

A Comparison of Edge Detectors in the Framework of Wake Pattern Modeling for Wind Turbines

YanJun Yan¹, James Z. Zhang², and Hayrettin Bora Karayaka³

^{1,3} *Department of Engineering and Technology, Western Carolina University, Cullowhee, NC, 28723, USA*
yyan@wcu.edu
hbkarayaka@email.wcu.edu

² *Department of Electrical and Computer Engineering, Kettering University, Flint, MI, 48504, USA*
jzhang@kettering.edu

ABSTRACT

To monitor wind turbine health, wind farm operators can take advantage of the historical SCADA (supervisory control and data acquisition) data to generate the wake pattern beforehand for each wind turbine, and then decide in real time whether observed reduction in power generation is due to wake or true faults. In our earlier efforts, we proposed an effective wake pattern modeling approach based on edge detector using Linear Prediction (LP) with entropy-thresholding, and smoothing using Empirical Mode Decomposition (EMD) on the wind speed difference plots. In this paper, we compare the LP based edge detector with two other predominant edge detectors, Sobel and Canny edge detectors, to quantitatively justify the appropriateness and effectiveness of the LP based edge detector in wind turbine wake pattern analysis. We generate a fused wake model for the turbine of interest with multiple neighboring turbines, and then analyze the wake effect on turbine power generation. With a fused wake pattern, we do not need to identify the individual source of wake any more. We expect that wakes cause reduced wind speed and hence reduced power generation, but we have also observed from the SCADA data that the wind turbines in wake zones tend to overreact when the wind speed is not yet close to the high-wind-shut-down threshold, which causes further power generation loss.

1. INTRODUCTION

With the rapid development of the wind energy industry, the modern wind farms consist of hundreds of wind turbines. The effect of wake (Burton et al., 2001) is that, when a tur-

bine extracts energy from the wind, it leaves behind a wake characterized by reduced wind speeds and increased levels of turbulence. Since wake adversely affects power generation, when a wind farm is being designed, the wind map information will be used to simulate the wake propagation for various configurations of wind turbines. The optimal locationing of the wind turbines should enable the turbines to generate the most power in some windy locations while incurring the least amount of wake. However, the co-existence of the large number of turbines complicates the air dynamics, and the practical constraints (such as property ownerships, road building capabilities to transport and build the wind turbines) restrict where the wind turbines can be built. As a result, wake effect is unavoidable at a certain wind direction for a certain turbine, although the likelihood is supposed to be minimized by proper wind farm planning. When wake happens, the power generation will be hampered. However, once the wind changes its direction and the downwind turbine is no longer in the wake region, the power generation will return to normal. Therefore, wake is not a fault of the wind turbine in itself, because the wind turbine can return to its full operational capacity as soon as it is no longer in wakes. Of course, while wake happens, the air turbulence will stress the wind turbine more than normal, which may lead to potential faults, but it is a gradual process, in a similar time scale as other wear and tear effects. When we observe reduced power generation, we need to identify the wake effect to separate such temporary power reduction from true faults.

There have been many efforts to model wake propagation across the wind farm, mostly for wind farm design purposes (Ainslie, 1988; Quarton & Ainslie, 1990; Hassan, 1992; Wu & Port-Agel, 2011; Nilsson, 2012). Such models are more accurate over an averaging of 30 degree bins than across 5 degree bins or 10 degree bins (Beaucage et al., 2012), because the wake effect in a wider angular range may cancel out to yield a better overall power prediction accuracy. Once

YanJun Yan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

the wind farms are built, the wake effects can be measured by sodar (Barthelmie et al., 2006), lidar (Lang & McKeogh, 2011), or remote sensing (Clive et al., 2012), in the entire atmosphere of the wind farm. Arthelmie et al. (Barthelmie et al., 2006) compared several state-of-the-art theoretical wake models with the sodar measurements to discover that the spread away from the wake model predictions is considerable even for relatively simple offshore single wake cases. Recently, Philippe Beaucage et al. proposed the Deep-Array Wake Model (DAWM) that agrees with the real data better than other models (such as Park (Jensen), Eddy Viscosity (Ainslie), and Computational Fluid Dynamics (CFD) models) for the offshore wind farms (Beaucage et al., 2012). For onshore applications, however, a better wake model is still of great interest in the wind energy community, particularly when the wake effect is caused by multiple sources and complex terrains other than the sea surface.

Once the wind turbines are built and the wind farm is established, the main interest in wake analysis is for operational monitoring purpose. However, the models mentioned in the previous paragraph are developed based on the physics of air dynamics, often under simplified assumptions, and they can not provide sufficiently fine resolution (Beaucage et al., 2012). Meanwhile, the models mentioned in the previous paragraph aim at estimating the entire wind farm's power generation to help the investors decide whether they want to build the wind farm or not. Once the wind farm is built, the wake models should aim at estimating the individual wind turbine's power generation to help monitor the wind turbine's health and repair it in a timely fashion, if needed. A wake model using SCADA data is discussed in this paper, which provides an operational model of wake effects for each turbine of interest on the wind farm. Using the same framework but individual data sets collected from each wind turbine, the generated wind pattern is customized for each wind turbine. This procedure is automatic and adaptive, which is desirable when handling a large number of wind turbines.

Yan et al. (Yan et al., 2009) proposed to model wake phenomena using SCADA data, where the major wake patterns were segmented using morphological imaging operators (dilation and erosion) to automate the wake pattern identification. Then on the wake patterns, a threshold based decision strategy was implemented to determine the existence of wakes. However, the morphological operators were restricted by the structuring element's shape and size, and the boundary detected by the morphological operators could be improved. Meanwhile, the threshold based approach needs *a priori* knowledge on how much the wake effect would influence the wind speed significantly, and such knowledge might not apply to different terrains. These issues were addressed in a recent paper (Yan & Zhang, 2014), where a solution of multiple steps was proposed: 1. In the wind speed difference versus wind direction scatter plot, the data points within the

wake pattern might still be unconnected, and hence a moving window based intensity map was used to connect such sparse data points. This intensity map was finer and more accurate than using morphological operators in (Yan et al., 2009). 2. The "valleys" in the wake data represented the wake region's width and severity. An edge detector was used to capture the characteristic pattern of the "valley", based on the relative depth calculated from the majority of the data, which was more adaptive than the threshold-based decision maker in (Yan et al., 2009). 3. A novel edge detector using Linear Prediction and Entropy Thresholding was used. The analysis in (Yan & Zhang, 2014) was qualitative without comprehensive comparison with other well-known edge detectors.

Therefore, the contributions of this paper include: 1. It presents a comparison of the LP edge detector with Canny edge detector and Sobel edge detector in the framework of wake pattern modeling. 2. It provides a comprehensive quantitative evaluation of all these approaches. 3. It includes discussions on the observations in wake pattern analysis and interpretations of the observations in each wake zone.

The rest of the paper is organized as follows. The previous work on wake pattern extraction is briefly explained in Section 2. The LP edge detector in this wake modeling framework is compared with other edge detectors in Section 3, where the properties of each approach are discussed. Section 4 demonstrates the wake pattern fusion of multiple neighboring turbines. Section 5 presents a metric to compare the three approaches. Based on the best wake pattern, we partition the data based on whether there is wake or not in Section 6, and interpret the data in Section 7. This paper is concluded in Section 8 with some future research areas identified.

2. WAKE REPRESENTATION AND INFORMATION EXTRACTION

In wind turbine's operation, wind speed measurement is crucial, because wind speed is the input to the wind turbine's control system, and the wind is the energy source that determines how much power can be generated.

