article is structured as follows: In Section 2 the optimization problem is formulated and Gaussian process regression is explained. In Section 3 the modifier adaptation using Gaussian process regression is presented and the numerical turbine

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

and wake models introduced. The approach is tested numerically ~~on several examples~~ in Section 4. The article ends with a conclusion.

2 Problem formulation

The optimization problem of the steady-state plant performance subject to constraints can be formulated as (Marchetti et al., 2016):

$$\mathbf{u}_p^* = \arg\min_{\mathbf{u}} \phi_p(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \tag{1a}$$

$$s.t.\ G_{p,j}(\mathbf{u}) := g_{p,j}(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \leq 0,\ \ j = 1, \ldots, n_g, \tag{1b}$$

$$\mathbf{u} \in \mathcal{U}, \tag{1c}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$ and $\mathbf{y}_p \in \mathbb{R}^{n_y}$ denote the plant input and output variables, respectively; $\mathbf{u} \in \mathbb{R}^{n_u}$ and $\mathbf{y}_p \in \mathbb{R}^{n_y}$ are the input-output pairs of the wind farm; $\phi_p : \mathbb{R}^{n_u} \to \mathbb{R}$ is the cost function to be minimized; $g_{p,j} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R},\ j = 1, \ldots, n_g$, are the inequality constraint functions; and $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ is the control domain, e.g. box constraints on the control inputs. Formulation (1) assumes that ϕ_p and $g_{p,j}$ as functions of $\mathbf{u}$, and $\mathbf{y}_p$ are exactly known. However, in any practical application the exact input-output map of the plant is unknown and instead an approximate model of the system is exploited for the optimization:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) \tag{2a}$$

$$s.t.\ G_j(\mathbf{u}) := g_j(\mathbf{u}, \mathbf{y}(\mathbf{u})) \leq 0,\ \ j = 1, \ldots, n_g, \tag{2b}$$

$$\mathbf{u} \in \mathcal{U}, \tag{2c}$$

where the quantities ϕ, $g_j(\mathbf{u}, \mathbf{y}(\mathbf{u}))$, $\mathbf{u}*$, and G_j refer to the inexact model counterparts of the true plant optimization problem in Eq. (1).

RTO takes advantage of the available measurements to compensate for plant-model mismatch and adapt the model-based optimization problem Eq. (2) to reach plant optimality.

The standard MA approach applies first-order correction terms that are added to the cost and constraint functions to match the necessary conditions of optimality upon convergence (Marchetti et al., 2009). Iteratively the following modified optimization problem is solved:

$$\hat{\mathbf{u}}_{k+1}^* = \arg\min_{\mathbf{u}} \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) + (\boldsymbol{\lambda}_k^{\phi})^T \mathbf{u} \tag{3a}$$

$$s.t.\ G_j(\mathbf{u}) + \varepsilon_{j,k} + (\boldsymbol{\lambda}_k^{G_j})^T (\mathbf{u} - \mathbf{u}_k) \leq 0,\ \ j = 1, \ldots, n_g, \tag{3b}$$

$$\mathbf{u} \in \mathcal{U}, \tag{3c}$$

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

where $\hat{\mathbf{u}}^*_{k+1}$ is the optimal solution at iteration $k+1$, the $\varepsilon_{j,k} \in \mathbb{R}$ are the zeroth-order modifiers for the constraints, and $\boldsymbol{\lambda}_k^{\phi}$ and $\boldsymbol{\lambda}_k^{G_j}$ are the first-order modifiers for the cost and constraints, respectively. The correction terms are given by:

$$\varepsilon_{j,k} := G_{p,j}(\mathbf{u}_k) - G_j(\mathbf{u}_k), \tag{4a}$$

$$(\boldsymbol{\lambda}_k^{\phi})^T := \frac{\partial \phi_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \phi}{\partial \mathbf{u}}(\mathbf{u}_k), \tag{4b}$$

$$(\boldsymbol{\lambda}_k^{G_j})^T := \frac{\partial G_{p,j}}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial G_j}{\partial \mathbf{u}}(\mathbf{u}_k). \tag{4c}$$

It is recommended to filter the input update $\hat{\mathbf{u}}^*_{k+1}$ to avoid excessive correction and reduce sensitivity to noise (Marchetti et al., 2016):

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{L}(\hat{\mathbf{u}}_{k+1} - \mathbf{u}_k), \tag{5}$$

with $\mathbf{L} = \mathrm{diag}(l_1, \ldots, l_{n_u})$, $l_i \in (0,1]$ where l_i may be reduced to help stabilize the iterations.

The MA scheme requires the estimation of the plant gradients at each RTO iteration, which is experimentally expensive and the main bottleneck for MA implementation in practice (Marchetti et al., 2016).

2.1 Gaussian processes

In this section we give a brief outline of GP regression ~~for our purposes~~, for more information ~~refer to~~ Rasmussen and Williams (2006). GP regression aims to identify an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ from data. Let the noisy observation of $f(\cdot)$ be given by:

$$y_k = f(\mathbf{u}_k) + \nu_k \tag{6}$$

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

We assume $f(\cdot)$ to follow a GP with 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). The SE covariance function can be expressed as follows:

$$k(\mathbf{u}_i, \mathbf{u}_j) = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{u}_i - \mathbf{u}_j)^T \boldsymbol{\Lambda}^{-1}(\mathbf{u}_i - \mathbf{u}_j)\right) \tag{7}$$

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

Assume we are given a training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ of size M consisting of M input vectors $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_M]^\mathsf{T}$ and corresponding observations $\mathbf{y} = [y_1, \ldots, y_M]^\mathsf{T}$ according to Eq. (6). From the GP distribution the data then follows a joint multivariate Gaussian distribution, which can be stated as:

$$p(\mathbf{y}|\mathbf{U}) = \mathcal{N}(\mathbf{0}, \mathbf{K} + \sigma_\nu^2 \mathbf{I}), \quad K_{ij} = k(\mathbf{u}_i, \mathbf{u}_j) \tag{8}$$

The hyperparameters $\boldsymbol{\psi} := [\sigma_f, \sigma_\nu, \lambda_1, \ldots, \lambda_{n_u}]^T$ are commonly unknown and hence need to be inferred from data. In this article the log marginal likelihood $p(\mathbf{y}|\mathbf{U})$ is used. Ignoring constant terms and factors, this can be stated as:

$$\mathcal{L}(\mathcal{D}, \Psi) = -\frac{1}{2}\mathbf{y}^T(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{y} - \frac{1}{2}\ln|\mathbf{K} + \sigma_\nu^2 \mathbf{I}|. \tag{9}$$

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The required maximum likelihood estimate is then given by $\hat{\boldsymbol{\psi}} \in \arg\max_{\boldsymbol{\psi}} \mathcal{L}(\mathcal{D}, \boldsymbol{\psi})$.

Next we require the predictive distribution of $f(\mathbf{u})$ at an arbitrary input $\mathbf{u}$, which can be found by the conditional distribution of $f(\mathbf{u})$ on the data distribution $p(\mathbf{y}|\mathbf{U})$. From the GP assumption this has a closed-form solution and can be stated as:

$$f(\mathbf{u})|\mathcal{D}, \hat{\boldsymbol{\psi}} \sim \mathcal{N}(\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}}), \sigma^2_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})) \tag{10}$$

$$\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}}) = \mathbf{k}^T(\mathbf{u})(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{y} \tag{11}$$

$$\sigma^2_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})) = \sigma_f^2 - \mathbf{k}^T(\mathbf{u})(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{k}(\mathbf{u}) \tag{12}$$

where $\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ can be seen as the GP prediction at $\mathbf{u}$ and $\sigma^2_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ as a corresponding measure of uncertainty to this prediction. The GP is a non-parametric model. The training data are explicitly required to construct the predictive distribution. For the above expression a matrix of size $M \times M$ must be inverted, which prohibits large data sets.

3 Methodology

3.1 Modifier Adaptation with Gaussian processes

The use of GPs in a MA approach to overcome the limitation of estimating the plant gradients was first proposed by de Avila Ferreira et al. (2018). The idea is to replace the zeroth- and first-order modifiers of the cost and constraints in (3) with GP regression terms. Since the wind farms considered in this article do not have inequality constraint functions they are not included in this section. However, inequality constraint functions can be easily incorporated into the method.

The training set of the GP to correct the objective function are the controlled inputs of the approximate model and the plant-model mismatch of the objective function. The new optimization problem of the MA scheme with GP modifiers (MA-GP) is

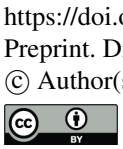
$$\hat{\mathbf{u}}^*_{k+1} = \arg\min_{\mathbf{u}} \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) + \mu^{\phi_p - \phi}_{\text{GP},k}(\mathbf{u}; \mathcal{D}_0, \hat{\boldsymbol{\psi}}_0), \qquad s.t. \quad \mathbf{u} \in \mathcal{U}, \tag{13}$$

where the plant-model mismatch of the cost function is modelled by $\mu^{\phi_p - \phi}_{\text{GP}}$. Similar to the original MA scheme the optimal input of Eq. (13) may be filtered with Eq. (5) to reduce the step-size and help stabilize the MA-GP scheme (del Rio Chanona et al., 2019). The whole MA-GP scheme is presented in Algorithm 1 and Fig. 1.

In Algorithm 1 the hyperparameters are updated if *HypOpt* is true. *HypOpt* is a user-defined condition, which allows to update the hyperparameter. The extrema are to update the hyerparameter each iteration or never. The hyperparameter update is usually the computational bottle-neck of the MA-GP algorithm. Especially for large data sets it can be expected that the hyperparameter do not change much from one iteration to the next. Therefore, it is reasonable to update the hyperparameters less frequent.

[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

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

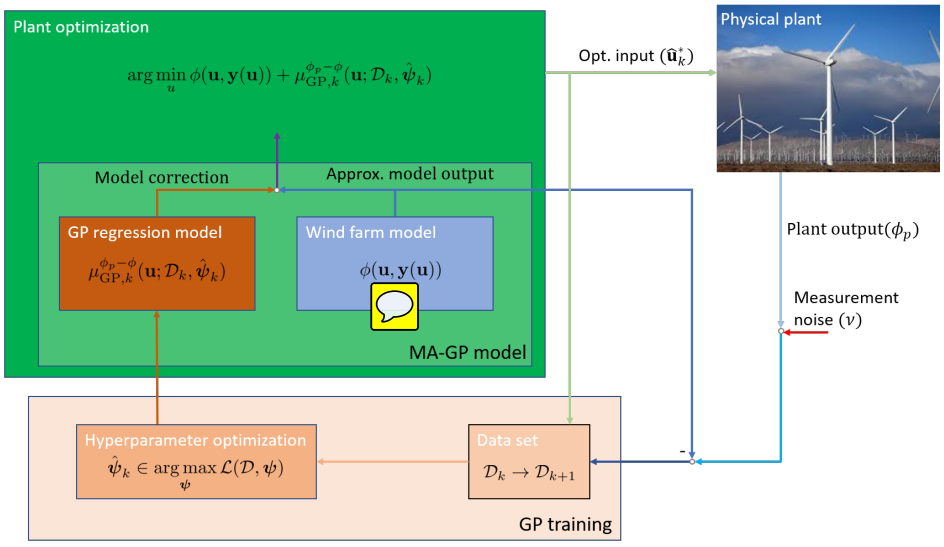


Figure 1. The basic idea of the MA-GP scheme for a wind farm. The GP regression model creates an input-output map of the control inputs to the plant-model mismatch. In the MA-GP model the GP regression model is used to correct the output of the approximate model. This MA-GP model is used in the optimization 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]

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

Initialisation: GP regression model $\mu_{\mathrm{GP}}^{\phi_p-\phi}$ and hyperparameters $\hat{\psi}_0$ found with (9) and data set $\mathcal{D}_0$; Optimal operation point of the approximate model $\mathbf{u}_0$.

for $k = 0,1,\ldots$ **do**

- Solve modified optimization problem Eq. (13);
- Filter new operating point $\mathbf{u}_{k+1}$ with Eq. (5);
- Evaluate approximate model at new operating point $\mathbf{u}_{k+1}$;
- Obtain measurements of cost function $\phi_p(\mathbf{u}_{k+1})$;
- Update the data set $\mathcal{D}_{k+1}$ with input $\mathbf{u}_{k+1}$ and output $y_{k+1} = \phi_p(\mathbf{u}_{k+1}) - \phi(\mathbf{u}_{k+1}) + \nu_{k+1}$. ;
- **if** *HypOpt* **then**
  - Update hyperparameters $\hat{\psi}_{k+1}$ with new data set $\mathcal{D}_{k+1}$ and Eq. (9) ;
- **end**
- Update GP regression term $\mu_{\mathrm{GP}}^{\phi_p-\phi}$ using $\mathcal{D}_{k+1}$ and hyperparameters $\hat{\psi}_{k+1}$;

end

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

3.2 Numerical turbine and wake models

The wind turbines in the wind farm are represented using the actuator disc theory, which couples the power and thrust coefficient, C_P and C_T (Burton et al., 2011)

$$C_P = 4a(1-a)^2, \tag{14}$$

$$C_T = 4a(1-a), \tag{15}$$

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 et al., 2016)

$$P = \frac{1}{2}\rho A C_P \cos\gamma^p u^3, \tag{16}$$

where A is the rotor area, ρ the air density and p a correction factor. 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 study it will be important to implement a "plant" and model, which are different from each other. Therefore, a second adjusted actuator disk turbine model is created. 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 a new model based on the actuator disk model. The equation for the thrust coefficient C_T is given by Eq. (15) while for the power coefficient C_P three new parameter are identified resulting in

$$C_P = 7.037a(0.625-a)^{1.364}. \tag{17}$$

The model fit is visualised in Fig. 2. Important in the numerical example is the different connection between thrust and power coefficients of both models (Fig. 2b). For the turbine dimensions the NREL 5-MW wind turbine is used (Jonkman et al., 2009). Consequently, the rotor diameter is $D = 136\,\mathrm{m}$ and the hub height $H_H = 90\,\mathrm{m}$.

