WIND FLOW MODELING UNCERTAINTY:

Theory and Application to Monitoring Strategies and Project Design*

UL Solutions
463 New Karner Road
Albany, NY 12205 USA
www.ul.com/openwind

*This report was originally published by AWS Truepower and was then acquired in 2016 by UL Solutions.

WIND FLOW MODELING UNCERTAINTY

Theory and Application to Monitoring Strategies and Project Design

Michael C Brower, Chief Technical Officer

Nicholas M Robinson, Director of Openwind

Santi Vila, Lead Engineer

February 2015

USA | SPAIN | INDIA | BRAZIL | CANADA

awstruepower.com | contactus@awstruepower.com

AWS Truepower®

ABSTRACT

The uncertainty of wind resource and energy production estimates is a critical element in wind project financing. Although wind flow modeling uncertainty is often a large contributor to the total uncertainty, it is rarely quantified rigorously. This can lead to an underestimation or overestimation of the financial risks of the project. In addition, the variation of the wind flow modeling uncertainty across a site is generally unknown and, as a consequence, is ignored in the process of designing a wind project. Consequently, the turbine layout may not be optimal, leading to larger-than-expected errors in energy production forecasts.

This report presents a theoretical framework for understanding wind flow modeling uncertainty and illustrates some applications of this framework in plant 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) quantify the variation of wind flow modeling uncertainty across a project site in a physically reasonable and statistically defensible way; (b) design monitoring campaigns to minimize the wind flow modeling uncertainty for a particular buildable area; and (c) optimize a wind project layout to maximize the PXX production (where XX is any confidence threshold such as 75%, 90% or 95%), as appropriate for its particular financing model.

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 or other measurement systems, but where there are no measurements, the wind flow model becomes the only source of resource information. At large, complex sites, in particular, wind flow modeling uncertainty can be a significant fraction - typically one-fourth or more – of the total uncertainty in the estimated energy production.[1]

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. Often, these estimates are not based on a thorough comparison of measurements and model predictions at similar sites, and thus represent more “educated guesswork” than quantitative analysis. Equally important, most assessments do not consider how the uncertainty may vary with position across a project area. It may be implicitly assumed that the uncertainty increases with distance from the nearest mast. However, the distance dependence is rarely quantified or tested against data.

The purpose of this research was to develop, validate, and implement a theoretical framework allowing wind flow modeling uncertainty to be taken into account in a rigorous, quantitative way. Among other applications, the approach described here should permit plant developers and consultants to deploy wind monitoring assets in a more cost effective manner, design more nearly optimal wind turbine layouts, and develop a more accurate understanding of the overall uncertainty in energy production.

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 physical distance over which the model must extrapolate from measurements, but by the degree to which wind conditions are likely to differ between points, which can be thought of as the distance in *resource space*, or simply, the *resource distance* (RD). Three examples help illustrate the point.

1. In nearly flat, featureless terrain, the wind resource normally varies little with position. Any errors related to wind flow modeling will probably also be small and increase only slowly with distance from the nearest meteorological tower.
2. In mountainous terrain, the wind resource varies rapidly with position, especially in directions of steep slope. An example is a mast located on top of a steep ridgeline (Figure 1). Predictions for turbines that are even a short distance off the ridge are likely to have greater uncertainty than predictions at more distant points along the ridge top (though both are likely to be less certain than predictions made in flat terrain).
3. In a coastal region, the direction of largest wind resource gradient tends to be perpendicular to the shoreline (Figure 2). This is also the direction a wind flow model will probably exhibit the largest errors. A developer would be ill-advised to rely on a mast located well inland to predict the wind resource at the coast, or to rely on a mast at the coast to predict the resource well offshore.

As these examples suggest, how well a model is likely to perform appears to depend on the underlying resource gradient as well as on the physical distance. The more two points differ in their wind characteristics, however far apart they are, the more difficult it is for the model to accurately predict the resource at one point based on measurements at the other. The combination of resource gradient and physical distance is the resource distance.

If this conceptual framework is correct, then the challenge is to devise a quantitative measure of wind resource

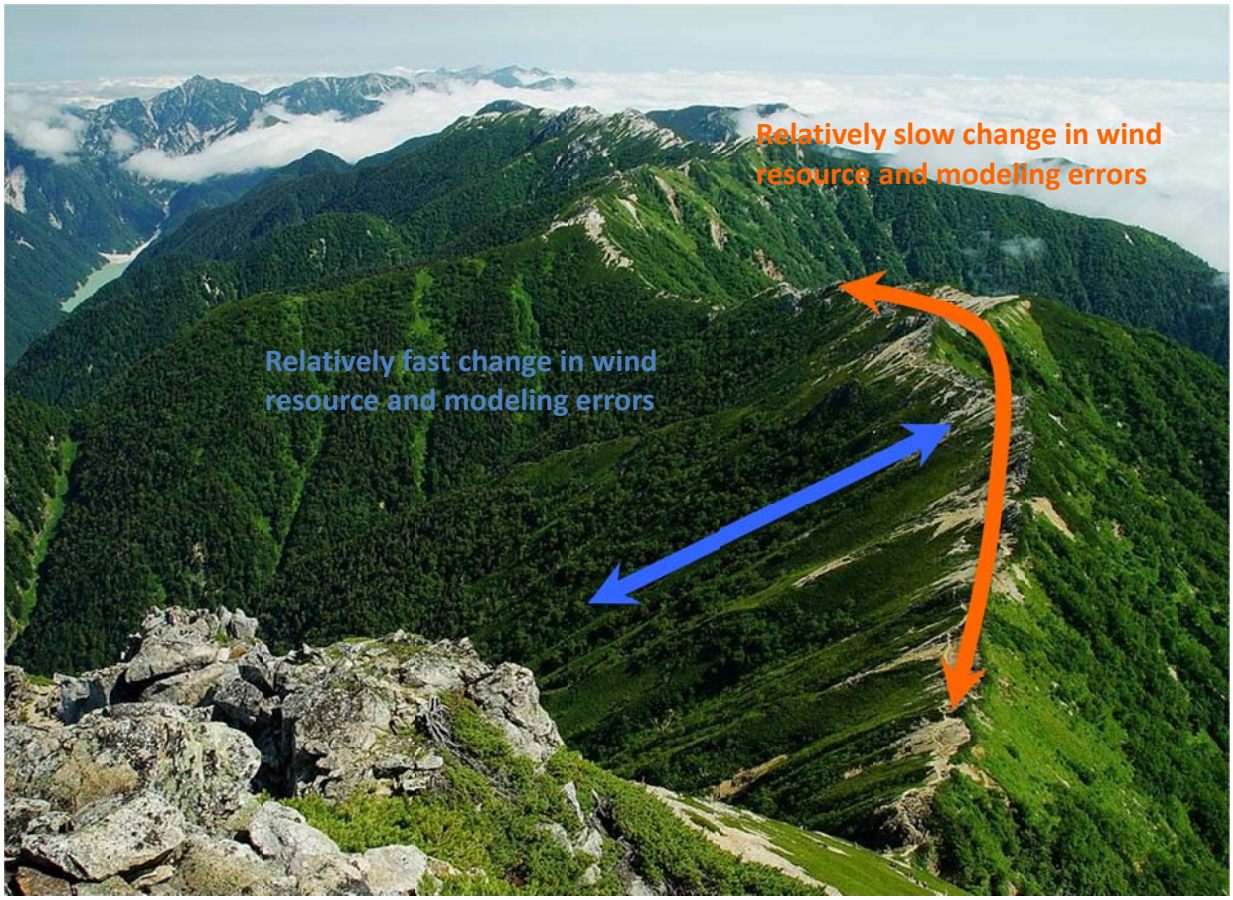

Figure 1. An illustration of how the wind flow modeling uncertainty is likely to increase more quickly in the direction of the most rapid change in the wind resource, in this case, directly down the slope from the ridge top.