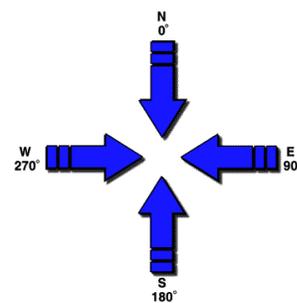


Figure 1. Compass coordinate to define the wind direction.

Wind direction is defined in the compass coordinate (US Dept of Commerce et al., 2002), as shown in Figure 1. In the wake

pattern, the x axis is the wind direction defined by Figure 1, the y axis is the normalized wind speed difference (Yan et al., 2009; Yan & Zhang, 2014) defined by

$$wsd = \frac{ws_1 - ws_2}{(ws_1 + ws_2)/2} = 2 - \frac{4}{ws_1/ws_2 + 1}, \quad (1)$$

where ws_1 is the wind speed of the 1st turbine (let say X), and ws_2 is that of the 2nd turbine (let say Y). The theoretical data range of wsd is -2 to 2 , but actual data range rarely goes beyond -1 or 1 . In wake analysis, we only care about the negative wsd value when the current turbine is in the wake of another turbine or obstruction, because the positive wsd value, when X blocks the wind to Y , is reflected in Y 's wake pattern, and hence does not matter in X 's wake analysis. The prominent features in this wake representation are the valley's width (wake angular span) and depth (wake intensity). With hundreds of turbines on the wind farm, an automated procedure of wake pattern analysis is desirable, where image processing is a handy tool to automate this procedure. The steps in this image processing procedure consist of intensity map generation, edge detection, envelope extraction, and envelope smoothing, to help estimate valley width and depth.

2.1. Intensity Map

Yan et al. proposed morphological imaging operators to segment the majority of the data (Yan et al., 2009), but the precision was limited by the shape and size of the structuring elements. To improve upon that approach and inspired by the fact that the data points were still sparse in the wake region (although much denser than the region with outliers), Yan and Zhang proposed the intensity map, a moving window based data intensity measure (Yan & Zhang, 2014).

2.2. LP Based Edge Detection Method

Linear prediction (LP) uses a linear model to predict future values of a discrete time signal using past and present values (Makhoul, 1975), while minimizing the error between the estimated value and the real value. In images, however, edges can be viewed as discontinuities in the given 2-D signal, and large errors in the estimated image represents the edge information of that image (Zhang & Punch, 2012). After applying the LP edge detection to an image, Yan and Zhang employed an entropy-based threshold to further eliminate the background information and retain the edge information (Yan & Zhang, 2014; Zhang & Punch, 2012).

2.3. Envelope Extraction and Smoothing using Empirical Mode Decomposition

Measurement of the valley width and the depth directly from an edge map is difficult given the irregular distribution of edge pixels. It is desirable to convert the edge map into a "time-series like" data sequence so that standard mathemat-

cal techniques can be applied to calculate the span of a segment as well as the extrema of the function. A lower envelope of an edge map is extracted by simply keeping the minimal intensity value at each corresponding direction of the wake. The challenge is that there may be excessive ringing effects caused by small variations in intensity around neighboring directions. Consequently, appropriate smoothing needs to be applied to the envelope to sufficiently remove ringing without significantly attenuate the amplitude. Due to its unique filtering characteristics, Yan and Zhang chose the Empirical Mode Decomposition method (ur Rehman & Mandic, 2011) (Flandrin et al., 2004) (Huang et al., 1998) for smoothing.

The basic concept of EMD is to identify proper time scales that reveals physical characteristics of the signals, and then decompose the signal into modes intrinsic to the function, which are referred to as Intrinsic Mode Functions (IMF). The following signal is used as the input to the algorithm to examine the output components for verification.

$$x(t) = \sin(2.5\pi t) + 0.1 \cos(50\pi t) + 0.8 \sin(5\pi t) \quad (2)$$

A signal defined in (2) is used as an example to show the procedure of EMD in this waking modeling framework, because this signal contains various known frequency components and it looks like a typical envelope observed in wake pattern analysis. The extracted IMFs are shown in figure 2 (a)-(c), which illustrate that the amplitude and frequency contents of the signal can be accurately extracted by EMD.

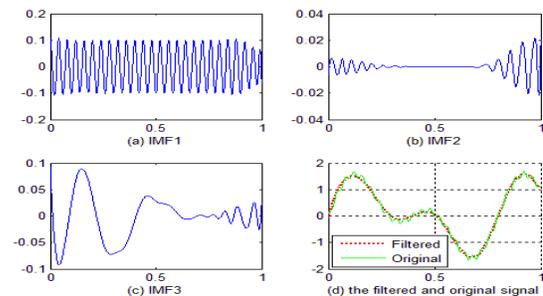


Figure 2. IMFs of $x(t)$

2.4. Wake Location and Severity Decision Maker

After the envelope of the scatter data is detected, the location and information of the "valleys" are determined. Given any point on the envelope (the boundary points are considered in a circular fashion), if the points on its left and right are higher than this point on the envelope, this point is considered a candidate of valley point. The candidates of peak points can be similarly identified. A segment of envelope with a valley point candidate in between two peak candidates is considered an oscillation. The "valleys" are oscillations that are significantly deeper than normal and are wide enough to be substantial (Yan & Zhang, 2014). Meanwhile, the "small dents" are consolidated into its neighboring "valleys" if the dents' depth is small. The selected valleys' depth indicates

the severity of the wakes, and the pixel breadth is converted into the corresponding angular width of the wakes.

3. COMPARISON WITH OTHER EDGE DETECTORS

The Canny edge detector (Canny, 1986) and Sobel edge detector (Farid & Simoncelli, 2004) are well established edge detectors, and they could be used in place of the LP edge detector. However, one needs to understand the mechanism and the purpose of the edge detector in our data processing to appreciate the differences in these edge detectors.

3.1. Sobel edge detector

The Sobel edge detector is one of the first-order differentiation based edge detectors (Belyaev, 2011), defined by

$$\Delta I_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \cdot I \quad (3)$$

$$\Delta I_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \cdot I \quad (4)$$

where I is an image, and its Sobel edge map is defined by

$$I_{edge} = \sqrt{\Delta I_x^2 + \Delta I_y^2} \quad (5)$$

3.2. Canny edge detector

The initial step in the Canny edge detector can be the Sobel edge detector (or other edge detectors), and then the pixels that are not within thin lines (regarded as not part of an edge) are removed (Canny, 1986). The Canny edge detector uses hysteresis to declare a pixel, p_d , with certain gradient value, d , to be an edge pixel ($E(p_d) = 1$) or not ($E(p_d) = 0$), based on the following rule:

$$E(p_d) = \begin{cases} 0, & \text{if } d < \tau_l \\ 1, & \text{if } \tau_l \leq d \leq \tau_u \text{ and } \exists E(\text{neighbor}(p_d)) = 1 \\ 0, & \text{if } \tau_l \leq d \leq \tau_u \text{ and } \forall E(\text{neighbor}(p_d)) = 0 \\ 1, & \text{if } d > \tau_u \end{cases} \quad (6)$$

where τ_u is an upper threshold, τ_l is a lower threshold, and a typical ratio between τ_u and τ_l is between 2 : 1 and 3 : 1. When $\tau_l \leq d \leq \tau_u$, the pixel, p_d , is declared an edge pixel only if it is connected to another edge pixel.

3.3. Results when using different edge detectors

Wake pattern analysis includes edge map generation, envelope extraction, EMD smoothing, and deep “valley” characteristic derivation. The LP edge detector can be replaced by Sobel and Canny edge detectors in this framework.

For a single turbine of interest with multiple neighboring tur-

bines, there are multiple wakes at different wind directions. Take real data as an example, turbine A is surrounded by twelve other turbines within 1000 meter radius, as shown in Figure 3, which is an extraction from a much larger wind farm.

