Data-driven Parameter Calibration in Wake Models

Bingjie Liu* and Eunshin Byon †
University of Michigan, Ann Arbor, Michigan, 48109, United States

Matthew Plumlee‡
Northwestern University, Evanston, Illinois, 60208, United States

Physical interactions among wind turbines, called wake effects, are known to be one of the significant factors that affect power generation performance in wind power systems. Among several wake modeling approaches, physics-based engineering models, such as Jensen's model, have been widely used due to their computational tractability. Although substantial efforts have been made to improve the accuracy of engineering wake models, few studies suggest calibrating the model parameters in the literature. We propose a new data-driven calibration approach for adjusting the model parameters using real operational data.

I. Nomenclature

C_t	=	thrust coefficient
D	=	rotor diameter
D_W	=	wake diameter
i	=	index of data records in a dataset
n	=	total number of data records in a dataset
t	=	turbine index
T	=	total number of turbines
x	=	downwind horizontal distance from an upstream turbine
u	=	free-flow wind speed
$u_\delta(x)$	=	wake-influenced wind speed at downstream distance x
y_{it}	=	power output from the t^{th} turbine at the i^{th} data record in a dataset
δ	=	wind speed deficit
θ	=	wake decay coefficient in Jensen's model
$\hat{\theta}_g$	=	globally calibrated wake decay coefficient in Jensen's model
$\hat{\theta}_l$	=	locally calibrated wake decay coefficient in Jensen's model
$\eta(u, t, \theta)$	=	power output from an engineering wake model for the t^{th} turbine at free-flow wind speed u

II. Introduction

As the scale of both wind turbines and wind farms grows, the wake effect becomes a significant factor when optimizing the performance in the wind industry. The process that a wind shade is cast by the upwind turbine to the downwind direction is called wake effect ([1]). The turbines at the upwind direction disturb and slow down the wind. The long wind trail that appears accordingly would cause both energy loss and increased structural/mechanical loads on the downstream turbine. Because of the limited land and budget constraints, many turbines in a wind farm are inevitably placed in the downstream position. Therefore, further expansion of wind energy application calls for accurate and efficient wake effect estimation.

Extensive studies have been conducted for understanding the wake effect. The main focus of wake effect models is to quantify the effect of wake on the power deficits at downstream turbines, where the power deficit refers to the decrease of power output generated from a downstream turbine, compared to that from a upstream turbine that faces free-stream wind.

*PhD candidate, Department of Industrial and Operations Engineering, 1773 IOE Building, 1205 Beal Avenue, Ann Arbor, MI 48109-2117
†Associate Professor, Industrial & Operations Engineering, 2773 IOE Building, 1205 Beal Avenue, Ann Arbor, MI 48109-2117
‡Assistant Professor, Industrial Engineering and Management Sciences, 2145 Sheridan Road, Tech, Evanston, IL 60208-3109

The approaches in the wake effect studies can be generally grouped into three categories - engineering models, computer fluid dynamics (CFD) models and data-driven statistical models. First, engineering models focus on developing analytic equations, based on physics-based principles. Tracing back to 1980s, engineering models, including Jensen's model [2], have been developed to estimate the wind speed deficit by applying the momentum equation. Later, Larsen model ([3]) and Frandsen model ([4]) were further developed to improve the estimation accuracy. These engineering models are usually intuitive and fast to run. However, they often rely on simplified assumptions in order to get an analytical solution and maintain the mathematical tractability. Such simplifications may cause inaccurate wake estimations. Second, with the advancement of numerical simulation models and computational technology, CFD models, which are based on the fluid mechanics, started to gain attentions ([5]). They provide sophisticated tools and have been applied to optimize turbine controls, e.g., yaw control, in the literature ([5]). However, the CFD models are usually complex and time-demanding, so their application has been limited to small-scale settings ([6]). Lastly, data-driven statistical models have been developed using real data collected from operational wind farms ([7–9]). This approach has benefits of quantifying wake effects in existing wind farms. However, it is not straightforward to apply its results to new wind farm settings (e.g., wind farms in different layouts, sizes, and/or terrains).

