Optimal Site Selection and Sizing of Distributed Utility-Scale Wind Power Plants

Michael R. Milligan  
National Renewable Energy Laboratory

Rory Artig  
Minnesota Department of Public Service

Presented at  
International Association of Energy Economists  
21st Annual International Conference  
Quebec City, Canada  
May 13–16, 1998

National Renewable Energy Laboratory  
1617 Cole Boulevard  
Golden, Colorado  80401-3393  
A national laboratory of the U.S. Department of Energy  
Managed by Midwest Research Institute  
for the U.S. Department of Energy  
under contract No. DE-AC36-83CH10093

Work performed under task number WE802020

April 1998
NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

Available to DOE and DOE contractors from:
Office of Scientific and Technical Information (OSTI)
P.O. Box 62
Oak Ridge, TN 37831
Prices available by calling (423) 576-8401

Available to the public from:
National Technical Information Service (NTIS)
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
(703) 487-4650

Printed on paper containing at least 50% wastepaper, including 20% postconsumer waste
1. INTRODUCTION

The utility industry in the United States is clearly moving towards a restructured market in which the traditional vertical integration of the industry may become all but obsolete. The wholesale markets for electricity will also undergo change, with a larger number of players and institutions and more complicated transactions and contract paths between seller and buyer. As electric market product unbundling occurs, sellers in the wholesale market for electricity will find it to their advantage to be able to specify the quantity of electricity available and the time of availability. Since wind power plants are driven by the stochastic nature of the wind itself, this can present difficulties to a wind plant operator. A wind plant operator who contracts for a sale of energy during periods of variable wind power output might be required to pay a significant penalty if the actual power diverges significantly from the level specified by contract. In previous work Milligan, Miller, and Chapman (1995) provided estimates of the benefit of accurate wind forecasting to the utility. To the extent that an accurate forecast is available, contract deviations, and therefore penalties, can be significantly reduced.

Generating capacity that is available during the peak is worth more than off-peak capacity. Even though we might have the ability to accurately forecast the availability of wind power, it might not be available during enough of the peak period to provide sufficient value. However, if the wind power plant is developed over geographically disperse locations, the timing and availability of wind power from these multiple sources could provide a better match with the utility's peak load than a single site.

There are several wind plants in various stages of planning or development in the United States. Although some of these are small-scale demonstration projects, significant wind capacity has been developed in Minnesota, with additional developments planned in Wyoming and Iowa. As these and other projects are planned and developed, there is a need to perform analysis of the value of geographically diverse sites on the efficiency of the overall wind plant.

In this paper, we use hourly wind-speed data from six geographically diverse sites collected by the Minnesota Department of Public Service to provide some insight into the potential benefits of disperse wind plant development. We provide hourly wind power from each of these sites to an electric reliability simulation model. This model uses generating plant characteristics of the generators within the state of Minnesota to calculate various reliability indices. Since we lack data on wholesale power transactions, we do not include them in our analysis, and we reduce the hourly load data accordingly. We present and compare results of our methods and suggest some areas of future research.
2. THE POTENTIAL BENEFIT OF WIND POWER PLANT GEOGRAPHIC DIVERSITY

Wind can be described as a stochastic process. As such, the power output from a wind power plant can vary substantially through time, and it is not controllable in the same way as conventional power plants. During lulls in the wind, electricity must be supplied by other resources. If geographically diverse wind sites can be chosen in such a way as to minimize the number or extent of wind power lulls, this can be beneficial because conventional resource use can be reduced accordingly. The extent of this reduction will influence the magnitude of fuel cost, operations and maintenance, and other costs to the utility. Of course the increase in wind power generation is not without additional costs, such as O&M.