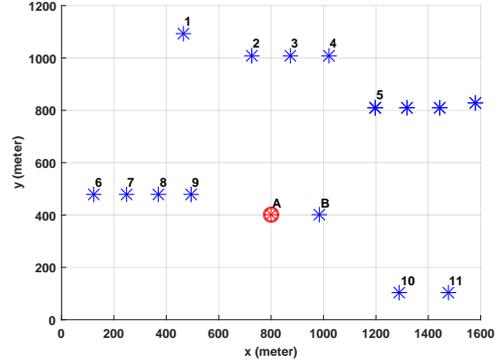


Figure 3. All the turbines surrounding turbine A within a 1000-meter radius.

The pair-wise wind speed difference versus wind direction plots are shown in Figures 4 and 5. The twelve pairs are shown in three blocks with four pairs in one block. The first row of each block shows the results from using LP edge detector, the second row using Sobel edge detector, and the third row using Canny edge detectors. Each column is for a pair of wind turbines. In each subplot, the x axis is the wind direction in degrees (from 0 to 360), and the y axis is the normalized wind speed difference from -1 to 1 as defined in (1). The magenta lines indicate the centers of the wakes.

The numerical results showing the wake angle regions and the intensity of the wake effect are listed in Table 1. The depth is a unit-less number: the closer it is to -1, the higher the wake intensity. The center, left and right angles of the wake regions are in degrees. If the detected valleys using three edge detectors are roughly around the same angles, they are aligned at the same rows for easy comparison.

As observed, Sobel edge detector renders similar results to those of the LP edge detector, but the edge extraction can be inaccurate to cause unnecessary oscillations and hence trivial wake regions, such as in Turbine pair ($A, 5$) in Figure 5. The Canny edge detector renders some edges that are so low that they include outlier data points. Canny edge detector is an optimal edge detector if the purpose is to extract objects from a natural scene (Canny, 1986), but given a scatter plot with patches of different data densities, Canny edge detector is not as effective as a gradient based edge detector, such as Sobel edge detector and our LP edge detector. Another caveat to use Sobel and Canny edge detectors is that when the envelope is spiky, the boundary condition in EMD smoothing may affect the result, as shown in the Sobel pair (A, B) in Figure 5, and Canny edge detector’s Sim IDW fusion result in Figure 6.

4. FUSION FOR A COMPREHENSIVE WAKE PATTERN

A complete understanding of the wake pattern of a central turbine demands an analysis from all of its neighbors, such as in Figures 4 and 5. The individual pairs' results are further fused by the four schemes (Yan et al., 2009):

1. Equal-weight fusion (EW), where all data scatters are combined with equal weight.
2. Inverse-distance-weight fusion (IDW), where the data scatters are combined with heavier weight on the closer neighbors.
3. Similarity clustering based EW fusion (Sim EW), where the similar clusters are fused first, and the clusters are treated equally.
4. Similarity clustering based IDW fusion (Sim IDW), where the clusters are weighted based on their distances to the centering turbine.

The numerical results showing the fused wake patterns using the proposed framework and the comparison of three edge detectors are listed in Table 2.

Figure 6 shows the fused wake pattern using the LP edge detector, the Sobel edge detector, and the Canny edge detector. The similarity-clustering based fusion preserves the wake intensity information better than the ones without clustering, and the equal-weight based fusion turns out to be more robust than the inverse-distance-weight scheme, which further

indicates that distance alone is not necessarily a stable indicator of wake intensity.

As mentioned earlier, Sobel and Canny edge maps can be spikier than LP edge map, yielding harder constraints on the boundary conditions for the EMD procedure to cause undesired smoothing results, such as in the right-most plot in the last row in Figure 6, where the filtered envelope using Canny edge detector fails to follow the edge map in certain segments. This did not happen for LP edge detector in our experiments, because LP edge detector uses entropy thresholding to effectively adjust its edge declaration threshold for each plot.

5. A METRIC TO EVALUATE THE WAKE PATTERN

To quantify the goodness of an approach to derive the wake pattern, we propose a numerical metric, besides the analysis earlier. The desired property of such a metric is that the derived wake pattern should follow the data accurately. Namely, the data points beneath the smoothed edge map should be reasonable few, with low density.

Take Figure 6 as an example: In each column of subplots, the same fused data plot is fed into the wake pattern analysis system. Depending on the choice of edge detector, the edge map is different. Different columns use different fused data. While comparing different edge detectors, we focus on each column of Figure 6.

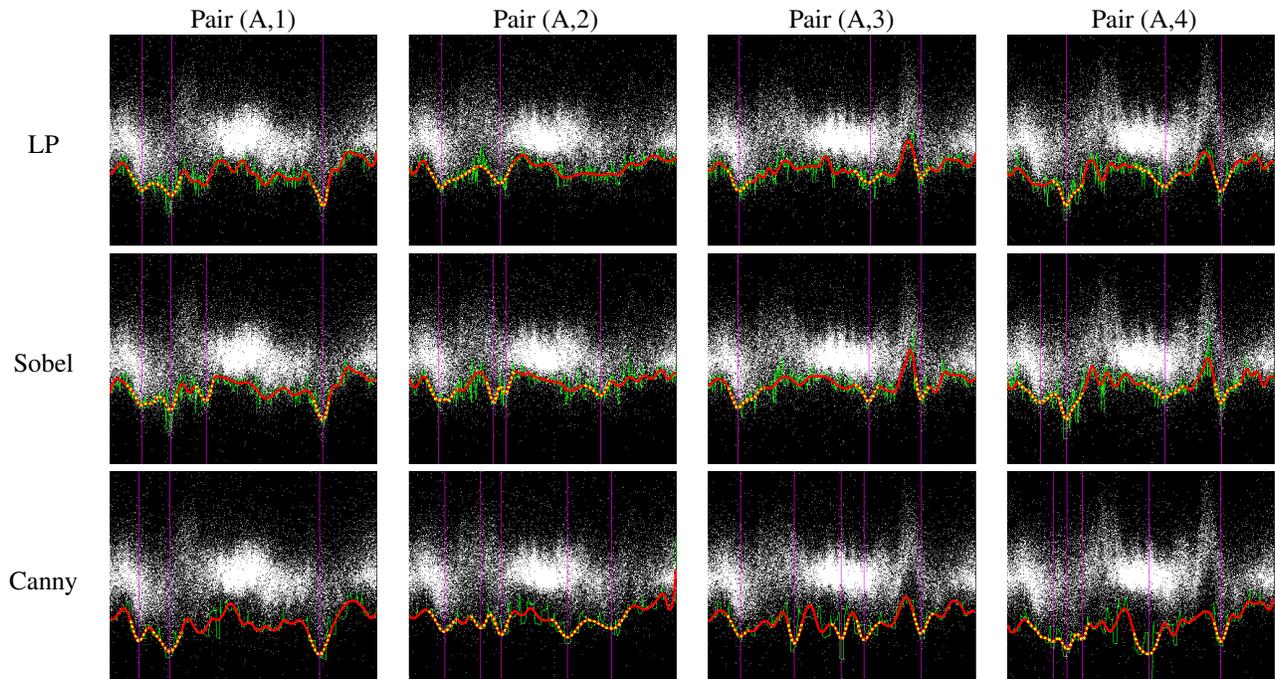


Figure 4. Results using our LP edge detector (in the first row), Sobel edge detector (in the second row), and the Canny edge detector (in the third row), between turbine *A* and its first four neighboring turbines within a 1000-meter radius. In each subplot of this figure and the next two figures, the *x* axis is the wind direction in degrees (from 0 to 360), and the *y* axis is the normalized wind speed difference from -1 to 1 as defined in (1).