The Gaussian wake model by Bastankhah and Porté-Agel (2014, 2016) is used to model the flow in the wind farm. The three-dimensional steady-state far wake velocity is assumed to be Gaussian distributed and can be estimated with

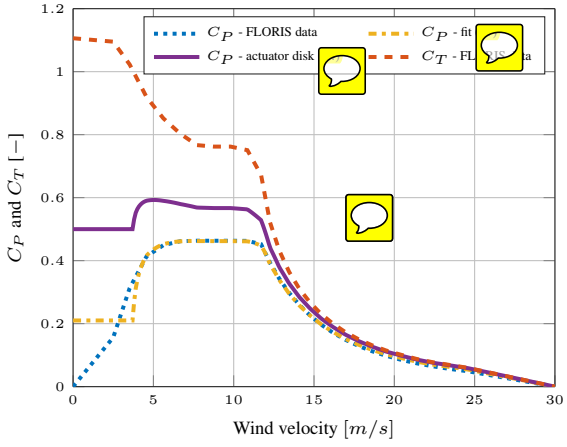
$$\frac{\bar{v}(x,y,z)}{\bar{v}_\infty} = 1 - Ce^{-0.5((y-\delta)/\sigma_y)^2} e^{-0.5((z-z_h)/\sigma_z)^2}, \tag{18a}$$

$$C = 1 - \sqrt{1 - \frac{C_T \cos\gamma}{8(\sigma_y \sigma_z / d^2)}}, \tag{18b}$$

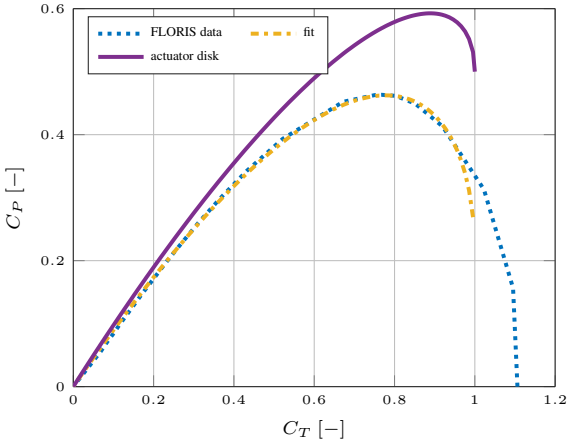
where z_h is the tower height, δ is the wake deflection, and σ_y and σ_z are the wake widths in lateral and vertical directions. An important variable for the model is the skew angle of the flow past a yawed turbine. The flow skew angle is approximated by

$$\theta \approx \frac{\alpha_1 \gamma}{\cos\gamma}\left(1 - \sqrt{1 - C_T \cos\gamma}\right), \tag{19}$$

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



(a) Comparison of C_P values in dependency of the wind velocity. The dashed line gives the corresponding thrust coefficient, which is the same for both models.



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

Figure 2. Comparison between data, the new model based on the actuator disk model and the actuator disk model. The thrust coefficients are kept smaller than one for the actuator disk models. The models give a different connection between thrust and power coefficients.

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. In the simulation study different values are chosen for this parameter in the plant and approximated model resulting in different optimal operating points.

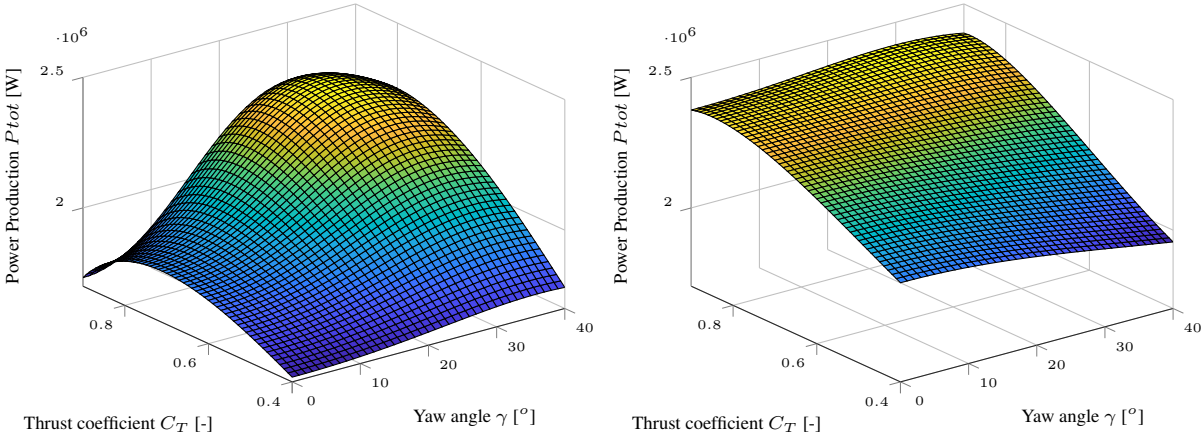
4 Numerical case study

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 amount 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 control inputs are constrained by box constraints with

$$0 \leq C_{T,i} \leq 0.95, \quad \text{and} \quad 0° \leq \gamma_i \leq 40°. \tag{20}$$

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).

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



(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.

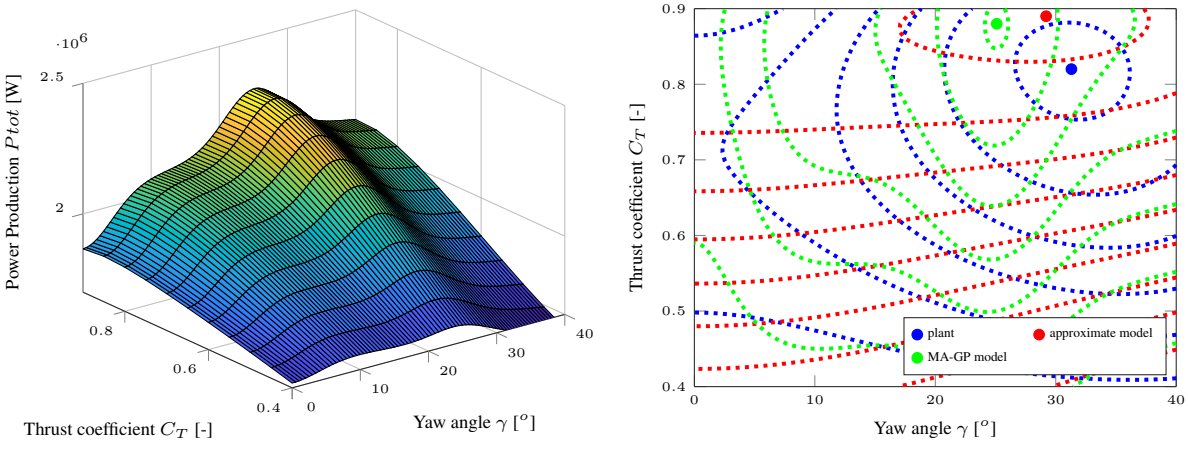
4.1 Two turbine case

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 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 $1.67\,\%$. Still, the model assume 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 the 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 operating point. Fig 5 shows the operating points of the first ten iterations. After four iterations the error in power production is about $0.2\,\%$ and after ten iterations it is $0.0009\,\%$. In addition, the contour lines of the objective function are well approximated. 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

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



(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.

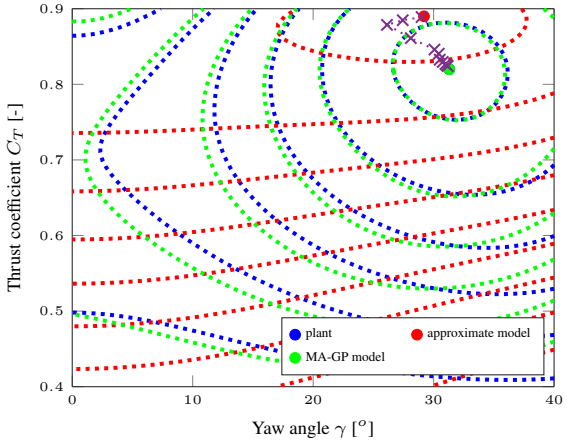
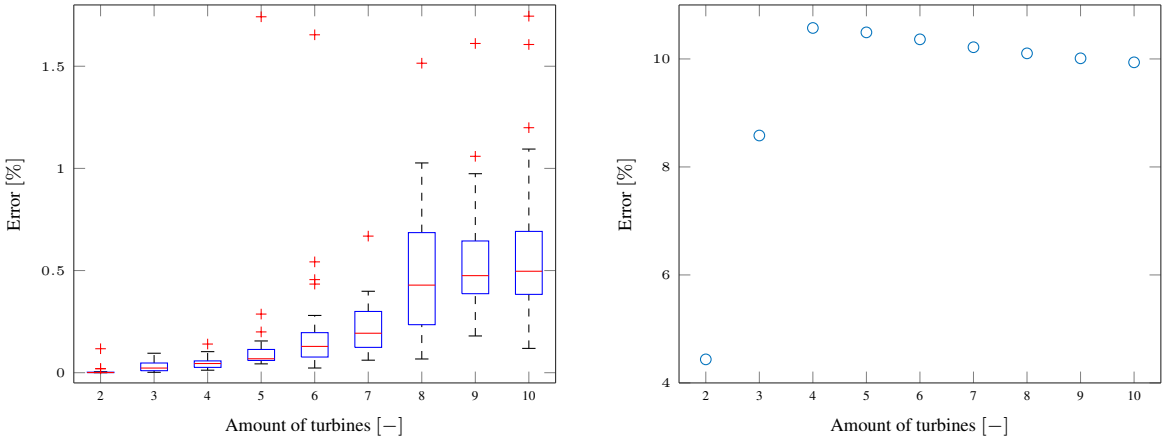


Figure 5. The contour plot of plant, approximate model and MA-GP model after ten iterations. The operating points of each iteration are marked with a cross.

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 is assumed that the measurements are noise-free. If noise is added to the power measurements the correct identification becomes more challenging and a larger training data set is necessary. A noise with a standard deviation of 50 kW is added to the measurement. The standard deviation is of the same size as the error in the power production of approximate model and plant at the optimal operating point. A training data set of 20 points is created. After ten iterations the error in the power production is about 0.6 %. The algorithm is able to converge. However, due to the measurement noise a

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



(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.

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 $0.35\,\%$.

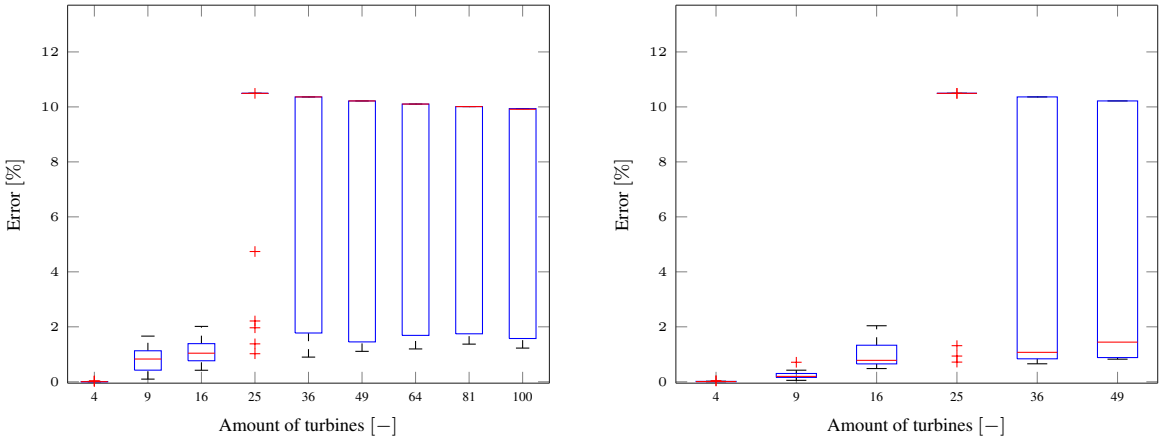
4.2 n turbine row case

In this subsection n turbines aligned in a row are optimized with the MA-GP algorithm. The size of the initial training set is chosen to be $n_d = 10n_u$, where n_u is the amount of control inputs. The operating points of the training set are randomly chosen. 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 statistic of the error after 25 iterations is shown in Fig. 6a. The error increases with the amount of turbines while it is almost zero for 2 to 4 turbines. Even though, the error increases with the amount of turbines the algorithm is able to reduce the model error significantly (Fig. 6b). It is not surprising that the error increases with the amount of control inputs. The control inputs are mapped to the total power output of the wind farm. With a large amount of control inputs the correct identification of this input-output map becomes more challenging, which increases the error in the MA-GP algorithm. Again, the error could be decreased with more data in the training set. 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 the data set. However, the computational time of the GP regression grows cubic with the amount of data. Therefore, at some point a trade-off between performance and computational time is necessary.

In contrast to purely model-free approaches, e.g. extremum seeking (Johnson and Fritsch, 2012) or MPPT (Gebraad et al.,

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.