Figure 2. Similar to Figure 1, but for a shoreline. Here it is not mainly the terrain that determines the gradient in the wind resource and wind flow modeling uncertainty, but the thermal and surface roughness contrast between land and water.

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.

(i) It should reduce to zero when the two points are the same.
(ii) It should always be non-negative, as it represents the width of an error distribution.
(iii) 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 features 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 frequencies. Variations in speed are often linked to topographic effects, such as acceleration over a ridge, as well as to 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 thermal effects, such as land-water temperature contrasts.

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,

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

The direction deviation formula is,

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

The sums are over *ND* wind directions, where *ND* is typically 12 or 16. f_i refers to the frequency of occurrence 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, however, depends on whether the wind flow model is sensitive to the main factors that are likely to affect the resource. 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 their uncertainty. The more sophisticated the model, in other words, the better the results of this approach are likely to be.

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 an independent 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 wind resource grid (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 first by distance, 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 3 show the relationships between the observed modeling uncertainty and SD and DD for our data sample. Overlaid on each plot is a fitted line of the form

$$\sigma = E\left(1 - B\frac{A}{x + A}\right) \tag{3}$$

where σ 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 determination 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 (where it approaches $E(1\text{-}B)$) and for very large values of x (where it approaches E), consistent with the constraints listed in the Conceptual Framework section.

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. This led to the following equation:

$$\sigma = E\left[1 - B\frac{A}{DD + A} - D\frac{C}{SD + C}\right] \tag{4}$$

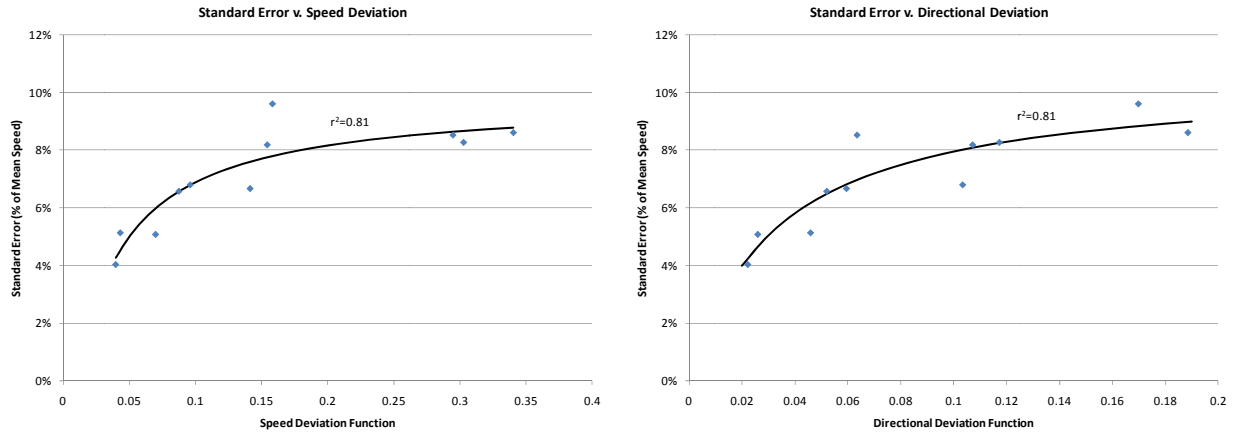

Figure 3. Plots of the modeling uncertainty, expressed as a percent of mean speed, as a function of SD (*left*) and DD (*right*). The fitted lines are linear in the reciprocal of SD and DD, as described in the text.

The values of the coefficients are as follows:

A	0.071
B	0.411

$$\begin{array}{ll} C & 0.097 \\ D & 0.423 \\ E & 0.121 \end{array}$$

The predicted uncertainty based on this equation is plotted against the observed uncertainty in Figure 4. Note that 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 value is interpreted as a residual measurement error. It implies that even if two masts are at exactly the same location, their measurements may differ by an amount of that order, reflecting differences in the performance and mounting of anemometers, among other factors. To obtain the modeling uncertainty alone, the residual should be removed as follows,

$$\sigma_{WFM} = \sqrt{\sigma^2 - 0.02^2} \tag{5}$$

so that the uncertainty goes to zero. The subscript WFM denotes the uncertainty due to the wind flow model alone. With this form of the equation, the measurement uncertainty (as well as other non-wind-flow-modeling uncertainties) can be accounted for separately according to the particular characteristics of the masts and monitoring equipment. The method of doing so is discussed later.

For large values of *SD* and *DD*, the uncertainty in Eq. 4 goes to 0.12. 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 actually about 10%.

