WIND FLOW MODELING UNCERTAINTY

Quantification and Application to Monitoring Strategies and Project Design

Michael C Brower, Chief Technical Officer

Nicholas M Robinson, Director of openWind

Erik Hale, Assistant Meteorology Manager

May 2010

ABSTRACT

The uncertainty of wind resource and energy production estimates is a critical element in wind project financing. Wind flow modeling uncertainty is an especially important contributor to the total uncertainty, though one that is rarely quantified rigorously. In addition, the variation of the wind flow modeling uncertainty across a site is often overlooked and, as a consequence, goes unaccounted for in the process of designing a wind project. The result is that the energy production estimate may not represent a true P50, and the uncertainty may not be minimized, leading to unnecessary risk for the project owner, investor, or lender.

This report presents a conceptual framework for understanding wind flow modeling uncertainty and illustrates some applications of this framework in the *openWind®* wind project design software. The uncertainty model is derived from an analysis of observed wind flow modeling errors for sites spanning a range of topographic and meteorological conditions. Our research shows that, with the appropriate model, it is possible to (a) reduce the uncertainty in energy production compared to current methods, (b) optimize wind project layouts to maximize the PXX production (where XX is any confidence threshold such as 75%, 90% or 95%), and (c) design monitoring campaigns to minimize the wind flow modeling uncertainty within a buildable area.

INTRODUCTION

The financial models on which wind project investments are based depend on a solid understanding of the uncertainty, or risk, in the underlying wind resource estimates. An important source of uncertainty - especially for large projects involving numerous wind turbines - is related to numerical wind flow modeling.

The main role of wind flow modeling is to estimate the wind resource at every proposed or potential wind turbine location so that the wind plant's overall production can be calculated and its design optimized. The modeling is constrained by wind resource measurements at one or more meteorological towers, but where there are no measurements, errors can occur. At large, complex sites, wind flow modeling uncertainty is a significant fraction - typically one-fourth or more - of the total uncertainty in the estimated energy production.

Despite the importance of wind flow modeling uncertainty, the wind industry currently lacks a rigorous analytical framework for accounting for this factor in project design and assessment. The most common practice is to assign a single uncertainty value to the modeling uncertainty, either for a group of turbines associated with a particular mast or for the project as a whole. Typically this value ranges from 3% to 10% of the mean wind speed, depending on the complexity of the terrain and the distance to the mast.

Usually, these estimates are not based on a thorough comparison of measurements and model predictions at sites representing a range of conditions. Equally important, they do not consider the variation in uncertainty with position across the project area, which can lead to placing turbines at locations where the resource seems attractive but is actually highly uncertain. These gaps can produce sizable disparities between expected and actual production once the plants are built.

The purpose of this research was to develop and validate an analytical framework that would permit wind flow modeling uncertainty to be taken into account in a rigorous, quantitative way for both wind monitoring campaign design and wind project design.

CONCEPTUAL FRAMEWORK

Most wind resource analysts recognize intuitively that wind flow modeling uncertainty is likely to vary across a project area. It is commonly assumed, for example, that modeling errors are smaller close to a meteorological mast than farther away; that is why project design guidelines call for placing turbines no farther than a certain distance, such as 1 or 2 km, from the nearest mast. But can this variation be quantified? And is distance the only, or even the most important, factor?

At the outset of this study, we hypothesized that wind flow modeling uncertainty is driven not so much by the distance over which the model must extrapolate from measurements, but by the degree to which wind conditions are likely to differ between points. Three examples help illustrate the point.

AWS Truepower®

1. In nearly flat, featureless terrain, errors related to wind flow modeling are likely to be small and largely independent of position. An example is a wind project in Iowa, where a single mast can provide a reasonably accurate resource estimate for the entire project, with only small adjustments for minor variations in terrain.
2. In steep, mountainous terrain, wind flow modeling errors are likely to grow rapidly with distance away from a mast, especially in directions of steep slope. An example is a mast located on top of a uniform, steep ridgeline. Predictions for turbines that are even a short distance off the ridgeline have greater uncertainty than predictions at other points along the ridge top.
3. In a coastal region, errors are likely to depend on distance from the coast because of the influence of differences in the thermal properties and surface roughness of land and water. One would not want to rely on a mast that is located well inland to predict the wind resource at the coast or offshore.

