https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

Open Access
WIND ENERGY SCIENCE
DISCUSSIONS

Multifidelity Multiobjective Optimization for Wake Steering Strategies

Julian Quick[1,2], Ryan N. King[2], Garrett Barter[2], and Peter E. Hamlington[1]

[1]University of Colorado, Boulder, CO, USA
[2]National Renewable Energy Laboratory, Golden, CO, USA

Correspondence: Julian Quick (julian.quick@colorado.edu)

Abstract. Wake steering is an emerging wind power plant control strategy where upstream turbines are intentionally yawed out of perpendicular alignment with the incoming wind, thereby "steering" wakes away from downstream turbines. However, trade-offs between the gains in power production and fatigue loads induced by this control strategy are the subject of continuing investigation. In this study, we present a multifidelity multiobjective optimization approach for exploring the Pareto front of trade-offs between power and loading during wake steering. An unsteady large-eddy simulation is used as the high-fidelity model, where an actuator line representation is used to model wind turbine blades, and a rainflow-counting algorithm is used to compute damage equivalent loads. A coarser simulation with a simpler loads model is employed as a supplementary low-fidelity model. A multifidelity Bayesian optimization is performed to iteratively learn both a surrogate of the low-fidelity model and an additive discrepancy function, which maps the low-fidelity model to the high-fidelity model. Each optimization uses the expected hypervolume improvement acquisition function, weighted by the total cost of a proposed model evaluation in the multifidelity case. The multifidelity approach is able to capture the logit function shape of the Pareto frontier at a computational cost that is only 30% of the single fidelity approach. Additionally, we provide physical insights into the vortical structures in the wake that contribute to the Pareto front shape.

1 Introduction

As wind energy systems have matured, plant-level control has emerged as a new paradigm, where groups of turbines are controlled in coordination to maximize collective power production. This is in contrast to more traditional control strategies, where individual turbines are controlled to maximize their own power production. A potentially promising form of such plant-level control is "wake steering," where upstream wind turbine yaw positions are intentionally misaligned from the incoming wind, "steering" the wake away from downstream turbines. A counter-rotating pair of vortices is generated by the rotating blades (Fleming et al., 2018; Martínez-Tossas and Branlard, 2020), and the direction of rotation of these vortices is determined by the direction of thrust of the wind turbine rotor, which is determined by the yaw offset direction. This allows the performance of wind power plants to be improved by diverting wakes away from downstream turbines

It is speculated that wake steering may produce more power while inducing less total fatigue on all turbines when compared to the baseline strategy of aligning each turbine with the incoming wind (Howland et al., 2019; Hulsman et al., 2020). However,

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

very few studies have quantified the trade-offs between power and damage. Hulsman et al. (2020) used an actuator line model to train polynomial chaos surrogates for optimization of a weighted sum of power and damage equivalent loads. Yin et al. (2020) present a multiobjective genetic algorithm for the maximization of power and minimization of total thrust. Damiani et al. (2018) performed a detailed analysis of a single wind turbine, noting that negative yaw offsets tended to increase fatigue loading more than positive yaw offsets. Wang et al. (2020) demonstrated the potential of individual pitch control to alleviate loads induced from intentionally offsetting the turbine yaw. Van Dijk et al. (2017) utilized the FLORIS and CCBlade tools to examine trade-offs between power produced and the edgewise and flapwise fatigue loading induced through wake steering. Lin and Porté-Agel (2020) utilize a large-eddy simulation (LES) framework to construct the Pareto set between power and flapwise bending moment loading through a comprehensive parameter sweep. While these studies all provide insights into the trade-offs between power and loading, there remains a need for an efficient optimization algorithm to quantify these trade-offs using computationally intensive simulations.

Despite its promise, plant-level control via wake steering involves complex physics and is challenging to model. While engineering wake models are remarkably accurate in power prediction, they have dubious accuracy when predicting fatigue loading, which high-fidelity models predict more accurately (Rinker et al., 2021). In this study, we propose a multifidelity multiobjective optimization framework to address this challenge and explore trade-offs between power and loading in wake steering strategies. In practice, power and loading will likely be optimized in real time using a singular weighted objective. The relative weights may be decided upon by exploring trade-offs between power and loading using multiobjective optimization to estimate the Pareto frontier. When searching for the Pareto set, an efficient algorithm must balance exploration and exploitation. Several models have been developed that may be used to study the effects of control strategies with various levels of mathematical detail and real-world accuracy (i.e., fidelity) (Annoni et al., 2018; Martínez-Tossas et al., 2019; Hulsman et al., 2020).

Multifidelity optimization exploits the correlation between low- and high-fidelity models to reduce the overall computational cost of optimization. For instance, Andersson and Imsland (2020) present a real-time modifier adaptation approach for wake steering design, where a Gaussian Process (GP) is used to iteratively learn the difference between observed operational data and the predictions of an engineering wake model. Ariyarit and Kanazaki (2017) present a two-objective bifidelity approach that iteratively builds a GP discrepancy function. Huang et al. (2006) and Rajnarayan et al. (2008) employ an augmented expected improvement formulation, including three factors to account for the correlation between the low- and high-fidelity models, the observed error, and the cost ration between the low- and high-fidelity models. It is not always clear when a proposed low-fidelity model is appropriate for use in multifidelity optimization, though it is common to assess candidate low-fidelity models by measuring their correlation with the high-fidelity model (Giselle Fernández-Godino et al., 2019).

The novelty of the present study is the application of this multifidelity technique to wind energy systems, resulting in new insights into wake steering flow physics. The present approach uses the low-fidelity model to first explore the full parameter space, then iteratively builds the low- and high-fidelity model surrogates to gain the most improvement in the Pareto front per model evaluation costs. While this framework is similar to those presented by Ariyarit and Kanazaki (2017) and Andersson

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

and Imsland (2020), the exact framework outlined here is new—and this is the first demonstration of any such approach in the context of wind energy systems.

2 Optimization Framework

A Bayesian framework for multifidelity multiobjective optimization is presented. Throughout this section, we assume that minimization of functions is the objective of the optimization procedure (as opposed to maximization).

This study employs GP models to approximate power and loading dynamics. A GP is a set of random variables, where all subsets form joint Gaussian distributions (Rasmussen and Williams, 2006). The true objective functions for power and loads, f_i, are approximated using individual Gaussian process surrogate models, $\boldsymbol{g}$, which are defined as:

$$g_i(\boldsymbol{\gamma}) \sim \mathcal{GP}\left[\mu_i(\boldsymbol{\gamma}), k_i(\boldsymbol{\gamma}, \boldsymbol{\gamma}')\right], \tag{1}$$

where $\boldsymbol{\gamma}$ is a vector of proposed yaw angles, $\boldsymbol{\gamma}'$ is an arbitrary vector of yaw angles, $\mu_i(\boldsymbol{\gamma})$ is the GP mean function, $k_i(\boldsymbol{\gamma}, \boldsymbol{\gamma}')$ is the GP kernel covariance function, and the index i refers to power ($i = 1$) or loading ($i = 2$) objectives.

We perform Bayesian inference on functions by conditioning the GP on a set of observed input-output pairs, $\mathcal{D}_i = \{\boldsymbol{\Gamma}, \boldsymbol{Y}_i\}$, where $\boldsymbol{\Gamma} = [\boldsymbol{\gamma}^{(1)}, \boldsymbol{\gamma}^{(2)}, \ldots]$ is a matrix of observed yaw offsets and $\boldsymbol{Y}_i = [f_i(\boldsymbol{\gamma}^{(1)}), f_i(\boldsymbol{\gamma}^{(2)}), \ldots]$ is a vector of observed power or loading outputs. The objective functions are normalized to have zero mean and unity variance during a preprocessing step. After conditioning on $\mathcal{D}_i$, we obtain Gaussian distributions at test locations $\boldsymbol{\gamma}^*$ with the following posterior estimates of the mean

$$\mu_i(\boldsymbol{\gamma}^* \mid \mathcal{D}_i) = \boldsymbol{k}_{i,*}^T \boldsymbol{K}_i^{-1} \boldsymbol{Y}_i \tag{2}$$

and variance

$$\sigma_i^2(\boldsymbol{\gamma}^* \mid \mathcal{D}_i) = k_i(\boldsymbol{\gamma}^*, \boldsymbol{\gamma}^*) - \boldsymbol{k}_{i,*}^T \boldsymbol{K}_i^{-1} \boldsymbol{k}_{i,*}, \tag{3}$$

where $\boldsymbol{k}_{i,*} = k_i(\boldsymbol{\Gamma}, \boldsymbol{\gamma}^*)$ is a vector and $\boldsymbol{K}_i = k_i(\boldsymbol{\Gamma}, \boldsymbol{\Gamma})$ is a matrix.