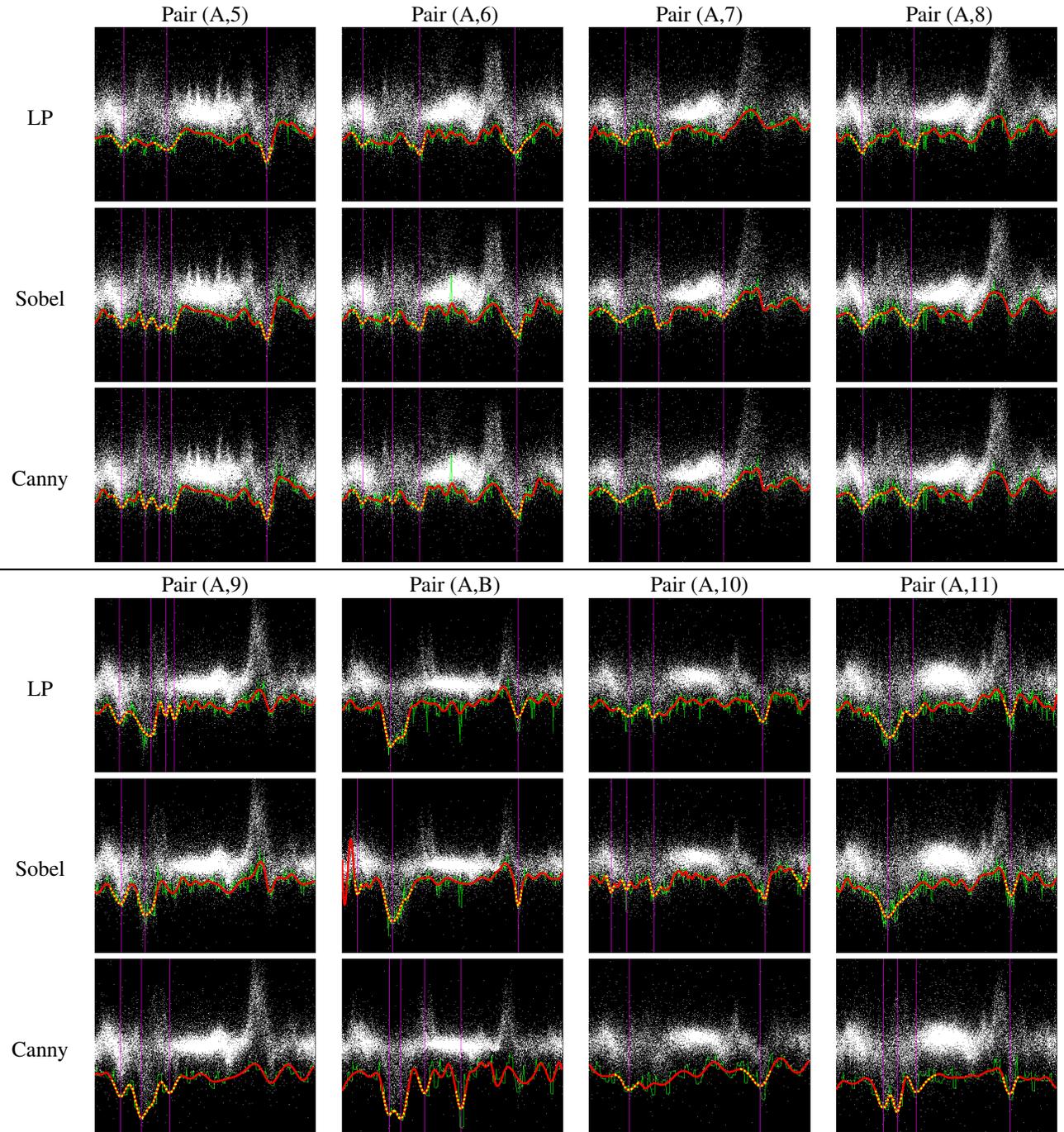


Figure 5. Results using our LP edge detector (in the first and fourth rows), Sobel edge detector (in the second and fifth rows), and the Canny edge detector (in the third and sixth rows), between turbine *A* and its last eight neighboring turbines within a 1000-meter radius.

Suppose that the data plot is a logical matrix, WSD , with M rows and N columns. The rows of WSD are partitions of the normalized wind speed difference values, and the columns are partitions of the wind directions. A logical 1 in the WSD matrix indicates the existence of data point(s) landing in that cell, and 0, otherwise. The data counting beneath the edge

map is hence

$$D_i = \sum_{j=edge_i}^M WSD_{ij}, \text{ with } i = 1, \dots, N \quad (7)$$

where the edge map is an N -dimensional row vector, with

Table 1. Numerical results of the wakes for individual turbine pairs. The depth is a unit-less number: the closer it is to -1, the higher the wake intensity. The center, left and right angles of the wake regions are in degrees. If the detected valleys using three edge detectors are roughly around the same angles, they are aligned at the same rows for easy comparison.

