ASSESSING THE REPRESENTATIVENESS OF WIND DATA
FOR WIND TURBINE SITE EVALUATION

D. S. Ranné
Pacific Northwest Laboratory
Richland, Washington 99352

and

R. B. Corotis
Northwestern University
Evanston, Illinois 60201

ABSTRACT

Once potential wind turbine sites (either for single installations or clusters) have been identified through siting procedures, actual evaluation of the sites must commence. This evaluation is needed to obtain estimates of wind turbine performance and to identify hazards to the machine from the turbulence component of the atmosphere. These estimates allow for more detailed project planning and for preliminary financing arrangements to be secured. The site evaluation process can occur in two stages: 1) utilizing existing nearby data, and 2) establishing and monitoring an onsite measurement program. Since step (2) requires a period of at least 1 yr or more from the time a potential site has been identified, step (1) is often an essential stage in the preliminary evaluation process. Both the methods that have been developed and the unknowns that still exist in assessing the representativeness of available data to a nearby wind turbine site are discussed in this paper. This paper then discusses how the assessment of the representativeness of available data can be used to develop a more effective onsite meteorological measurement program.

1.0 INTRODUCTION

Successful site selection activities, described in another paper at this conference [1], should lead to the identification of one or more "candidate" turbine or turbine cluster installation sites. Once these candidate sites are identified, evaluation of the wind characteristics at the sites must commence to allow for project planning and for preliminary financing arrangements. This paper discusses our current state
of knowledge of the site evaluation process as it relates to installation of large machine clusters, and points out areas where additional work is required.

In general, candidate sites will not have sufficient onsite documentation of wind characteristics to be completely adequate for planning the installation. The developer of the installation must then make use of existing data sources, and determine the representativeness of those data to the candidate site. A number of techniques may have to be applied, including numerical simulations, statistical procedures, and even subjective judgment, to estimate the wind characteristics at the candidate sites. These estimates may give early information on the projected performance of a turbine or turbine cluster at the candidate site; although the lack of onsite information will probably mean that the developer can not place the highest possible confidence on the performance projections. In addition, there may be certain turbulent characteristics of the atmosphere at the candidate site that only an onsite measurement program could identify adequately.

Another important reason for the developer or planner to assess the representativeness of existing data to a candidate site is that a more intelligent and cost-effective onsite measurement strategy can be developed. In some cases, nearby data may be sufficiently representative to eliminate the need for onsite measurements at all; in other cases, an extensive measurement system (including multiple towers if it is a large cluster site) may be necessary before sufficient information can be derived from a site to allow the project to proceed.

Basically, the following must be taken into consideration in assessing the representativeness of an existing data set to a nearby candidate site:

1. The degree to which the data represents the interannual variability of diurnal, seasonal, and annual wind statistics at the site.

2. The degree to which the data represents actual conditions at the site given geographical and topographical variability between the data source and the site.

3. The degree to which there would be impacts on the representativeness of existing data occurring due to modifications in the lower atmospheric boundary layer as a result of an installation of a large cluster of turbines.

The focus of the first section of this paper is on (1) and (2); understanding of (3) is very limited at present and should be a major thrust of research in the future. Following this first section, the results of recent research in determining the most appropriate onsite measurement strategy for site evaluation will be presented.
2.0 DETERMINING THE REPRESENTATIVENESS OF EXISTING DATA SETS

2.1 Climatological Adjustment of Short-Term Data

In many cases, decisions to develop a wind turbine project will have to be based on the existence of wind data that covers only a short period of time. The data may have been collected by the developer of the wind project, so that most likely there would not have been time to wait until a full climatology of the wind characteristics at the site has been established. In other cases, nearby data collected for some other purpose may be available, but again will not have had sufficient length of record to discern interannual variability. Since the wind project developer, as well as finance institutions that provide the developer with working capital, need to know variations in year-to-year cash flow that the project will experience, and in particular need to be aware of potentially serious cash flow situations that would occur during one or several consecutive low-wind years, knowledge of the interannual variability is extremely important in the planning stages of a project.