Among the three approaches, this paper uses the engineering modeling approach and focuses on improving the engineering model. The engineering model can be regarded as an inexact computer model. Outputs from the model may show a specific pattern of estimation deviation from real data ([10]). The estimation accuracy of existing engineering models can be improved by calibrating the model parameters. For example, in the Jensen's model, the commonly used wake decay coefficient, θ, is 0.075 and 0.04 for land-based and offshore wind farms, respectively ([11, 12]). However, Peña and Rathmann [13] show that the wake decay coefficient should be set differently from the recommended value. Specifically they suggest using the ratio of upstream-undisturbed friction velocity to undisturbed hub-height wind speed in order to adjust to an infinite row of wind turbines. Peña et al. [14] also obtain closer estimation of wind speed to real data by adjusting the wake decay coefficient at $\theta = 0.038$ for a land-based wind farm. These studies indicate that the estimation accuracy of engineering models can be improved by taking advantage of parameter calibration in the model.

Existing studies on the model calibration ([13, 14]) focus on calibrating the parameters globally ([15]). Here, the global calibration implies that it finds a unique parameter value that can be applied in any input conditions. However, some recent studies ([8, 9]) discuss that power deficits differ, depending on the environment condition. Specifically, power deficits are heterogeneous over the range of wind speeds due to the control mechanism to regulate power outputs from wind turbines. Therefore, the global approach, even with the carefully calibrated parameter, may cause the engineering model to generate biased outputs under some wind conditions.

Considering the heterogeneous pattern of wake, this study proposes a new wind-dependent calibration method in the engineering wake models. Although the proposed methodology can be applicable to any engineering models, we employ the Jensen wake effects model to illustrate the proposed method. Our implementation results indicate that when the wake decay coefficient in the Jensen's model is calibrated locally, depending on the incoming free-flow wind condition, the model outputs get closer to data, compared to the cases where the commonly used value ($\theta = 0.04$ or 0.075) or the globally calibrated parameter is used.

The remainder of this paper is organized as follows. Section III briefly reviews the Jensen's model and discusses the proposed wind-dependent calibration method. Section IV implements the proposed calibration method using data from an operational wind farm. Finally, Section V summarizes the study and provides future direction.

III. Wind-dependent Parameter Calibration

The original Jensen's model focuses on the wake caused by a single turbine ([2]). Later, it is extended to estimate wind deficits due to multiple wakes using the sum of square of velocity deficits, following the procedure described in Katic et al. [11]. Below we summarize the simple wake model due to its simplicity, but the proposed methodology can be applied to the multiple wake case.

In the Jensen's wake model ([2]), the wake can be characterized as a top-hat shape shown in Figure 1. The wake diameter, D_W, along with downwind horizontal distance, x, from an upstream turbine can be calculated as

$$D_W = D(1 + 2\theta \frac{x}{D}), \tag{1}$$

where D represents a rotor diameter and θ is the wake decay coefficient which serves as the parameter in the model (this is the parameter we will calibrate in the Jensen's model).

The wake-influenced area is defined as the area where downstream lateral distance is less than the wake diameter. In

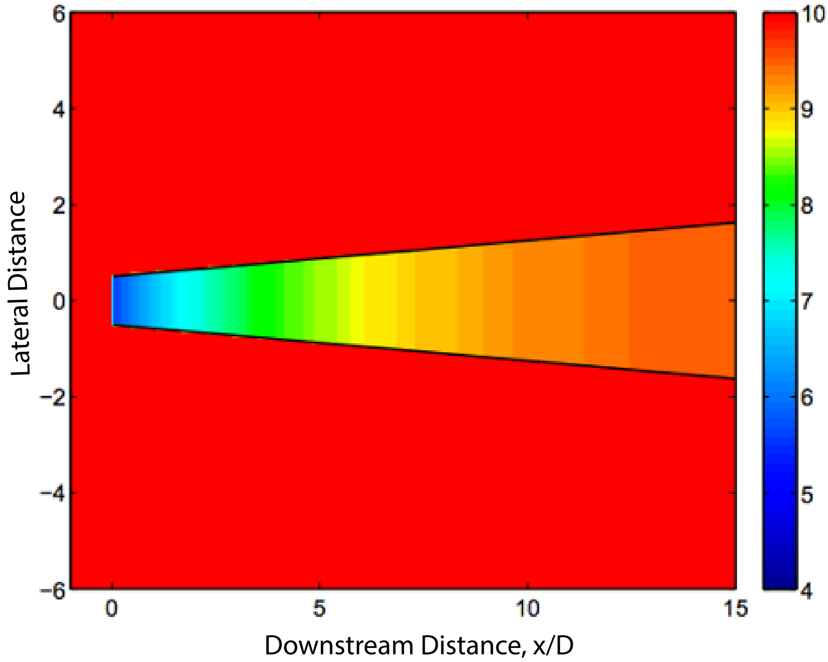

Fig. 1 Wake-influenced area in Jensen's model (excerpted and slightly modified from [16])