The kernel covariance function encodes prior knowledge about structural properties of the underlying signal, like smoothness, periodicity, and stationarity. In this study, we employ an anisotropic radial basis function kernel for k_i, given by

$$k_i(\boldsymbol{\gamma}, \boldsymbol{\gamma}') = \exp\left[\frac{1}{2}\left(\sum_{j=1}^{\dim(\boldsymbol{\gamma})} \frac{(\gamma_j - \gamma_j')^2}{l_{ij}^2}\right)^{1/2}\right], \tag{4}$$

where the correlation scale, l_{ij}, is estimated by maximizing the log-marginal likelihood function (Pedregosa et al., 2011).

2.1 Single-Fidelity Approach

In Bayesian optimization, an acquisition function is defined to maximize a metric representing both exploration and exploitation (Shahriari et al., 2015). A popular acquisition function for single-objective optimization is the expected improvement,

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

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

which is the expected value of an improvement function (Zhan and Xing, 2020) with respect to the predicted uncertainty of a GP. This acquisition function is employed by the Efficient Global Optimization (EGO) algorithm (Jones et al., 1998). The improvement function, I, quantifies the improvement in the objective function for a new evaluation, as compared to the best sampled objective, and is zero if the new objective does not outperform all of the previously sampled points. This results in

$$I(\boldsymbol{\gamma}) = \max\left[f^* - f(\boldsymbol{\gamma}), 0\right], \tag{5}$$

where f^* is the minimum sampled value and $f(\boldsymbol{\gamma})$ is the function value, which is generally unknown and must be predicted by a GP.

There is a range of potential outcomes from sampling a new point, and the GP framework conveniently estimates this uncertainty. These uncertainties are used to compute the expected value of the improvement function. It is important that the improvement function contains the maximum function; otherwise, there would be no exploration of regions of larger uncertainty. Other acquisition functions available include the knowledge gradient (Ghoreishi and Allaire, 2018), expected quantile improvement (He et al., 2017; Picheny et al., 2013), improved expected improvement (Qin et al., 2017), entropy search (Hennig and Schuler, 2012), and minimization of the predictor (Andersson and Imsland, 2020).

The expected improvement may be extended to a multiobjective context. This is done by introducing a hypervolume function, H, which measures the volume of a given Pareto front, A, using a reference point, $\boldsymbol{r}$. The expected hypervolume improvement (EHVI), introduced in Emmerich et al. (2006), is the multiobjective counterpart to the expected improvement acquisition function used in the EGO algorithm and is given by

$$EHVI(\boldsymbol{\gamma}, \boldsymbol{g}) = \mathbb{E}_{\boldsymbol{g}(\boldsymbol{\gamma})}\{HVI[\boldsymbol{f}(\boldsymbol{\gamma})]\}. \tag{6}$$

Here, $\mathbb{E}_{\boldsymbol{g}(\boldsymbol{\gamma})}[\cdot]$ represents the expectation with respect to a normal distribution, $\boldsymbol{g}(\boldsymbol{\gamma})$, and is expressed as

$$\mathbb{E}_{\boldsymbol{g}(\boldsymbol{\gamma})}\{HVI[\boldsymbol{f}(\boldsymbol{\gamma})]\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} HVI([P,L]) \frac{1}{\sigma_1(\boldsymbol{\gamma})\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{P-\mu_1(\boldsymbol{\gamma})}{\sigma_1(\boldsymbol{\gamma})}\right]^2} \frac{1}{\sigma_2(\boldsymbol{\gamma})\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{L-\mu_2(\boldsymbol{\gamma})}{\sigma_2(\boldsymbol{\gamma})}\right]^2} dP dL, \tag{7}$$

where μ_1 and σ_1 are the mean and standard deviation, respectively, of the powers modeled by g_1, and μ_2 and σ_2 are the mean and standard deviation, respectively, of loads modeled by g_2. The hypervolume improvement (HVI) indicator function is the multiobjective counterpart to the improvement indicator function. It is given as

$$HVI([P,L]) = H(A \cup \{[P,L]\}) - H(A), \tag{8}$$

where A is an estimated Pareto frontier, $A \cup \{[P,L]\}$ is a new Pareto set that potentially includes $[P,L]$, H is a hypervolume function, which measures the volume spanned by the Pareto set of objective functions relative to a reference point, $\boldsymbol{r}$, which must not be dominated by $A \cup \{[P,L]\}$. The Pareto frontier is defined as the set of all function values that are not strictly

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

dominated by other function values. The formal definition is

$$A = \left\{ \boldsymbol{y}' \in \{\boldsymbol{y} \in \mathbb{R}^{\dim(\boldsymbol{f})} : \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{\gamma}), \boldsymbol{\gamma} \in \Omega_{\boldsymbol{\gamma}}\} : \{\boldsymbol{y}'' \prec \boldsymbol{y}', \boldsymbol{y}'' \neq \boldsymbol{y}'\} = \emptyset \right\}, \tag{9}$$

where $\Omega_{\boldsymbol{\gamma}}$ is the set of allowable yaw offsets, $\prec$ denotes Pareto dominance, and $\emptyset$ is the empty set.

The hypervolume, H, measures the extent of the Pareto set as the volume of the Pareto-dominated space bounded by a reference point, $\boldsymbol{r}$, namely,

$$H(A) = \mathrm{Vol}\left(\{\boldsymbol{y} \in \mathbb{R}^{\dim(\boldsymbol{f})} | \boldsymbol{y}' \in A \prec \boldsymbol{y} \text{ and } \boldsymbol{y} \prec \boldsymbol{r}\}\right). \tag{10}$$

In practice, the Pareto set is computed by filtering a set of discrete inputs so that only nondominated points remain,

$$A = \left\{ \boldsymbol{y}' \in \boldsymbol{Y} : \{\boldsymbol{y}'' \prec \boldsymbol{y}', \boldsymbol{y}'' \neq \boldsymbol{y}'\} = \emptyset \right\}, \tag{11}$$

where $\boldsymbol{Y}$ is a matrix of observed function values. This filters out observed samples that are Pareto dominated by other observed samples.

Although these ideas may also be extended to more objectives, assuming two objectives simplifies the problem. Assuming only two objectives, the observed Pareto set is defined as $A \approx (\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots, \boldsymbol{y}^n)$ such that $y_1^1 < y_1^2 < \cdots < y_1^n$. The hypervolume is estimated using rectangular quadrature as

$$H(A) \approx \sum_{i=1}^{n-1} (y_1^{i+1} - y_1^i)(r_2 - y_2^i) + (r_1 - y_1^n)(r_2 - y_2^n), \tag{12}$$

where n is the number of points in the given Pareto set.

In this study, the EHVI is approximated through Monte Carlo simulation by

$$EHVI(\boldsymbol{\gamma}, \boldsymbol{g}) \approx \frac{1}{N_s} \sum_{k=1}^{N_s} \left[H\left(A \cup \{\mathcal{N}[\boldsymbol{\mu}(\boldsymbol{\gamma}), \boldsymbol{\sigma}(\boldsymbol{\gamma})]\}^{(k)}\right) - H(A) \right], \tag{13}$$

where N_s is the number of Monte Carlo samples and $\{\mathcal{N}[\boldsymbol{\mu}(\boldsymbol{\gamma}), \boldsymbol{\sigma}(\boldsymbol{\gamma})]\}^{(k)}$ is draw k from the GP model of power and loading, $\boldsymbol{g}$.