A number of studies have been made to adjust short-term measurements to long-term climatological values. These studies include statistical procedures for relating short-term data to long-term records that are near the site, such as National Weather Service stations, time series analysis of short-term data in comparison with nearby long-term weather records, studies of synoptic climatology patterns, and use of numerical models.

2.1.1 Wind Variability Statistics

Extensive research into using statistical and time series procedures for determining the representativeness of short-term to long-term (e.g., 30 yr) means, and the variance of the mean has been made [2,3,4,5,6]. These studies are summarized by Hiester and Pennell [7].

Corotis [5], for example, assesses the reliability of a computed mean speed with a confidence interval. If \( \bar{V} \) and \( \sigma_v \) are the computed mean and hourly standard deviation of wind speed for a period of time, such as a season or a year, then there is a probability of \((1-\alpha)\) that the true long-term mean wind speed for that period of time is bounded by:

\[
\text{Probability} = \bar{V} + k_{(1-\alpha)/2} \sigma_v/\sqrt{n}
\]

where \( k_{(1-\alpha)/2} \) is the standard normal deviate evaluated at a cumulative level of \((1-\alpha/2)\), and \( n \) is the number of equivalent independent hourly readings on which \( \bar{V} \) is based. Values of \( k_{(1-\alpha)/2} \) are readily available from standard statistics textbooks. This study, as well as others referenced above, all conclude that the following general guideline applies: "...the climatic mean wind speed will be within \pm 10\% of a single mean wind speed observation with about 90\% confidence." In general, this same guideline applies to a single season of measurements being within \pm 10\% of the climatic seasonal average with 90\% confidence.
Obviously, this guideline does not apply to all sites. In fact, the accuracy of the estimate at any given site depends on the coefficient of variation of the site annual mean wind speed \( \sigma_a/\bar{V} \), where \( \bar{V}_a \) is the climatic mean and \( \sigma_a \) is the standard deviation of annual means. Information on how \( \sigma_a/\bar{V}_a \) can provide more precise error ranges on a long-term estimate from a single monthly or annual value is demonstrated in [3]. Figure 1 shows their results of the average distribution of monthly and annual wind variations for a survey of 40 National Weather Service stations. This figure shows that on the average a single observation of annual mean wind speed falls within the interval of approximately 0.8 \( V \) and 1.16 \( V \) with 90% certainty. They then looked at the distribution of \( \sigma_m/\bar{V}_m \) for monthly values and \( \sigma_a/\bar{V}_a \) for annual values. The variability distributions of various percentiles of \( \sigma/\bar{V} \) values are shown in Figure 2 (for annual) and Figure 3 (for monthly). Thus, for sites where it is assumed extreme variability occurs, the 90 percentile \( \sigma/\bar{V} \) values would be within about \( \pm 15\% \) of a given single annual or \( \pm 18\% \) of a monthly value.

![Figure 1. Average distribution of monthly and annual mean winds about the long-term mean (from [3]).](image)

This information is also important in developing an onsite measurement strategy which can give an indication of the long-term mean using only intermittent measurements. These strategies will be discussed in Section 3.2.1.

2.1.2 Climatological Adjustment Using a Reference Station

A number of attempts have been made to develop reliable techniques for adjusting short-term (season or year) data to climatological values by

(a) In this report, subscript "a" refers to annual values, and subscript "m" refers to monthly values.

176
FIGURE 2. AVERAGE DISTRIBUTION OF ANNUAL-MEAN WINDS ABOUT THE LONG-TERM MEAN FOR VARIOUS $\sigma/V$ VALUES, ASSUMING A GAUSSIAN DISTRIBUTION OF $V/V$ (FROM [3]).