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

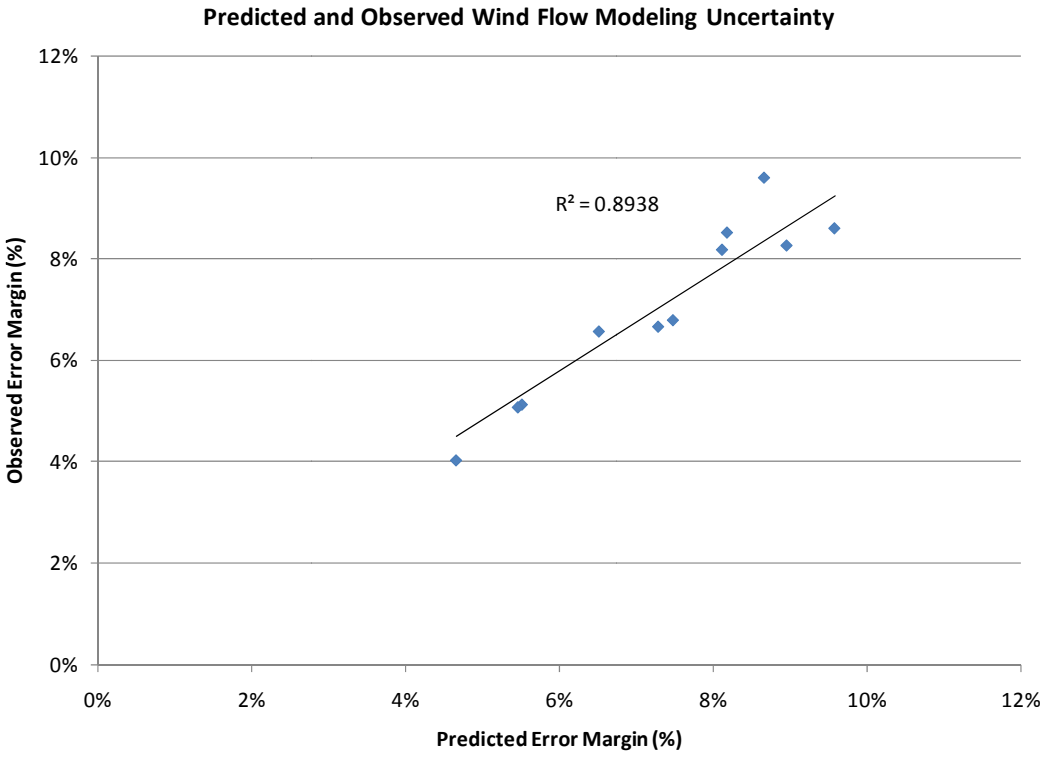

Figure 4. Plot of the predicted and observed modeling uncertainty, expressed as a percent of mean speed, according to Eq. 4.

IMPLEMENTATION

The results of the foregoing research have been implemented in the Openwind software. The following sections describe technical aspects of the implementation.

Uncertainty Maps Derived from a Single Mast

The first and simplest application of the wind flow modeling uncertainty is creating an uncertainty map for a single measurement tower. Such maps are created by using Eq. 1 and Eq. 2 to derive the directional and speed deviations with respect to the tower location for every point in the area, and then by using Eq. 4 and Eq. 5 to derive the corresponding modeling uncertainty. A set of four such maps, along with the corresponding topographic map, is shown in Figure 5.

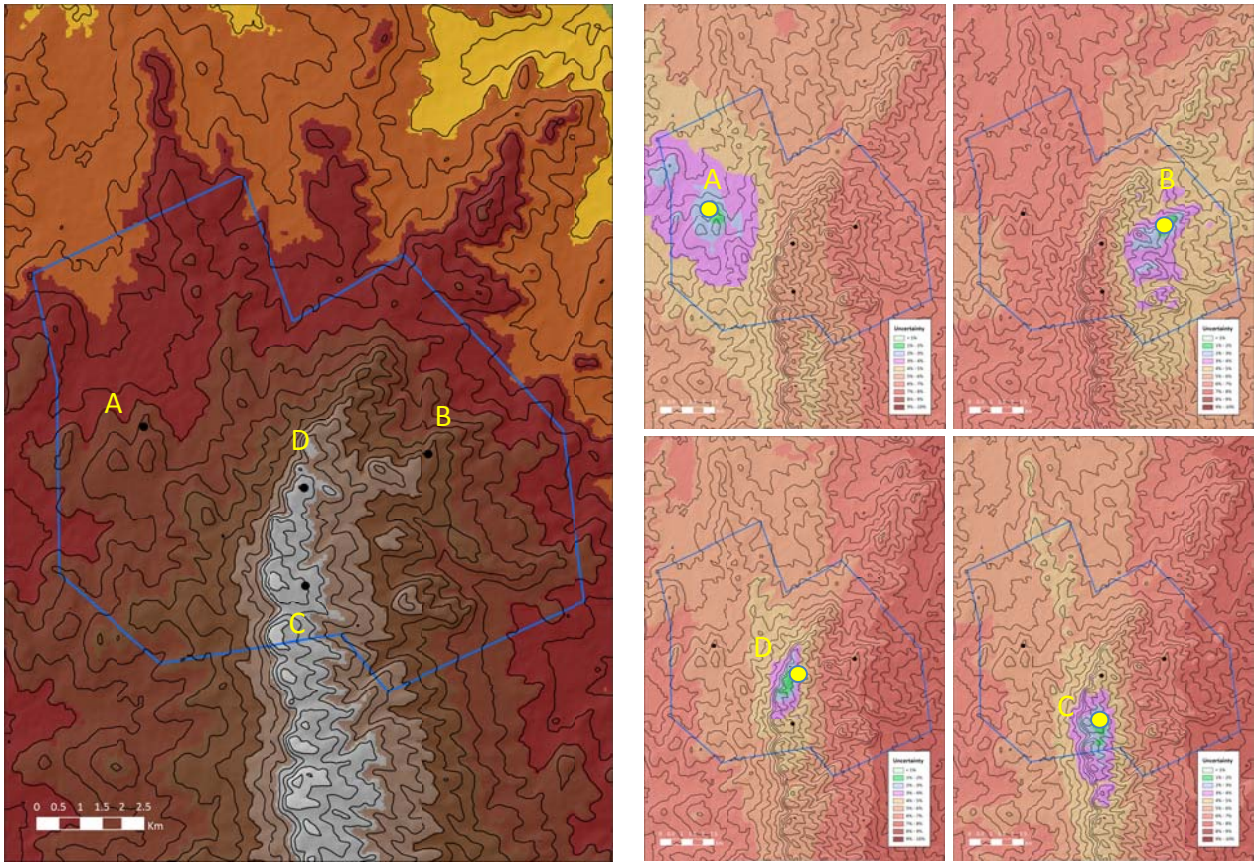

Figure 5. An example of four uncertainty maps created from four separate masts at a wind project site. The map on the left shows the topography: a ridgeline (light brown to white) tapering into a plain (dark brown to red). Four masts are shown as black dots and labeled A, B, C, and D. In the four panels on the right, the uncertainty map generated for each mast is shown. Blue to green colors represent smaller uncertainties, while pink to red colors represent larger uncertainties.

Note that in each of the four uncertainty maps, the area of lower uncertainty tends both to be close to the reference tower and to follow the contours of terrain that is similar to where the tower is located. For example, for towers A and B, located on either side of the ridge, the area of relatively low uncertainty is on the same side of the ridge as the corresponding tower, and the uncertainty increases much more rapidly towards the ridge than along it. Similarly, for towers C and D, the area of relatively low uncertainty follows the ridge top. These results are consistent with the conceptual model discussed earlier (see Figure 1).

Combining Resource and Uncertainty Estimates from Different Masts

In recent years it has become common to install more than one meteorological tower at a wind project site. Typically in such cases, a separate wind flow modeling run or model adjustment is performed for each mast. Each run presents a somewhat different picture of the wind resource variation across the site. For any particular point, how should the different estimates be combined, and what is the uncertainty of the final result?

The general approach we use to answering these questions has been developed to handle situations involving multiple measurements of various kinds – for example, an observatory measuring the diameter of a planet on different nights, or several laboratories measuring the effectiveness of a new drug in treating a disease in different groups of subjects. In each case, it is possible to combine the results of well-designed experiments to derive a new (and usually more accurate) estimate of the parameter being sought.