	LP					Sobel					Canny				
	depth	center	left	right	width	depth	center	left	right	width	depth	center	left	right	width
Pair (A.1)	-0.48	44.7	25.7	57.9	32.3	-0.45	43.9	24.0	67.9	43.9	-0.62	40.5	29.0	53.8	24.8
	-0.53	84.4	57.9	97.7	39.7	-0.49	82.8	67.9	92.7	24.8	-0.72	81.9	59.6	96.8	37.2
	-0.63	288.0	274.8	294.6	19.9	-0.40	130.8	115.9	140.7	24.8	-0.58	287.2	274.8	295.4	20.7
Pair (A.2)	-0.46	44.7	26.5	92.7	66.2	-0.40	40.5	24.8	66.2	41.4	-0.53	48.8	27.3	65.4	38.1
	-0.41	123.3	96.0	141.5	45.5	-0.41	114.2	105.1	123.3	18.2	-0.49	96.8	83.6	110.1	26.5
						-0.56	125.0	110.1	138.2	28.1	-0.58	213.5	192.0	242.5	50.5
						-0.33	130.8	123.3	142.3	19.0	-0.50	273.1	242.5	297.9	55.5
Pair (A.3)	-0.49	43.0	27.3	58.8	31.4	-0.47	41.4	28.1	77.8	49.7	-0.57	44.7	27.3	56.3	29.0
	-0.41	218.5	197.0	235.9	38.9	-0.41	216.8	198.6	230.9	32.3	-0.64	116.7	108.4	127.5	19.0
	-0.37	286.3	278.1	297.9	19.9	-0.39	286.3	279.7	297.9	18.2	-0.60	179.6	166.3	189.5	23.2
						-0.60	210.2	200.3	232.6	32.3	-0.54	287.2	276.4	303.7	27.3
Pair (A.4)	-0.62	80.3	62.1	101.8	39.7	-0.42	45.5	15.7	64.5	48.8	-0.62	62.9	28.1	70.3	42.2
	-0.45	213.5	173.0	234.2	61.2	-0.58	80.3	64.5	103.5	38.9	-0.69	81.1	70.3	95.2	24.8
	-0.49	288.8	278.1	302.9	24.8	-0.37	212.7	183.7	238.3	54.6	-0.61	101.8	95.2	111.7	16.6
						-0.42	288.8	279.7	315.3	35.6	-0.74	191.2	171.3	206.1	34.8
Pair (A.5)	-0.39	48.0	25.7	67.0	41.4	-0.38	44.7	31.4	57.9	26.5	-0.48	44.7	19.9	57.9	38.1
	-0.42	118.3	97.7	142.3	44.7	-0.39	82.8	73.7	94.3	20.7	-0.48	110.9	91.0	144.8	53.8
	-0.55	280.6	269.8	288.8	19.0	-0.40	105.9	94.3	114.2	19.9	-0.40	125.8	114.2	139.0	24.8
						-0.51	281.4	269.8	288.8	19.0	-0.60	282.2	265.7	293.0	27.3
Pair (A.6)	-0.36	34.8	16.6	48.8	32.3	-0.33	34.8	19.0	43.0	24.0	-0.63	72.0	55.5	91.9	36.4
	-0.47	126.6	102.6	135.7	33.1	-0.33	82.8	73.7	95.2	21.5	-0.57	119.2	91.9	138.2	46.3
	-0.45	282.2	261.5	310.3	48.8	-0.40	126.6	114.2	137.4	23.2	-0.75	282.2	269.8	295.4	25.7
						-0.49	285.5	267.3	295.4	28.1	-0.50	62.9	47.2	87.7	40.5
Pair (A.7)	-0.34	59.6	39.7	88.5	48.8	-0.32	53.0	18.2	91.0	72.8	-0.46	111.7	87.7	163.0	75.3
	-0.38	113.4	88.5	122.5	33.9	-0.38	114.2	105.1	122.5	17.4					
						-0.26	219.3	207.7	243.3	35.6					
Pair (A.8)	-0.45	43.0	24.0	55.5	31.4	-0.40	43.9	18.2	75.3	57.1	-0.46	6.6	1.7	16.6	14.9
	-0.38	127.5	116.7	143.2	26.5	-0.36	123.3	96.8	149.0	52.1	-0.58	31.4	23.2	43.0	19.9
											-0.51	81.1	63.7	91.0	27.3
											-0.49	124.1	101.0	137.4	36.4
											-0.49	151.4	141.5	160.6	19.0
Pair (A.9)	-0.44	40.5	24.8	56.3	31.4	-0.44	43.9	26.5	57.1	30.6	-0.50	287.2	278.9	294.6	15.7
	-0.59	91.9	62.9	101.0	38.1	-0.57	82.8	70.3	101.0	30.6	-0.59	43.0	29.0	57.9	29.0
	-0.36	115.9	106.8	122.5	15.7						-0.83	77.0	57.9	105.9	48.0
	-0.40	129.9	122.5	141.5	19.0						-0.55	123.3	109.2	143.2	33.9
Pair (A.B)	-0.71	79.5	67.0	109.2	42.2	-0.33	25.7	21.5	33.1	11.6	-0.79	77.8	67.0	85.2	18.2
						-0.64	82.8	66.2	115.0	48.8	-0.84	96.0	85.2	107.6	22.3
	-0.37	287.2	275.6	302.1	26.5	-0.46	288.0	282.2	293.0	10.8	-0.55	134.9	124.1	145.7	21.5
Pair (A.10)	-0.36	67.0	48.0	92.7	44.7	-0.31	37.2	28.1	43.0	14.9	-0.73	194.5	186.2	201.1	14.9
	-0.38	105.9	92.7	115.0	22.3	-0.28	62.1	53.8	68.7	14.9	-0.51	66.2	43.0	82.8	39.7
	-0.42	283.0	265.7	293.8	28.1	-0.34	106.8	97.7	119.2	21.5	-0.46	278.9	246.6	297.1	50.5
						-0.39	286.3	276.4	293.0	16.6					
Pair (A.11)	-0.60	88.5	54.6	111.7	57.1	-0.26	350.1	331.0	357.5	26.5	-0.71	77.8	64.5	89.4	24.8
	-0.36	125.8	114.2	145.7	31.4	-0.59	84.4	55.5	122.5	67.0	-0.77	100.1	89.4	108.4	19.0
	-0.36	284.7	273.1	294.6	21.5	-0.35	285.5	275.6	295.4	19.9	-0.53	131.6	115.9	163.0	47.2

each element, $edge_i$, indicating the row number of the wake pattern's edge point at that particular column i in WSD . WSD_i is the i th column of WSD .

Then we can examine the mean and variance of the vector D of each approach to see which approach can achieve both a smaller mean and a smaller variance of its D value.

In our particular partitions of WSD , we have $M = 343$, and $N = 435$, and hence the resultant mean and variance of D are in a particular data range in Table 3. If the partitions of WSD are of different dimensions from our choices, the absolute values of such mean and variance will be different, but the relative relationships between different approaches will be similar.

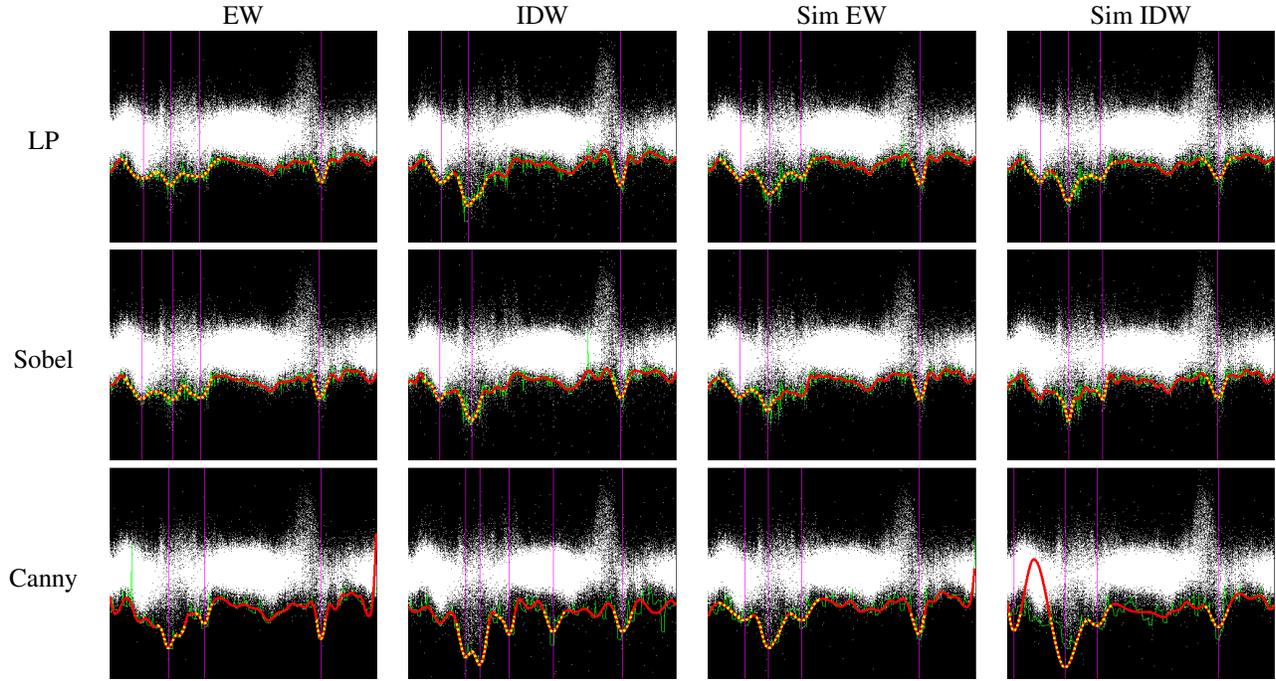


Figure 6. Four fusion schemes using our LP edge detector (in the first row), Sobel edge detector (in the second row), and the Canny edge detector (in the third row) to generate the overall wake pattern of turbine A .

Table 2. Numerical results of the fused wake pattern (EW: equal weight, IDW: inverse distance weight, Sim: similarity). The depth is a unit-less number: the closer it is to -1, the higher the wake intensity. The center, left and right angles of the wake regions are in degrees. If the detected valleys using three edge detectors are roughly around the same angles, they are aligned at the same rows for easy comparison.