One of the first comprehensive studies to address the issue of geographical diversity of wind plants was done by Kahn (1979), who used California wind and utility data. He found that reliability does increase as a function of geographic dispersal, but this increase is limited by the geographic wind diversity and the barrier of large wind plant penetrations relative to the conventional generator mix. Kahn also points out that wind sites that are uncorrelated will generally provide better combined reliability than sites that are highly correlated, in absolute value. However, Kahn’s analysis ignores the fact that two or more wind regimes with significantly different time-scale properties can both provide the same correlation with utility load (see Milligan and Artig, 1998). A study by Brower (1993) found some benefits to distributed wind development in Minnesota, but the benefits were somewhat constrained by the relatively high wind-speed correlation between wind sites.

3. STATE OF MINNESOTA DATA COLLECTION PROJECT

The wind resource data used in this study were collected through the Minnesota Department of Public Service's (DPS) wind resource assessment programs and the DPS/U.S. Department of Energy (DOE) Tall Tower Wind Shear Study.

DPS has conducted wind resource assessment since the early 1980's, providing utilities, developers, and other interested persons with wind data collected at sites around the state. Since the programs began, DPS has expanded and improved the data collection process by adding new monitoring sites and more sophisticated equipment.

The monitoring sites which provided data for this paper are equipped with cellular data loggers that automatically send the collected information to a base station computer located in DPS offices. These sites use existing communication towers and have monitoring levels at 30, 50, and 70 meters above ground level. Two anemometers are mounted at each level, one on each side of the tower. This configuration has a number of advantages. It reduces the wind shadow effect the tower would have on the data if only one anemometer were used at each level; it provides a degree of redundancy at each level so that the failure of one sensor does not eliminate the data collection at that level; and it provides the opportunity to do sensor-to-sensor calibration and helps diagnose potential problems with sensors. Each tower is also equipped with wind vanes at the 30 and 70 meter levels. In addition to the internal logger temperature, some of the sites are equipped with external temperature probes mounted at approximately 4 meters above ground level.

In 1996, DPS, in cooperation with DOE, installed four advanced monitoring sites as part of the DPS/DOE Tall Tower Wind Shear Study. These sites use existing communication towers and have monitoring levels at 10, 30, 40, 50, 60, and 70 meters above ground. As with the sites described above, each monitoring level has two anemometers, one mounted on each side of the tower. Each tower is also equipped with wind directional sensors at the 10, 30, 60, and 70 meter levels.
4. WIND SITE SELECTION

We chose six wind sites that, in our judgement, did the best job of representing the diversity of climatology in the state. The sites we selected are Alberta, Becker City, Brewster, Crookston, Currie, and Luverne. They appear in Figure 1, and are identified by the first two letters of the respective site name. One of the unique features of the Minnesota DPS data collection effort is that wind data is collected to a height of 70 meters. This made it possible for us to use the power curve of a modern utility-scale turbine at a hub-height of 65 meters. A comparison of output at different hub-heights appears in Milligan and Artig (1998). We used wind data for one year beginning November, 1995.

For many utilities, simply maximizing wind energy capture during the system peak will not necessarily result in maximum benefit. The reason is that generating units’ ramp-rates and minimum-run levels may not allow for full utilization of the wind power if there is significant variation in the timing of wind power delivery to the grid. In addition, it is possible that the ability to accurately forecast hourly wind speeds and the consequent hourly wind power output is somewhat impaired by wind sites with high hourly variability. For a more thorough discussion of the relationship between wind forecasting and wind capacity credit, see Milligan, Miller, and Chapman (1995). For these reasons, it may be to the utility’s advantage to install wind plants and select among sites in such a way that hour-to-hour output variations between wind plants are reduced, while still obtaining as much wind power output as possible during the peak period.

Figures 2 and 3 illustrate some potential benefit from different sites. The graphs are based on real wind data which was used to calculate output of fictitious wind power plants, each with 100 MW of installed capacity. The hourly wind power is calculated by taking actual wind-speed data and applying calculating power output based on a utility-scale wind turbine, as described more fully below. Figure 2 shows a 48-hour period for three of the sites used in this study: Alberta, Currie, and Luverne. During the first day, the highest output comes from the Currie site, and Alberta and Luverne show low power output. The second day, wind power output at Currie is low, whereas power output at both Alberta and Luverne reach maximum rated output (less losses). Power output at Luverne drops about 4-5 hours earlier than at Alberta late in the second day. Figure 3 shows another view of multiple-site wind power output. This 24-hour period shows a general correlation between sites, yet one can also detect time-lags of 2-3 hours early and late in the day. These variations are somewhat typical of the data used in this study and illustrate the potential benefit of geographically diverse wind power plants.