In applying this approach, we treat the wind resource estimate at any point derived from one mast as a single "measurement." In the following section, we start with the simple case where each such measurement is assumed to be independent of every other measurement.[2] Later, we address the more general case of semi-dependent, or partially correlated, estimates.

Independent Estimates

We start by supposing that there are two independent measurements of a quantity e, referred to as e_1 and e_2, with associated uncertainties σ_1 and σ_2. (For example, e could be the average wind speed.) The best estimate of e is assumed to be a weighted average of the two estimates,

$$e = w_1 e_1 + w_2 e_2 = w_1 e_1 + (1 - w_1) e_2 \tag{6}$$

where the last term takes advantage of the fact that in a weighted average, the weights must sum to one:

$$w_1 + w_2 = 1 \tag{7}$$

The questions posed at the start of this section boil down to this: What weights in the above equations will produce the value of e with the smallest uncertainty, and what is that uncertainty? It is easiest to consider the variance of the estimates (the square of their uncertainty) rather than uncertainty itself. The equation for the variance of e corresponding to Eq. 6 is

$$\sigma^2 = w_1^2 \sigma_1^2 + (1 - w_1)^2 \sigma_2^2 \tag{8}$$

To find the value of w_1 where the total variance is minimized, we take the derivative of this equation with respect to w_1 and set the result to zero. With some reorganization, the result is

$$w_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \frac{\dfrac{1}{\sigma_1^2}}{\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2}} \tag{9}$$

(Though the first form of the equation is simpler, the second provides more insight into what is going on, as will be apparent below.) The corresponding value of the second weight, w_2, is given by the same equation with subscripts 1 and 2 switched. After inserting this expression into Eq. 6, we have

$$e = \frac{\dfrac{e_1}{\sigma_1^2} + \dfrac{e_2}{\sigma_2^2}}{\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2}} \tag{10}$$

In other words, the optimal weights are inversely proportional to the squared uncertainty of each measurement. (The denominator is a normalization factor which ensures that the weights sum to one.) The corresponding blended uncertainty is derived by replacing Eq. 9 in Eq. 8 and reorganizing:

$$\sigma^2 = \frac{(\sigma_1\sigma_2)^2}{\sigma_1^2+\sigma_2^2} = \frac{1}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}} \tag{11}$$

Note that when the two uncertainties are equal, the equations reduce to the familiar case where the optimal estimate is the average of the two independent measurements, and the combined uncertainty is the uncertainty of one measurement divided by the square root of 2.

These results are easily generalized to many measurements, as the first pair of measurements can be combined to create a new estimate, which can then be combined with the next measurement to create another new estimate, and so on until the last measurement is reached. The general result is:

$$e = \frac{\sum \frac{e_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} \tag{12}$$

$$\sigma^2 = \frac{1}{\sum \frac{1}{\sigma_i^2}} \tag{13}$$

where e_i and σ_i are the ith estimates of the parameter and its uncertainty, respectively. It can be shown that the blended uncertainty is never larger (and is usually smaller) than the smallest 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.")

It is interesting to take a moment here to reflect on the meaning of these equations. At points where the wind resource predicted by a mast has a low uncertainty compared to that of other masts, that mast will be given a relatively large weight in the final estimate. Conversely if the resource predicted by the mast has a relatively high uncertainty, the mast will be given comparatively little weight. This makes intuitive sense.

Furthermore, this approach provides a fresh perspective on the common practice of weighting estimates from different masts according to the inverse of the distance squared to each mast. Now we can see that this method is the same as assuming that the uncertainty of each estimate is strictly proportional to the distance to the corresponding mast. This assumption, as we have noted, fails to account for the important influence of the wind resource gradient on the wind flow modeling uncertainty.

Partially Correlated Estimates

While the equations described above are fairly simple, they have a significant drawback: the uncertainty always decreases with the addition of more towers. It is easy to imagine situations where this would not happen. For example, one could place two towers right next to each other. Since they measure exactly the same resource, the second tower would add no value (except to reduce the measurement error), and should result in no decrease in the modeling uncertainty at points away from the towers. Eq. 13, however, says that it would.

The reason the equations fail in this situation is that the measurements are not, in fact, independent. In this example, the modeling errors from the two masts would be perfectly correlated, meaning that if an estimate at a point based on one of the masts were biased high or low, then an estimate based on the other mast would be biased in the same direction, and by exactly the same amount.

Other situations can be imagined where the correlation would be, if not 1, at least significantly greater than zero. For example, suppose only tower C was initially present in Figure 5. Would the addition of a tower at point D

reduce the uncertainty at, say, point A by as much as the square root of two? Probably not, since both C and D measure the ridge-top resource rather than the resource in the valley, and hence their valley predictions may be biased in a similar way.

Finding the optimal blend of partially correlated wind resource estimates, though more complicated, is based on the same technique used for independent estimates. Starting with the relatively simple case of two towers, the equation for the combined variance is as follows (compare with Eq. 8):

$$\sigma^2 = w_1^2\sigma_1^2 + (1-w_1)^2\sigma_2^2 + 2w_1(1-w_1)r_{12}\sigma_1\sigma_2 \tag{14}$$

Here, r_{12} is the correlation of the two measurements, and the expression $r_{12}\sigma_1\sigma_2$ is their covariance, sometimes written $cov(1,2)$. Note that when the covariance is positive, the combined variance is larger than if the measurements were independent. This shows directly that the combined uncertainty of two partially correlated measurements is always larger than the combined uncertainty if they are independent (assuming the correlation is positive).

After taking the derivative of this expression with respect to the weight and setting the result equal to zero, we can solve for the weight that produces the lowest uncertainty in the combined estimate. The result is

$$w_1 = \frac{\sigma_2^2 - r_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2r_{12}\sigma_1\sigma_2} = \frac{\frac{1}{\sigma_1^2} - \frac{r_{12}}{\sigma_1\sigma_2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} - \frac{2r_{12}}{\sigma_1\sigma_2}} \tag{15}$$

In the special case where $r_{12}=1$ and $\sigma_1=\sigma_2$, the weights w_1 and w_2 are undefined, as the denominator in Eq. 15 goes to zero. No matter what values we give them, however, Eq. 14 states that the total uncertainty equals the uncertainty of one measurement alone; in other words, when perfectly correlated, the second measurement adds no value.