	LP					Sobel					Canny				
	depth	center	left	right	width	depth	center	left	right	width	depth	center	left	right	width
EW	-0.40	46.3	16.6	64.5	48.0	-0.43	43.9	21.5	62.1	40.5	-0.71	80.3	65.4	102.6	37.2
	-0.46	82.8	64.5	99.3	34.8	-0.43	86.1	62.1	100.1	38.1	-0.48	128.3	116.7	139.0	22.3
	-0.38	121.7	99.3	152.3	53.0	-0.41	123.3	100.1	139.9	39.7	-0.48	128.3	116.7	139.0	22.3
	-0.44	285.5	269.8	296.3	26.5	-0.41	282.2	273.9	297.1	23.2	-0.62	284.7	278.9	291.3	12.4
IDW	-0.44	45.5	24.0	62.1	38.1	-0.43	43.0	22.3	52.1	29.8	-0.79	77.8	65.4	86.1	20.7
	-0.66	81.1	67.0	107.6	40.5	-0.62	86.1	68.7	101.0	32.3	-0.87	96.8	86.1	108.4	22.3
											-0.58	135.7	117.5	143.2	25.7
											-0.56	195.3	182.9	211.9	29.0
	-0.46	285.5	274.8	293.8	19.0	-0.41	285.5	273.9	297.1	23.2	-0.55	288.0	276.4	301.2	24.8
Sim EW	-0.41	44.7	16.6	58.8	42.2	-0.41	43.9	14.1	59.6	45.5	-0.44	50.5	15.7	62.9	47.2
	-0.55	83.6	63.7	110.9	47.2	-0.53	81.1	59.6	91.9	32.3	-0.71	81.9	66.2	106.8	40.5
	-0.39	125.8	112.6	144.8	32.3						-0.45	126.6	111.7	151.4	39.7
	-0.45	285.5	273.9	293.8	19.9	-0.43	284.7	274.8	294.6	19.9	-0.57	283.9	273.1	291.3	18.2
											-0.54	9.9	3.3	19.0	15.7
Sim IDW	-0.44	45.5	21.5	59.6	38.1	-0.62	83.6	67.0	92.7	25.7	-0.89	78.6	58.8	110.1	51.3
	-0.61	82.8	66.2	113.4	47.2	-0.42	129.1	120.0	135.7	15.7	-0.50	121.7	110.1	141.5	31.4
	-0.40	125.8	113.4	147.3	33.9										
	-0.40	284.7	261.5	307.0	45.5	-0.41	283.9	269.0	297.9	29.0	-0.51	284.7	267.3	298.8	31.4

As noted earlier, each column in Table 3 is corresponding to a different fusion scheme with different data plot, so when we compare the three approaches, we compare them only within each column. Table 3 shows that the LP approach achieves a small mean and a small variance in all the columns (In some columns, the Canny approach can achieve a smaller mean, but with a much larger variance, due to the overshooting of its edge maps, which is not desirable). Overall, the LP approach

is the best among the three.

Then comparing the four columns in Figure 6, we see that the fourth column, the fused plot based on similarity based inverse distance weighting, preserves the wake intensity the best, and hence preferred.

The same metric can be applied to individual wind speed difference plots in Figures 4 and 5, but omitted to save space.

Table 3. The mean and variance comparison of metric D (a good approach achieves both small μ and small σ). Each column is corresponding to a different fused data set, so the comparison is within each column.

μ	EW	IDW	Sim EW	Sim IDW
LP	1.7678	2.1195	1.8391	1.9011
Sobel	2.2552	3.2069	2.2161	2.4483
Canny	1.292	0.7839	1.2552	5.308
σ	EW	IDW	Sim EW	Sim IDW
LP	3.3354	3.7368	3.3105	3.4626
Sobel	4.8172	7.123	3.8933	4.5751
Canny	47.6035	15.308	33.1721	283.3565

6. DATA PARTITION BASED ON THE WAKE PATTERN

The relationship between wind speed and power is most widely presented by a power curve, with wind speed along the x -axis and the generated power along the y -axis, as illustrated in Figure 7. When the wind speed is lower than the wind turbine start-up threshold, $T1$, the power that would be needed to keep the turbine running exceeds the generated power and hence the turbine is often braked to be not in operation. When the wind speed is higher than $T1$ but less than the constant-power threshold, $T2$, the generated power increases with an increasing wind speed. However, when the wind speed exceeds $T2$, the wind turbine starts to generate constant power at its maximum capacity, with a nearly constant rotating speed. If the wind speed gets even higher, to be above the shut-down threshold, denoted by $T3$, the wind turbine is often shut down because such high wind speed may cause the wind turbine to over-rotate or over-heat to damage itself. Given different manufacturers of wind turbines, the values of the wind speed thresholds may differ, but the general idea and control logic are similar. In this paper, the generated power is normalized to be 1 as its maximum value, and the wind speed is presented in proportion to the thresholds of $T1$, $T2$, and $T3$. The nominal power curve provided by the manufacturer provides a baseline of the desired power generation performance. The nominal power curve is typically a piece-wise connected line, and the middle region between $T1$ and $T2$ is nonlinear but monotonically increasing. The actual power curve data, however, is not so regular, with variations above or below the nominal power curve, as shown in Figure 7.

The best wake pattern obtained by the LP edge detector and the fourth fusion scheme is taken out from Table 2 to form Table 4, with a new last column showing the percentage of the data within each wake zone relative to the entire data set we have analyzed. As reported in Table 4, there are four wake zones of turbine A . To illustrate the wake effect on power generation, we take an arbitrary segment of data (the wind speed and the generated power of turbine A) in various scenarios: when the wind direction is not in the wake zones or in

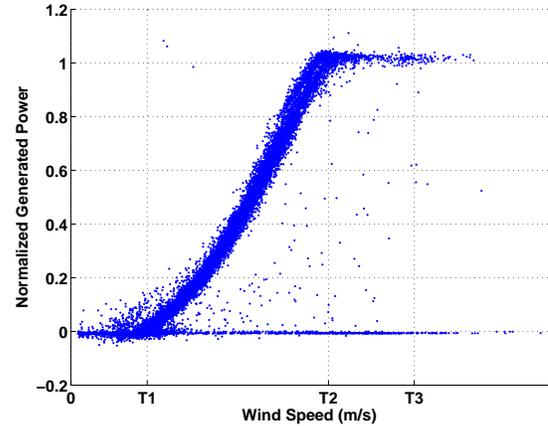


Figure 7. Real data power curve of turbine A .

either of the four wake zones.

Table 4. The best fused wake pattern of turbine A .

wake zone	depth	center	left	right	percentage
1	-0.44	45.5	21.5	59.6	12.2%
2	-0.61	82.8	66.2	113.4	7.7%
3	-0.40	125.8	113.4	147.3	7.6%
4	-0.40	284.7	261.5	307.0	4.5%

If the wind farm is thoughtfully designed with optimal wind turbine locationing, the wake effects should be minimized along the most predominant wind directions in that region. In our two and a half years of data, there are about 68% data points not in the wake zones, 12.2% in wake zone 1, 7.7% in wake zone 2, 7.6% in wake zone 3, and 4.5% in wake zone 4. This data proportion roughly indicates the likelihood when wakes may happen (roughly 32% of the time), which is not negligible but not significant, either. To present these time series in a legible way, an arbitrary segment of 800 consecutive data points is chosen from each scenario. The other segments are visually examined and they show similar trends between wind speed and power.

Using the wind speed thresholds as references, we present the wind speed measurements and generated power from the SCADA data as two time-series in the same plot. The segment not in the wake zone is presented in Figure 8 (a). The data segments from each of the four wake zones are presented in Figure 8 (b)-(e). The left axis shows the wind speed in blue dashed line, and the right axis shows the normalized power in solid color line. We align the data pair as follows:

- Power value of 0 is aligned with wind speed threshold $T1$, since when the wind speed is below $T1$, the wind turbine is braked and does not generate power.
- Power value of 1 is aligned with wind speed threshold

T_2 , since the wind turbine will produce constant power when the wind speed is above T_2 (and below T_3).