Once the EHVI is estimated, it must be maximized. This is not necessarily trivial, as the EHVI computation is complicated and difficult to vectorize. The EHVI optimum may be determined using a grid search, random sampling, direct optimization, or surrogate-based optimization.

2.2 Multifidelity Approach

The multifidelity approach introduces cheaper, less-accurate representations of the high-fidelity model, which allow for greater control between exploration and exploitation in the Bayesian optimization. Samples of the low-fidelity model are adaptively re-

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

fined throughout the optimization as a cheap means for exploration of the high-fidelity function space. Throughout this section, we assume a known hierarchy of model fidelities, $(f_k^1, f_k^2, \ldots, f_k^N)$, where f_k^1 is the lowest-fidelity model of power/loading, N is the number of different fidelity models, and f_k^N is the highest-fidelity model of power/loading. When a model is evaluated at a point, $\boldsymbol{\gamma}$, we assume that all lower-fidelity models will also be evaluated at this point.

The lowest-fidelity model, f_k^1, is approximated using a GP, g_k^1, resulting in the following output distribution:

$$g_k^1(\boldsymbol{\gamma}) \sim \mathcal{N}\left[\mu_k^1(\boldsymbol{\gamma}), \sigma_k^1(\boldsymbol{\gamma})\right], \tag{14}$$

where μ_k^1 and σ_k^1 are the mean and standard deviations, respectively, associated with the lowest-fidelity power/loading model.

Higher-fidelity models, $f_k^i(\boldsymbol{\gamma})$, are approximated using additive discrepancy functions that map the next lowest fidelity function, $f_k^{i-1}(\boldsymbol{\gamma})$, to $f_k^i(\boldsymbol{\gamma})$:

$$f_k^i(\boldsymbol{\gamma}) \approx f_k^{i-1}(\boldsymbol{\gamma}) + \delta_k^i(\boldsymbol{\gamma}) \; \forall i > 1, \tag{15}$$

where $\delta_k^i(\boldsymbol{\gamma})$ is the discrepancy function associated with objective k and fidelity i,

$$\delta_k^i(\boldsymbol{\gamma}) \sim \mathcal{N}\left[\mu_k^i(\boldsymbol{\gamma}), \sigma_k^i(\boldsymbol{\gamma})\right] \; \forall i > 1, \tag{16}$$

where μ_k^i and σ_k^i are the mean and standard deviations, respectively, modeled by the discrepancy function GP associated with fidelity i.

New GPs are defined to extend the EHVI to a multifidelity context. No matter which fidelity is to be sampled next, the ultimate goal is to minimize the highest-fidelity function, so each GP is constructed to predict the high-fidelity output. However, GPs associated with lower-fidelity models should not take into account uncertainties associated with higher-fidelity models, as these uncertainties will not be collapsed if the lower-fidelity model is sampled. Sampling the highest-fidelity model must take all sources of surrogate uncertainty into account, as a high-fidelity model evaluation will be associated with sampling all lower-fidelity models at the same point. So, new GP models, $\boldsymbol{h}^i$, are constructed to predict the high-fidelity output while encoding different uncertainty information. The GPs associated with each fidelity are defined as

$$h_k^i(\boldsymbol{\gamma}) \sim \mathcal{N}\left\{\sum_{j=1}^{N} \mu_k^j(\boldsymbol{\gamma}), \sqrt{\sum_{j=1}^{i}\left[\sigma_k^j(\boldsymbol{\gamma})\right]^2}\right\}. \tag{17}$$

Putting the above formulations together, it is natural to define a multifidelity multiobjective acquisition function as the ratio of EHVI per evaluation cost:

$$J(\boldsymbol{\gamma}, l) = \frac{EHVI(\boldsymbol{\gamma}, \boldsymbol{h}^l)}{\sum_{j=1}^{l} C_j}, \tag{18}$$

where J is the optimization objective, to be maximized with respect to yaw offsets, $\boldsymbol{\gamma}$, and model fidelity, l, in each optimization iteration, and C_l is the computational cost associated with model l.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

In this study, we examine the bifidelity case ($N = 2$), where

$$h_k^1(\boldsymbol{\gamma}) \sim \mathcal{N}\left[\mu_k^1(\boldsymbol{\gamma}) + \mu_k^2(\boldsymbol{\gamma}), \sigma_k^1(\boldsymbol{\gamma})\right] \tag{19}$$

and

$$h_k^2(\boldsymbol{\gamma}) \sim \mathcal{N}\left\{\mu_k^1(\boldsymbol{\gamma}) + \mu_k^2(\boldsymbol{\gamma}), \sqrt{[\sigma_k^1(\boldsymbol{\gamma})]^2 + [\sigma_k^2(\boldsymbol{\gamma})]^2}\right\}. \tag{20}$$

The bifidelity workflow is visualized in Figure 1.

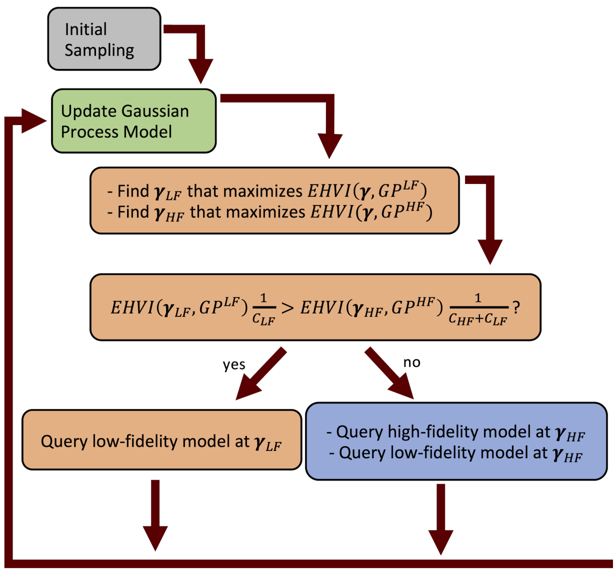

Figure 1. Workflow visualization for the bifidelity optimization case. $\boldsymbol{h}^1$ is represented by GP^{LF}, $\boldsymbol{h}^2$ is represented by GP^{HF}, C_1 is represented by C_{LF}, and C_2 is represented by C_{HF}.

3 Numerical Approach

3.1 Flow Modeling

This study used the WindSE framework (National Renewable Energy Laboratory, 2021) to model flow within the wind power plant. We investigate a two-turbine case, with a single wind direction and speed, where the turbines are spaced 7 rotor diameters apart and the wind direction is such that the front turbine directly wakes the back turbine. The inflow boundary is modeled using a logarithmic profile with a hub-height wind speed of 7.5 m/s. The top, side, and outflow boundaries are specified as no-stress boundaries, and the ground is specified as a no-slip boundary. We consider turbine representations of the IEA 3.4 MW reference turbine (Bortolotti et al., 2019), with hub heights of 120 m and rotor diameters of 130 m. The turbine blades are represented as actuator lines with 15 force nodes. The domain is represented with a $2{,}260 \times 2{,}000 \times 520$ m^3 mesh, corresponding to a 301-s flow-through time. The mesh is refined near the center of the domain and where the turbines are located. A target Courant–Friedrichs–Lewy condition of 0.98 is specified.