the wake-influenced area, the decreased wind speed is calculated as

$$u_\delta(x) = u\left[1 - \frac{1-\sqrt{1-C_t}}{(1+2\theta\frac{x}{D})^2}\right], \tag{2}$$

where u represents the free-stream wind speed that an upstream turbine faces and C_t is the thrust coefficient of a turbine. Then, wind speed deficit, which refers to the decreased proportion of free-stream speed, can be expressed as

$$\delta(x) = 1 - \frac{u_\delta(x)}{u} = \frac{1-\sqrt{1-C_t}}{(1+2\theta\frac{x}{D})^2} \tag{3}$$

The wake decay parameter, θ, in the Jensen's model plays an important role in estimating the wake-influenced (or downstream) wind speed $u_\delta(x)$. Figure 2 shows that as the distance between the upstream and downstream turbines gets larger, the downstream turbine faces smaller wind speed deficits, $\delta(x)$. Another important aspect is that as the wake decay parameter decreases, the deficit increases (see Equation (2)). This implies that when the wake is substantial, a smaller value should be used for the wake decay coefficient, θ.

To calibrate the wake decay coefficient, we use a dataset collected from a real wind farm. In the dataset, each i^{th} data record, $i = 1, \ldots n$, includes the power output, y_{it}, from the tth turbine, $t = 1, \cdots, T$, and the free-flow wind speed, u_i, collected at the meteorological tower (or mast) when the mast is not under wake. Let $\eta(u_i, t, \theta)$ denote the wake model's power output from the t^{th} turbine when the free-flow wind speed is u_i. Because the Jensen's model only calculates the incoming wind speed (that is, $u_\delta(x)$ in (2)) of each turbine but does not produce the power output, we combine the Jensen's model and power curve to obtain $\eta(u_i, t, \theta)$.

Our goal is to tune the parameter, θ in the wake model, $\eta(u_i, t, \theta)$, so that the wake model can explain the observed data better. In other words, we aim at obtaining the best value that minimizes the deviation between the wake model output and actual observed data. One possible approach is to globally calibrate the parameter by minimizing the squared deviation between the data and model outputs, that is,

$$\hat{\theta}_g = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^{T} \sum_{i=1}^{n} (y_{it} - \eta(u_i, t, \theta))^2 . \tag{4}$$

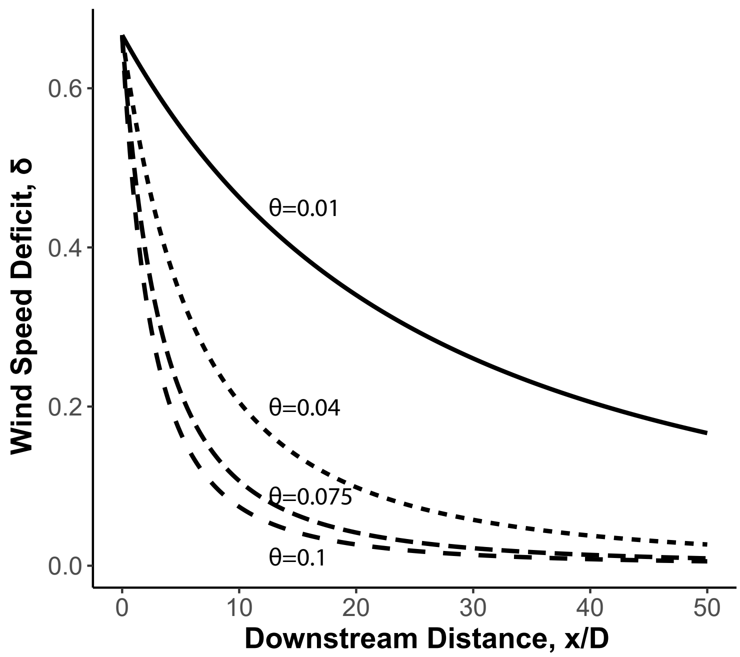

Fig. 2 Relationship between wake decay parameter, θ, and wind speed deficit, $\delta(x)$, in the Jensen wake model. The x-axis denotes the downstream distance relative to the turbine diameter.