- T_3 is not aligned with any power value, but when the wind speed is above T_3 , we expect the power to be 0, since the wind turbine would have been shut down at such high wind speed to avoid damaging itself.

Figure 8 (a) clearly demonstrates an agreement between the data and the desired control logic as explained above. Figure 8 (b)-(e) indicate a similar agreement, but with exceptions: When the wind speed is between T_2 and T_3 , sometimes the generated power is 1, but other times the wind turbine is shut down with a generated power of 0, such as in wake zone 1 between time index 520 to 550, in wake zone 2 between time index 275 to 293, and in wake zone 3 between time index 278 and 329. These exceptions warrants more research into the wind turbine behaviors in the wake zones.

7. INTERPRETATION OF THE WAKE PATTERN

The wake pattern in Table 4 is replotted in Figure 9 without data points. We want to connect the data with the specific wind farm that we are studying, and hence we overlay the wake pattern with the farm layout, which requires a polar representation of the wake pattern. If we were to convert Figure 9 to a polar format directly, the “base DC term”, or the common variation in the wind speed difference, up to about -0.2 , will “overshadow” the variations due to wake effect in the polar plot, so we first subtract this DC term, and then scale it up to overlay with the turbine layout, to generate Figure 10. Figure 10 is an exaggerated wake pattern in the sense that the relative size of the rose pedal in the polar plot does not indicate the severity of the wake in proportion, but in relative sense. The angular range of each wake zone, whose numerical values are already presented in Table 2, is also marked by magenta dashed lines (wake zones 2 and 3 share one boundary).

As observed from Figure 10, the rose pedals indicate the directions from which wind comes towards turbine A that causes turbine A to see less wind speed. Along northeast, east, southeast, and west directions, turbine A sees significant wakes due to those neighboring wind turbine clusters. On the other hand, at this wind farm, the north and south winds are most predominant, and this wake pattern confirms that turbine A is least affected by those wind directions in terms of wakes. It verifies that the locations of the wind turbines are properly designed.

Wakes are influenced by distance, as turbine A 's nearest neighbor, turbine B , causes the most significant wakes to turbine A . However, wakes are affected by more than just the distance factor. The wake phenomenon involves various aspects of air flow dynamics, and the northern wind turbine cluster is observed to not affect turbine A as much as other neighboring turbines.

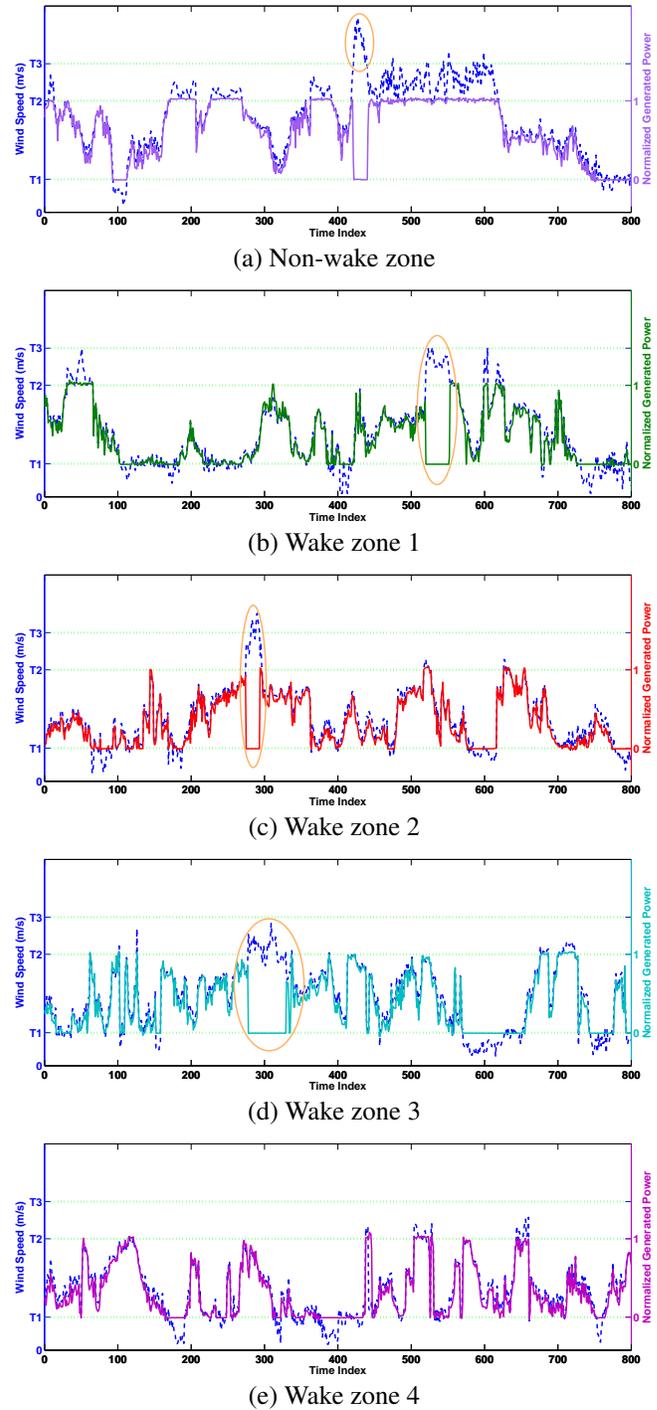


Figure 8. Data segments from non-wake zone (a) and four wake zones (b-e) of turbine A . Without wakes, the turbine is shut down when the wind speed exceeds threshold T_3 . With wakes, however, (as observed in three out of four wake zones), the turbine is shut down at relatively low wind speeds.

When the wind blows from the west, significant wakes are observed in the pair-wise comparison between turbine A and turbines 1, 5, and 6. It does not imply that turbines 1, 5, and

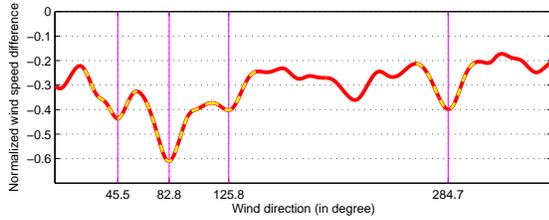


Figure 9. Extracted wake pattern of turbine A.

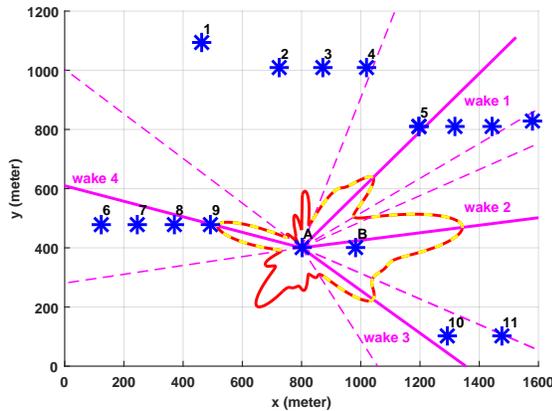


Figure 10. Exaggerated wake pattern overlaid with turbine layout. Note that the rose pedal length indicates relative severity of wakes, but not in proportion, since the DC term in Figure 9 is removed.

6 are the causes of the wakes seen by turbine A. Actually, the whole cluster (of turbines 6, 7, 8 and 9) is likely to be the cause, but not just a single turbine 6. However, turbines 7, 8 and 9 are also in the wakes of turbine 6, and hence the wind speed differences between turbine pair A and 7, etc, are not as significant as that between turbine pair A and 6. Wake analysis only indicates that the turbine of interest is in wake, but does not imply which turbine(s) is/are the cause of the wake.

Note that turbines 1, 5, and 6 are all located at the west-most position in turbine chain clusters, and hence when the wind is blowing from the west, they see the highest wind speed.