The simulations solve the filtered conservation of mass and Navier-Stokes equations given by

$$\nabla_{\boldsymbol{x}} \cdot (\rho \tilde{\boldsymbol{u}}) = 0 \tag{21}$$

and

$$\frac{D\tilde{\boldsymbol{u}}}{Dt} = -\frac{1}{\rho}\nabla_{\boldsymbol{x}}\tilde{p} + \left(\frac{\mu}{\rho} + \nu_t\right)\nabla^2_{\boldsymbol{x}}\tilde{\boldsymbol{u}}, \tag{22}$$

where $\frac{D}{Dt}$ is the material derivative, $\tilde{\boldsymbol{u}}$ is the velocity, t is time, $\boldsymbol{x}$ is the spatial location, $\nabla_{\boldsymbol{x}}$ is the spatial gradient, ρ is the density, $\tilde{p}$ is the pressure, μ is the dynamic viscosity, and ν_t is the turbulent viscosity. The density is specified as $\rho = 1$ kg m^{-3}, and the dynamic viscosity is specified as $\mu = 1.8 \times 10^{-5}$ kg m^{-1} s^{-1}. The turbulent viscosity, ν_t, is modeled using the Smagorinsky–Lilly LES model as

$$\nu_t = C_s^2 \Delta^2 |\boldsymbol{S}|, \tag{23}$$

where $C_s = 0.17$, Δ is the grid cell size, and $\boldsymbol{S}$ is the strain rate tensor given by

$$\boldsymbol{S} = \frac{1}{2}\left[\nabla_{\boldsymbol{x}}\tilde{\boldsymbol{u}} + (\nabla_{\boldsymbol{x}}\tilde{\boldsymbol{u}})^T\right], \tag{24}$$

and $|\boldsymbol{S}| = (2\boldsymbol{S} : \boldsymbol{S})^{1/2}$.

Low- and high-fidelity models were developed for this study using the WindSE framework. A Cartesian discretization of the computation domain is specified, where the grid is refined twice in the wake region as well as near the turbine rotors. Each high-fidelity simulation is run to 1,200 s using Taylor-Hood elements (Ern and Guermond, 2004). The power and loading results only use information from the final 600 s of simulation time. The low-fidelity model was selected using the same grid as the high-fidelity model, but runs to 400 simulation seconds and uses piecewise linear elements. The power is averaged after 300 s, and the loading is estimated using the front and back turbine moments past 300 s. Using 8 processors, the high-fidelity model was measured to take 5.4 hours to run, and the low-fidelity model was measured to take 0.24 hours to run. This corresponds to a cost ratio of approximately 0.05. The flow fields produced by the low- and high-fidelity models are compared in Figure 2.

3.2 Objective Functions

The objective of the optimization is to minimize negative power, $-P$, and loading, L, with respect to each turbine yaw offset such that the yaw offset angles, $\boldsymbol{\gamma}$, are bounded between $-30°$ and $30°$.

To compute power and loading, the actuator line model outputs forcing history for each node on each blade on each turbine. These forces are used to compute the time-dependent flapwise bending moment as

$$M(t) = \int_0^R F(r,t)\, r\, dr, \tag{25}$$

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

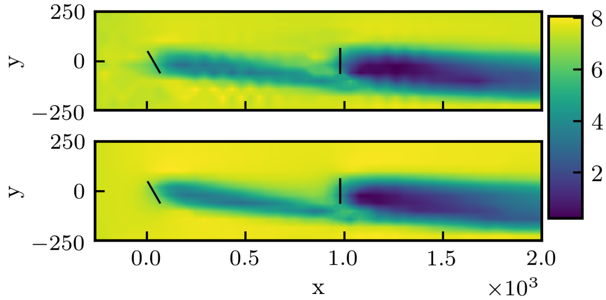

Figure 2. Time-averaged velocity magnitude fields at the turbine hub height associated with the low- and high-fidelity models. The top panel shows the flow field associated with the low-fidelity model and the bottom panel shows the flow field associated with the high-fidelity model. In both cases, the front turbine is offset by 30° and the back turbine has no yaw offset. Brighter colors correspond to faster velocity magnitudes. Turbine positions are shown with dark lines.

where M is the flapwise bending moment, r is the radial location along the blade, R is the radius of the blade, and F is the force normal to the rotor plane. In practice, force is computed by integrating interpolated thrust and lift coefficients along the blade span, and the differential forces are multiplied by the length to the rotor center and summed.

Power and loading are quantified, discarding an initial transient period. Power is computed as the average total power after the initial transient period. Loading is quantified by summarizing the time history of the flapwise bending moment of one blade in the front and back turbines after the same initial transient period. Using the high-fidelity model, loading is computed as the sum of damage equivalent loads (DELs) (International Electro-technical Commission, 2015) associated with the front and back turbine flapwise bending moments. Each DEL is computed using the rainflow counting algorithm as

$$DEL(M) = \left(\sum_{i=1}^{100} R_i^m \frac{c_i}{\Delta t} \right)^{1/m}, \tag{26}$$

where i loops through each cycle found using the rainflow counting algorithm, R_i is a load range, c_i is the number of cycles associated with the ith range bin of the moment load spectrum, Δt is the time elapsed in seconds, and m is the Wöhler Exponent. In this study, m is set as 10, and R_i and c_i are computed using the fatpack Python package (Frøseth and Capponi, 2021), utilizing 100 loading bins. The loading objective is computed as the sum of the flapwise bending moment DELs associated with the front and back turbines as

$$\hat{L}^{\mathrm{HF}} = DEL(M^{\mathrm{front}}) + DEL(M^{\mathrm{back}}), \tag{27}$$

where M^{front} is the moment associated with the front turbine and M^{back} is the moment associated with the back turbine, measured in Newton-meters. The power is measured in megawatts. The loading objective, L, is normalized to be negative and on a similar scale to power, as $L = \hat{L}/10^7 - 10$, where $\hat{L}$ is the load prior to normalization.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

3.3 Optimization Implementation

Here we use a simple optimization approach for simplicity of demonstration. In each iteration, the $\boldsymbol{l}$ correlation scale parameter in Eq. (4) is selected based on the maximum likelihood function (Pedregosa et al., 2011), with a lower bound of $5°$ and an upper bound of $30°$. The reference point, $\boldsymbol{r}$, in Equation 12 is specified as $(0°, 0°)$. In our formulation, we maximize $J(\boldsymbol{\gamma}, l)$ using a grid-based search for the maximum value. The grid is evenly spaced with 31 inputs per yaw offset dimension. Individual grids are considered for all values of l. The $EHVI$ is computed using Monte Carlo sampling with 1,200 samples taken from the GP. The EHVI may also be computed through numerical quadrature (Emmerich et al., 2011; Hupkens et al., 2015). When dealing with inputs of larger dimension, the EHVI may be maximized using an optimization algorithm, such as a genetic algorithm or the EGO approach. After the optimization, the Pareto set was refined using B-spline interpolation in SciPy (Virtanen et al., 2020).

3.3.1 Initial Sampling

Initial sampling points are selected using a heuristic approach, where an assumed kernel is used to progressively minimize the standard deviation of the predictor. The simplest approach to initializing the optimization procedure is to randomly sample the low-fidelity and discrepancy functions. Random initial sampling may drastically affect the optimization results, so a deterministic and symmetric sampling strategy is used as a test case. An isotropic kernel is used with a correlation scale of $10°$. The GP model is initialized with the point $(0°, 0)$. A 100-by-100 grid of inputs ranging between $-30°$ and $30°$ degrees is used to find the next point that minimizes uncertainty in the predicted variance. This process was repeated iteratively to generate 100 points. These points were used to naively estimate the optimal power and loading, and Pareto hypervolume, as a reference. The first 5 points were used as the initial high-fidelity samples and the first 20 points are used as the initial low-fidelity samples. This heuristic sampling approach is also used to generate 100 samples for use in a correlation analysis and as a naive, baseline approach to searching for the Pareto set.