FIGURE 3. AVERAGE DISTRIBUTION OF MONTHLY MEAN WINDS ABOUT THE LONG-TERM MEAN FOR VARIOUS $\sigma/V$ VALUES, ASSUMING A GAUSSIAN DISTRIBUTION OF $V/V$ (FROM [3]).
comparison with nearby long-term data. These techniques range from evaluating the ratios of the data sets to applying statistical or time series climatological relations [8, 3, 5, 9]. For example, Corotis [5] suggests use of the statistical relation:

\[
\bar{V}_c = V_c - \rho (\bar{V}_r - \bar{V}_r) \frac{\sigma_c}{\sigma_r}
\]  \[2\]

where \(\bar{V}_c\) = estimated long-term mean at the candidate site

\(V_c\) = observed mean at the candidate site

\(V_r\) = observed mean at the reference site during the period of candidate site measurements

\(\bar{V}_r\) = observed long-term mean at the reference site

\(\rho\) = spatial cross-correlation between sites during period of candidate site measurements

\(\sigma_c, \sigma_r\) = hourly standard deviations for candidate and reference sites, respectively.

However, because investigations have shown that monthly and annual cross correlations between nearby stations is relatively low, even in simple terrain, ample evidence exists to show that adjustment of annual or monthly mean wind speeds using relations such as (2) or the simple ratio method does not result in significant and reliable improvement in estimates of climatic means over using the guidelines presented in the previous section.

Recent and ongoing research show that synoptic climatology methods may also provide useful ways of characterizing short-term data sets to a long-term climatology. For example, Ossenbrugen et al. [9], in a study of offshore wind power potential, compared the 3-yr Boston Lightship records with the 31-yr Boston Logan Airport records by stratifying the latter into various categories of "wind seasons". By applying statistical tests between comparative data sets, and accounting for serial correlation between observations, the authors were able to adjust the Lightship data to the longer term Logan Airport data to within specified confidence limits.

2.2 Adjusting Nearby Data to a Site

2.2.1 Numerical Methods

In many cases, there will be no data available on the candidate site at all. This leaves the wind project developer with the dual problem of first interpolating nearby data to his site accurately, and then adjusting the interpolated values to long-term climatological statistics. Recently several studies were completed to produce methods of estimating
wind characteristics at a candidate site using only available nearby National Weather Service data [10,11,12,13]. These studies ranged from the application of sophisticated numerical objective analysis models to simpler interpolation techniques. All of the studies incorporated the effects of topography on wind flow between the National Weather Service stations and the site. Topography is incorporated in the numerical schemes by introducing topographic features on the computational grid and enforcing mass conservation in the calculations. The interpolation scheme is adjusted for topography by utilizing weighting factors.

A sample of the results of a test of one of the numerical methods is shown here. This method is an improvement of a modeling technique developed by Bhumralkar [13] following a concept of atmospheric flow modeling in the lower atmosphere developed by Sherman [14]. The details of the methodology and test procedure are described by Endlich et al. [15]. Basically, the model operates in a sigma coordinate system. The wind flow model is used to solve the continuity equation for each set of representative wind vectors (eigenvectors). A simulated wind history can then be produced at the candidate site, and from this simulation the statistics of the winds at the candidate site can be estimated.

The results of a test of the model at Clayton, New Mexico where nearby National Weather Service data is used as model input and compared with simultaneous wind measurements at the Clayton MOD-0A site is shown in Figures 4 and 5. Figure 4 gives a map of the location of the Clayton site, as well as locations of the available National Weather Service sites. Figure 5 gives comparisons of seasonal and diurnal values of simulations and observations.

In general, the simulated results are lower than observation for this case. Tests at other locations give similar results—only occasionally do simulations actually give higher statistics than observations. Despite these deficiencies, the simulations produce seasonal and diurnal patterns comparable with the observations. This does not necessarily occur at some of the more complex terrain sites, where National Weather Service stations are in valleys and candidate sites are on mountain tops. Nevertheless, modifications to this modeling scheme have produced improved results from earlier tests. This approach can be useful for estimating wind characteristics at a site with no known wind information. Results of these model applications should only be used as a means of providing preliminary guidance on the wind characteristics at a candidate site.