(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.

2013), is the MA-GP the algorithm able to find a near optimal point in one iteration. The MA-GP model is already a better representation of the plant after the initial training than the approximate model. Nonetheless, measurements close to the optimum can help to refine the MA-GP model.

4.3 $n \times n$ turbine grid case

In this subsection the turbines in the wind farm a arranged in a $n \times n$ grid. The wind direction is aligned with the rows of the grid, interaction between parallel rows is neglectable. Consequently, the wind farm consist of n turbine rows each containing n turbines. The distance between turbines is $5\,\mathrm{D}$. The identification of the power production of this wind farm layout becomes more challenging. The input space increased and the sensitivity of inputs onto the total power production of the wind farm become similar.

Again the size of the initial training set is chosen to depend linearly on the size of the amount of control inputs with $n_d = 10n_u$. Otherwise the setup is the same as in the turbine row case.

The error after 25 iterations is shown in Fig. 7a. Again the algorithm converges for a small amount of turbines. 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 initial conditions, which is defined by the optimal operation point of the model (Fig. 6b)[2]. 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 is equal to the percentage in initial error of the grid.

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The problems to identify the plant model correctly with a larger inputs space are not surprising. The sample density decreases drastically for larger inputs spaces. 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. The plant optimization takes less than 10 s. Consequently, the initial data set could be increased to improve the performance of the larger wind farms.

The increase of the initial training set improves the convergence of the method for both small and large inputs spaces (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, it also has to be pointed out that the training of the hyperparameters in the GP regression scales cubic with the amount of data. Obviously this limits the size of the initial training set. Otherwise the approach becomes quickly computational infeasible. In case of an initial set of $n_d = 10n_u$ and a wind farm with 49 turbines the median time for the hyperparameter optimization is about 3 s. The maximum computational time in the 625 hyperparameter optimization is about 60 s. In case of an initial set of $n_d = 30n_u$ the median optimization time is about 50 s while the maximum optimization time is about 23 min. In these cases the optimization algorithm did not converge to an optimum and the maximum amount of iterations until termination was performed. The optimization time could be reduces by limiting the number of iterations. It is expected that it will not influence the performance since the objective function value in cases the optimization did not converge to an optimum did not change for most of the iterations.

If the MA-GP algorithm for the larger wind farms converges to an optimum it usually takes first a few iterations, where the wind farm is operated at the model optimum point, 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 actual 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.

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.

5 Conclusions

The modifier-adaptation approach with Gaussian processes applied to wind farm control is presented. It is a real-time optimization strategy, which corrects optimization model errors 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 exceptionally well for small input spaces. Here the GP regression is able to correct the model almost perfectly. Consequently, operating points very close

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

to the real optimum are found in the optimization. For larger input spaces, on the other hand, the error increases. 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.

In future work the performance of the method for large inputs spaces has to be improved. Several ideas are possible to achieve it:

- 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.
- Extend the algorithm with an exploration part. This can be achieved, for example, by including the variance of the GP regression model in the optimization.
- Include the single turbine power measurements in the identification of the GP regression model. In such a multi-input and multi-output approach the sensitivities of control inputs to the single outputs increase. The model identification should benefit from the approach. Moreover, it is expected that a smaller data set is necessary to achieve the same performance as with the in the article presented multi-inputs and single-output approach. The idea is pursued in Andersson et al. (2020a) with very promising results in increasing the accuracy of the approach with a smaller initial data set.

In addition, the sensitivity of the approach to measurements and inputs noise has to be investigated. 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).

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. This work was supported by OPWIND, RCN project no. 268044.

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

References

Adaramola, M. and Krogstad, P.-Å.: Experimental investigation of wake effects on wind turbine performance, Renewable Energy, 36, 2078–2086, 2011.

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 - [submitted], 2020b.

Andersson, L. E., Doekemeijer, B., van der Hoek, D., and Imsland, L.: Adaptation of Engineering Wake Models using Gaussian Process Regression and High-Fidelity Simulation Data, in: TORQUE 2020 - [submitted], 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.

Corten, G. and Schaak, P.: Heat and flux: Increase of wind farm production by reduction of the axial induction, in: Proceedings of the European Wind Energy Conference, 2003.

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.

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

del Rio Chanona, E., Alves Graciano, J., Bradford, E., and B., C.: Modifier-Adaptation Scheme Employing Gaussian Processes and Trust Regions for Real-Time Optimization, in: Proc. of the 12th IFAC Symposium on Dynamics and Control of Process Systems, Florianopolis, Brazil, 2019.

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.

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.

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. and Van Wingerden, J.: A control-oriented dynamic model for wakes in wind plants, in: Journal of Physics: Conference Series, vol. 524, p. 012186, IOP Publishing, 2014.

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.

Horvat, T., Spudić, V., and Baotić, M.: Quasi-stationary optimal control for wind farm with closely spaced turbines, in: MIPRO, 2012 Proceedings of the 35th International Convention, pp. 829–834, IEEE, 2012.

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.

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

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 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.

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

Medici, D.: Experimental studies of wind turbine wakes: power optimisation and meandering, Ph.D. thesis, KTH, 2005.

https://doi.org/10.5194/wes-2020-18
Preprint. Discussion started: 4 February 2020
© Author(s) 2020. CC BY 4.0 License.

Munters, W. and Meyers, J.: Effect of wind turbine response time on optimal dynamic induction control of wind farms, in: Journal of Physics: Conference Series, vol. 753, p. 052007, IOP Publishing, 2016.

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

Park, J., Kwon, S., and Law, K. H.: Wind farm power maximization based on a cooperative static game approach, in: Active and Passive Smart Structures and Integrated Systems 2013, vol. 8688, p. 86880R, International Society for Optics and Photonics, 2013.

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

Rotea, M. A.: Dynamic programming framework for wind power maximization, IFAC Proceedings Volumes, 47, 3639–3644, 2014.

Schepers, J. and Van der Pijl, S.: Improved modelling of wake aerodynamics and assessment of new farm control strategies, in: Journal of Physics: Conference Series, vol. 75, p. 012039, IOP Publishing, 2007.

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

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, 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.

Wagenaar, J., Machielse, L., and Schepers, J.: Controlling wind in ECN's scaled wind farm, Proc. Europe Premier Wind Energy Event, pp. 685–694, 2012.

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. ~~Real-time optimization (RTO) covers a family of optimization methods that incorporate process measurements in the optimization to drive the real process (plant) to optimal performance while guaranteeing constraint satisfaction. Modifier Adaptation (MA) introduces zeroth and first-order correction terms (bias and gradients) for the cost and constraint functions. Instead of updating the plant model, in MA the optimization problem is updated directly from data guaranteeing to meet the necessary condition of optimality upon convergence. The main burden of the MA approach is the estimation of the first-order modifiers of the cost and constraint functions at each RTO iteration. Finite-difference approximation is the most common approach that requires at least nu + 1 steady-state operation points to estimate the gradients, where nu is the number of control inputs. Obtaining these can require a long convergence time. For this reason, this work considers the use of Gaussian process (GP) regressionto estimate the plant-model mismatch based on plant measurements, and replace the usual modifiers by these high order regression functions. GP~~ 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 the surrogate model. 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~~well known in the machine learning community. The approach is tested on several numerical test cases simulating wind farms. It is shown that the approach is able to correct the model and converges to the plant optimal point. Several improvements for large inputs spaces, which is a challenging problem for the approach presented in the article, ~~ , 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 ~~the~~ wind turbines in a wind farm are operated ~~at their individual optimal operating point. This~~ 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* ~~wind farm control since the interactions between turbines are not taken into account. However, it~~ since it only focuses on the operation of an individual wind turbine. It is expected that ~~the greedy control strategy leads to sub-optimal performance of the wind farm (Steinbuch et al., 1988; Johnson and Thomas, 2009; Barthelmie et al., 2~~

~~. A coordinated wind farm controller, which takes the wake interactions between turbines in a wind farm into account , may result in a superior performance compared to the greedy wind farm controller.~~ 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 Thoma
.

The two main wind farm control strategies are axial induction control ~~, e.g. Steinbuch et al. (1988); Corten and Schaak (2003); Horvat et al.~~ and wake steering ~~control, e. g. Medici (2005); Adaramola and Krogstad (2011); Wagenaar et al. (2012); Park et al. (2013); Gebraad and V~~ ~~.~~ (Kheirabadi and Nagamune, 2019). The idea behind ~~the former~~ 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 ~~and the power production of the downwind turbine changes~~decreases. The target net effect is an overall increase of the power production and possibly ~~an~~ a decrease of fatigue loads. However, ~~recent studies suggest~~ 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~~(Schepers and Van der Pijl, 2007; Campagnolo et al., 20~~
.

~~The currently~~ Currently the more promising wind farm control strategy using ~~steady-state~~ 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 ~~promising~~ encouraging results were conducted by Fleming et al. (2017, 2019); Howland et al. (2019). In these experiments lookup tables with optimal yaw settings ~~of each turbine are created with help of an~~ depending on the wind conditions were created using steady-state ~~model. Hence the wind farm is operated in an~~ models. The look-up tables were not updated using plant measurements. Therefore, these approaches can be seen as open-loop~~control setting~~.

The steady-state ~~wake models used in~~ surrogate models must be simple to allow optimization but also accurate to permit good performance of the model-based ~~control are usually relatively simple . They estimate the velocity deficit in wakes. For a long time one of~~ controller. The development of surrogate models is an active research field. One of the most popular wake models ~~was~~ 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). ~~Hence, real time optimization (RTO) , which incorporates plant measurements to improve the performance of the wind farm controller, is extremely useful for this process.~~

~~Probably one of the most intuitive RTO strategies is the "two-step" approach. Here, first~~ 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 ~~are updated, and then new control inputs are computed based on the updated model . The two steps refer to the parameter optimization and control input optimization, which are performed sequentially (Marchetti et al., 2016)~~ of the surrogate model to improve closed-loop control was proposed by Doekemeijer et al. (2019b). However, if the ~~two-step approach cannot guarantee plant optimality upon convergence if the~~ parametric model is structurally incorrect ~~(Marchetti et al., 2016)~~ parameter estimation is not able to remove the mismatch between surrogate model and the plant. An example that an improved parameterisation of ~~the steady-state wake~~ a surrogate model was not able to remove the mismatch between a low order model and a high fidelity model ~~of wake~~ 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.

~~In contrast, modifier adaptation (MA) corrects the cost and constraint functions of the optimization problem directly, and reaches, under suitable assumptions, true plant optimality upon convergence (Marchetti et al., 2009). The bottleneck of the MAapproach~~ 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 ~~the estimation of the gradients of the objective and constraint functions at each RTO iteration. Finite difference approximation is one of the most common approaches that requires $n_u + 1$ steady-state operation points to estimate the gradients, where n_u is the amount of control inputs. These can lead to a long convergence time, especially for processes with high dimensional input spaces. Therefore, in this work Gaussian process (GP) regression is combined with MA (de Avila Ferreira et al., 2018; ?)~~used. GP is a probabilistic, non-parametric modelling technique well known in the machine learning community (Rasmussen and Williams, 2006). The ~~GP regression model estimates the plant-model mismatch using plant measurements.Then the GP model is used~~ 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 ~~original optimization problem and by this improve the optimization of the plant inputs~~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~~and Gaussian process regression is explained. .~~ 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 ~~on several examples~~ 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

~~The optimization problem of the~~ Model-based wind farm optimization usually employs a steady-state ~~plant performance subject to constraints can~~ 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 ~~(Marchetti et al., 2016):~~

$$\begin{aligned} \mathbf{u}_p^* &= \arg\min_{\mathbf{u}} \phi_p(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \\ s.t.\ G_{p,j}(\mathbf{u}) &:= g_{p,j}(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \leq 0,\ \ j = 1, \ldots, n_g, \\ \mathbf{u} &\in \mathcal{U}, \end{aligned}$$

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

where $\mathbf{u} \in \mathbb{R}^{n_u}$ ~~and $\mathbf{y}_p \in \mathbb{R}^{n_y}$~~ denote the plant input ~~and output variables, respectively; $\mathbf{u} \in \mathbb{R}^{n_u}$~~ 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 $\mathbf{y}_p \in \mathbb{R}^{n_y}$ ~~are the input-output pairs of the wind farm; $\phi_p : \mathbb{R}^{n_u} \to \mathbb{R}$ is the cost function to be minimized; $g_{p,j} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R},\ j = 1, \ldots, n_g,$ are the inequality constraint functions; and~~ $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ is the control domain, e.g. box constraints on the control inputs. ~~Formulation (1) assumes that ϕ_p and $g_{p,j}$ as functions of $\mathbf{u}$, and $\mathbf{y}_p$ are exactly known. However, in any practical application the exact input-output map of the plant is unknown and instead an approximate~~

The challenge of optimizing the power production of a wind farm is that only an approximate surrogate model of the ~~system is exploited for the optimization:~~

$$\begin{aligned} \mathbf{u}^* &= \arg\min_{\mathbf{u}} \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) \\ s.t.\ G_j(\mathbf{u}) &:= g_j(\mathbf{u}, \mathbf{y}(\mathbf{u})) \leq 0,\ \ j = 1, \ldots, n_g, \\ \mathbf{u} &\in \mathcal{U}, \end{aligned}$$