These equations can be extended to any number *N* of measurements (or towers). The solution involves creating a system of *N* equations representing the derivatives of the total variance with respect to the weight given to each mast, which can be written,

$$\frac{\partial(\sigma^2)}{\partial w_k} = \sum_i \sum_j \frac{\partial(w_i w_j)}{\partial w_k} r_{ij}\sigma_i\sigma_j = 0 \tag{16}$$

subject to the constraint,

$$\sum_i w_i = 1$$

(Note that when $i=j$, then $r_{ij}=1$ and $\sigma_i\sigma_j=\sigma_i^2$.) The constraint reduces the number of equations to *N*-1, rendering the system solvable through matrix inversion. In matrix notation, after grouping all terms containing w on one side and all other terms on the other, we have

$$\boldsymbol{A}w = \boldsymbol{b} \tag{17}$$

For *N*=3 (i.e., three masts), for example, the matrix ***A*** and vector ***b*** are as follows:

$$A = \begin{bmatrix} \sigma_1^2 + \sigma_3^2 - 2r_{13}\sigma_1\sigma_3 & \sigma_3^2 + r_{12}\sigma_1\sigma_2 - r_{23}\sigma_2\sigma_3 \\ \sigma_3^2 + r_{12}\sigma_1\sigma_2 - r_{23}\sigma_2\sigma_3 & \sigma_2^2 + \sigma_3^2 - 2r_{23}\sigma_2\sigma_3 \end{bmatrix} \tag{18}$$

$$b = \begin{bmatrix} \sigma_3^2 - r_{13}\sigma_1\sigma_3 \\ \sigma_3^2 - r_{23}\sigma_2\sigma_3 \end{bmatrix}$$

Once ***A*** and ***b*** are known, then the solution for the vector of weights is given by

$$\boldsymbol{w} = \boldsymbol{A}^{-1}\boldsymbol{b} \tag{19}$$

following standard matrix inversion methods.

Correlation Coefficients

It remains to determine the correlation (or more precisely, the covariance) between any two towers for a given point. Without this information, we cannot create or solve the system of equations in Eq. 16.

Lacking a general model of the correlations of wind flow modeling errors, we adopt a practical approach based on the concept of resource distance. This extends the logic of the earlier example, in which two towers located on the same ridgeline were used to predict the resource in the valley. We posit that if the wind resources at two masts are much more similar to one another than either is to that of the point in question, then the modeling errors for the two masts are likely to be correlated at that point. Conversely if the resource of either mast is more similar to that of the point than to that of the other mast, or if all three points have different resources, then the modeling errors are likely to be uncorrelated.

Resource distance (*RD*) is measured by the combined directional and speed deviations *DD* and *SD* between the masts and between each mast and point. In a manner analogous to the *x* and *y* components of distance on a plane, SD and DD are treated as orthogonal, and the total RD is given by:

$$\mathrm{RD}_{\mathrm{ij}} = \left(\mathrm{SD}_{\mathrm{ij}}^2 + \mathrm{DD}_{\mathrm{ij}}^2\right)^{1/2} \tag{20}$$

The correlation coefficient is then given by the ratio

$$r_{ij}^{(p)} = \frac{min\left(RD_{ip}^2, RD_{jp}^2\right)}{\left[min\left(RD_{ip}^2, RD_{jp}^2\right) + RD_{ij}^2\right]} \tag{21}$$

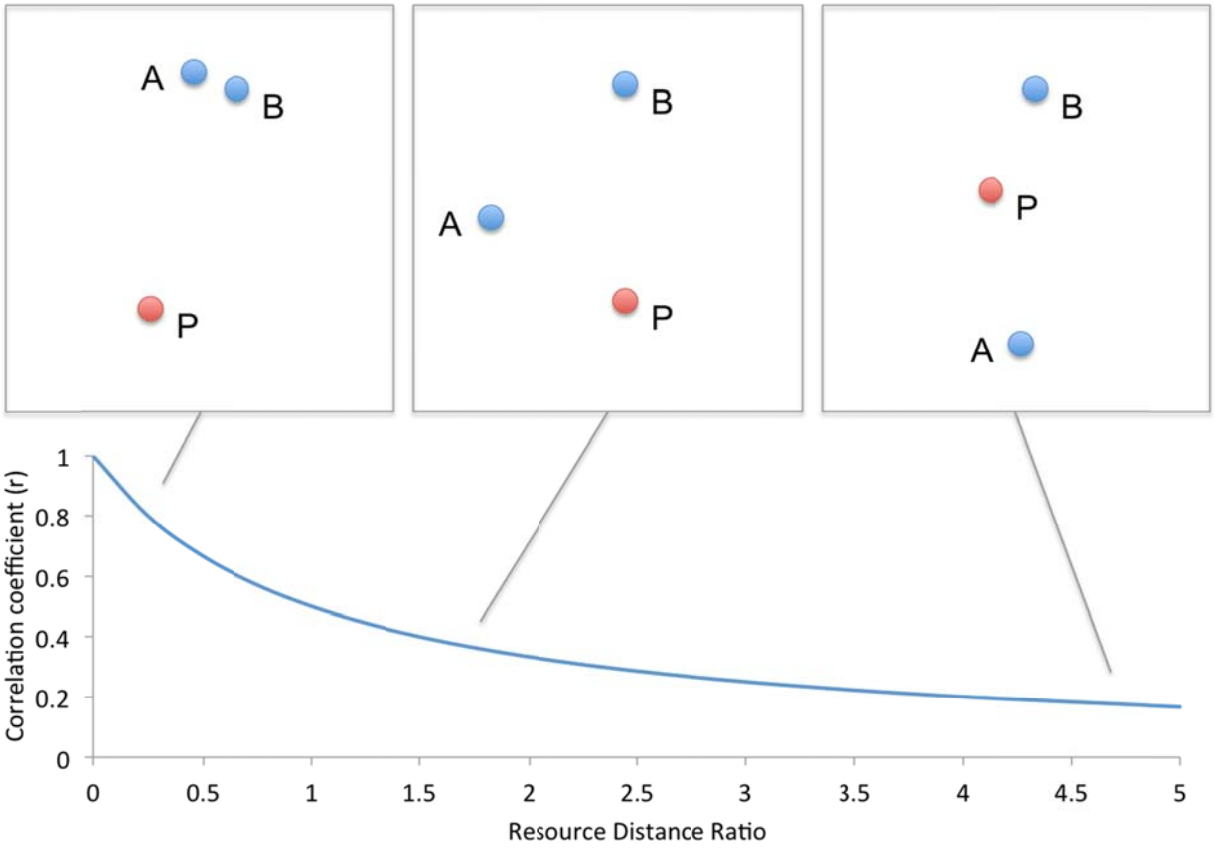

Figure 6. Three illustrations of different arrangements of two towers, *A* and *B*, relative to a point *P*, and the corresponding estimate of the correlation coefficient. The resource distance ratio is the ratio of the RD^2 between masts to the smallest RD^2 between the point and either mast. For further explanation, see the text.

Here, p is an index representing the point (e.g., a grid coordinate), and i and j represent a pair of masts. In effect, the *RD* between the masts establishes the scale against which the *RD* between the point and the most similar (smallest *RD*) mast is compared. When the RD between masts is much larger than the RD between the point and its most similar mast, then r_{ij} will be close to zero. When the opposite is true, then r_{ij} will be close to one.

The concept is illustrated in Figure 6. On the left of this figure, the masts are assumed to be much closer to each other than they are to the point, with the result that modeling errors from the two masts are likely to be highly correlated (r ~1). On the right, the point is between the two towers, and the modeling errors are assumed to be mostly independent. The situation in the middle is between the two extremes. Note that, for convenience, these illustrations show the physical positions of the towers and point. However, it is the resource distance between them that counts.