2.2.2 Use of Spatial Data Arrays

Suppose that some knowledge of the variance of wind speed is available in the region around a site, and that it is desirable to determine the representativeness of nearby data sets to a site. This type of information may be available, for example, if some preliminary site survey work had been done over the entire region in which the candidate site had been selected. Such data are often available in areas where a known resource exists and extensive measurements were established
FIGURE 4. LOCATION OF THE CLAYTON, NM MOD-OA SITE AND NATIONAL WEATHER SERVICE STATIONS USED IN THE SRI MODEL (FROM [15])

FIGURE 5. SEASONAL AND DIURNAL WIND SPEED CURVES FOR CLAYTON, NM AS SIMULATED BY THE SRI WINDFLOW MODEL AND AS MEASURED AT THE SITE (FROM [15])
because of an interest in wind energy development. Under these circumstances the principles of Section 2.1.1 can be applied to establish confidence limits on the representativeness of a given data station to the site. Rewriting Equation [1]:

\[ \bar{V}_s = k \left( \frac{\sigma_s}{\sqrt{n}} \right) \]  \hspace{1cm} [3]

where \( \bar{V}_s \) = spatial average of mean wind speed values
\( \sigma_s \) = standard deviation of individual site mean values to the spatial average.

The total standard deviation, accounting for spatial and temporal variability, is given by:

\[ \sigma_T = \left( \sigma_v^2 + \sigma_s^2 \right)^{\frac{1}{2}} \]  \hspace{1cm} [4]

Based on guidelines established from a survey of National Weather Service stations around the U.S. by Corotis [2]:

\[ \sigma_v = 0.1 \frac{\bar{V}}{k_{90\%}} \]  \hspace{1cm} [5]

Then, for one year of measurements:

\[ 90\% \text{ confidence} = \bar{V} \pm 0.1 \left( \frac{\sigma_T}{\sigma_v} \right) \]  \hspace{1cm} [6]

Thus, Equation [6] states that, in general, inclusion of spatial variability of mean wind speeds with climatological adjustment increases our uncertainty of the true long-term mean wind speed. Conversely, we have less confidence that the true long-term mean wind speed is \( \pm 10\% \) of the nearby measured mean winds.

For other confidence limits, [6] can be written in a more general form:

\[ \text{Confidence interval (x)} = \bar{V} \pm 0.1 \left( \frac{\sigma_T}{\sigma_v} \right) \left( \frac{k_x}{k_{90\%}} \right) \]  \hspace{1cm} [7]

3.0 ESTABLISHING A SITE MEASUREMENT STRATEGY

In the previous sections we have explored a number of techniques that can be used to estimate wind characteristics at a potential turbine or turbine cluster site where little or no onsite data exist. However, in many cases the information gained from employing these techniques is adequate only for initial planning purposes. It will not provide sufficient reliability or detail on wind characteristics at the proposed
site to complete detailed cluster design and to evaluate the economics of the cluster performance. Consequently, representative onsite measurements will still be required. In this section, we will explore factors that need to be considered for an effective onsite measurement strategy and review some recent research intended to shed a better light on a meaningful site evaluation process.

3.1 Current Site Evaluation Technology

In broad terms, a site evaluation strategy should be designed to build an information based on what is already known about a site. It should be structured to acquire needed information in a cost-effective and efficient manner. Guidelines have been published in documents such as siting handbooks [6,7], and experience is being gained through meteorological measurements at the Department of Energy's candidate site and turbine test site programs. The current strategy at DOE's candidate sites is described here as well as in reference [17].

Three levels of high response wind speed and direction cup-and-vane type sensors are installed on a 48.8-m guyed meteorological tower at heights of 9.1, 30.5, and 45.7 m. Although current technology is producing large wind turbines with heights considerably taller than these towers, the decision was made to install a shorter tower to conserve costs due to the large number of sites. (Currently, there are 34, including 6 with large wind turbines installed by DOE for field testing). Furthermore, it is assumed that the three levels of measurements will allow a reasonably accurate extrapolation of wind characteristics upward to a higher hub height. In addition, costs of purchasing, installing, and maintaining meteorological towers taller than 61 m (200ft) increase significantly because of FAA requirements to install safety beacon lights. Nevertheless, at a given site, a tower with sensors at hub height or above, particularly in areas of complex terrain, may well be worth the additional costs so that improved estimates of turbine or turbine cluster performance can be obtained.