~~where the quantities ϕ, $g_j(\mathbf{u}, \mathbf{y}(\mathbf{u}))$, $\mathbf{u}*$, and G_j refer to the inexact model counterparts of the true plantoptimization problem in Eq. (1). RTO takes advantage of the available measurements to compensate for plant-model mismatch and adapt the model-based optimization problem Eq. (??) to reach plant optimality.~~ 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 ~~approach applies first-order correction terms that are added to~~ adds first order modifiers to correct the gradient of the ~~cost and constraint functions to match the necessary conditions of optimality upon~~

convergence (Marchetti et al., 2009). Iteratively the following modified optimization problem is solved:

$$\hat{\mathbf{u}}^*_{k+1} = \arg\min_{\mathbf{u}} \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) + (\boldsymbol{\lambda}^{\phi}_k)^T \mathbf{u}$$

$$s.t. \quad G_j(\mathbf{u}) + \varepsilon_{j,k} + (\boldsymbol{\lambda}^{G_j}_k)^T (\mathbf{u} - \mathbf{u}_k) \leq 0, \quad j = 1, \ldots, n_g,$$

$$\mathbf{u} \in \mathcal{U},$$

where $\hat{\mathbf{u}}^*_{k+1}$ is the optimal solution at iteration $k+1$, the $\varepsilon_{j,k} \in \mathbb{R}$ are the zeroth-order modifiers for the constraints, and $\boldsymbol{\lambda}^{\phi}_k$ and $\boldsymbol{\lambda}^{G_j}_k$ are the first-order modifiers for the cost and constraints, respectively. The correction terms are given by:

$$\varepsilon_{j,k} := G_{p,j}(\mathbf{u}_k) - G_j(\mathbf{u}_k),$$

$$(\boldsymbol{\lambda}^{\phi}_k)^T := \frac{\partial \phi_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \phi}{\partial \mathbf{u}}(\mathbf{u}_k),$$

$$(\boldsymbol{\lambda}^{G_j}_k)^T := \frac{\partial G_{p,j}}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial G_j}{\partial \mathbf{u}}(\mathbf{u}_k).$$

It is recommended to filter the input update $\hat{\mathbf{u}}^*_{k+1}$ to avoid excessive correction and reduce sensitivity to noise (Marchetti et al., 2016) :

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{L}(\hat{\mathbf{u}}_{k+1} - \mathbf{u}_k),$$

with $\mathbf{L} = \text{diag}(l_1, \ldots, l_{n_u})$, $l_i \in (0, 1]$ where l_i may be reduced to help stabilize the iterations. The MA scheme requires the surrogate model. However, the estimation of the plant gradients at each RTO iteration , which in each iteration is experimentally expensive and the main bottleneck for 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 3.1. Thereafter, in Section 3.2, 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 regressionfor our purposes, for more information refer to consult Rasmussen and Williams (2006). GP regression aims to identify identifies an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ from data. Let the noisy observation It is assumed that the noisy observations of $f(\cdot)$ be 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^2_ν, $\nu_k \sim \mathcal{N}(0, \sigma^2_\nu)$.
We assume In GP regression, $f(\cdot)$ to follow a GP with 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 $\mathbf{\Lambda} = \text{diag}(\lambda_1^2, \ldots, \lambda_{n_u}^2)$ is a scaling matrix.

~~Assume we are given a training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ of size M consisting of M input vectors $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_M]^\mathsf{T}$ and corresponding observations $\mathbf{y} = [y_1, \ldots, y_M]^\mathsf{T}$ according to Eq. (2). From the GP distribution the data then follows a joint multivariate Gaussian distribution, which can be stated as:~~

$$p(\mathbf{y}|\mathbf{U}) = \mathcal{N}(\mathbf{0}, \mathbf{K} + \sigma_\nu^2 \mathbf{I}), \quad K_{ij} = k(\mathbf{u}_i, \mathbf{u}_j)$$

~~The hyperparameters $\boldsymbol{\psi} := [\sigma_f, \sigma_\nu, \lambda_1, \ldots, \lambda_{n_\mathbf{u}}]^T$ are commonly unknown and hence need to be inferred from data. In this article the log marginal likelihood $p(\mathbf{y}|\mathbf{U})$ is used. Ignoring constant terms and factors, this can be stated as:~~

$$\mathcal{L}(\mathcal{D}, \Psi) = -\frac{1}{2}\mathbf{y}^T(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{y} - \frac{1}{2}\ln|\mathbf{K} + \sigma_\nu^2 \mathbf{I}|.$$

~~The required maximum likelihood estimate is then given by $\hat{\boldsymbol{\psi}} \in \arg\max_{\boldsymbol{\psi}} \mathcal{L}(\mathcal{D}, \boldsymbol{\psi})$. Next we require the~~ Due to the GP assumption the predictive distribution of $f(\mathbf{u})$ at an arbitrary input $\mathbf{u}$ ~~, which can be found by the conditional distribution of $f(\mathbf{u})$ on the data distribution $p(\mathbf{y}|\mathbf{U})$. From the GP assumption this~~ given the training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ has a closed-form solution~~and can be stated as:~~

$$f(\mathbf{u})|\mathcal{D}, \hat{\boldsymbol{\psi}} \sim \mathcal{N}(\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}}), \sigma_{\text{GP}}^2(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}}))$$

$$\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}}) = \mathbf{k}^T(\mathbf{u})(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{y}$$

$$\sigma_{\text{GP}}^2(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})) = \sigma_f^2 - \mathbf{k}^T(\mathbf{u})(\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1}\mathbf{k}(\mathbf{u})$$

~~where~~ . The resulting mean $\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ can be seen as the GP prediction at $\mathbf{u}$ and the variance $\sigma_{\text{GP}}^2(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ as a corresponding measure of uncertainty to this prediction. ~~The GP is a non-parametric model. The~~

The performance of the GP is dependent on hyperparameters $\hat{\boldsymbol{\psi}}$. 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 ~~the above expression~~ this a matrix of size $M \times M$ must be inverted, ~~which prohibits~~ where M is the number of measurements. Clearly, this makes large data sets challenging.

4 ~~Methodology~~

3.1 Modifier Adaptation with Gaussian processes

The use of GPs in a MA approach to overcome In the MA-GP approach the limitation of estimating the plant gradients was first proposed by de Avila Ferreira et al. (2018). The idea is to replace the zeroth- and first-order modifiers of the cost and constraints in (??) with GP regression terms. Since the wind farms considered in this article do not have inequality constraint functions they are not included in this section. However, inequality constraint functions can be easily incorporated into the method. The training set of the GP to correct the objective function are the controlled inputs of the approximate model and the plant-model mismatch of the objective function. The new 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 scheme with GP modifiers (MA-GP) is becomes

$$\hat{\mathbf{u}}^*_{k+1} = \arg\min\max_{\mathbf{u}} \phi P(\mathbf{u},()) + \mu_{\mathrm{GP},k}^{\phi_p-\phi}(\mathbf{u};\mathcal{D}_{0k},\hat{\psi}_{0k}), \qquad s.t. \quad \mathbf{u}\in\mathcal{U}, \tag{4}$$

where the plant-model mismatch of the cost function is modelled by $\mu_{\mathrm{GP}}^{\phi_p-\phi}$. Similar to the original MA scheme the optimal input of Eq. (4) may be filtered with Eq. (5) to reduce the step-size and help stabilize the MA-GP scheme (?). The whole MA-GP scheme is presented in Algorithm 1 and Fig. 1. 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.

In Algorithm 1 the hyperparameters are updated if The MA-GP approach for wind farm optimization is visualised in Fig.

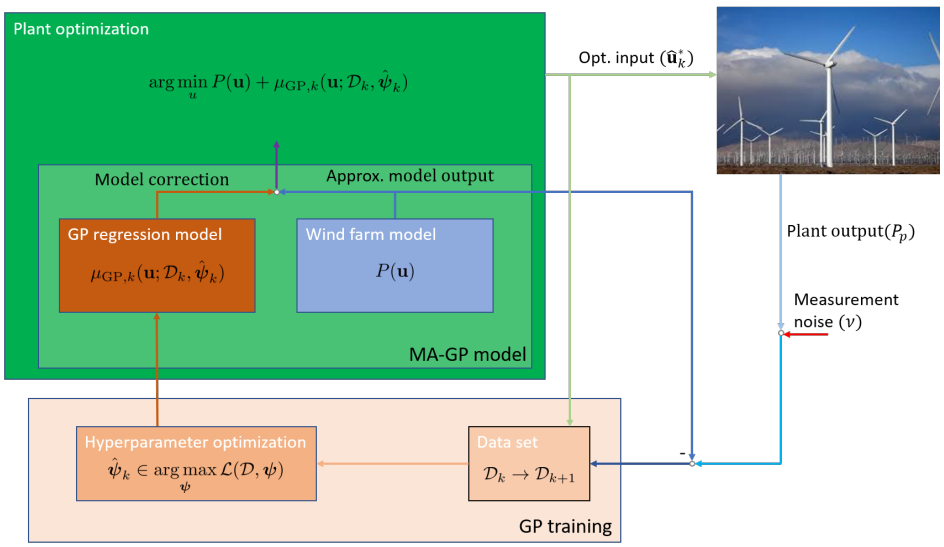


Figure 1. The basic idea of the MA-GP scheme for a wind farm. The GP regression model creates an input-output map of the control inputs to the plant-model mismatch. In the MA-GP model the GP regression model is used to correct the output of the approximate model. This MA-GP model is used in the optimization 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]

[1] The wind farm picture is by Erik Wilde from Berkeley, CA, USA , *Wind turbines in southern California 2016,*

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~~. *HypOpt* , which~~ is a user-defined condition~~, which allows to update the hyperparameter. The extrema are to update the hyerparameter each iteration or never. The~~ . The hyperparameter update is usually the computational bottle-neck of the MA-GP algorithm. ~~Especially~~ We observed that especially for large data sets ~~it can be expected that the hyperparameter~~ the hyperparameters do not change much from one iteration to the next. Therefore, ~~it is reasonable~~ the hyperparameters can be updated less frequent to decrease computational delay. However, it is recommend to update the hyperparameters ~~less frequent.~~ as often as possible.

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

3.2 Numerical turbine and wake models

A turbine and wake model are necessary to create a model of a wind farm. The wind turbines ~~in the wind farm~~ are represented using the actuator 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 et al., 2016)

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

where A is the rotor area, ρ the air density~~and~~ , 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

[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

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 study it will be important to implement a "plant" and model, which are different from each other. ~~Therefore, a~~ 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 ~~a new model based on the actuator disk~~ 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 resulting in

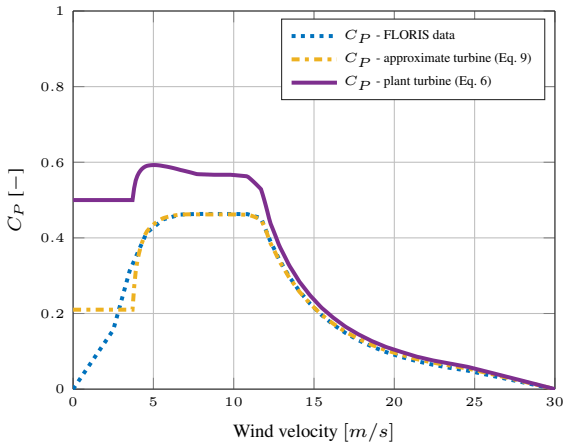
$$C_P = 7.037a(0.625 - a)^{1.364}. \tag{9}$$

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 ~~both models~~ 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 ~~$D = 136\,\text{m}$~~ $D = 126.4\,\text{m}$ and the hub height $H_H = 90\,\text{m}$.

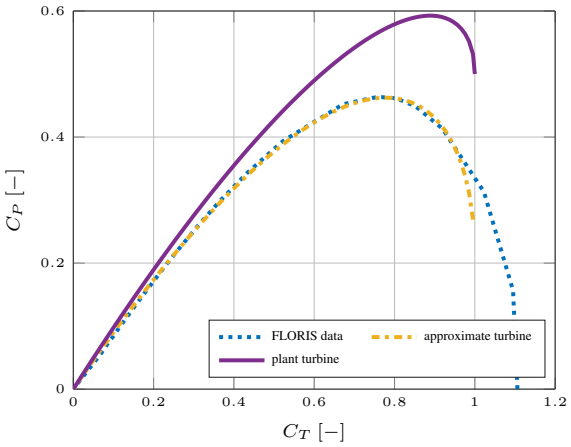
The Gaussian wake model by Bastankhah and Porté-Agel (2014, 2016) is used to model the flow in the wind farm. The three-dimensional steady-state far wake velocity ~~is assumed to be~~ deficit is Gaussian distributed and can be estimated ~~with~~ by

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

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



(a) Comparison of C_P values in dependency of the wind velocity.~~The dashed line gives the corresponding thrust coefficient, which is the same for both models.~~



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

Figure 2. Comparison between data, the ~~new model based on the actuator disk model~~ plant turbine and the ~~actuator disk~~ approximate turbine model. The ~~thrust coefficients are kept smaller than one for the actuator disk~~ models ~~. The models~~ give ~~a~~ different ~~connection~~ connections between thrust and power coefficients.