The underlying assumption of Equation (4) is that the best parameter, $\hat{\theta}_g$, does not depend on the input wind speed. However, this assumption may cause biased calibration under some input wind speeds, when the parameter should depend on the input value. A recent study by You et al. [8] discusses that, as wind speed increases, the effect of wake on the power deficit of a downstream turbine effects increase, then the effect starts to decrease around the mid wind speed range due to the pitch control algorithm to regulate the power generation under high wind speeds. Those heterogeneous wake effects suggest that the wake decay coefficient, θ, should be calibrated locally, depending on the wind speed.

If we know a specific pattern of the parameter over the input space, e.g., from physical laws or domain knowledge, we can use a parametric form in our calibration process. For example, if we know the appropriate pattern of the wake decay coefficient should be quadratic over the input wind speed, u, we can employ $\theta(u) = \beta_0 + \beta_1 u + \beta_2 u^2$ and find the best β_0, β_1 and β_2 to minimize the loss function in (4). In many cases such specific forms may not be available.

Another way is to use local data points. Suppose we want to obtain the best θ at a specific input wind speed, u. Ideally, if we have a sufficient number of data records at u (i.e., if many u_i's in the dataset are the same as u), we can get θ by applying Equation (4) at those points. However, in practice it is impossible to have such data points in every u between the cut-in and cut-out wind speeds.

Our study proposes a nonparametric calibration that uses the information from a local neighborhood. The basic idea is to assign a higher weight on the data points near u in finding the best parameter at the specific wind speed u. Specifically, we locally calibrate the parameter as follows.

$$\hat{\theta}_l(u) = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} w(u_i, u) \sum_{t=1}^{T} (y_{it} - \eta(u_i, t, \theta))^2 , \tag{5}$$

where $w(\cdot, \cdot)$ denotes the weight function whose value gets higher as the difference between the u_i and u gets smaller. Note that the proposed non-parametric calibration estimation takes advantage of weighted errors in order to emphasize the influence from the neighbor.

In the proposed non-parametric calibration method, we can employ different weight functions for $w(\cdot, \cdot)$. In our implementation, we use the following Gaussian kernel,

$$w(u_i, u) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(u_i - u)^2}{2\sigma^2}\right), \tag{6}$$

where σ controls the smoothness of the calibrated parameter value over the input space.

IV. Case Study

We implement the proposed calibration method, using data collected from a wind farm. The wind farm contains more than thirty turbines and the layout is regular. Due to the data confidentiality, we omit detailed information about the wind farm. The dataset includes the 10-minute average wind speeds measured at the mast as an input and the 10-minute average power outputs from the turbines. Because the input in the Jensen's model denotes the free-stream incoming wind speed, we select the data records when the mast is not under wake.

As discussed earlier, the Jensen's model does not estimate the power output of each turbine, but generates the incoming wind speed at each turbine. Thus, we build the power curve that relates the wind speed and the power generation, using data from the upstream turbines. Specifically we employ a polynomial function in estimating the power curve for the wind farm and assume that every turbine exhibits the same power curve. The fitted power curve is shown in Figure 3. Then, the power output from the wake engineering model, $\eta(\cdot)$, is obtained by applying the estimated incoming wind speed from the Jensen's model to the power curve. Note that due to the data confidentiality, we normalize the power output with the rated power of turbines.