As these examples suggest, how well a model is likely to perform appears to be linked to the amount of underlying variation in the wind resource. The more two points differ in their wind characteristics, the more difficult it is for the model to accurately predict the resource at one point based on measurements at the other.

If this conceptual framework is correct, then the challenge is to devise a quantitative measure of wind resource differences between points, and to link that measure to the wind flow modeling uncertainty. To be useful, such a parameter should have at least the following general characteristics. It should reduce to zero when the two points are the same. It should be non-negative, as it represents the width of an error distribution. Last, it should be bounded by some maximum value representing the global range of potential wind modeling errors.

In addition, the parameter should capture some key feature of the wind climate representing, in this context, what we mean when we say that two points experience "similar" or "different" wind conditions. Although there are many possibilities, two such features are certainly the directional mean speeds and the directional frequency distributions. Variations in speed are often linked to topographic effects, such as acceleration over a ridge, as well as changes in land cover. Changes in direction may also be caused by topography, such as when a flow is channeled through a mountain gap. In addition, both speed and direction can be altered by mesoscale effects independent of terrain, such as the land-water influences mentioned before.

Based on these considerations, we proposed two possible measures of wind resource variation, one based on the predicted directional speed ratios between points, and the other on the predicted differences in directional frequencies between points. The speed deviation formula is,

$$SD = \left[\sum_{i=1}^{ND} f_i^{(r)} \left(\frac{v_i^{(t)}}{v_i^{(r)}} - 1 \right)^2 \right]^{1/2}$$

(Eq. 1)

The direction deviation formula is,

$$DD = \left[\sum_{i=1}^{ND} \left(f_i^{(t)} - f_i^{(r)} \right)^2 \right]^{1/2}$$

(Eq. 2)

The sums are over ND wind directions, where ND is typically 12 or 16. f_i refers to the frequency of occurrence (scaled between zero and one) of direction i, and v_i is the mean speed for that direction. The superscripts t and r refer to the target and reference points, respectively. The reference can be thought of as a mast, while the target is a proposed turbine location. However, the formulas can be applied to any two points.

It is noteworthy that the speed and direction deviations in these formulas are derived entirely from the wind flow modeling, without explicit reference to the terrain, land cover, or other characteristics of the region. This conveniently eliminates the need to run other models or perform other analyses to estimate the uncertainty. Whether the approach works well depends on whether the wind flow model is sensitive to the main factors that are likely to affect modeling errors. For example, a simple model that cannot simulate gap flows or mountain-valley circulations will not only make poor predictions in and around these features, it will also underestimate of the uncertainty. The more complete the model, the better the results are likely to be.

AWS Truepower®

EXPERIMENTAL DESIGN

We set out to validate this conceptual framework through a statistical analysis of data from a large number of tall towers at diverse wind project sites. The main objectives were (a) to determine if there is a statistically significant relationship between SD and DD and wind flow modeling uncertainty, and (b) to devise a suitable functional form to represent that relationship. Along the way, we wanted to test distance as a predictor of error and compare its performance against the other parameters.

We identified 74 masts at seven project sites around the United States. At least one year of validated wind measurements were available for each mast. The data were corrected to the long-term wind climate through a standard MCP process, and tabular (TAB) files summarizing the frequencies in different speed and direction bins were created. The sites span a range of conditions, including steep mountainous terrain in the Northeast, flat-topped mesas in Texas, and broad mountain gaps in the West. For each project site we created a wind resource grid (WRG) using the SiteWind system. SiteWind employs a combination of mesoscale modeling at 1.2 km resolution and microscale modeling at 50 m resolution for a representative sample of historical days. The modeling is done initially without reference to on-site measurements; this produces what we call a raw WRG. The raw WRG files contain, for each grid point and for each of 12 wind directions, the estimated frequency of occurrence and Weibull scale and shape factors (A and k) of the speed distribution.