The effect of this approach on a real-world case is illustrated in Figure 7 for the same site as shown in Figure 5. As is to be expected, the uncertainty has been reduced at all points, and the areas of low uncertainty now represent a blend of the corresponding areas from the original maps.

Towards a Complete Uncertainty Model

The final stage of the Openwind implementation was to extend the method to encompass other sources of uncertainty, which can include sensor calibration, shear, MCP, and other factors. This is not the place for a detailed discussion of these sources, which is covered well in other references. Instead we will content ourselves with a simple treatment that can be adapted to many situations.

For convenience, we assume that any source of uncertainty can be classified as being one of two types: uncorrelated or correlated.[3] Uncorrelated sources are independent between measurement points. Examples might

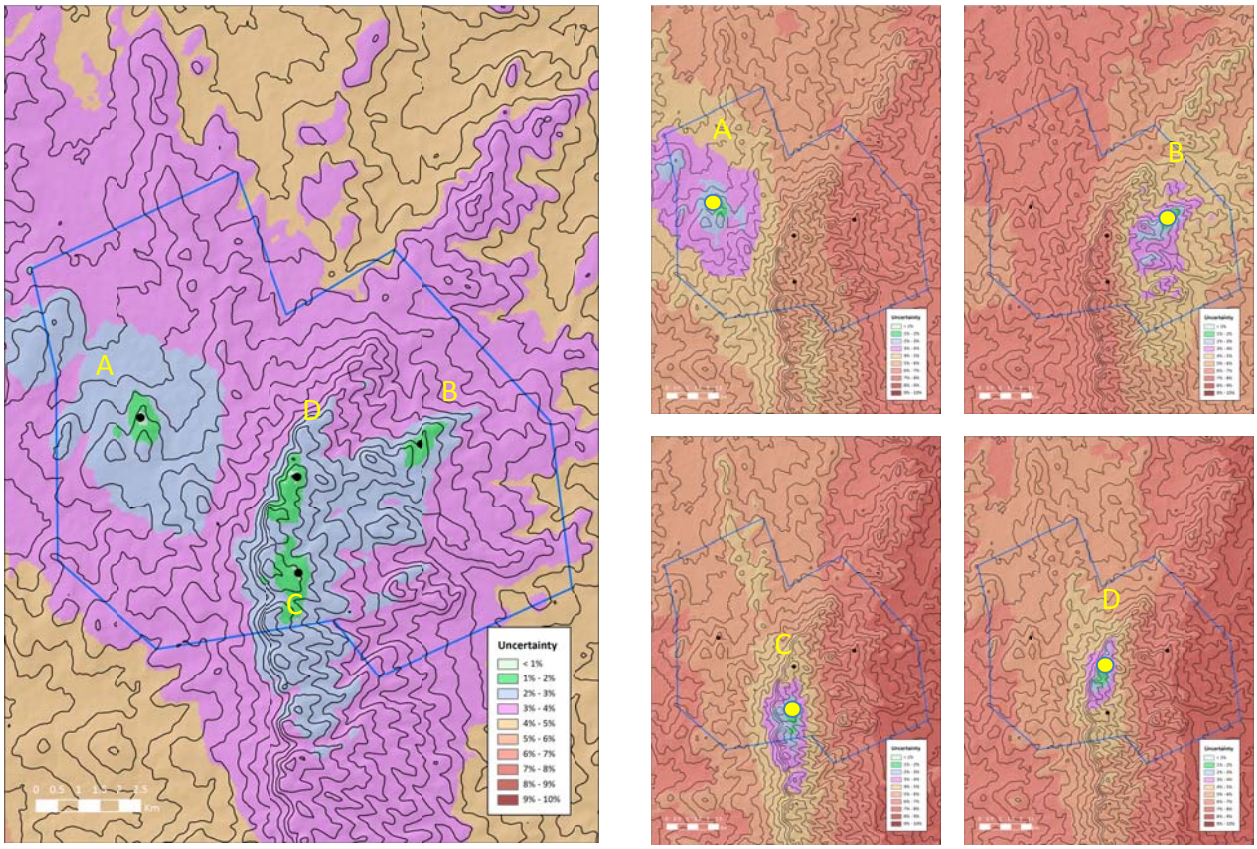

Figure 7. The map on the left is the blended uncertainty of mean wind speed estimates from the four masts A, B, C, and D, created using the method described here. The individual uncertainty maps are shown on the right (and are the same as those shown in Figure 5).

include measurement errors due to variations in the sensitivity of anemometers and other instruments. (Note, however, that some measurement errors *can* be correlated between towers if the measurements share a common characteristic, such as the same instrument mounting.) Correlated sources of uncertainty affect all measurements in the same way. An example of a source of correlated uncertainty might be the MCP process. In most cases, the towers at a site span a similar period of record and are adjusted to the same long-term reference. Thus, any MCP error caused by the choice of period or reference is likely to affect all towers more or less equally.

Both types of uncertainty are incorporated easily into the framework described in previous sections. For uncorrelated uncertainties, this is accomplished by increasing the variance of each measurement while keeping the covariances unchanged. For a given tower *i*, the combined variance is

$$\sigma_i^2 = \sigma_{i,WFM}^2 + \sigma_{i,U}^2 \tag{22}$$

where the subscript *WFM* refers to the wind flow modeling uncertainty and the subscript *U* refers to all other uncorrelated sources of uncertainty. The effect of adding the uncorrelated uncertainty is to increase the variances relative to the covariances in the ***A*** and ***b*** matrices in Eq. 17. This, in turn, changes the weights assigned to the masts. In this way, the relative importance of each mast in the final resource estimate depends not only on its wind flow modeling uncertainty, but on other components of uncertainty such as the mast height, instrument quality, period of record, and so on. Including the uncorrelated uncertainty term has the added virtue of making the matrix ***A*** in Eq. 17 less prone to singularities. For both reasons, it is always recommended to include the uncorrelated uncertainty when using the uncertainty model in the Openwind software.

A consequence of including uncorrelated uncertainties is that the weight given to a single mast may never reach 100% *even at that mast's location*. This is contrary to common practice. When only the wind flow modeling uncertainty is considered, the uncertainty associated with a particular mast goes to zero at the mast, and therefore that mast becomes the only one that counts in the blended resource estimate at that point. However, when other uncorrelated uncertainties are included, the uncertainty does not go to zero, and thus the blended resource may not quite match the observed at that tower. This is discomforting at first, but is the logical consequence of the fact that no measurement is perfect, and all measurements, within reason, have some value.

The correlated uncertainty is dealt with by combining it with the final blended uncertainty, i.e.,

$$\sigma^2 = \sigma_{WFM+U}^2 + \sigma_C^2 \tag{23}$$

This component increases all uncertainties without changing the relative weights assigned to different masts.

APPLICATIONS

With the implementation of the uncertainty model described in the previous section, it is possible to consider a number of applications beyond simply creating uncertainty maps. Here we discuss two: cost-effective monitoring campaign design, and optimizing layouts to consider resource uncertainty.