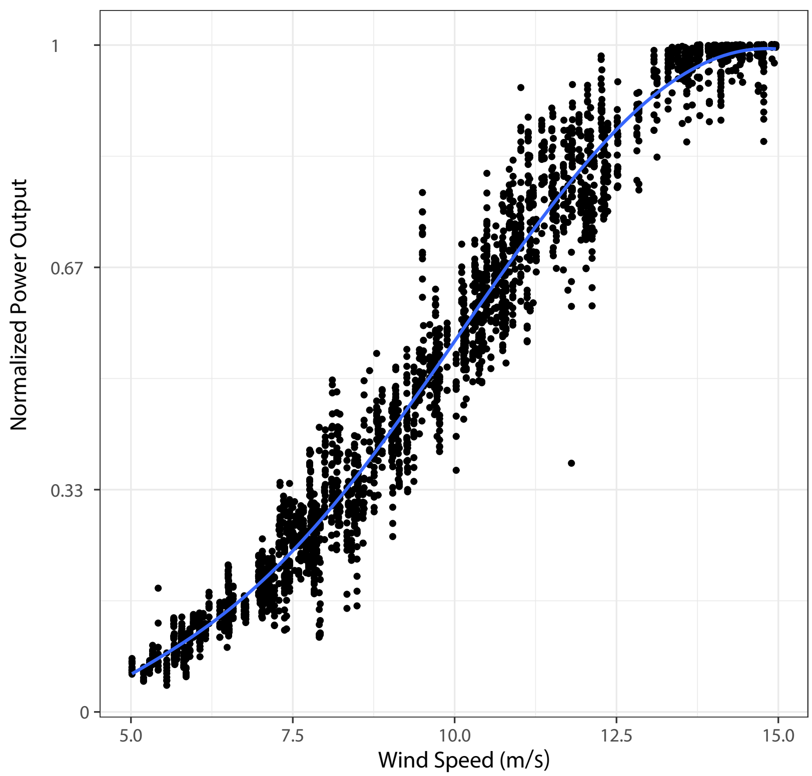

Fig. 3 Scatter plot and power curve

Figure 4 shows the localy calibrated coefficient value, $\hat{\theta}_l(u)$, over the input free-flow wind speed u. We observe the wind-dependent patterns from the estimation of optimal parameter. Specifically, we observe a U-shape, i.e., the calibrated value, $\hat{\theta}_l(u)$, is small in the mid speed range, whereas it is relatively high under low or high wind speeds. Recall that from the discussion on the influence of parameter θ in Jensen model in Section III, a smaller wake decay parameter implies more severe wake effects. We provide interpretations on the U-shape pattern as follows.

- When the incoming wind is weak, wind deficits caused by upstream turbines can be easily recovered back to the original wind speed because the energy loss is not significant. This phenomenon explains a large value under low wind speeds.
- As the wind speed increases toward the rated speed, wind speed recovery is not easily made and upstream turbines extract energy as much as possible, leading to significant power deficits at downstream turbines. Therefore, $\hat{\theta}_l(u)$