A software program was written to read the raw WRG file and TAB files for each site and then perform a number of calculations. Starting with an arbitrary mast (the reference), the program predicts the wind resource for each of the other masts (the targets) at the same project site. This is done, in the usual way, by calculating a speed-up ratio for each direction from the raw WRG file, applying it to the observed mean speed for that direction from the reference TAB file, and then taking the sum of the predicted speeds over all directions weighted by the observed reference frequencies. Finally, the resulting mean speeds are compared with the observed means to derive the bias, or error, for each target mast. The process is then repeated using each of the other masts as the reference in turn.

The result was a set of predicted mean speeds and errors for 990 mast pairs in all. Arguably, since a given pair of towers is represented twice in the data set (once as reference-target and once as target-reference), the number of truly independent samples is only half as large. However, using one mast as the reference and the other as the target is not exactly the same as doing the reverse, especially if the directional frequencies differ between them. Thus, it was decided to retain the duplicate pairs of masts in the data set both to maximize the sample size and to avoid inadvertently biasing the results for any given pair through the choice of reference mast.

In the next stage of the process, a table of the observed and predicted speeds and errors, along with distance, SD, and DD, was created in an Excel workbook. It was decided to bin the data by direction, SD, and DD, with four bins for each parameter. Bins with less than 10 samples (mast pairs) were eliminated. For those bins with ten or more samples, the mean error and standard deviation of errors were calculated, along with the mean values of distance, SD, and DD. The mean error was virtually zero, which is not surprising given the design of the experiment. For a sufficient sample size, the standard deviation of errors represents the uncertainty in the modeling predictions for the given combination of distance, DD, and SD.

Through a multiple regression analysis, it was then determined that distance is not a significant independent predictor of the standard deviation of errors. Instead, it appears to operate primarily through the other two parameters, both of which are significantly correlated with it. Consequently, the data were re-binned by SD and DD alone.

The scatter plots in Figure 1 show the relationships between the observed modeling uncertainty and SD and DD. Overlaid on each plot is a fitted line of the form $y = E\left(1 - B\dfrac{A}{x+A}\right)$, where y is the uncertainty, x is the speed or direction deviation, and A, B, and E are constants of the fit. Although there is some scatter, the curves fit the data quite well, with an r^2 coefficient of 0.81 in each case. Other functional forms, including logarithmic and exponential forms, also fit the data well. The virtue of the form we chose is that it behaves well near zero and for very large values of x.

AWS Truepower®

A multiple regression analysis demonstrated that both SD and DD are independently significant predictors of the error. Thus it was decided to create a combined equation using both parameters. Since there was insufficient data for a reliable combined fit, we weighted each deviation parameter equally. After some experimentation, we arrived at the following equation:

$$\sigma_{M,r} = E\left[1 - B\frac{A}{DD + A} - D\frac{C}{SD + C}\right]$$ (Eq. 3)
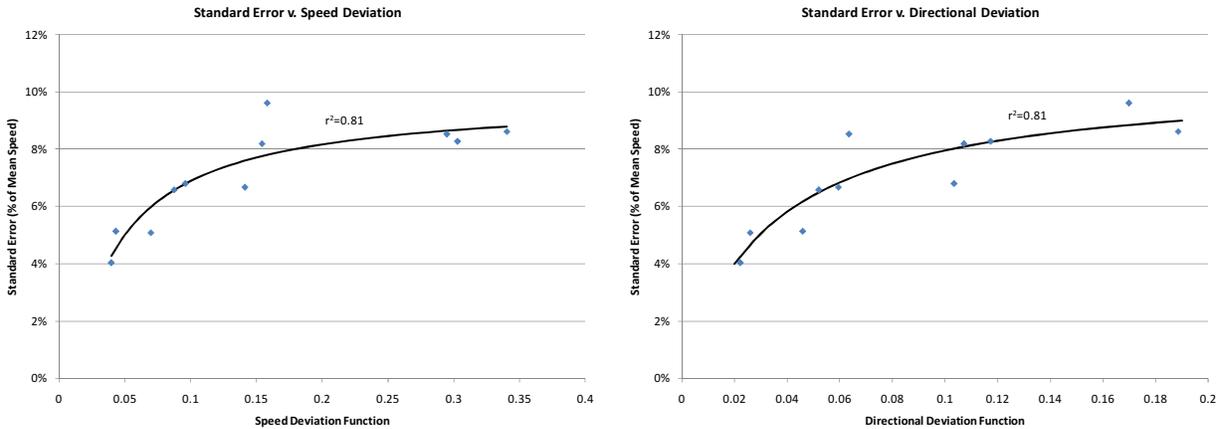