Similarly, when the wind blows from the east, significant wakes are observed in the pair-wise comparison between turbine A and turbines 4, 9, B, and 11. Note that turbines 4, 9, B, and 11 are all located at the east-most position in turbine chain clusters, and hence when the wind is blowing from the east, they see the highest wind speed.

Back to Figure 8, the actual power generation in the wake zones suggests that when the wind speed is above T_2 but below T_3 , sometimes turbine A over-reacts and decides to shut down, otherwise it could have generated power at its full capacity. Meanwhile, we need to point out that the wakes cause non-fault power generation loss, and the wake effects are expected to be minimized via optimal wind turbine locationing.

However, not all wind farms are designed in this desired manner, and wakes can cause significant loss in power generation.

8. CONCLUSIONS AND FUTURE WORK

In an optimal wind farm design, the wake effects should be minimized, but never eliminated, because one can not control the wind direction and wake is bound to happen at a certain wind direction, when multiple turbines or obstacles exist.

For operational monitoring purpose, when reduced power generation is observed, it is desirable to generate a wake model for each wind turbine on the farm to exclude wake from true faults. The SCADA data is used to represent the 2D data (normalized wind speed difference versus wind direction), similar to an image, and then one can use an edge detector to discern the pattern in the data to capture the wake effect. In this paper, the Linear Prediction (LP) and Entropy Thresholding based edge detector are thoroughly compared with popularly used Sobel and Canny edge detectors to extract the wake pattern, and both qualitative and quantitative analysis illustrate that the LP edge detector is more accurate and more robust than the other edge detectors. The Canny edge detector is optimal to detect edges in natural images, however, since the wake data representation is a scatter plot, without natural edges, Canny edge detector tends to retain the unnecessary edge points. Also, the entropy thresholding is effective for the LP edge detector to adjust its edge pixel declaration, but the Sobel or Canny edge detectors may yield spiky envelopes, which cause difficulty in setting the boundary condition in the EMD smoothing.

The wake zones identified from the wake pattern are further used to segment the power data. Power analysis is carried out in non-wake zone and in each of the four wake zones of this particular turbine of interest. With a reduced wind speed in the wake zones, the power is less than what would have been generated, which is intuitive. But from the power data, we can see that this wind turbine still generates the amount of power that it is expected to generate at that reduced wind speed, so it is healthy. An unexpected observation is that when the measured wind speed is between the constant power threshold T_2 and the shut-down threshold T_3 , the wind turbine is expected to generate constant power at its maximum capacity, but sometimes it is shut down instead, which causes further power generation loss.

In the future, since wake effect is closely related to turbulence and turbine height, we plan to partition the data by turbulence or turbine heights to provide more detailed information and more accurate prediction on the wake pattern.

ACKNOWLEDGMENT

The authors thank AWS-Truepower for providing the SCADA data. This work was supported by Western Carolina Univer-

sity Faculty Research and Creative Activities Award 2014-2015, and the Center for Rapid Product Realization and the Kimmel School Award 2014-2016.

REFERENCES

- Ainslie, J. F. (1988). Calculating the flowfield in the wake of wind turbines. *Journal of Wind Engineering and Industrial Aerodynamics*, 27(1-3), 213-224.
- Barthelmie, R. J., Larsen, G. C., Frandsen, S. T., Folkerts, L., Rados, K., Pryor, S. C., ... Schepers, G. (2006). Comparison of wake model simulations with offshore wind turbine wake profiles measured by sodar. *Journal of Atmospheric and Oceanic Technology*, 23(7), 888-901.
- Beaucage, P., Brower, M., Robinson, N., & Alonge, C. (2012). Overview of six commercial and research wake models for large offshore wind farms. In *EWEA 2012*.
- Belyaev, A. (2011). On implicit image derivatives and their applications. In *Proceedings of the british machine vision conference* (pp. 72.1-72.12). BMVA Press.
- Burton, T., Sharpe, D., Jenkins, N., & Bossanyi, E. (Eds.). (2001). *Wind energy handbook*. Wiley.
- Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8*(6), 679-698.
- Clive, P. J. M., Dinwoodie, I., & Quail, F. (2012). Direct measurement of wind turbine wakes using remote sensing. *SgurrEnergy Ltd Technical Report*.
- Farid, H., & Simoncelli, E. (2004). Differentiation of discrete multidimensional signals. *Image Processing, IEEE Transactions on*, 13(4), 496-508.
- Flandrin, P., Rilling, G., & Goncalves, P. (2004). Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2), 112-114.
- Hassan, U. (1992). A wind tunnel investigation of the wake structure within small wind turbine farms. *Department of Energy in ETSU, UK*.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. Roy. Soc. Lond. A*, 903-1005.
- Lang, S., & McKeogh, E. (2011). Lidar and sodar measurements of wind speed and direction in upland terrain for wind energy purposes. *Remote Sensing*, 3, 1871-1901.
- Makhoul, J. (1975). Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63(4), 561-580.
- Nilsson, K. (2012). Numerical computations of wind turbine wakes and wake interaction: Optimization and control. *Trita-MEK, KTH, Mechanics*(2012:18), vi, 56.
- Quarton, D., & Ainslie, J. (1990). Turbulence in wind turbine wakes. *Wind Engineering*, 14(1).
- ur Rehman, N., & Mandic, D. P. (2011). Filter bank property of multivariate empirical mode decomposition. *Signal Processing, IEEE Transactions on*, 59(5), 2421 - 2426.
- US Dept of Commerce, National Oceanic and Atmospheric Administration, & National Weather Service. (2002). <http://www.ndbc.noaa.gov/educate/seabreeze.shtml>. *NDBC Science Education Pages*.
- Wu, Y.-T., & Port-Agel, F. (2011). Large-eddy simulation of wind-turbine wakes: Evaluation of turbine parametrizations. *Boundary-Layer Meteorology*, 138(3), 345-366.
- Yan, Y., Kamath, G., Osadciw, L. A., Benson, G., Legac, P., Johnson, P., & White, E. (2009). Fusion for modeling wake effects on wind turbines. In *Proceedings of 12th international conference on information fusion*. Seattle, Washington, USA.
- Yan, Y., & Zhang, J. (2014). Using edge-detector to model wake effects on wind turbines. In *IEEE international conference on prognostics and health management 2014*.
- Zhang, J. Z., & Punch, A. J. (2012). Pattern extraction of sonar images using an lpc edge detector with entropy thresholding. *OCEANS, 2012*, 1-8.

BIOGRAPHIES

Yan Jun Yan received her B.S. and M.S. degrees in Electrical Engineering from Harbin Institute of Technology (China), and the M.S. degree in Applied Statistics and the Ph.D. degree in Electrical Engineering from Syracuse University. She is an assistant professor in engineering and technology at Western Carolina University. Her research interests are statistical signal processing, diagnostics, and particle swarm optimization.

James Z. Zhang received his B.S.E.E. degree from Hunan University (China), the M.A. degree in Telecommunications from Indiana University-Bloomington, and the M.S.E. and the Ph.D. degrees from Purdue University-West Lafayette, both in Electrical Engineering. He is a professor in electrical and computer engineering at Kettering University. His current research interests include signal processing techniques and their applications, wireless communication systems, signal design and modulation schemes for wireless communication.

H. Bora Karayaka received his B.S. and M.S. degrees in Control and Computer Engineering from Istanbul Technical University (Istanbul, Turkey), in 1987 and 1990 respectively, and his Ph.D. degree in Electrical Engineering from The Ohio State University (Columbus, OH, USA), in 2000. He is currently an Assistant Professor in the Department of Engineering and Technology, Western Carolina University. His research interests are power engineering, ocean wave energy, identification, modeling and control for electrical machines and smart grid.