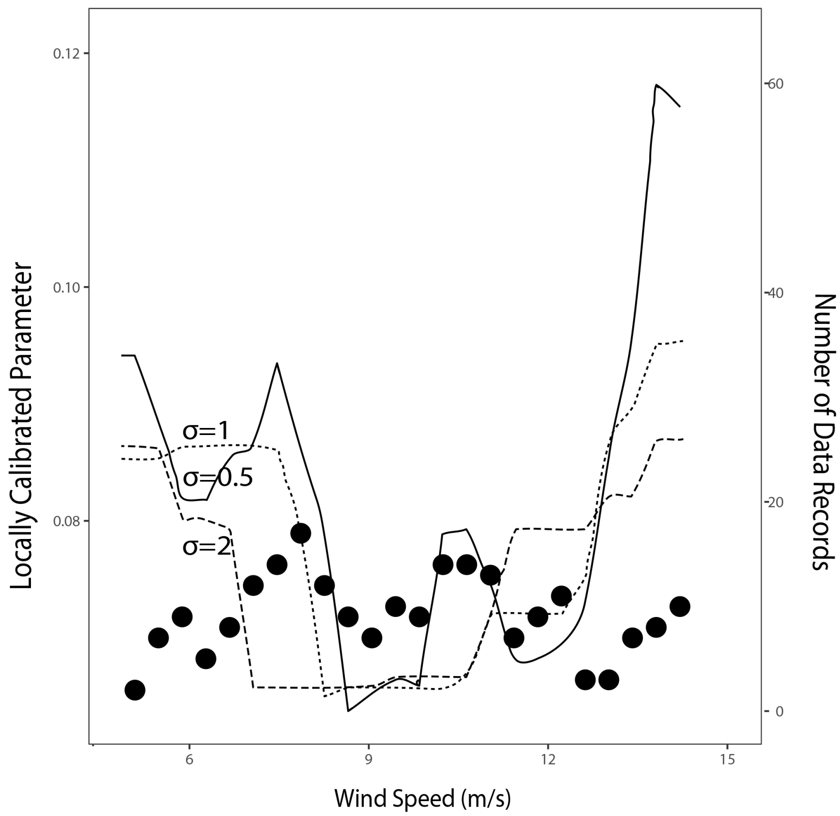

Fig. 4 Locally calibrated wake decay parameter. Multiple lines are obtained from different σ values in the weight function. Solid dots represent the number of data records. The right y-axis denotes the number of data records.

decreases as wind speed increases up to about 9 m/s.

- Under mid to high wind speeds, turbines operate pitch controls to protect the turbine structure. Therefore, upstream turbines do not extract the maximum level of energy, which weakens the wake effects and enables downstream turbines to generate the rated power. This explains the upward pattern between mid and high wind speed interval.

In summary, the wake effect becomes the most significant under the mid-speed regime. The valley of the curve represents this wake pattern. This pattern is also related to the power curve shown in Figure 3 where the curve changes from the convex to concave pattern at around 9 m/s. Up to 9 m/s, a small change in the wind speed leads to a large change in power output, which is associated with the downward wake decay coefficient. Then, the change in the power output slows down, corresponding to the upward pattern in $\hat{\theta}_l(u)$ when wind speed exceeds about 9 m/s.

We also note that a careful selection of σ in the weight function (or Gaussian kernel) is needed. Depending on the value of σ, we obtain different calibration curves of $\hat{\theta}_l(u)$. In general, a larger σ leads to a smoother curve. However, a too large σ may result in a relatively flat line. Selecting an appropriate value for σ is the subject of our future study.

We compare the performance of our approach with alternative approaches. We use the following root mean squared error (RMSE) to quantify the estimation accuracy of each method.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\sum_{i=1}^{t}(y_{it} - \hat{y}_{it})^2}{nT}}, \tag{7}$$

where $\hat{y}_{it}$ is the estimated power output from the wake engineering model.

Table 1 compares RMSEs from different settings of θ in the Jensen's model. Recall that $\theta = 0.075$ and $\theta = 0.04$ are the recommended values for land-based and offshore wind farms, respectively ([11, 12]). The global calibration approach, which finds the optimal coefficient globally in (4), uses $\hat{\theta}_g = 0.0794$ in this dataset. The proposed local calibration approach generates smaller errors (see the last three rows with different σ values) than the two recommended settings and the global calibration.

Table 1 Comparison of RMSEs from different parameter settings

Wake decay coefficient	RMSE
$\theta = 0$ **(no wake)**	0.19884
$\theta = 0.075$	0.05650
$\theta = 0.04$	0.07379
Global calibration ($\hat{\theta}_g = 0.794$)	0.05634
Local calibration with $\sigma = 0.5$	0.05548
Local calibration with $\sigma = 1$	0.05569
Local calibration with $\sigma = 2$	0.05606

To further evaluate the superiority of the local calibration approach over the global approach, we conduct statistical t-tests and evaluate the significance of the differences in RMSEs between the global and local calibration approaches. Specifically, we employ the paired and one-sided t-test. We obtain the p-value of less than 0.00001, indicating that the local calibration approach provides smaller estimation errors than the global approach in our case study.

V. Conclusion

In this study, we develop a wind-dependent calibration method that locally estimates wake decay coefficient in the engineering wake model. The proposed local calibration approach takes a nonparametric procedure without assuming any specific form of parameter pattern over the input space. Our case study suggests that our approach has a great potential to improve the estimation accuracy of engineering models.

In the future, we plan to apply the proposed idea to other engineering models including Larsen model ([3]) and Frandsen model ([4]). Moreover, we would like to test the proposed approach with multiple datasets collected from a wide rage of settings, e.g., offshore vs. land-based wind farms, regular vs. irregular layouts, small-, medium- and large-size wind farms when such datasets become available to us. More sophisticated power curve functions, e.g., spline model ([17]) or non-parametric curves ([18]), will be considered in our future study. Moreover, we plan to quantify the estimation uncertainty in the proposed local calibration. The outcomes of this research can be applied to the optimization of a wind farm layout, with the premise that the proper evaluation of wake effects will enable us to accurately estimate power deficits in downstream turbines and thus better optimize the power output from a whole wind farm, given the distribution of environmental factors.