At DOE's candidate sites, instantaneous samples of wind speed and direction are recorded once every 2 min on a digital cassette data logger. For most data loggers available on the market, this allows at least 2 wk of measurements to be made before the tape must be removed and replaced. Computerized monthly summaries of each site's wind characteristics are prepared from these cassettes. The summaries include such information as data recovery rates, mean wind speed and resultant wind direction for each level, peak gusts, diurnal wind characteristics, frequency distribution, power law coefficients by direction, and turbulent intensities.

3.2 Research on Alternative Measurement Strategies

3.2.1 Intermittent Measurement Strategies

In cases where funds for site measurements are limited and numerous sites need to be evaluated, Ramsdell et al. [6] have investigated the
value of intermittent measurement strategies at a site, where measurement equipment would be installed for only a few months out of the year, and then reinstalled at other sites. The equipment would be rotated among the sites so that eventually all sites would have a few months of measurements each year. Such a strategy may be particularly significant for small machine applications, especially in cases where the time required to obtain sufficient measurements to have climatological significance is not critical.

The authors define a relative uncertainty as a way of relating short-term measurements at a site to long-term climatological averages. The relative uncertainty is the ratio of the standard deviations of estimates of long-term mean wind using individual monthly averages of a specific intermittent measurement strategy to that of a continuous measurement strategy. Based on an examination of 40 data sets, Figure 6 shows how the relative uncertainty decreases as the duration of monthly measurements increases. The figure shows that little reduction in uncertainty occurs beyond about 24 mo of measurements.

![Graph showing relative uncertainty vs. duration of measurement program in months.](image)

**Figure 6. Relative Uncertainties of Continuous Measurement Strategies (from [6]).**

Figure 7, also taken from Ramsdell et al. [6] shows how an intermittent measurement strategy can actually decrease the relative uncertainty of establishing a long-term climatology at the site. For an equal amount of time of equipment usage, the authors show that the samples obtained using an intermittent strategy are in essence independent, thus improving the certainty of having acquired a representative sample of data for establishing a long-term mean wind speed at the site. Of course, an intermittent strategy requires more actual time to develop a climatology. Thus, if an intermittent strategy is employed, Figure 6 shows that nearly twice the measurement period is required to obtain the same relative uncertainty than if a continuous strategy were used.

The study by Ramsdell et al. also confirms the conclusion of the climatological representativeness of short-term measurements discussed
in Section 2.1.1. For the data sets used by the authors, a relative uncertainty of 0.35 for 12 mo of continuous measurements was obtained. Figure 8 gives the relationship between relative uncertainty and percent error in estimates of the long-term mean from their data. At a relative uncertainty of 0.35, the figure shows that there is a 90% confidence that the long-term mean wind speed \( \bar{u} \) within \( \pm 10\% \) of the continuous measurements for 1 yr.

3.2.2 Site Evaluation at DOE's Turbine Test Sites

At the DOE MOD-0A 200 kW wind turbine test sites, the meteorological towers that had been installed prior to the installation of the turbine for site evaluation purposes have been retained to support this phase of the research program. At each of the sites the towers are located less than 1 m from the turbine, and are situated such that they are measuring essentially the free-stream meteorological conditions at the site. Besides the routine measurement program discussed earlier, turbine output parameters are simultaneously recorded on the cassette data logger. These parameters include turbine electrical power output, nacelle yaw error, and wind speed and direction recorded from the

(a)Located at Clayton, New Mexico; Block Island, Rhode Island; Culebra, Puerto Rico; and Kahuku, Oahu, Hawaii.
nacelle anemometer. In addition, the data logger has the capability of recording short bursts of high frequency data (∼ 1 sample/sec). This capability allows research into the most appropriate sampling strategy (average time per sample and frequency of samples) that should be undertaken as part of a site evaluation process at candidate sites.