3.3.2 Low-Fidelity Loading Model

The low-fidelity model was selected to have a low cost ratio and high correlation in predicted power and loading when compared to the high-fidelity model. We used 100 samples obtained using the heuristic sampling method described in Section 3.3.1 to test the correlation. In practice, this correlation test would not be part of the optimization procedure and reasonably accurate models would be identified based on past experience and/or expert opinion. The correlation analysis revealed a correlation of 0.976 between the low- and high-fidelity power predictions. Using the DEL as the load proxy in the low-fidelity model yielded a low correlation between the low- and high-fidelity models. We explored other potential loading proxies—applying the proxy to both the front and back turbine moment histories then summing the results—and the results of each correlation analysis are

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

presented in Table 1. We used the proxy associated with the highest-measured correlation,

$$\hat{L}^{\mathrm{LF}} = \mu_t\left[M^{\mathrm{front}}(t)\right] + 5\sigma_t\left[M^{\mathrm{front}}(t)\right] + \mu_t\left[M^{\mathrm{back}}(t)\right] + 5\sigma_t\left[M^{\mathrm{back}}(t)\right], \tag{28}$$

where μ_t and σ_t are the mean and standard deviation, respectively, with respect to time.

Table 1. Correlations observed between high-fidelity DEL and different loading proxies of the low-fidelity model using 100 heuristic samples.

Proxy	DEL	μ	σ	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	$\mu+4\sigma$	$\mu+5\sigma$	$\mu+6\sigma$
Correlation	0.742	0.103	0.800	0.479	0.745	0.857	0.892	0.899	0.896

4 Results and Discussion

4.1 Pareto Set Computation

The convergence of the single-fidelity and multifidelity optimization approaches are compared in Figure 3. The dashed lines show the hypervolume, best-sampled load, and best-sampled power found from 100 sampled points using the heuristic sampling approach, which the single-fidelity and multifidelity approaches both outperformed. The EHVI associated with the multifidelity approach was generally lower than the EHVI associated with the single-fidelity approach. The multifidelity approach took less than one-third as much total time to estimate the optimal power and loading compared to the single-fidelity approach. After five initial samples, the single-fidelity approach locates the optimal power after two iterations, but requires 35 iterations to locate the optimal load. The single-fidelity approach converges after 40 equivalent high-fidelity samples, including the five initial samples. The multifidelity approach requires 12 equivalent high-fidelity iterations to find the optimal power and loading. The optimal power is achieved with 26° offset in the front turbine and 2° offset in the back turbine. The optimal loading is achieved with 22° offset in the front turbine and $-30°$ offset in the back turbine.

Figure 4 compares the Pareto sets found using 50 equivalent high-fidelity evaluations utilizing the single-fidelity and multifidelity approaches. The multifidelity approach captures five Pareto points and the single-fidelity approach captures four Pareto points. The estimated Pareto sets are very similar, although the single-fidelity approach captures more of the Pareto set close to the optimal power and the multifidelity approach captures more of the Pareto set close to the optimal loading. The results of the single-fidelity and multifidelity approaches are combined to show a single Pareto set, which has a shape similar to a logit function.

The Pareto set resulting from the combination of the single-fidelity and multifidelity approaches was interpolated to create refinement samples using B-spline interpolation (Virtanen et al., 2020) with 10 interpolation points. To pick up more of the Pareto set, this interpolation was offset in the γ_1 direction by $-2°$, $-1°$, $1°$, and $2°$ when creating the input refinement set. The Pareto set resulting from these additional refinement samples is visualized in Figure 5. The resulting Pareto set has three more

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

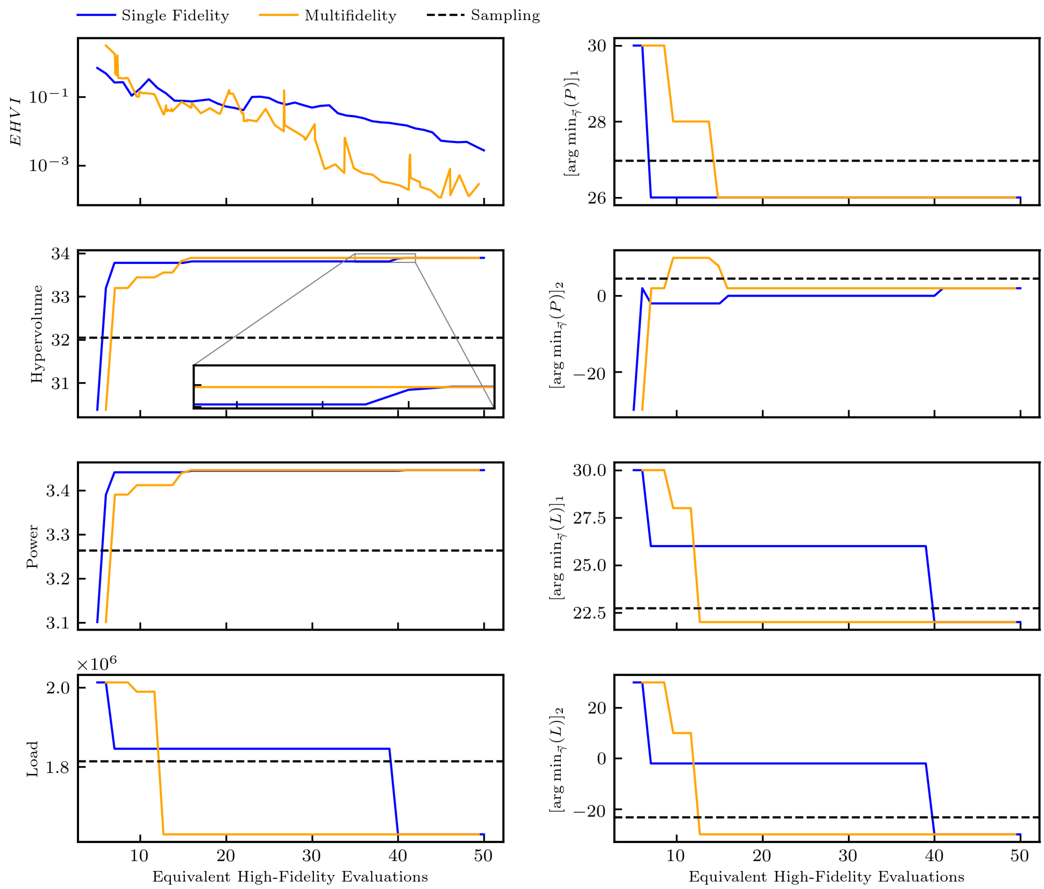

Figure 3. Convergence history of the single-fidelity and multifidelity approaches. The left plots show the EHVI, hypervolume, and best-observed power and loading. The right plots show the yaw configurations associated with the best-observed power and loading.

points than the Pareto set found combining the single-fidelity and multifidelity approaches. The optimization algorithm did not originally fill in these points because they reside in a relatively flat portion of the Pareto set (i.e., dP/dL is small), where adding points would not be expected to increase the Pareto set hypervolume based on the rectangular quadrature employed in this study. Adding these points increased the hypervolume of the discovered Pareto set by a very small amount, on the order of 0.002%. Even when using such a refined set of inputs, there are several points where the Pareto set jumps from one yaw position to another; attention would be needed for an operator to control the turbines to operate along the Pareto set.