Acknowledgments

This work was supported by the National Science Foundation (Grant No. CMMI-1362513, CMMI-1536924, and IIS-1741166) and the University of Michigan MCubed Grant.

References

[1] Vermeer, L., Sørensen, J. N., and Crespo, A., "Wind turbine wake aerodynamics," *Progress in aerospace sciences*, Vol. 39, No. 6-7, 2003, pp. 467–510.

[2] Jensen, N. O., *A note on wind generator interaction*, 1983.

[3] Larsen, G. C., *A simple wake calculation procedure*, 1988.

[4] Frandsen, S., "On the wind speed reduction in the center of large clusters of wind turbines," *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 39, No. 1, 1992, pp. 251 – 265.

[5] Andersen, S. J., Sørensen, J. N., Ivanell, S., and Mikkelsen, R. F., "Comparison of engineering wake models with CFD simulations," *Journal of physics: Conference series*, Vol. 524, 2014, p. 012161.

[6] Rados, K., Larsen, G., Barthelmie, R., Schlez, W., Lange, B., Schepers, G., Hegberg, T., and Magnisson, M., "Comparison of wake models with data for offshore windfarms," *Wind Engineering*, Vol. 25, No. 5, 2001, pp. 271–280.

[7] Hwangbo, H., Johnson, A. L., and Ding, Y., "Spline Model for Wake Effect Analysis: Characteristics of Single Wake and Its Impacts on Wind Turbine Power Generation," *to appear in IISE Transactions*, 2017. doi:10.1080/24725854.2017.1370176.

[8] You, M., Byon, E., Jin, J., and Lee, G., "When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms," *IISE Transactions*, Vol. 49, No. 1, 2017, pp. 84–95.

[9] You, M., Liu, B., Byon, E., Huang, S., and Jin, J. J., "Direction-dependent Power Curve Modeling for Multiple Interacting Wind Turbines," *to appear in IEEE Transactions on power systems*, 2017. doi:10.1109/TPWRS.2017.2737529, URL `http://ieeexplore.ieee.org/abstract/document/8006285/`.

[10] Keck, R.-E., "Validation of the standalone implementation of the dynamic wake meandering model for power production," *Wind Energy*, Vol. 18, No. 9, 2015, pp. 1579–1591.

[11] Katic, I., Højstrup, J., and Jensen, N. O., "A simple model for cluster efficiency," *European Wind Energy Association*, 1986.

[12] DTU Wind Energy, "Wind resources for energy production of wind turbines," , 2017. URL `http://http://www.wasp.dk/wasp#details__wakeeffectmodel`.

[13] Peña, A., and Rathmann, O., "Atmospheric stability-dependent infinite wind-farm models and the wake-decay coefficient," *Wind Energy*, Vol. 17, No. 8, 2014, pp. 1269–1285.

[14] Peña, A., Réthoré, P.-E., and Laan, M. P., "On the application of the Jensen wake model using a turbulence-dependent wake decay coefficient: the Sexbierum case," *Wind Energy*, 2015.

[15] Barthelmie, R. J., and Pryor, S., "An overview of data for wake model evaluation in the Virtual Wakes Laboratory," *Applied energy*, Vol. 104, 2013, pp. 834–844.

[16] Renkema, D. J., "Validation of wind turbine wake models," *Master of Science Thesis, Delft University of Technology*, Vol. 19, 2007.

[17] Lee, G., Byon, E., Ntaimo, L., and Ding, Y., "Bayesian spline method for assessing extreme loads on wind turbines," *The Annals of Applied Statistics*, Vol. 7, No. 4, 2013, pp. 2034–2061.

[18] Byon, E., Choe, Y., and Yampikulsakul, N., "Adaptive learning in time-variant processes with application to wind power systems," *IEEE Transactions on Automation Science and Engineering*, Vol. 13, No. 2, 2016, pp. 997–1007.