Figure 1. Plots of the modeling uncertainty, expressed as a percent of mean speed, as a function of SD (*left*) and DD (*right*). The fitted lines are linear in the reciprocal of SD and DD, as described in the text.

The predicted uncertainty based on this equation is plotted against the observed uncertainty in Figure 2. The r^2 coefficient has increased from 0.81 to 0.89.

The uncertainty according to this formula goes to 0.02 when *SD=DD=0*. This residual value is interpreted as a residual measurement error, denoted by the subscript *r* in Eq. 3. When the formula is applied to a real project, the residual should be removed so that the uncertainty goes to zero, as follows:

$$\sigma_M = \sqrt{\sigma_{M,r}^2 - 0.02^2}$$ (Eq. 4)

With this form of the equation, the measurement uncertainty (as well as other uncertainties) can be accounted for separately according to the particular characteristics of the masts and monitoring equipment.

For large values of SD and DD, the uncertainty goes to 0.12 when SD and DD become very large. This is the global uncertainty referred to earlier, although it is rarely reached in practice. The maximum error band for the binned cases evaluated in this study was about 10%.

It should be stressed that this formula is specific to AWS Truepower's wind modeling system, SiteWind. Other wind flow models may exhibit different behavior.
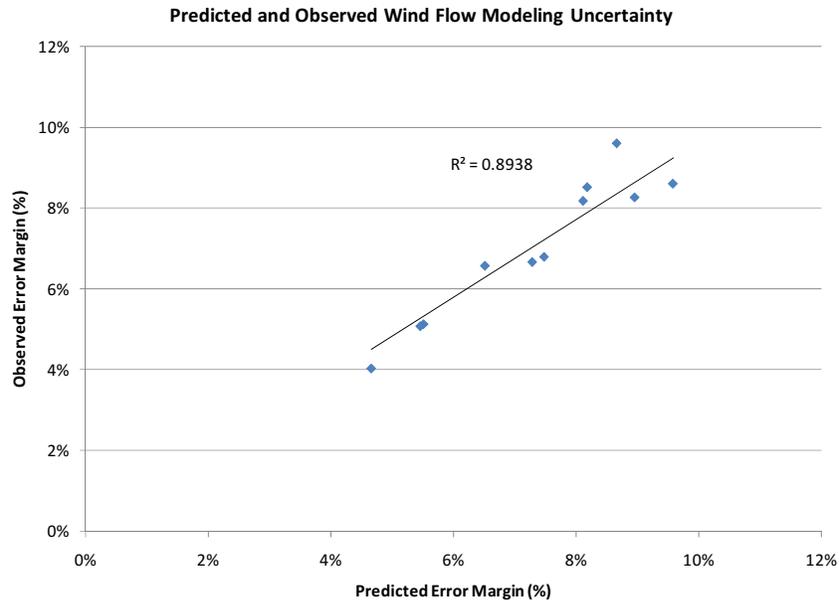
AWS Truepower®

Predicted and Observed Wind Flow Modeling Uncertainty

Figure 2. Plot of the predicted and observed modeling uncertainty, expressed as a percent of mean speed, according to Eq. 3.

APPLICATIONS

Having a good estimate of wind flow modeling uncertainty is of clear value in assessing risks for wind plant financing. The approach presented here can meet this need. For a given project layout, the uncertainty in wind resource can be calculated and combined with other sources of uncertainty, and from that the PXX production (where XX is a confidence threshold such as 75%, 90%, or 95%) can be derived.

In addition, the ability to map the spatial variation of the uncertainty opens the door to a number of other, highly useful applications in wind project development. Some of these applications have been implemented in the *openWind* software, and are described in this section.