Variations within a given wind plant are well-known, providing some measure of diversity within even a relatively small geographic area. The extent to which this intra-site variability would influence geographic optimization is unclear. However, some of this intra-site variability is because of local turbulence in what is known as the turbulent scale of the relative spectral intensity of the wind-speed (Stull, 1988). For electrical reliability, we must especially be concerned with the

Figure 1. Wind sites chosen for this study
5. MODELING

Electrical reliability is a function of customer demand and the characteristics of the various generators. Utilities experience a pronounced peak period, often during several hours of the day during a particular season. A utility can sometimes dramatically increase its generator reliability by installing peaking units to generate power when it is most needed: during the peak hours. Even though these peaking units might be available at night, their availability at night would likely have a negligible effect on system reliability. Likewise, a wind plant that delivers a significantly higher annual energy output does not necessarily significantly contribute to system reliability. What is needed is for the wind output to occur at times of otherwise high-risk periods during system peak.

Kahn's (1978) discussion centers around the statistical correlation between the various wind-plant sites. Sites with high positive correlation will provide higher output during the same time periods, whereas sites with high negative correlation will be complementary. When the first site is providing a high level of electrical output to the grid, the second site will likely be idle. Conversely, when the first site is not producing electricity, it is likely that the second site is. Milligan and Artig (1998) borrow from the analysis in Milligan and Parsons (1997), applying several methods that do not require the use of a production cost or reliability model.

Reliability of the electrical system can be calculated by a production cost/reliability model. The model we used is a load-duration curve model, Elfin, and is produced by the Environmental Defense Fund. The model uses hourly electrical load data and generation data to calculate the optimal mix of generating resources required to serve the load. The hourly loads are arranged by subperiod and sorted, then arranged into a cumulative probability distribution. Generating resources are then matched to the load based on merit order, which is usually least-cost dispatch. Since each generator has a probability of failure, its dispatch is uncertain. The model takes this into account during the dispatch simulation, and various calculated indices can then summarize system reliability. The two most common reliability measures are loss-of-load expectation (LOLE) and expected energy not served (ENS). Figure 4 shows a
normalized load duration curve for the state of Minnesota. After all available generation has been dispatched, there may be a small spike at the top of the curve. The vertical distance from the curve to the 100% level of peak represents the expected loss of load, and the area of the spike representing the energy unserved. More details can be found in Milligan (1996a).

The precise form of the optimal wind-plant reliability problem may vary according to the customer loads, wind sites, and characteristics of the utility system. For example, a combination of wind regimes that all exhibit diurnal variation may provide the utility an opportunity to select a combination of sites that together have a high probability of offsetting system peak requirements. Such a scenario might involve calculating various probability levels of wind generation during several peak hours every day of the peak month(s). Alternatively, when the wind resources do not follow a pronounced diurnal pattern, the utility might be more interested in looking at the overall probability levels of wind generation during the month, without necessarily allowing for a repetitive daily pattern in wind generation.

Choosing the best combination of wind sites can be done with a number of objective functions, depending on what the decision-makers believe is most relevant. Among these are (1) least-cost combinations of wind sites, (2) wind sites that minimize load swings during system peak, or (3) most reliable wind sites. The goals of least-cost production and most-reliable production are usually not consistent with each other. Reliability must be traded off against cost because a perfectly reliable system (if one were to exist) would not be cost-effective. Likewise, a least-cost solution might result in a generating system that does not possess sufficient reliability. For this analysis, we chose to pursue an optimization based on reliability. The results can be easily extended to capacity credit or other parameters of interest.