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. ~~In the simulation study different values are chosen for this parameter in the plant and approximated model resulting in different optimal operating points~~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 ~~amount~~ 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 error in the power production is given by

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

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

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

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

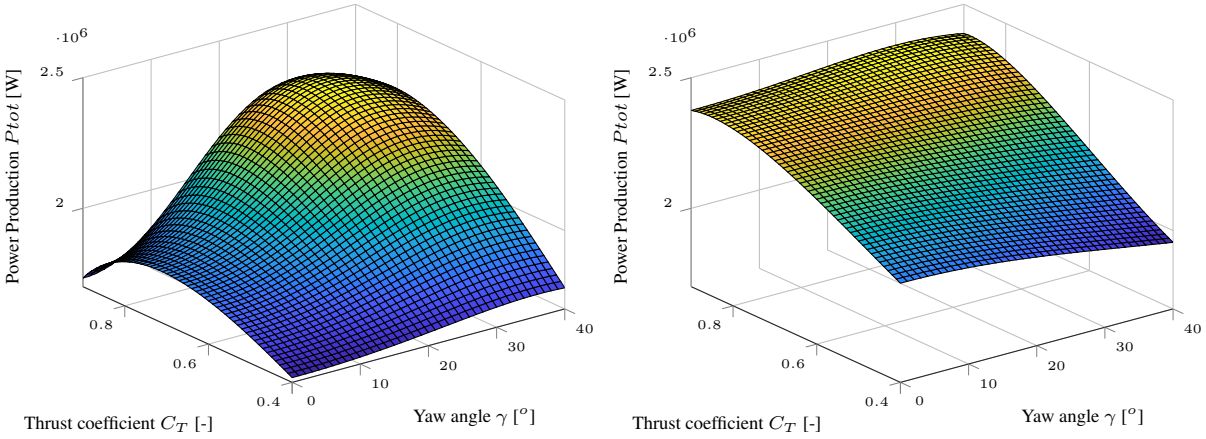
The yaw angles γ_i are constrained to positive yaw angles since the Gaussian wake model is symmetric. Asymmetry as in a real wind farm is not represented in the models used in this article. If the MA-GP approach is applied to a real wind farm it would be unnecessary to constrain the yaw angle to positive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



(a) Surface of objective function of plant.

(b) Surface of objective function of approximate model.

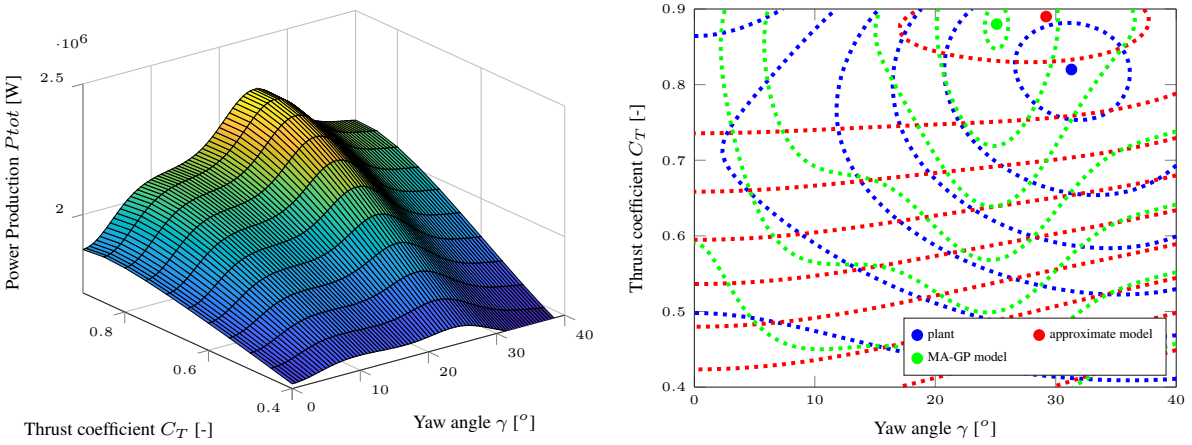
Figure 3. The power production of plant and approximate model in dependency of the control inputs of the upwind turbine.

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 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 ~~1.67 %~~$\Theta = 1.67\%$. Still, the model ~~assume~~ 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 hyperparameters 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 operating point ~~. Fig 5shows the operating points of the first ten iterations~~(Fig 5). After four iterations the ~~error in power production~~ 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



(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.

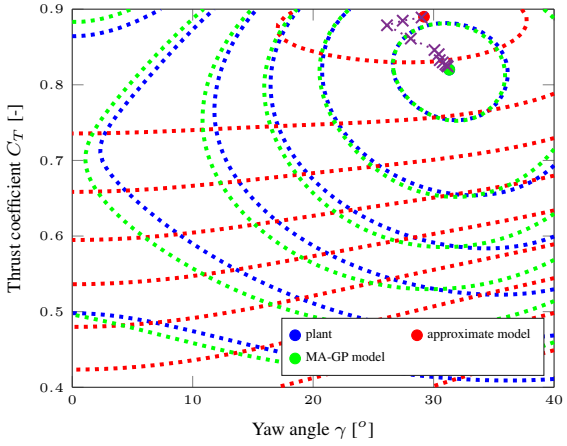
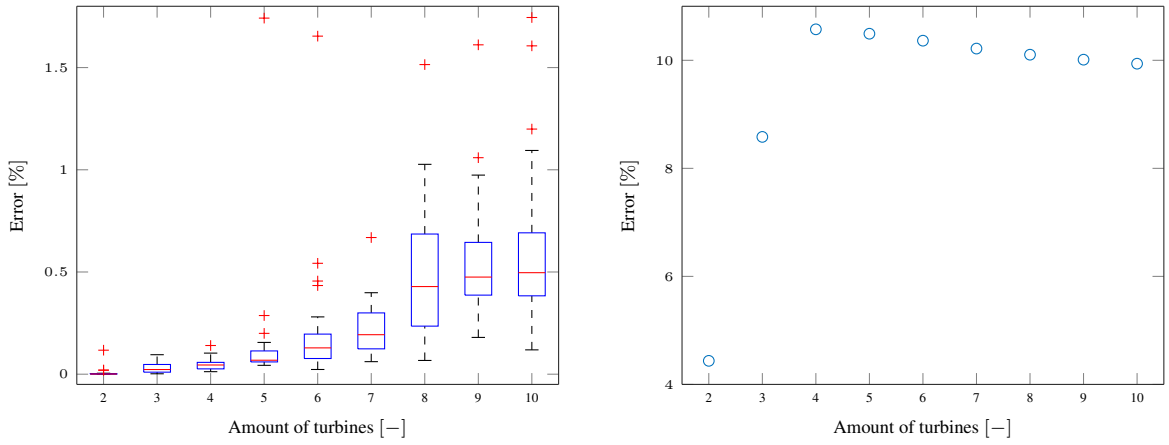


Figure 5. The contour plot of plant, approximate model and MA-GP model after ten iterations. The operating points of each iteration are marked with a cross.

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 ~~is~~ was assumed that the measurements are noise-free. If noise is added to the power measurements the correct identification becomes more challenging and a larger training data set is necessary. A noise with a standard deviation of 50 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 ~~approximate model and plant~~ plant and approximate model



(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.

at the optimal operating point of the plant. A training data set of 20 points is created. After ten iterations the ~~error in the power production~~ relative error Θ is about $0.6\,\%$. The algorithm 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 $0.35\,\%$.

4.2 n turbine row case

In this subsection the optimization of n turbines aligned in a row ~~are optimized with the~~ 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 ~~algorithm. The~~ 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. The operating points of the training set are ~~randomly chosen~~ 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 ~~statistic of the error after 25 iterations is shown in Fig. 6a. The error~~ error increases with the amount of turbines while it is almost zero for 2 to 4 turbines ~~. Even though, the error increases with the amount of turbines the algorithm is able to reduce the model error significantly~~ (Fig. ~~6b). It is not surprising that the error increases with the amount of control inputs. The control inputs are mapped to the total~~ 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 ~~wind farm. With a large amount of control inputs the correct~~

identification of this plant. It makes it challenging to correctly identity the input-output mapbecomes more challenging, which increases the error in the MA-GP algorithm. Again, the error could . The error can be decreased with more data in the training set. 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. However, the computational time of the GP regression grows cubic with the amount of data. Therefore, at some point a trade-off between performance and computational time is necessary.

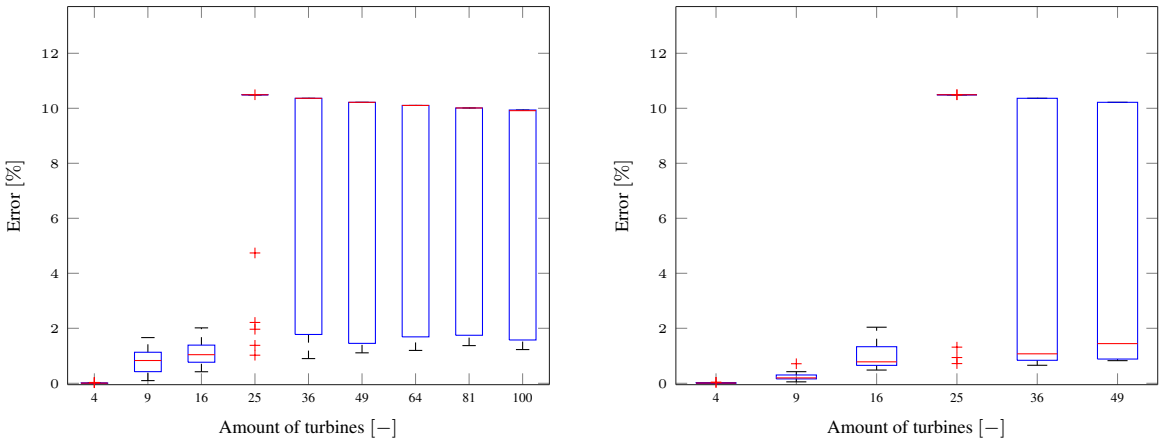
In Assuming a sufficient large initial training set the MA-GP approach is able to find the near optimal point in one iteration since the approach basically just improves the surrogate model. This stands in contrast to purely model-free approaches, e.g. extremum seeking (Johnson and Fritsch, 2012) or MPPT (Gebraad et al., 2013), is the which usually need several iterations to find an optimum. Moreover, after the initial training the MA-GP the algorithm able to find a near optimal point in one iteration. The MA-GP model is already a better representation usually represents of the plant after the initial training better than the approximate model. Nonetheless, measurements close to the optimum of the MA-GP model can help to refine the MA-GP model improve the model further.

4.3 $n \times n$ turbine grid case

In this subsection the turbines in the wind farm a optimization of a wind farm with turbines arranged in a $n \times n$ grid . The wind direction is aligned with the rows of the grid. Interaction between parallel rows is neglectablewith a spacing of $5\,\mathrm{D}$ is presented. Consequently, the wind farm consist of n turbine rows each containing n turbines. The distance between turbines is 5 D. The identification of the power production of this wind farm layout becomes more challenging. The input space increased and the sensitivity of inputs onto the total power production of the wind farm become similar. wind direction is aligned with the rows 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 $n_d = 10n_u$. Otherwise the setup is the same as in the turbine row caseand the MA-GP approach is tested on 25 Monte Carlo simulations. The error after 25 iterations is shown in Fig. 7a. 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 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.

The problems to identify the plant modelcorrectly with a larger inputs space are not surprising. The sample density decreases drastically for larger inputs spacesIf 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

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



(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.

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 ~~initial data set could be increased~~ 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 ~~for both small and large inputs spaces~~ (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, ~~it also has to be pointed out that~~ the training of the hyperparameters in the GP regression scales cubic with the amount of data. Obviously this ultimately limits the size of the ~~initial training set . Otherwise the approach becomes quickly~~ training set since the approach can become computational infeasible. ~~In case of an initial set of $n_d = 10n_u$ and a wind farm with 49 turbines the median time for the hyperparameter optimization is about 3 s. The maximum computational time in the 625 hyperparameter optimization is about 60 s. In case of an initial set of $n_d = 30n_u$ the median optimization time is about 50 s while the maximum optimization time is about 23 min. In these cases the optimization algorithm did not converge to an optimum and the maximum amount of iterations until termination was performed. The optimization time could be reduces by limiting the number of iterations. It is~~

expected that it will not influence the performance since the objective function value in cases the optimization did not converge to an optimum did not change for most of the iterations.

If the MA-GP algorithm for the larger wind farms converges to an optimum it usually takes first a few iterations, where the wind farm is operated at the model optimum point, 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 actual 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.

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. (2020d)[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.

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).

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

[3] The article is submitted to TORQUE 2020

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.

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 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 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.
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

The modifier-adaptation approach with Gaussian processes applied to wind farm control is presented. It is a real-time optimization strategy, which corrects ~~optimization model errors~~ 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 ~~exceptionally~~ 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. 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.

In future work the performance of the method for large inputs spaces has to be improved . Several ideas are possible to achieve itThe 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 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 regression model in the optimization.
- Include the single turbine power measurements in the identification of the GP regression model. In such a multiple-input and multiple-output approach the sensitivities of control inputs to the single outputs increase. The model identification should benefit from the approach. Moreover, it is expected that a smaller data set is necessary to achieve the same performance as with the in the article presented multiple-inputs and single-output approach. The idea is pursued in Andersson et al. (2020a) with very promising results in increasing the accuracy of the approach with a smaller initial data set.

In addition, the An important investigation is the sensitivity of the approach to measurementsand inputs noise has to be investigated, 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)Andersson et al. (2020d).

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

Adaramola, M. and Krogstad, P.-Å.: Experimental investigation of wake effects on wind turbine performance, Renewable Energy, 36, 2078–2086, 2011.

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., and Imsland, L.: Adaptation of Engineering Wake Models using Gaussian Process Regression and High-Fidelity Simulation Data, in: TORQUE 2020 - [submitted], 2020c.

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, 2020d.