The multifidelity approach was successful in quantifying the trade-offs between loading and power, and was shown to be more efficient than its single-fidelity counterpart. From the presented results, we find that loading may be reduced by 4% while only reducing the optimal power by 0.3%. Table 2 shows the power and front and back turbine DELs associated with several strategies. Slight adjustments to the back turbine angle result in substantial differences in the back turbine loading. These small changes in yaw position adjust the turbine thrust away from the flow, reducing the total thrust imparted on the back turbine.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

Open Access WIND ENERGY SCIENCE DISCUSSIONS

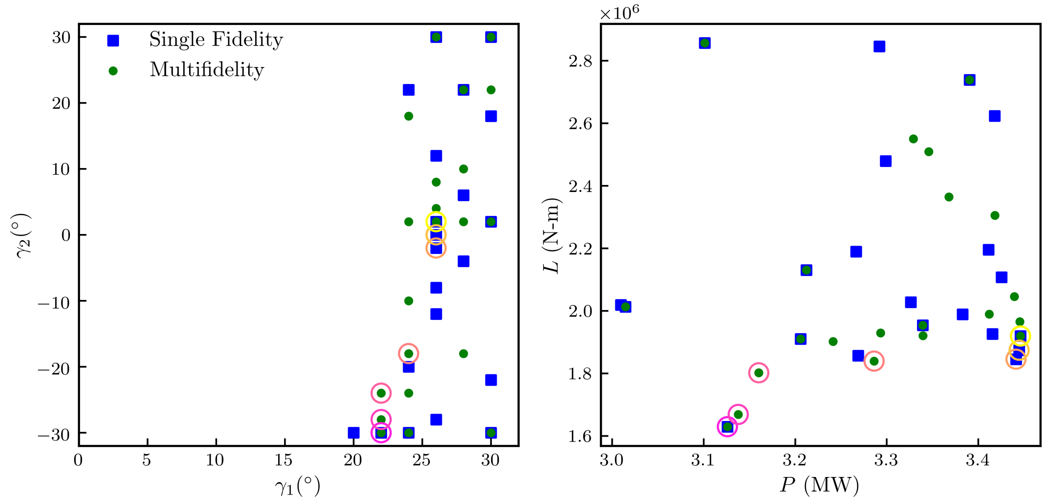

Figure 4. Sampled inputs and outputs associated with power greater than 3 MW and loads less than 3 MN-m. Points associated with the single-fidelity approach are shown using blue square markers, and points associated with the multifidelity approach are shown using green circular markers. A Pareto set constructed from the single-fidelity and multifidelity results is highlighted with hollow circles, where darker (magenta) circles correspond to Pareto points with lower loads and lower powers.

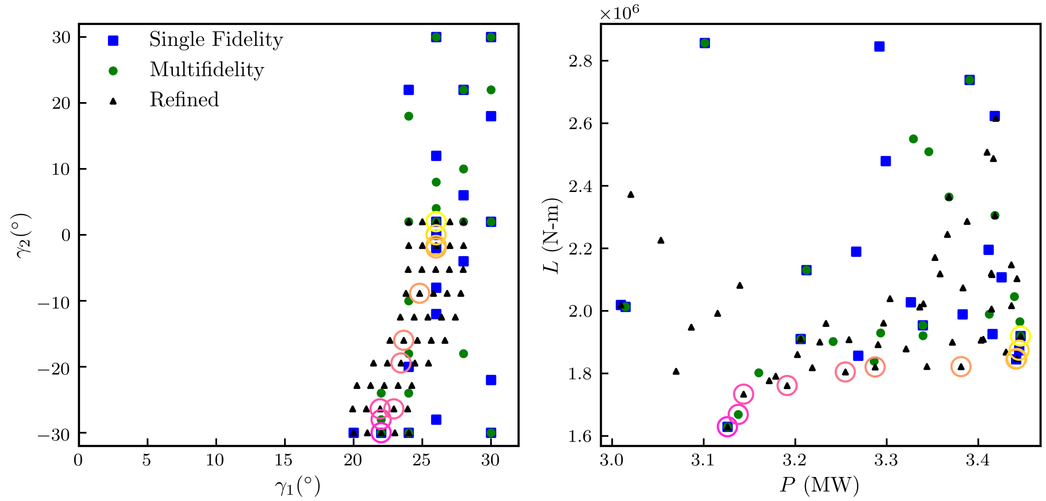

Figure 5. Sampled inputs and outputs associated with power greater than 3 MW and loads less than 3 MN-m. Points associated with the single-fidelity approach are shown using blue square markers and points associated with the multifidelity approach are shown using green circular markers. Refinement points are shown as black triangles. A Pareto set constructed from the single-fidelity, multifidelity, and refinement samples is highlighted with hollow circles, where darker (magenta) circles correspond to Pareto points with lower loads and lower powers.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

Table 2. Observed power and loads for various yaw configurations

γ_1 (°)	γ_2 (°)	Power (MW)	Load (M-Nm)	Front Turbine DEL (M-Nm)	Back Turbine DEL (M-Nm)
0	0	2.24	2.77	0.10	2.67
30	0	3.24	3.53	0.77	2.75
−30	0	2.98	5.59	0.93	4.66
26	2	3.45	1.92	0.61	1.31
26	0	3.44	1.87	0.61	1.27
26	−2	3.44	1.85	0.61	1.24
26	−30	3.25	2.17	0.61	1.57
22	0	3.26	3.55	0.52	2.07
22	−30	3.12	1.63	0.52	1.11

4.2 Flow Physics Insights

Figure 6 shows the flow fields associated with neutral, $-30°$, and $+30°$ yaw offsets in the front turbine, with the back turbine aligned with the wind direction. When the front turbine is offset, two structures are produced: a pair of counter-rotating vortices as well as a coherent structure that is drawn from the boundary layer. The direction of vortex rotation is determined by the direction of thrust the turbine imparts on the incoming air. Induced vortices generally rotate in the opposite direction from the blades that generated them, and the location of the vortices is determined by their rotational direction and the direction of blade rotation. The upper vortex associated with the positive yaw offset is lower in elevation than the upper vortex produced by the negative yaw offset. The upper vortex also drifts less in the crossflow direction when using the positive yaw offset than when using a negative yaw offset. The bottom vortex drifts similarly in both the positive and negative offset cases. All this amounts to a larger and more extreme velocity deficit encroaching on the back turbine when using the negative yaw offset, rather than the positive yaw offset, resulting in more loading and less power.

Time-averaged flow fields associated with the optimal power, $\boldsymbol{\gamma} = (26°, 2°)$, and optimal loading, $\boldsymbol{\gamma} = (22°, -30°)$, solutions are compared in Figure 7, which shows vertical slices of the flow field before and after the flow reaches the back turbine. The lesser front turbine yaw offset angle in the $\boldsymbol{\gamma} = (22°, -30°)$ case results in less lateral movement of the wake, and the wake structure has greater overlap with the back turbine than in the $\boldsymbol{\gamma} = (26°, 2°)$ case. The stronger vortical motion resulting from the $\boldsymbol{\gamma} = (26°, 2°)$ case results in the boundary layer being convected further inward. This boundary layer structure also impacts the back turbine less in the $\boldsymbol{\gamma} = (22°, -30°)$ case because of the reduced back turbine projected area. As the wake convects past the back turbine, additional vorticity is added to the flow. In the $\boldsymbol{\gamma} = (22°, -30°)$ case, the boundary layer structure appears to be pushed back down by the rotation of the bottom vortex.

Figure 8 shows time histories of the flapwise bending moment associated with the front and back turbines for various wake steering strategies. The spikes in the back turbine loading history are caused by the wake impacting the back turbine. The baseline strategy, $\boldsymbol{\gamma} = (0°, 0°)$, has a lower mean value in the back turbine load than the other offset strategies, but has greater