Blending Wind Speeds and Resource Grids from Different Masts

A common wind project design challenge is to combine the wind resource estimates derived from more than one mast. In a typical situation, a separate wind flow modeling run is performed for each mast. Each run presents a different picture of the wind resource variation across the site. For a particular location, which run is to be trusted? If more than one appears to be reliable, how should the estimates be combined?

The solution to this problem starts with creating an uncertainty map, or raster, for each mast from a raw WRG file. An example is shown in Figure 3. The lowest uncertainty (yellow-green) on this map is found near the mast on terrain with a similar elevation and orientation. The greatest uncertainty (pink-red) is seen in the valleys to the south of the mast, as well as with increasing distance offshore.

With an uncertainty map like this for each mast, the problem of blending wind resource estimates from different masts can be solved in a statistically rigorous way. For a given location within the project area, the wind modeling prediction associated with a particular mast can be treated as an independent measurement of the wind resource. According to statistics (and under certain restrictions described below), when there is more than one independent measurement of a parameter, the best estimate of its true value is a weighted average of the measurements, where the weights are inversely proportional to the uncertainty squared. This is captured in the following equation,
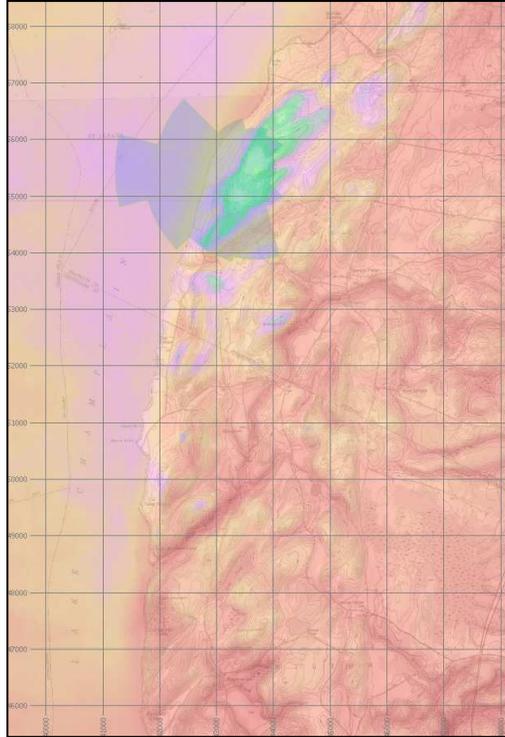
AWS Truepower®

Figure 3. Map of wind flow modeling uncertainty based on a single mast at a low-elevation coastal site. Relatively low uncertainties are indicated by yellow and green colors, relatively high uncertainties by pink and red.

$$e = \frac{\sum \dfrac{e_i}{\sigma_i^2}}{\sum \dfrac{1}{\sigma_i^2}}$$

(Eq. 5)

The uncertainty associated with this blended estimate is given by

$$\sigma^2 = \frac{1}{\sum \dfrac{1}{\sigma_i^2}}$$

(Eq. 6)

The blended uncertainty is smaller than any of the constituent uncertainties, and is in fact the smallest possible uncertainty that can be obtained by combining the measurements. (This is what is meant by the best estimate.)

By combining equations 3 and 4 with equations 5 and 6, one can derive a blended wind speed map or WRG based on any number of individual WRG files from different masts, and calculate the associated uncertainty. Figure 4 shows the difference between one such uncertainty-weighted blended map, based on two masts, with a map derived from an inverse-distance-squared interpolation. Also shown are two turbine layouts designed to maximize the energy production under each blended WRG.