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.

Corten, G. and Schaak, P.: Heat and flux: Increase of wind farm production by reduction of the axial induction, in: Proceedings of the European Wind Energy Conference, 2003.

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.

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), 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.

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.

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, 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. and Van Wingerden, J.: A control-oriented dynamic model for wakes in wind plants, in: Journal of Physics: Conference Series, vol. 524, p. 012186, IOP Publishing, 2014.

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.

Horvat, T., Spudić, V., and Baotić, M.: Quasi-stationary optimal control for wind farm with closely spaced turbines, in: MIPRO, 2012 Proceedings of the 35th International Convention, pp. 829–834, IEEE, 2012.

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.

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 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.

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 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.

Medici, D.: Experimental studies of wind turbine wakes: power optimisation and meandering, Ph.D. thesis, KTH, 2005.

Munters, W. and Meyers, J.: Effect of wind turbine response time on optimal dynamic induction control of wind farms, in: Journal of Physics: Conference Series, vol. 753, p. 052007, IOP Publishing, 2016.

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., and Law, K. H.: Wind farm power maximization based on a cooperative static game approach, in: Active and Passive Smart Structures and Integrated Systems 2013, vol. 8688, p. 86880R, International Society for Optics and Photonics, 2013.

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

Rotea, M. A.: Dynamic programming framework for wind power maximization, IFAC Proceedings Volumes, 47, 3639–3644, 2014.

Schepers, J. and Van der Pijl, S.: Improved modelling of wake aerodynamics and assessment of new farm control strategies, in: Journal of Physics: Conference Series, vol. 75, p. 012039, IOP Publishing, 2007.

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

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, 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.

Wagenaar, J., Machielse, L., and Schepers, J.: Controlling wind in ECN's scaled wind farm, Proc. Europe Premier Wind Energy Event, pp. 685–694, 2012.

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. [..[1]][..[2]]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 the surrogate model. 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 [..[3]]non-parametric modelling technique[..[4]], 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 [..[5]]wind turbines in a wind farm are operated [..[6]]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* [..[7]]since it only focuses on the operation of an individual wind turbine. It is expected that [..[8]]a wind farm control strategy that takes the interaction between turbines into account can improve the overall

[1]removed: Real-time optimization (RTO) covers a family of optimization methods that incorporate process measurements in the optimization to drive the real process (plant) to optimal performance while guaranteeing constraint satisfaction. Modifier Adaptation (MA) introduces zeroth and first-order correction terms (bias and gradients) for the cost and constraint functions. Instead of updating the plant model, in MA the optimization problem is updated directly from data guaranteeing to meet the necessary condition of optimality upon convergence.

[2]removed: The main burden of the MA approach is the estimation of the first-order modifiers of the cost and constraint functions at each RTO iteration. Finite-difference approximation is the most common approach that requires at least nu + 1 steady-state operation points to estimate the gradients, where nu is the number of control inputs. Obtaining these can require a long convergence time. For this reason, this work considers the use of Gaussian process (GP) regressionto estimate the plant-model mismatch based on plant measurements, and replace the usual modifiers by these high order regression functions. GP

[3]removed: ,

[4]removed: well known in the machine learning community. The approach is tested on several numerical test cases simulating wind farms. It is shown that the approach is able to correct the model and converges to the plant optimal point. Several improvements for large inputs spaces, which is a challenging problem for the approach presented in the article,

[5]removed: the

[6]removed: at their individual optimal operating point. This

[7]removed: wind farm control since the interactions between turbines are not taken into account. However, it

[8]removed: the greedy control strategy leads to sub-optimal performance of the wind farm (Steinbuch et al., 1988; Johnson and Thomas, 2009; Barthelmie et al., 2009). A coordinated wind farm controller, which takes the wake interactions between turbines in a wind farm into account , may result in a superior performance compared to the greedy wind farm controller.

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 [..[9]]and wake steering [..[10]](Kheirabadi and Nagamune, 2019). The idea behind [..[11]]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 [..[12]]decreases. The target net effect is an overall increase of the power production and possibly [..[13]]a decrease of fatigue loads. However, [..[14]]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[..[15]].

[..[16]]Currently the more promising wind farm control strategy using [..[17]]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 [..[18]]encouraging results were conducted by Fleming et al. (2017, 2019); Howland et al. (2019). In these experiments lookup tables with optimal yaw settings [..[19]]depending on the wind conditions were created using steady-state [..[20]]models. The look-up tables were not updated using plant measurements. Therefore, these approaches can be seen as open-loop[..[21]].

The steady-state [..[22]]surrogate models must be simple to allow optimization but also accurate to permit good performance of the model-based [..[23]]controller. The development of surrogate models is an active research field. One of the most popular wake models [..[24]]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,

[9] removed: , e.g. Steinbuch et al. (1988); Corten and Schaak (2003); Horvat et al. (2012); Rotea (2014); Munters and Meyers (2016)
[10] removed: control, e. g. Medici (2005); Adaramola and Krogstad (2011); Wagenaar et al. (2012); Park et al. (2013); Gebraad and Van Wingerden (2014).
[11] removed: the former
[12] removed: and the power production of the downwind turbine changes
[13] removed: an
[14] removed: recent studies suggest
[15] removed: (Schepers and Van der Pijl, 2007; Campagnolo et al., 2016; Bartl and Sætran, 2016; Annoni et al., 2016)
[16] removed: The currently
[17] removed: steady-state
[18] removed: promising
[19] removed: of each turbine are created with help of an
[20] removed: model. Hence the wind farm is operated in an
[21] removed: control setting
[22] removed: wake models used in
[23] removed: control are usually relatively simple . They estimate the velocity deficit in wakes. For a long time one of
[24] removed: was

they are just vague approximations of a complex phenomena that is, in fact, not well understood (Veers et al., 2019). [..[25]][..[26]]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 [..[27]]of the surrogate model to improve closed-loop control was proposed by Doekemeijer et al. (2019b). However, if the [..[28]]parametric model is structurally incorrect [..[29]]parameter estimation is not able to remove the mismatch between surrogate model and the plant. An example that an improved parameterisation of [..[30]]a surrogate model was not able to remove the mismatch between a low order model and a high fidelity model [..[31]]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.

[..[32]]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 [..[33]]used. GP is a probabilistic, non-parametric modelling technique well known in the machine learning community (Rasmussen and Williams, 2006). The [..[34]]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 [..[35]]surrogate model in a flexible manner (de Avila Ferreira et al., 2018) and improve the performance of the wind farm controller.

[25] removed: Hence, real time optimization (RTO) , which incorporates plant measurements to improve the performance of the wind farm controller, is extremely useful for this process.

[26] removed: Probably one of the most intuitive RTO strategies is the "two-step" approach. Here, first

[27] removed: are updated, and then new control inputs are computed based on the updated model . The two steps refer to the parameter optimization and control input optimization, which are performed sequentially (Marchetti et al., 2016)

[28] removed: two-step approach cannot guarantee plant optimality upon convergence if the

[29] removed: (Marchetti et al., 2016)

[30] removed: the steady-state wake

[31] removed: of wake

[32] removed: In contrast, modifier adaptation (MA) corrects the cost and constraint functions of the optimization problem directly, and reaches, under suitable assumptions, true plant optimality upon convergence (Marchetti et al., 2009). The bottleneck of the MAapproach

[33] removed: the estimation of the gradients of the objective and constraint functions at each RTO iteration. Finite difference approximation is one of the most common approaches that requires $n_u + 1$ steady-state operation points to estimate the gradients, where n_u is the amount of control inputs. These can lead to a long convergence time, especially for processes with high dimensional input spaces. Therefore, in this work Gaussian process (GP) regression is combined with MA (de Avila Ferreira et al., 2018; **?**)

[34] removed: GP regression model estimates the plant-model mismatch using plant measurements.Then the GP model is used

[35] removed: original optimization problem and by this improve the optimization of the plant inputs

The article is structured as follows: In Section 2 the optimization problem is formulated[..[36]]. 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 [..[37]]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

[..[38]]Model-based wind farm optimization usually employs a steady-state [..[39]]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 [..[40]]

[..[41]] [..[42]]

[..[43]] [..[44]]

[..[45]]

$$\mathbf{u}_\mathsf{p}^* = \arg\max_{\mathbf{u}} \mathsf{P}_\mathsf{p}(\mathbf{u}), \quad \mathbf{u} \in \mathcal{U}, \quad \mathbf{u} = [\mathsf{u}_1^\mathsf{T}, \mathsf{u}_2^\mathsf{T}, \ldots, \mathsf{u}_\mathsf{N}^\mathsf{T}]^\mathsf{T}, \tag{1}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$ [..[46]]denote the plant input [..[47]]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 [..[48]]$\mathcal{U} \subseteq \mathbb{R}^{n_u}$ is the control domain, e.g. box constraints on the control inputs. [..[49]]

The challenge of optimizing the power production of a wind farm is that only an approximate surrogate model of the [..[50]]

[..[51]] [..[52]]

[..[53]] [..[54]]

[..[55]]

[36] removed: and Gaussian process regression is explained.

[37] removed: on several examples

[38] removed: The optimization problem of the

[39] removed: plant performance subject to constraints can

[40] removed: (Marchetti et al., 2016):

[46] removed: and $\mathbf{y}_p \in \mathbb{R}^{n_y}$

[47] removed: and output variables, respectively; $\mathbf{u} \in \mathbb{R}^{n_u}$

[48] removed: $\mathbf{y}_p \in \mathbb{R}^{n_y}$ are the input-output pairs of the wind farm; $\phi_p : \mathbb{R}^{n_u} \to \mathbb{R}$ is the cost function to be minimized; $g_{p,j} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $j = 1, \ldots, n_g$, are the inequality constraint functions; and

[49] removed: Formulation (1) assumes that ϕ_p and $g_{p,j}$ as functions of $\mathbf{u}$, and $\mathbf{y}_p$ are exactly known. However, in any practical application the exact input-output map of the plant is unknown and instead an approximate

[50] removed: system is exploited for the optimization:

[..[56]][..[57]]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 [..[58]]adds first order modifiers to correct the gradient of the [..[59]]

[..[60]]

[..[61]] [..[62]]

[..[63]]

[..[64]]

[..[65]] [..[66]]

[..[67]] [..[68]]

[..[69]] [..[70]]

[..[71]]

[..[72]]

[..[73]][..[74]]surrogate model. However, the estimation of the plant gradients [..[75]]in each iteration is experimentally expensive and the main bottleneck [..[76]]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 3.1. Thereafter, in Section 3.2, the turbine and wake models used in the case study are explained.

[56] removed: where the quantities ϕ, $g_j(\mathbf{u}, \mathbf{y}(\mathbf{u}))$, $\mathbf{u}*$, and G_j refer to the inexact model counterparts of the true plantoptimization problem in Eq. (1).

[57] removed: RTO takes advantage of the available measurements to compensate for plant-model mismatch and adapt the model-based optimization problem Eq. (**??**) to reach plant optimality.

[58] removed: approach applies first-order correction terms that are added to

[59] removed: cost and constraint functions to match the necessary conditions of optimality upon convergence (Marchetti et al., 2009). Iteratively the following modified optimization problem is solved:

[64] removed: where $\hat{\mathbf{u}}^*_{k+1}$ is the optimal solution at iteration $k+1$, the $\varepsilon_{j,k} \in \mathbb{R}$ are the zeroth-order modifiers for the constraints, and $\boldsymbol{\lambda}_k^{\phi}$ and $\boldsymbol{\lambda}_k^{G_j}$ are the first-order modifiers for the cost and constraints, respectively. The correction terms are given by:

[71] removed: It is recommended to filter the input update $\hat{\mathbf{u}}^*_{k+1}$ to avoid excessive correction and reduce sensitivity to noise (Marchetti et al., 2016):

[73] removed: with $\mathbf{L} = \text{diag}(l_1, \ldots, l_{n_u})$, $l_i \in (0, 1]$ where l_i may be reduced to help stabilize the iterations.

[74] removed: The MA scheme requires the

[75] removed: at each RTO iteration , which

[76] removed: for

3.1 Gaussian processes

In this section we give a brief outline of GP regression[..77], for more information [..78]consult Rasmussen and Williams (2006). GP regression [..79]identifies an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ from data. [..80]It is assumed that the noisy observations of $f(\cdot)$ [..81]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, $\nu_k \sim \mathcal{N}(0, \sigma_\nu^2)$.
[..82]In GP regression, $f(\cdot)$ [..83]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 $\mathbf{\Lambda} = \text{diag}(\lambda_1^2, \ldots, \lambda_{n_u}^2)$ is a scaling matrix.
[..84]

[..85][..86][..87]

[..88]

[..89]

---

[77] removed: for our purposes

[78] removed: refer to

[79] removed: aims to identify

[80] removed: Let the noisy observation

[81] removed: be

[82] removed: We assume

[83] removed: to follow a GP with

[84] removed: Assume we are given a training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ of size M consisting of M input vectors $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_M]^\mathsf{T}$ and corresponding observations $\mathbf{y} = [y_1, \ldots, y_M]^\mathsf{T}$ according to Eq. (2). From the GP distribution the data then follows a joint multivariate Gaussian distribution, which can be stated as:

[88] removed: The hyperparameters $\boldsymbol{\psi} := [\sigma_f, \sigma_\nu, \lambda_1, \ldots, \lambda_{n_\mathbf{u}}]^T$ are commonly unknown and hence need to be inferred from data. In this article the log marginal likelihood $p(\mathbf{y}|\mathbf{U})$ is used. Ignoring constant terms and factors, this can be stated as:

[..[90]][..[91]]Due to the GP assumption the predictive distribution of $f(\mathbf{u})$ at an arbitrary input $\mathbf{u}$ [..[92]]given the training dataset $\mathcal{D} = \{\mathbf{U}, \mathbf{Y}\}$ has a closed-form solution[..[93]]

[..[94]]

[..[95]]

[..[96]]

[..[97]]. The resulting mean $\mu_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ can be seen as the GP prediction at $\mathbf{u}$ and the variance $\sigma^2_{\text{GP}}(\mathbf{u}; \mathcal{D}, \hat{\boldsymbol{\psi}})$ as a corresponding measure of uncertainty to this prediction. [..[98]]

The performance of the GP is dependent on hyperparameters $\hat{\boldsymbol{\psi}}$. 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 [..[99]]this a matrix of size $M \times M$ must be inverted, [..[100]]where M is the number of measurements. Clearly, this makes large data sets challenging.