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

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

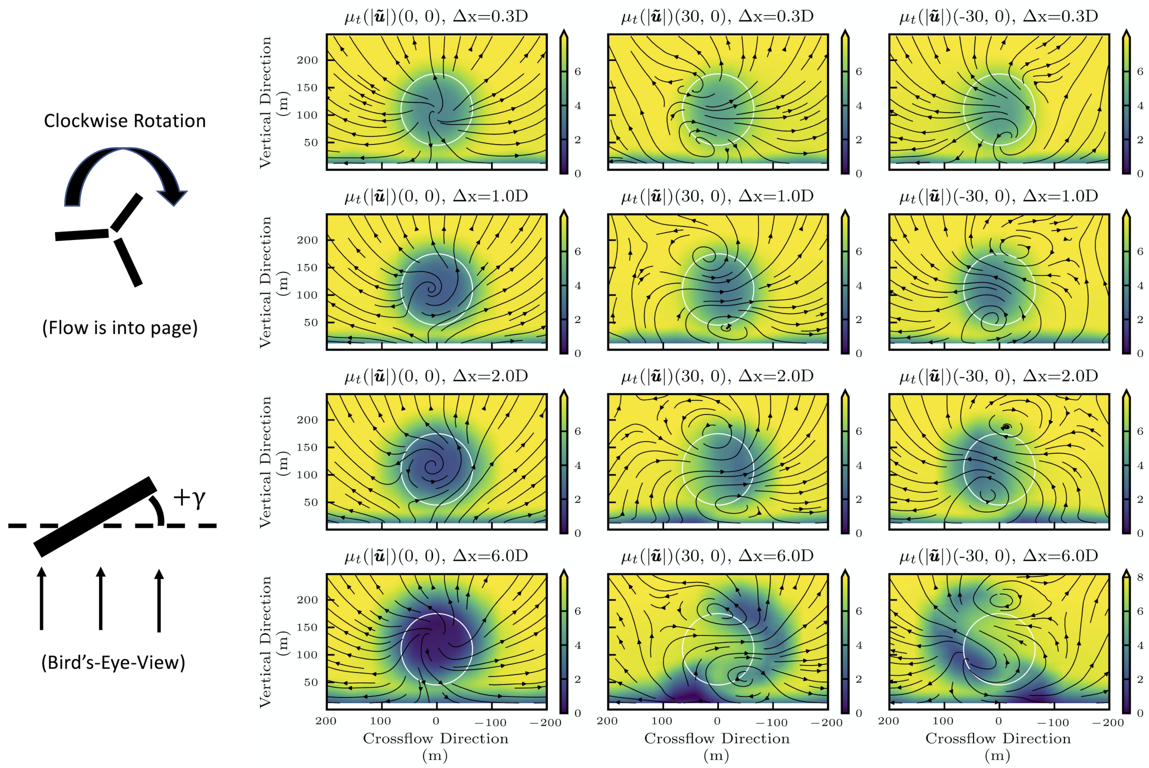

Figure 6. Flow fields associated with the extreme and neutral offsets in the front turbine, viewed from upstream. The Δx term indicates the distance downstream from the front turbine in terms of rotor diameters. Brighter colors show faster velocities. Streamlines show the direction of the crossflow and vertical velocity components. In each plot, the vertical and crossflow location of the back turbine is shown as a white circle. The turbines rotate clockwise when viewed from upstream. A diagram is shown on the left depicting the direction of positive yaw offset when viewing the turbine from above.

variation due to wake effects. This causes a larger DEL than many of the other strategies examined (see Table 2). When $\boldsymbol{\gamma} = (-30°, 0°)$, the back turbine shows greater downward spikes in loading than when $\boldsymbol{\gamma} = (30°, 0°)$, because of the greater velocity deficit discussed above. The $\boldsymbol{\gamma} = (26°, 0°)$ case yields smaller downward spikes associated with the back turbine loading than when $\boldsymbol{\gamma} = (30°, 0°)$, because the strength of the counter-rotating vortices is such that the structure convected from the boundary layer does not impact the back turbine as adversely. The $\boldsymbol{\gamma} = (22°, 0°)$ offset case has larger downward spikes in the back turbine loading than the $\boldsymbol{\gamma} = (26°, 0°)$ offset case, because the latter steers the wake further away from the back turbine. When the back turbine is offset to $-30°$, the back turbine thrust and associated moments are generally reduced. With this extreme back turbine yaw offset, there is less variation in the back turbine loading when the front turbine is offset by 22° than when it is offset by 26°, because the former case results in greater variation of velocity across the back turbine rotor plane.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

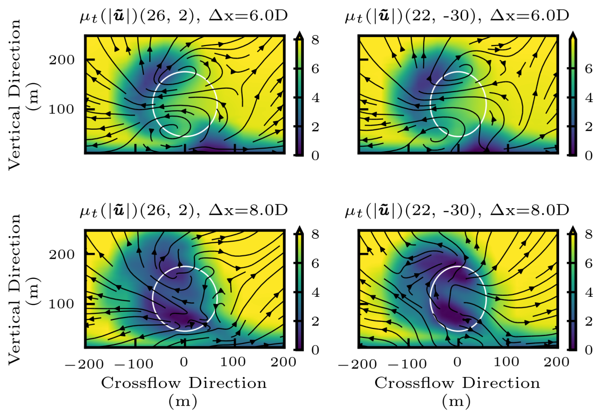

Figure 7. Time-averaged flow fields associated with the optimal power (left) and loading (right) found by the optimization, viewed from upstream, 6 rotor diameters (top) and 8 rotor diameters (bottom) away from the front turbine. Brighter colors indicate faster velocity magnitudes. In each plot, the vertical and crossflow locations of the back turbine are shown as a white ellipse.

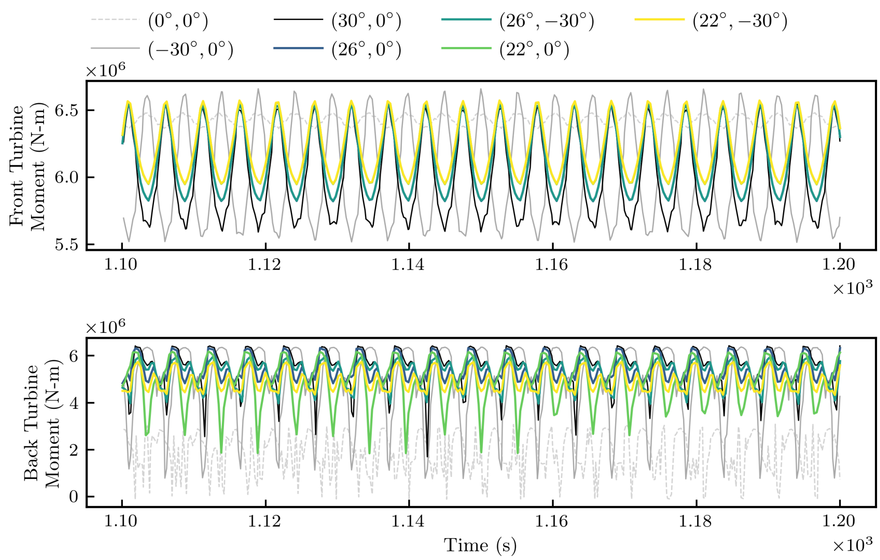

Figure 8. Loading histories associated with different yaw offset strategies (values of $\boldsymbol{\gamma}$).

5 Conclusions

This paper has demonstrated a multifidelity multiobjective optimization approach for wake steering strategies. Actuator line simulations were carried out using the WindSE tool, using a coarser simulation as the low-fidelity model. The high-fidelity loading was characterized as the sum of flapwise bending moment DELs on blades on the front and back turbines. Character-

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

izing the low-fidelity loading with a DEL resulted in a relatively low correlation between the low- and high-fidelity loading predictions, so a different low-fidelity surrogate was developed with a higher correlation.

The multifidelity multiobjective optimization approach was effective in exploring the trade-offs between loading and power when developing a wake steering design. Convergence was achieved in the multifidelity optimization case after approximately 30% as many equivalent high-fidelity model evaluations as in the single-fidelity case. Exploring the solutions in the final Pareto sets guided insights into the fundamental flow physics. A positive front turbine yaw offset is more effective at reducing loading and increasing power than a negative yaw offset because the counter-rotating vortices produce a greater velocity deficit in the downstream wake. The boundary layer is convected by the counter-rotating vortices, adversely affecting loading, and this may be avoided using less-extreme front turbine yaw offsets. Slightly modifying the back turbine yaw offset reduced loading by 4% and only reduced power by 0.3%. Greater offsets in the back turbine also led to less overall loading, with significantly less power generation.

It is well known that yaw position errors can adversely affect the performance of wake steering strategies. This is especially true when it comes to turbine loading. A 30° yaw offset is already an aggressive strategy, and unfavorable yaw position errors may result in even more aggressive yaw offsets in practice. Yaw offset errors are generally extreme in lower wind speeds, which is when wake steering strategies are most efficient at increasing power. Previous work examined the potential of considering yaw error uncertainties in the wake steering optimization problem (Quick et al., 2017, 2020). The multifidelity optimization approach presented in this paper could conceivably be extended to optimization under uncertainty, using the final GP models to propagate yaw position uncertainty, and potentially even modifying the EHVI definition to include uncertainty information.