The distance-weighted method places more weight on the mast nearest a turbine, whereas the uncertainty-weighted method places more weight on the mast whose location is topographically most similar to the turbine location. In this instance, most of the onshore area is topographically more similar to the southern mast than to the northern mast. Since the southern mast produces a somewhat greater wind resource estimate than does the northern mast, the blended wind resource is generally greater with the uncertainty-weighted method than with the distance-weighted method. The increase is particularly marked in the north, resulting in a larger number of turbines being placed there in the uncertainty-weighted layout.
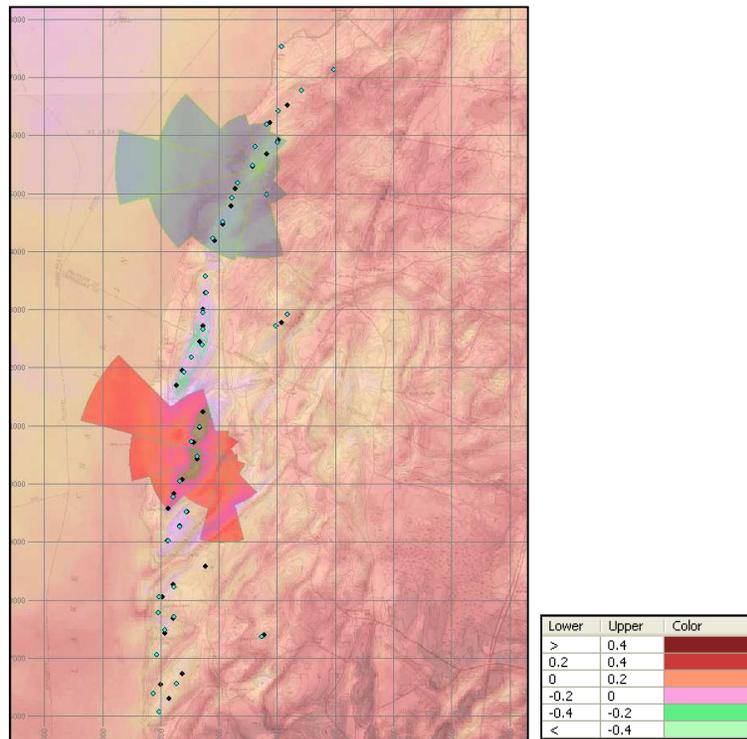
Lower	Upper	Color
>	0.4	
0.2	0.4	
0	0.2	
-0.2	0	
-0.4	-0.2	
<	-0.4	

Figure 4. Difference in the blended annual mean wind speed maps based on two masts, as predicted by the uncertainty-based method and by an inverse distance-squared interpolation method. Areas where the uncertainty-weighted method produces an increase in the mean wind speed are colored red. Optimized layouts are shown for both methods (distance-weighted: black, uncertainty-weighted: blue).

Equations 5 and 6 do not always hold. A key assumption behind them is that the different estimates represent random samples drawn from a well-behaved distribution of possible values. As a practical matter, this assumption can break down in two ways. First, there may be a sampling bias due to the placement of masts in topographically similar locations – for example, only along ridgelines. This bias is similar to what one would find if one tried to measure the average height of all men by looking only at basketball players: no matter how many measurements were taken, one would never approach the correct value. Second, the modeling results for a single mast can sometimes be very atypical of the results that would be obtained from a large number of masts at random locations. Through sheer bad luck, the mast may be an outlier of the distribution.

There are tests that can detect these problems. A bracketing test can determine whether the prospective turbine locations fall within the envelope of wind conditions represented by the masts. (This is analogous to setting distance limits so that no turbine falls outside the area covered by the masts.) The second type of problem can be addressed by comparing the deviations between the predicted and observed speeds for a given reference mast and for each target mast to the uncertainties associated with the same reference mast. If the deviations lie well outside the expected error range, the reference mast may be an outlier.

There are no established rules for handling such situations. It may be reasonable in some cases to exclude turbines from areas where it is suspected the method breaks down, i.e., where the uncertainty is likely to be underestimated, but in other cases that may not be feasible. Other options include increasing the modeling uncertainty and excluding certain masts.