4 [..[101]]

3.1 Modifier Adaptation with Gaussian processes

[..[102]]In the MA-GP approach the limitation of [..[103]][..[104]]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 [..[105]]MA-GP [..[106]]becomes

$$\hat{\mathbf{u}}^*_{k+1} = \arg[..^{107}]\max_{\mathbf{u}}[..^{108}]\mathrm{P}(\mathbf{u}[..^{109}][..^{110}])[..^{111}] + \mu_{\text{GP},k}[..^{112}](\mathbf{u}; \mathcal{D}[..^{113}]_{\mathrm{k}}, \hat{\boldsymbol{\psi}}[..^{114}]_{\mathrm{k}}), \qquad s.t. \quad \mathbf{u} \in \mathcal{U}, \tag{4}$$

[90] removed: The required maximum likelihood estimate is then given by $\hat{\boldsymbol{\psi}} \in \arg\max_{\boldsymbol{\psi}} \mathcal{L}(\mathcal{D}, \boldsymbol{\psi})$.

[91] removed: Next we require the

[92] removed: , which can be found by the conditional distribution of $f(\mathbf{u})$ on the data distribution $p(\mathbf{y}|\mathbf{U})$. From the GP assumption this

[93] removed: and can be stated as:

[97] removed: where

[98] removed: The GP is a non-parametric model. The

[99] removed: the above expression

[100] removed: which prohibits

[101] removed: Methodology

[102] removed: The use of GPs in a MA approach to overcome

[103] removed: estimating the plant gradients was first proposed by de Avila Ferreira et al. (2018). The idea is to replace the zeroth- and first-order modifiers of the cost and constraints in (**??**) with GP regression terms. Since the wind farms considered in this article do not have inequality constraint functions they are not included in this section. However, inequality constraint functions can be easily incorporated into the method.

[104] removed: The training set of the GP to correct the objective function are the controlled inputs of the approximate model and the plant-model mismatch of the objective function. The new

[105] removed: MA scheme with GP modifiers (

[106] removed:) is

where the plant-model mismatch of the cost function is modelled by [..[115]]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. [..[119]]The MA-GP approach for wind farm optimization is visualised in Fig. 1. The power output of the surrogate model

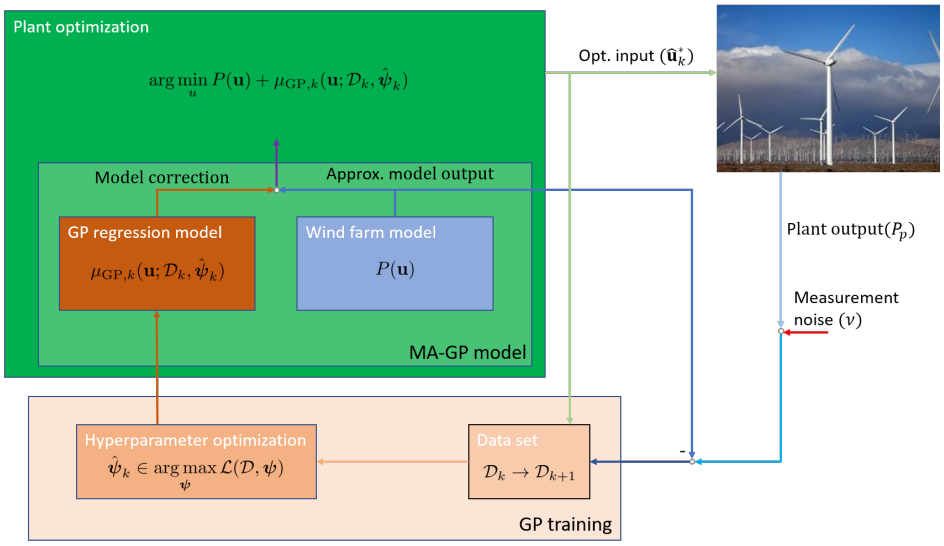


Figure 1. The basic idea of the MA-GP scheme for a wind farm. The GP regression model creates an input-output map of the control inputs to the plant-model mismatch. In the MA-GP model the GP regression model is used to correct the output of the approximate model. This MA-GP model is used in the optimization 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 [116]

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

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

[115] removed: $\mu_{GP}^{\phi_p - \phi}$. Similar to the original MA scheme the optimal input of Eq. (4) may be filtered with Eq. (5) to reduce the step-size and help stabilize the MA-GP scheme (**?**). The whole MA-GP scheme is presented in Algorithm 1 and Fig. 1.

[116] [..[117]][..[118]]

[119] removed: In Algorithm 1 the hyperparameters are updated if

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[..[120]], which is a user-defined condition[..[121]]. The hyperparameter update is usually the computational bottle-neck of the MA-GP algorithm. [..[122]]We observed that especially for large data sets [..[123]]the hyperparameters do not change much from one iteration to the next. Therefore, [..[124]]the hyperparameters can be updated less frequent to decrease computational delay. However, it is recommend to update the hyperparameters [..[125]]as often as possible.

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

Algorithm 1: Basic MA-GP scheme [..[126]](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

3.2 Numerical turbine and wake models

[120] removed: . *HypOpt*

[121] removed: , which allows to update the hyperparameter. The extrema are to update the hyerparameter each iteration or never. The

[122] removed: Especially

[123] removed: it can be expected that the hyperparameter

[124] removed: it is reasonable

[125] removed: less frequent.

[125] 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

A turbine and wake model are necessary to create a model of a wind farm. The wind turbines [..[127]]are represented using the actuator 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 et al., 2016)

$$P = \frac{1}{2}\rho A C_P \cos\gamma^p [..^{128}] v^3, \tag{8}$$

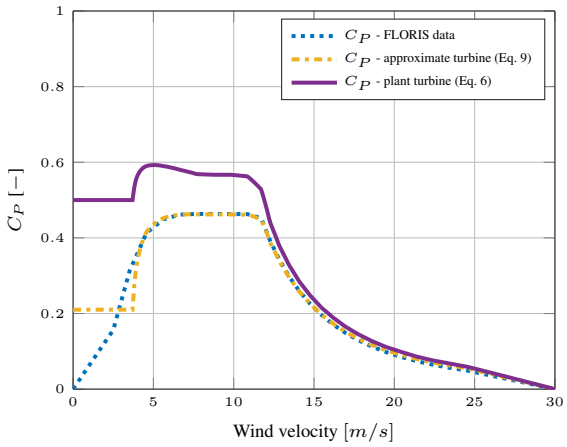
where A is the rotor area, ρ the air density[..[129]], p a correction factor and v is the wind velocity. In actuator disc theory $p = 3$ (Burton et al., 2011). However, based on large-eddy simulations, the turbine power yaw misalignment has been shown to match the output when $p = 1.88$ for the NREL 5MW turbine (Annoni et al., 2018), which we will use in this article. In the numerical study it will be important to implement a "plant" and model, which are different from each other. [..[130]]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 [..[131]]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 resulting in

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

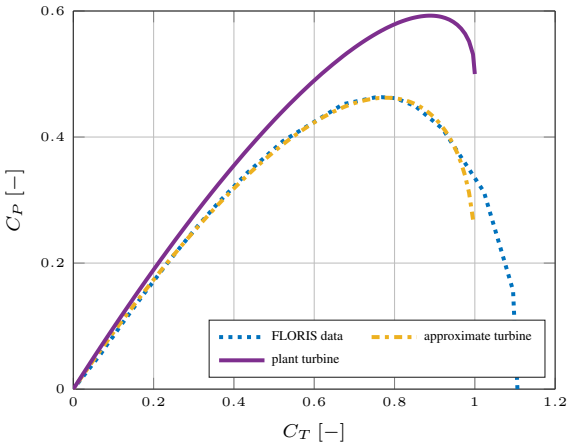
The approximate turbine model fit is visualised in Fig. 2. Important in the numerical example is the different connection between thrust and power coefficients of [..[132]]plant and approximate turbine model (Fig. 2b). For the turbine dimensions the NREL 5-MW wind turbine is used (Jonkman et al., 2009). Consequently, the rotor diameter is [..[133]]$D = 126.4\,\mathrm{m}$ and the hub height $H_H = 90\,\mathrm{m}$.

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

[127] removed: in the wind farm
[129] removed: and
[130] removed: Therefore, a
[131] removed: a new model based on the actuator disk
[132] removed: both models
[133] removed: $D = 136\,\mathrm{m}$



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



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

[a]removed: The dashed line gives the corresponding thrust coefficient, which is the same for both models.

Figure 2. Comparison between data, the [..[134]]plant turbine and the [..[135]]approximate turbine model. The [..[136]]models [..[137]]give [..[138]]different [..[139]]connections between thrust and power coefficients.

three-dimensional steady-state far wake velocity [..[140]]deficit is Gaussian distributed and can be estimated [..[141]]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. [..[142]]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.

[140]removed: is assumed to be

[141]removed: with

[142]removed: In the simulation study different values are chosen for this parameter in the plant and approximated model resulting in different optimal operating points

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 [..[143]]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 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

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

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.

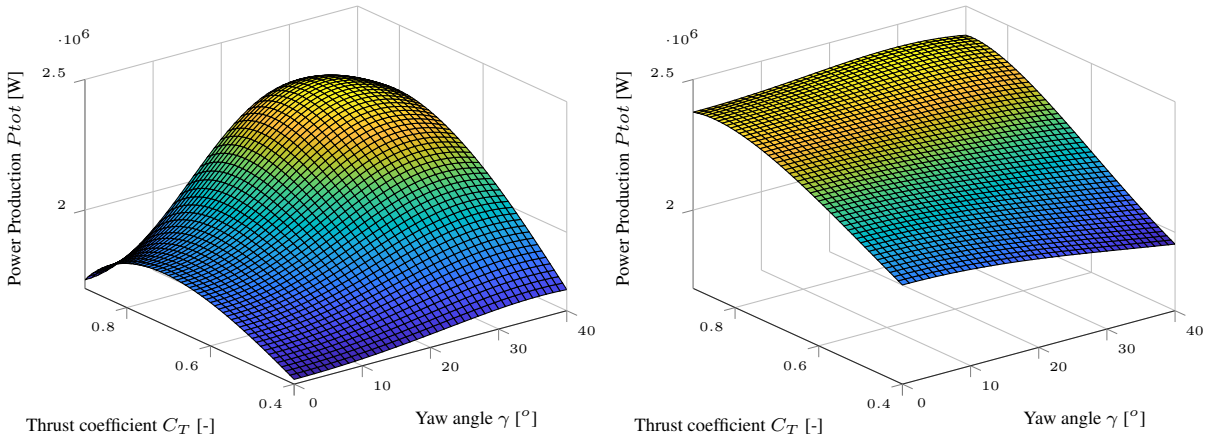
4.1 Two turbine case

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 point. The turbine row is facing the wind and the spacing between turbines is $5\,\mathrm{D}$. The power production of the wind

[143] removed: amount

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



(a) Surface of objective function of plant.

(b) Surface of objective function of approximate model.

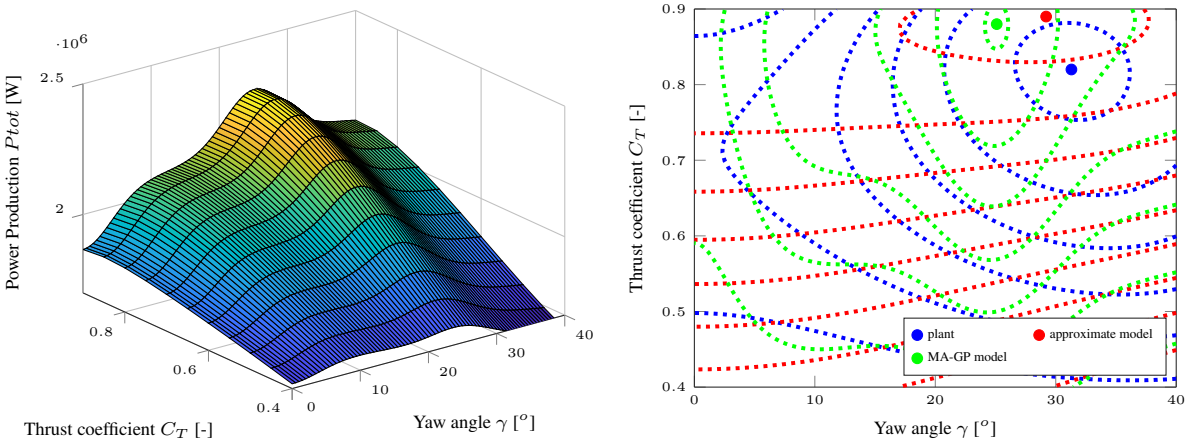
Figure 3. The power production of plant and approximate model in dependency of the control inputs of the upwind turbine.