A drawback of the presented approach is that it requires sequential high-fidelity model evaluations. In practice, it is often feasible to evaluate a high-fidelity model several times in parallel, and the greatest expense is time needed to run the optimization. This framework may be extended to allow for parallel function evaluations. A simple approach is to use predictions of the GP as stand-ins for future model evaluations, iteratively using these points to construct the next iteration of the GP and the associated EHVI (Ginsbourger et al., 2010). Yang et al. (2019) propose dividing the input space into separate regions for parallelization of EHVI optimization. Another intuitive approach could be to include refinement points during each iteration. Refinement points could be selected using the Pareto set predicted by the GP models or interpolated along the observed Pareto set.

In future work, this framework should be applied to a larger array of turbines. While this presents additional complications in maximizing the EHVI, we anticipate there will be even greater cost savings from the multifidelity approach as the number of turbines increases. Additionally, the framework should be extended to allow for optimization under uncertainty, as it is not realistic to assume perfect control of wind turbine yaw positions. Finally, the framework should incorporate more lower-fidelity models and be combined with layout optimization to realize the full benefits of multifidelity multiobjective wake steering optimization.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

6 Acknowledgments

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. This research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory.

References

Andersson, L. E. and Imsland, L.: Real-time optimization of wind farms using modifier adaptation and machine learning, Wind Energy Science, 5, 885–896, 2020.

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, 2018.

Ariyarit, A. and Kanazaki, M.: Multi-Fidelity Multi-Objective Efficient Global Optimization Applied to Airfoil Design Problems, Applied Sciences, 17, https://doi.org/10.3390/app7121318, 2017.

Bortolotti, P., Dykes, K., Merz, K., Sethuraman, L., Verelst, D., Zahle, F., et al.: IEA Wind Task 37 on systems engineering in wind energy, WP2-Reference Wind Turbines.: IEA Wind Task, 37, 2019.

Damiani, R., Dana, S., Annoni, J., Fleming, P., Roadman, J., van Dam, J., and Dykes, K.: Assessment of wind turbine component loads under yaw-offset conditions, Wind Energy Science, 3, 173–189, 2018.

Emmerich, M. T., Giannakoglou, K. C., and Naujoks, B.: Single-and multiobjective evolutionary optimization assisted by Gaussian random field metamodels, IEEE Transactions on Evolutionary Computation, 10, 421–439, 2006.

Emmerich, M. T., Deutz, A. H., and Klinkenberg, J. W.: Hypervolume-based expected improvement: Monotonicity properties and exact computation, in: 2011 IEEE Congress of Evolutionary Computation (CEC), pp. 2147–2154, IEEE, 2011.

Ern, A. and Guermond, J.-L.: Theory and practice of finite elements, vol. 159, Springer, 2004.

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, 2018.

Frøseth, G. T. and Capponi, L.: fatpack, https://github.com/Gunnstein/fatpack, 2021.

Ghoreishi, S. F. and Allaire, D.: A Fusion-Based Multi-Information Source Optimization Approach using Knowledge Gradient Policies, https://doi.org/10.2514/6.2018-1159, 2018.

Ginsbourger, D., Le Riche, R., and Carraro, L.: Kriging is well-suited to parallelize optimization, in: Computational intelligence in expensive optimization problems, pp. 131–162, Springer, 2010.

Giselle Fernández-Godino, M., Park, C., Kim, N. H., and Haftka, R. T.: Issues in deciding whether to use multifidelity surrogates, AIAA Journal, 57, 2039–2054, 2019.

He, X., Tuo, R., and Wu, C. J.: Optimization of multi-fidelity computer experiments via the EQIE criterion, Technometrics, 59, 58–68, 2017.

Hennig, P. and Schuler, C. J.: Entropy search for information-efficient global optimization, The Journal of Machine Learning Research, 13, 1809–1837, 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.

Huang, D., Allen, T. T., Notz, W. I., and Miller, R. A.: Sequential kriging optimization using multiple-fidelity evaluations, Structural and Multidisciplinary Optimization, 32, 369–382, 2006.

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, 5, 309–329, 2020.

Hupkens, I., Deutz, A., Yang, K., and Emmerich, M.: Faster exact algorithms for computing expected hypervolume improvement, in: international conference on evolutionary multi-criterion optimization, pp. 65–79, Springer, 2015.

https://doi.org/10.5194/wes-2021-152
Preprint. Discussion started: 18 January 2022

International Electro-technical Commission: IEC 61400-13: Wind turbines-Part 13: Measurement of mechanical loads, International Electro-technical Commission (IEC), Geneva, 2015.

Jones, D. R., Schonlau, M., and Welch, W. J.: Efficient global optimization of expensive black-box functions, Journal of Global optimization, 13, 455–492, 1998.

Lin, M. and Porté-Agel, F.: Power Maximization and Fatigue-Load Mitigation in a Wind-turbine Array by Active Yaw Control: an LES Study, in: Journal of Physics: Conference Series, vol. 1618, p. 042036, IOP Publishing, 2020.

Martínez-Tossas, L. A. and Branlard, E.: The curled wake model: equivalence of shed vorticity models, in: Journal of Physics: Conference Series, vol. 1452, p. 012069, IOP Publishing, 2020.

Martínez-Tossas, L. A., Annoni, J., Fleming, P. A., and Churchfield, M. J.: The aerodynamics of the curled wake: a simplified model in view of flow control, Wind Energy Science, 4, 127–138, 2019.

National Renewable Energy Laboratory: WindSE, https://github.com/NREL/WindSE, 2021.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Picheny, V., Ginsbourger, D., Richet, Y., and Caplin, G.: Quantile-based optimization of noisy computer experiments with tunable precision, Technometrics, 55, 2–13, 2013.

Qin, C., Klabjan, D., and Russo, D.: Improving the expected improvement algorithm, in: Advances in Neural Information Processing Systems, pp. 5381–5391, 2017.

Quick, J., Annoni, J., King, R., Dykes, K., Fleming, P., and Ning, A.: Optimization under uncertainty for wake steering strategies, in: Journal of physics: Conference series, vol. 854, p. 012036, IOP Publishing, 2017.

Quick, J., King, J., King, R. N., Hamlington, P. E., and Dykes, K.: Wake steering optimization under uncertainty, Wind Energy Science, 5, 413–426, 2020.

Rajnarayan, D., Haas, A., and Kroo, I.: A multifidelity gradient-free optimization method and application to aerodynamic design, in: 12th AIAA/ISSMO multidisciplinary analysis and optimization conference, p. 6020, 2008.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian processes for machine learning, Adaptive computation and machine learning, MIT Press, Cambridge, Mass, 2006.

Rinker, J. M., Soto Sagredo, E., and Bergami, L.: The Importance of Wake Meandering on Wind Turbine Fatigue Loads in Wake, Energies, 14, 7313, 2021.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N.: Taking the human out of the loop: A review of Bayesian optimization, Proceedings of the IEEE, 104, 148–175, 2015.

Van Dijk, M. T., van Wingerden, J.-W., Ashuri, T., and Li, Y.: Wind farm multi-objective wake redirection for optimizing power production and loads, Energy, 121, 561–569, 2017.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, C., Campagnolo, F., and Bottasso, C.: Does the use of load-reducing IPC on a wake-steering turbine affect wake behavior?, in: Journal of Physics: Conference Series, vol. 1618, p. 022035, IOP Publishing, 2020.

Yang, K., Palar, P. S., Emmerich, M., Shimoyama, K., and Bäck, T.: A multi-point mechanism of expected hypervolume improvement for parallel multi-objective bayesian global optimization, in: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 656–663, 2019.

Yin, X., Zhang, W., Jiang, Z., and Pan, L.: Data-driven multi-objective predictive control of offshore wind farm based on evolutionary optimization, Renewable Energy, 2020.

Zhan, D. and Xing, H.: Expected improvement for expensive optimization: a review, Journal of Global Optimization, 78, 507–544, 2020.