Several hours of high speed data collected at the Clayton, New Mexico MOD-OA site on August 22, 1980 at a time when the turbine was operating between the cut-in and rated wind speed values is being examined as part of this site evaluation research effort. Preliminary results of this analysis are presented here to give an indication of the type of sampling strategy that might be appropriate at a site being considered for a turbine or turbine cluster installation.

Figure 9 shows the autocorrelation of 1-sec average values obtained from the three levels of the meteorological tower. In general there is an exponential decrease in autocorrelation with time between samples, occurring more rapidly at the lower levels where turbulent fluctuations in the surface layer due to friction are greater. As the correlation approaches 0, the samples become more and more independent. Since only independent samples enter into a climatological average, this figure implies that, when collecting instantaneous (i.e., 1-sec) samples to obtain a climatology of wind characteristics at a site, no more than
one sample per minute is necessary, and one sample every 2 min is adequate. This is comparable to the sampling strategy used at DOE's candidate sites.

As the sample averaging time is increased, the reduction in autocorrelation with lag time occurs at a slower rate. This is exemplified in Figure 10, which shows data from the 30-m level of the meteorological tower. The obvious conclusion from this figure is that a longer period of time between recording intervals is required as the averaging time increases to obtain independence between samples. It is interesting, however, that the autocorrelations converge at a sampling interval of 3 to 4 min, regardless of averaging time.

It is also of interest to examine cross correlation of measurements between the tower and the turbine parameters. Cross-correlation analysis sheds light on how various sampling strategies at the tower can explain variation of turbine power output. Figure 11 compares measurements at the 30-m level on the tower with measurements from the nacelle anemometer. The displacement in lag from the center of the figures shows the effect of travel time for small scale eddies from the tower to the turbine (the tower is approximately 100 m upwind of the turbine). As the averaging time increases, the cross correlation increases since smaller scale eddies are averaged out of the calculations. However, even for relatively long averaging periods, a perfect correlation is not obtained.

The same pattern is evident, but correlations are significantly lower, when the tower values are compared with turbine power output (Figure 12).
FIGURE 10. AUTOCORRELATION FOR VARIOUS AVERAGING INTERVALS AT THE 30-m LEVEL OF THE METEOROLOGICAL TOWER AT CLAYTON, NM, AUGUST 22, 1980

The obvious conclusion is that the variance in turbine power output cannot be entirely explained by variances in short-term wind speed fluctuations from the type of anemometers used. Other parameters, such as wind direction, are also important. In addition, longer averaging times may also be appropriate.

The preliminary conclusion from all this is terms of site evaluation measurement strategies is that near-instantaneous samples once every few minutes are adequate for obtaining a site wind climatology, but some type of sample averaging is appropriate to better estimate turbine energy production. In addition, other parameters besides wind speed fluctuations are needed to better relate wind observations to turbine performance. These parameters are probably wind direction fluctuations and perhaps fluctuations in atmospheric density. Nonmeteorological factors, such as the interface between the turbine and the electric grid, may also come into play.

4.0 SUMMARY AND CONCLUSIONS

Perhaps the best way to summarize this paper is by an examination of Figure 13. The goal is for a wind project developer to have an accurate understanding of the performance of the turbine or turbine cluster on his proposed site. The "performance" includes energy production from the array as well as operation and maintenance costs. Determination of this performance depends on a large extent on...
an appropriate site evaluation measurement strategy. This paper has discussed how such a measurement strategy can be defined. First, a number of techniques are available to utilize existing data or short term (including intermittent) onsite measurements to obtain general knowledge of wind characteristics at the site. But in order to estimate cluster performance, actual onsite measurement programs must be established that reflects knowledge of turbine operating strategies, as well as factors relating to the interface between the turbines and the electric utility system. This nonmeteorological information can most likely be provided by manufacturers and utility personnel.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy under contract DE-AC06-76RLO 1830. The authors gratefully acknowledge the efforts of D. L. Atkin and J. M. Reeder during the preparation of the manuscript for publication.

REFERENCES