Cost-Effective Monitoring Campaign Design

The optimal monitoring campaign should minimize the uncertainty for the most productive possible layout. Understanding the spatial variation of wind flow modeling uncertainty can allow this to be accomplished in a systematic and cost-effective way.

Figure 8 illustrates how the process can unfold in practice. At this site, the developer at first assumes that the best wind resource is along the upwind (southern) edges of the mesa, and so initially deploys monitoring masts there. Subsequent wind flow modeling based on these towers indicates the downwind (northern) edges may have as good or better resource. However, the uncertainty map on the left shows that this prediction is relatively uncertain, and should be confirmed. Placing masts in the locations indicated on the right substantially reduces the uncertainty over a large part of layout.

Something like the decision-making process described here could easily occur without uncertainty maps. An experienced analyst would likely recognize that the wind resource in the indicated area should be confirmed by measurement. The key advantage of uncertainty mapping is that the effects on uncertainty of changes in a

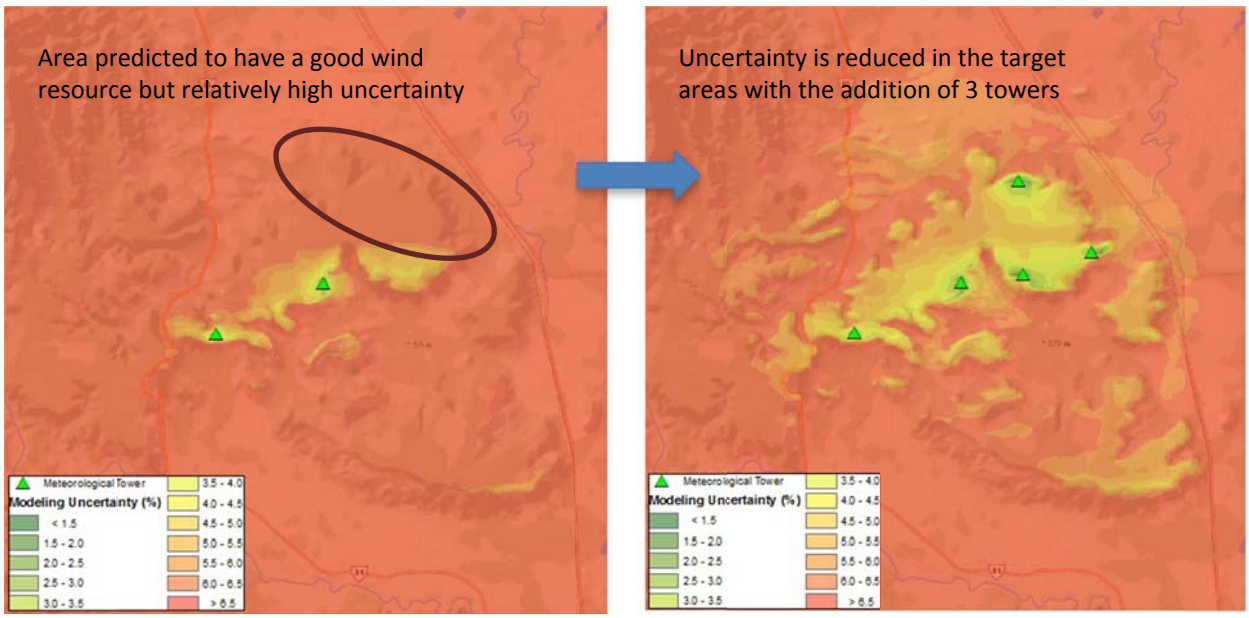

Figure 8. Illustration of how uncertainty maps can be used to inform monitoring campaign design. *Left:* Initial monitoring focused on the upwind (southern) edges of this mesa. The wind resource map (not shown) suggested the northern edge would be windier, but the uncertainty there was relatively high (ellipse). *Right:* After the installation of 3 more towers, the uncertainty was substantially reduced across the mesa, and especially in the areas of interest.

monitoring campaign design can be quantified and used to determine, in an objective way, when and where additional measurements might be cost effective. Without uncertainty maps based on rigorous analysis, this type of trade off must be done subjectively, and different analysts will likely arrive at different conclusions.

In addition, uncertainty maps provide useful information on the optimal placement of measurement systems to achieve the greatest reduction in uncertainty. This is in contrast to the common practice of employing guidelines based on the maximum distance between turbines and towers in different types of terrain. The blind application of such guidelines can lead to false confidence that there are enough towers for a given layout and that the towers are optimally placed. For example, one could satisfy the distance requirement by placing all towers along a ridge top, even though many turbines might be placed well off the ridge. The same goes for coastal sites, where placing towers at different distances from the shore will produce a lower overall uncertainty than placing the same towers parallel to the shore, albeit the same distance apart.

A tool for using uncertainty maps in monitoring campaign design has been implemented in the Openwind software.[4] The program first calculates the blended uncertainty based on existing measurements (which could be from a single tower). Then it searches all allowed locations to find the one that would produce the largest decrease in the combined uncertainty for a particular layout or buildable area assuming an additional tower (or remote sensing system) were placed there. Once that location is chosen, the program recalculates the combined uncertainty, and then searches for another location that would produce the next largest decrease in uncertainty, and so on, until it is told to stop.

In this way, the tool can help the developer determine the smallest number of towers needed to achieve a desired level of uncertainty. Or, knowing the cost of each additional measurement and armed with a suitable financial model for the project, the developer can find the point of diminishing returns where the value of the next incremental reduction in uncertainty is outweighed by the cost of the next tower.

Figure 9 illustrates both approaches. In this chart, the blue curve represents a typical relationship between the number of towers deployed at a site and the estimated total uncertainty in plant production (including uncorrelated and correlated errors not associated with the wind flow model). After two towers, the curve levels out markedly as each additional tower produces a smaller decrease in the wind flow modeling uncertainty, and as the non-wind-flow-related uncertainties begin to dominate. If the goal is to reduce the uncertainty below, say, 8% (implying a debt service coverage ratio of 1.23 based on the P99 – a typical criteria employed by banks), the chart

indicates it can be accomplished with 4 towers, assuming they are placed in an optimal way according to the uncertainty model.

The red curve shows the corresponding incremental cost:benefit ratio, i.e., the cost of each additional additional tower (including equipment, installation, maintenance, and analysis, assumed to be $100,000 per tower) divided by the expected benefit of that tower (assumed to be $2,000,000 per percentage point decrease in uncertainty, based on a simple financial model, with a 25% chance of project success). If the goal is not to exceed an incremental cost:benefit ratio of 1.0, this is satisfied with 7 optimally-placed towers.

Figure 9 is merely an illustration. Every project will differ.

Optimizing Layouts to Maximize PXX

A second interesting application of the uncertainty mapping method is to optimize turbine layouts to maximize production at a desired confidence level, rather than merely to maximize the expected, or P50, production. This approach may be desirable if a lender or investor discounts the expected production when determining the financing terms. The aim is to maximize the value of the project under the lenders' or investors' own assumptions.