AWS Truepower®

Optimizing Layouts to Maximize PXX

Given a spatially varying uncertainty, one can optimize the layout to maximize production at any confidence level. This may be desirable if a lender or investor discounts the expected production when determining the financing terms. The goal is to maximize the value of the project under the lenders' or investors' own assumptions. For the experiment shown in Figure 5, a 40-turbine project layout was optimized in two ways: to maximize the P50 production, and to maximize the P99 production. Because the uncertainty is smaller in the northern part of the region, the P99 layout has more turbines in the north than does the P50 layout. The impacts on energy production at different confidence levels are tabulated in Table 2. The P99 layout offers a lower expected production, but a greater P99 production. Depending on the financing terms, this project design may have a higher value.
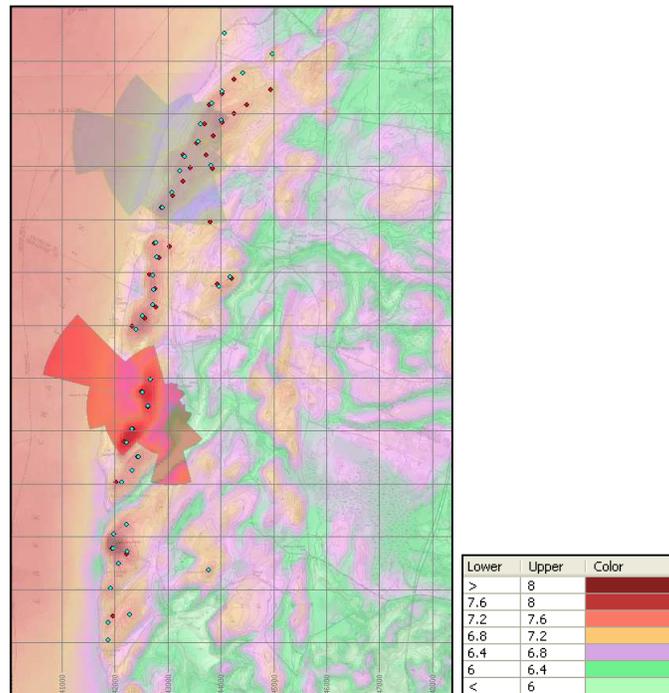

Lower	Upper	Color
>	8	
7.6	8	
7.2	7.6	
6.8	7.2	
6.4	6.8	
6	6.4	
<	6	

Figure 5. Predicted annual mean wind speed overlaid by two turbine layouts, one designed to maximize the P50 production (blue points), and the other to maximize the P99 production (red points). The P99 layout includes more turbines in the north than the P50 layout because the uncertainty is smaller there.

Table 2. Mean speed and energy production for layouts optimized to the P50 and P99 wind resource.

Quantity	P50 Layout	P99 Layout	Difference
P50 Mean Speed (m/s)	7.45	7.40	-0.7%
P99 Mean Speed (m/s)	5.83	5.94	1.9%
P50 Net Energy (GWh)	189.9	186.2	-1.9%
P99 Net Energy (GWh)	130.3	132.2	1.5%

Smart Monitoring Campaign Design

Uncertainty maps like that in Fig. 4 reveal where the uncertainty is greatest, and therefore can be used to show where additional monitoring masts are most desirable. This makes them a valuable tool for "smart" monitoring campaign design.

The ideal monitoring campaign should minimize the uncertainty for the most productive possible layout. Understanding the spatial variation of wind flow modeling uncertainty allows this to be done in a systematic way.

AWS Truepower®

In *openWind*, this concept has been implemented in the following way. First, the program calculates the blended uncertainty for all existing masts. Then it searches all allowed locations to find the one that would produce the greatest decrease in the combined uncertainty if a mast were placed there. In practice, this is best done for a preliminary layout or after first limiting the buildable area to windy sites. Otherwise, new masts may be proposed in clearly undesirable locations, such as valleys, merely because their topography is very different that of the existing masts.

CONCLUSIONS AND DISCUSSION

This research has established the following:

1. Spatial variation in the wind flow modeling uncertainty is an important factor to consider in designing both monitoring campaigns and wind projects.

2. The concept of wind resource similarity is useful for understanding the spatial variation of wind flow modeling uncertainty.

3. A model relating observed wind flow modeling errors to a measure of wind resource similarity exhibits a high degree of skill at several sites encompassing numerous reference-target mast pairs. This relationship, however, is specific to the SiteWind model.

4. An appropriate analytical model of uncertainty can be applied in wind farm design software such as *openWind®* to support smarter wind monitoring campaign design and wind project design.

ACKNOWLEDGEMENTS