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 [..[144]]$\Theta = 1.67\%$. Still, the model [..[145]]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

[144] removed: $1.67\,\%$

[145] removed: assume



(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.

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 operating point [..[146]](Fig 5). After four iterations the [..[147]]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 [..[148]](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 [..[149]]was assumed that the measurements are noise-free. If noise is added to the power measurements the correct identification becomes more challenging and a larger training data set is necessary. A noise with a standard deviation of $50\,\mathrm{kW}$ 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 [..[150]]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 [..[151]]relative error

[146] removed: . Fig 5shows the operating points of the first ten iterations
[147] removed: error in power production
[148] removed: .
[149] removed: is
[150] removed: approximate model and plant
[151] removed: error in the power production

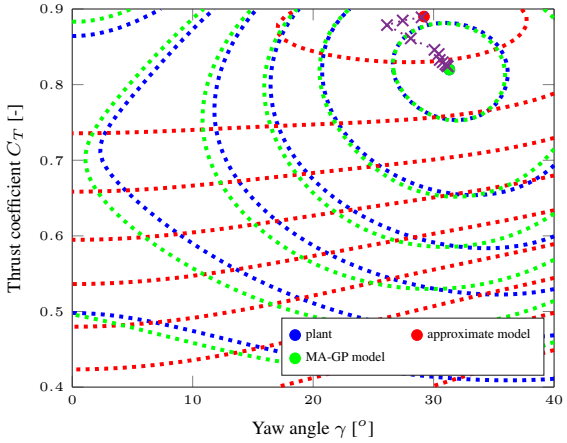


Figure 5. The contour plot of plant, approximate model and MA-GP model after ten iterations. The operating points of each iteration are marked with a cross.

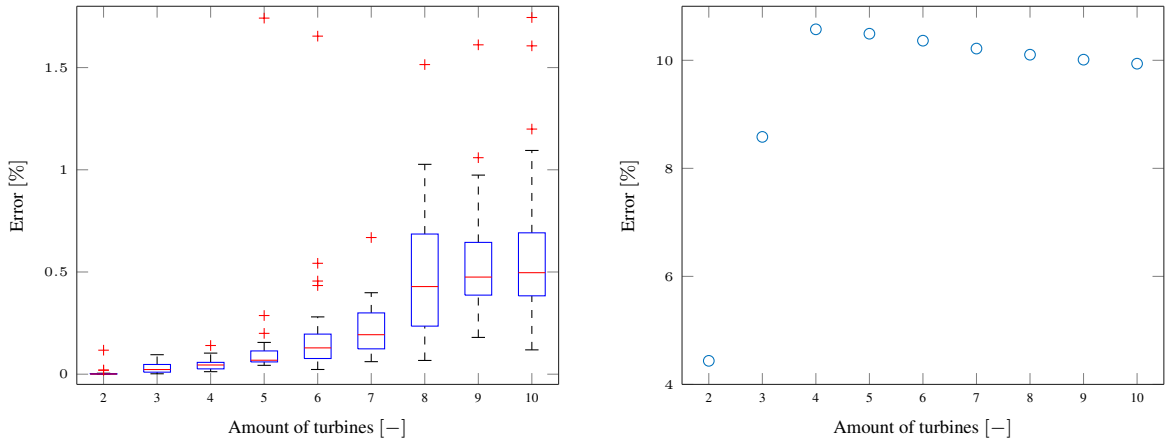
Θ is about $0.6\,\%$. The algorithm is able to converge. However, due to the measurement noise a small error remains after ten iterations. The error can be easily decreased with a larger initial data set, e.g. with a training set of 30 points the error after ten iterations is about $0.35\,\%$.

4.2 n turbine row case

In this subsection the optimization of n turbines aligned in a row [..[152]]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 [..[153]]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. The operating points of the training set are [..[154]]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 [..[155]]error increases with the amount of turbines while it is almost zero for 2 to 4 turbines [..[156]](Fig. [..[157]]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 [..[158]]plant. It makes it challenging to correctly identity the input-output map[..[159]]. The error can be decreased

[152] removed: are optimized with the
[153] removed: algorithm. The
[154] removed: randomly chosen
[155] removed: statistic of the error after 25 iterations is shown in Fig. 6a. The error
[156] removed: . Even though, the error increases with the amount of turbines the algorithm is able to reduce the model error significantly
[157] removed: 6b). It is not surprising that the error increases with the amount of control inputs. The control inputs are mapped to the total
[158] removed: wind farm. With a large amount of control inputs the correct identification of this
[159] removed: becomes more challenging, which increases the error in the MA-GP algorithm. Again, the error could



(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.

with more data in the training set. 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. [..[160]] [..[161]]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. extremum seeking (Johnson and Fritsch, 2012) or MPPT (Gebraad et al., 2013), [..[162]]which usually need several iterations to find an optimum. Moreover, after the initial training the MA-GP [..[163]]model [..[164]]usually represents of the plant [..[165]]better than the approximate model. Nonetheless, measurements close to the optimum of the MA-GP model can help to [..[166]]improve the model further.

[160]removed: However, the computational time of the GP regression grows cubic with the amount of data. Therefore, at some point a trade-off between performance and computational time is necessary.

[161]removed: In

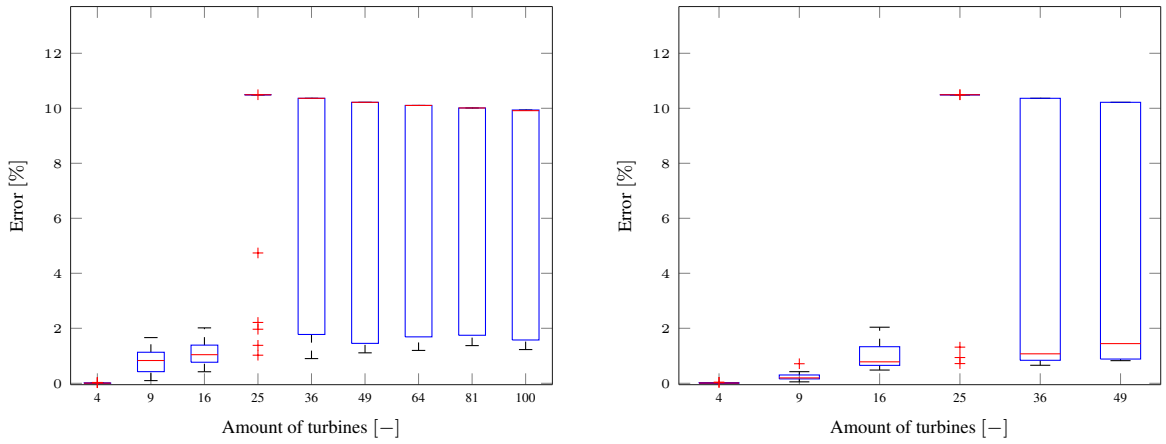
[162]removed: is the

[163]removed: the algorithm able to find a near optimal point in one iteration. The MA-GP

[164]removed: is already a better representation

[165]removed: after the initial training

[166]removed: refine the MA-GP model



(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.

4.3 $n \times n$ turbine grid case

In this subsection the [..[167]]optimization of a wind farm with turbines arranged in a $n \times n$ grid [..[168]]with a spacing of $5\,\mathrm{D}$ is presented. Consequently, the wind farm consist of n turbine rows each containing n turbines. The [..[169]]wind direction is aligned with the rows 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 $n_d = 10n_u$ [..[170]]and the MA-GP approach is tested on 25 Monte Carlo simulations.

[..[171]]Again the algorithm converges for a small amount of turbines [..[172]](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 initial conditions, which is defined by the optimal operation point of the model (Fig. 6b)[173]. 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.

[167] removed: turbines in the wind farm a

[168] removed: . The wind direction is aligned with the rows of the grid. Interaction between parallel rows is neglectable

[169] removed: distance between turbines is $5\,\mathrm{D}$. The identification of the power production of this wind farm layout becomes more challenging. The input space increased and the sensitivity of inputs onto the total power production of the wind farm become similar.

[170] removed: . Otherwise the setup is the same as in the turbine row case

[171] removed: The error after 25 iterations is shown in Fig. 7a.

[172] removed: .

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

[..[174]]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 $15\,\mathrm{s}$. In some rare cases it took about $5\,\mathrm{min}$. 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 $10\,\mathrm{s}$. Consequently, the [..[175]]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 [..[176]](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, [..[177]]the training of the hyperparameters in the GP regression scales cubic with the amount of data. Obviously this ultimately limits the size of the [..[178]]training set since the approach can become computational infeasible. [..[179]]

[..[180]]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

[174] removed: The problems to identify the plant modelcorrectly with a larger inputs space are not surprising. The sample density decreases drastically for larger inputs spaces

[175] removed: initial data set could be increased

[176] removed: for both small and large inputs spaces

[177] removed: it also has to be pointed out that

[178] removed: initial training set . Otherwise the approach becomes quickly

[179] removed: In case of an initial set of $n_d = 10n_u$ and a wind farm with 49 turbines the median time for the hyperparameter optimization is about $3\,\mathrm{s}$. The maximum computational time in the 625 hyperparameter optimization is about $60\,\mathrm{s}$. In case of an initial set of $n_d = 30n_u$ the median optimization time is about $50\,\mathrm{s}$ while the maximum optimization time is about $23\,\mathrm{min}$. In these cases the optimization algorithm did not converge to an optimum and the maximum amount of iterations until termination was performed. The optimization time could be reduces by limiting the number of iterations. It is expected that it will not influence the performance since the objective function value in cases the optimization did not converge to an optimum did not change for most of the iterations.

[180] removed: If the MA-GP algorithm for the larger wind farms converges to an optimum it usually takes first a few iterations, where the wind farm is operated at the model optimum point, 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 actual 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.

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. (2020d)[181] 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.

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).

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 $0°$, $45°$, $90°$, $135°$, $180°$, $225°$, $270°$, 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.

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 effects.

[181] The article is submitted to TORQUE 2020

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 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.
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

The modifier-adaptation approach with Gaussian processes applied to wind farm control is presented. It is a real-time optimization strategy, which corrects [..[182]]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 [..[183]]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. 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.

[..[184]]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 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.

[182] removed: optimization model errors

[183] removed: exceptionally

[184] removed: In future work the performance of the method for large inputs spaces has to be improved . Several ideas are possible to achieve it

– Extend the algorithm with an exploration part. This can be achieved, for example, by including the variance of the GP regression model in the optimization.

– [..[185]]

[..[186]]An important investigation is the sensitivity of the approach to measurements[..[187]], 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 [..[188]]Andersson et al. (2020d).

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)

[185] removed: Include the single turbine power measurements in the identification of the GP regression model. In such a multiple-input and multiple-output approach the sensitivities of control inputs to the single outputs increase. The model identification should benefit from the approach. Moreover, it is expected that a smaller data set is necessary to achieve the same performance as with the in the article presented multiple-inputs and single-output approach. The idea is pursued in Andersson et al. (2020a) with very promising results in increasing the accuracy of the approach with a smaller initial data set.

[186] removed: In addition, the

[187] removed: and inputs noise has to be investigated

[188] removed: Andersson et al. (2020c)

References

Adaramola, M. and Krogstad, P.-Å.: Experimental investigation of wake effects on wind turbine performance, Renewable Energy, 36, 2078–2086, 2011.

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., and Imsland, L.: Adaptation of Engineering Wake Models using Gaussian Process Regression and High-Fidelity Simulation Data, in: TORQUE 2020 - [submitted], 2020c.

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, 2020d.

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.

Corten, G. and Schaak, P.: Heat and flux: Increase of wind farm production by reduction of the axial induction, in: Proceedings of the European Wind Energy Conference, 2003.

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.

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), 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.

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.

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, 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. and Van Wingerden, J.: A control-oriented dynamic model for wakes in wind plants, in: Journal of Physics: Conference Series, vol. 524, p. 012186, IOP Publishing, 2014.

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.

Horvat, T., Spudić, V., and Baotić, M.: Quasi-stationary optimal control for wind farm with closely spaced turbines, in: MIPRO, 2012 Proceedings of the 35th International Convention, pp. 829–834, IEEE, 2012.

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.

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 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.

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 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.

Medici, D.: Experimental studies of wind turbine wakes: power optimisation and meandering, Ph.D. thesis, KTH, 2005.

Munters, W. and Meyers, J.: Effect of wind turbine response time on optimal dynamic induction control of wind farms, in: Journal of Physics: Conference Series, vol. 753, p. 052007, IOP Publishing, 2016.

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., and Law, K. H.: Wind farm power maximization based on a cooperative static game approach, in: Active and Passive Smart Structures and Integrated Systems 2013, vol. 8688, p. 86880R, International Society for Optics and Photonics, 2013.

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

Rotea, M. A.: Dynamic programming framework for wind power maximization, IFAC Proceedings Volumes, 47, 3639–3644, 2014.

Schepers, J. and Van der Pijl, S.: Improved modelling of wake aerodynamics and assessment of new farm control strategies, in: Journal of Physics: Conference Series, vol. 75, p. 012039, IOP Publishing, 2007.

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

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, 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.

Wagenaar, J., Machielse, L., and Schepers, J.: Controlling wind in ECN's scaled wind farm, Proc. Europe Premier Wind Energy Event, pp. 685–694, 2012.