In the experiment shown in Figure 10, a 40-turbine project layout was optimized in two ways: (a) to maximize the P50 production (blue points), and (b) to maximize the P99 production (red points). Because the uncertainty (not shown) is smaller in the northern part of the region, the P99 layout has more turbines in the north than does the P50 layout.

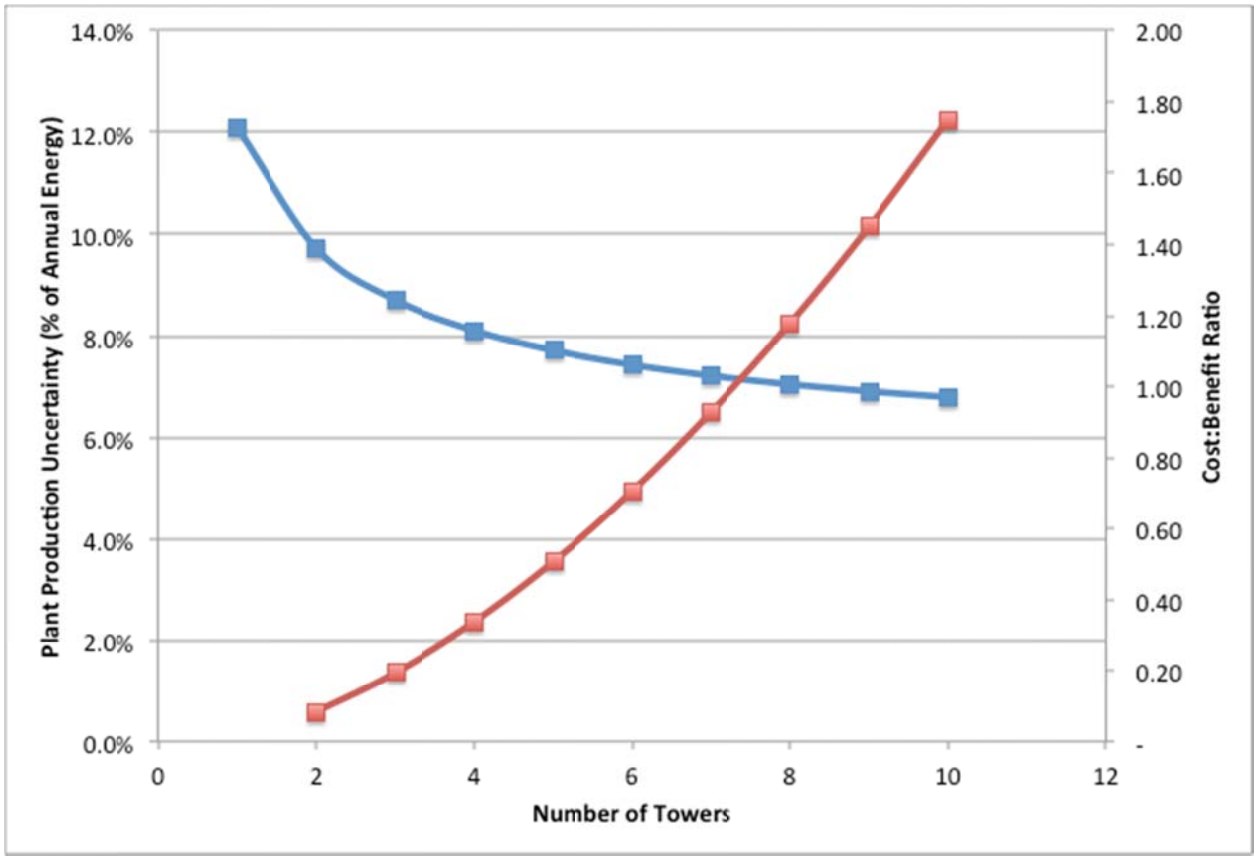

Figure 9. Typical dependence of plant production uncertainty (blue curve) and marginal cost-benefit ratio (red curve) on the number of towers or other measurement systems for a large wind project site.

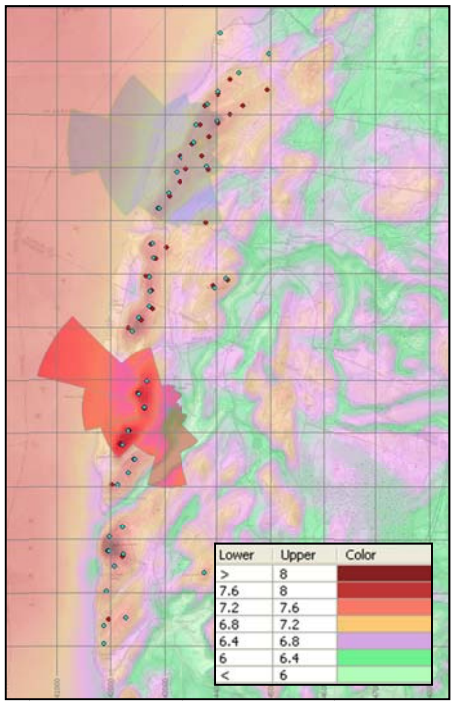

Figure 10. Predicted annual mean wind speed overlaid by two turbine layouts, one designed to maximize the P50 production (blue points), and the other to maximize the P99 production (red points). The P99 layout includes more turbines in the north than the P50 layout because the uncertainty is expected to be smaller there based on the similarity of the resource compared to that of the tower.

The impacts on energy production at different confidence levels for the same case are tabulated in Table 1. The P99 layout offers a lower P50 production, but a greater P99 production. Depending on the financing terms, this project design may have a significantly higher value.

Table 1. Mean speed and energy production for the two layouts shown in Figure 10.

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%

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. The uncertainty model derived from this research has been implemented in the Openwind software as a tool to support cost-effective wind monitoring campaign design, improve production uncertainty estimates, and optimize wind plant layouts.

ACKNOWLEDGEMENTS

Julien Bouget and Erik Hale contributed useful ideas and insights to this work.

[1] *Wind Resource Assessment: A Practical Guide to Developing a Wind Project,* Chapter 14, Michael C. Brower (ed.), Wiley 2012.
[2] Estimates or measurements are said to be independent if there is no tendency for their errors to be correlated with one another, i.e., to be wrong in the same direction.
[3] In general, it is possible to split every source of uncertainty into correlated and uncorrelated components. For the equations described here to work, the correlated components must be the same, i.e., they must apply to all measurements equally. Thus the problem reduces to identifying a single correlated uncertainty for all towers, and a set of uncorrelated uncertainties, one for each tower. The uncorrelated uncertainty for each tower may itself be a combination of uncorrelated uncertainties, but they do not need to be defined individually.
[4] The process is best carried out for a specified turbine layout, or after first limiting the buildable area to prospective turbine locations. Otherwise, new masts may be proposed in clearly undesirable locations, such as valleys, merely because their resource is very different from that of the existing towers.