5.1. Marginal Reliability Method

The first method is based on traditional marginal analysis. Although this approach is widely used in economics, it is difficult to apply to the problem at hand. The difficulty is that the marginal reliability of a wind plant is potentially different in each hour of the year, depending on customer load and other generating resource availability. For example, we could calculate the reliability of wind power plants at Brewster and Currie. If Brewster appears to be the best choice based on its marginal reliability, we can add 100 MW. However, now that we have 100 MW at Brewster, the marginal reliability of Currie will be very different than during the original comparison with Brewster. We used this method, and we compare the results to those of our preferred methods, below.

Since we do not have specific price information on wind development and production at these sites, we assume that the installed cost in $/kW or $/MW is the same at all sites. However, there is a difference in efficiency and reliability between sites because of the different winds experienced at each of the 6 different locations. We can describe the reliability level as a function of installed MW at each site:

\[ c_r = \Phi(x_r) \]  \hspace{1cm} (1)
where \( c_i \) = reliable capacity, as measured as a function of \( 1/r \), \( r \) is the reliability measure of choice: either LOLE or ENS, \( x_i \) is the rated installed power capacity of the wind plant, and \( i \) is the subscript of the wind plant location, and \( 1 \leq i \leq 6 \). Finding the optimal mix of resources assuming the same price/kW at each site implies that we choose the quantity of wind resources up to the point at which the marginal products of each site are equivalent (Varian, 1978). Written in terms of partial derivatives we have

\[
\frac{\partial \Phi(x_i)}{\partial x_i} = \frac{\partial \Phi(x_j)}{\partial x_j}
\]

(2)

\( \forall i \) and \( j \) such that \( 1 \leq i, j \leq 6 \).

Figures 5 and 6 show the reliability curves for two sites, selected to illustrate variations in reliability. The y-axis shows reliability, as measured by the marginal energy reliability index which is the marginal ENS, scaled to the interval (0,1) for convenience. From the diagram it is apparent that Brewster is clearly the more desirable site based on its marginal reliability. Becker was not chosen by this (or any other) algorithm.

For each of the six wind sites, we made a series of model runs with Elfin, adding 25 MW at a time from each site until the maximum of 500 MW was "built" at each of the sites. We then chose the combination of sites that satisfied equation (2), using the marginal ENS as the reliability measure, given the constraint of 500 MW total installed wind capacity. The optimal choice from this method is 75 MW at Alberta, 150 MW at Brewster, 125 MW at Currie, and 150 MW at Luverne. The algorithm did not select either the Becker or Crookston sites because the associated reliability curves are much lower than those of the four selected sites.

![Figure 5. Marginal reliability curve for Becker](image1)

![Figure 6. Marginal reliability curve for Brewster](image2)

5.2. Optimization with Elfin

This approach uses the production-cost/reliability model in such a way as to do a step-wise modification of the net remaining load after each incremental wind plant is built. This method could perhaps be best understood by referring to Figure 7. There are two possible variations of this approach. The first uses LOLE as the optimization parameter. The second variation uses ENS as the optimization parameter. The step in the diagram that illustrates the choice of site with the best reliability parameter instructs us to choose the wind site with the lowest parameter, either LOLE or ENS, because higher values of LOLE and ENS represent less reliable systems. The step entitled "Build an X MW wind plant" implies that the incremental size of the plant to simulate building can be varied. In our case, we chose X = 25 MW as a reasonable trade-off between accuracy and model run-time. Smaller values of X might be more...
accurate, although given the relative scale of X to the hourly loads, we don't think so. Large values of X compromise the results because the optimization algorithm is restricted to large increments of wind capacity, possibly overshooting a better mix of sites.

Following the outline of Figure 7, this is the optimization. This discussion focuses on the use of ENS as the reliability parameter, but the process is the same when we use LOLE as the optimization parameter. First, we run Elfin without any wind plants. The next step is to run Elfin for a block of 25 MW of installed wind capacity at each site, separately. We compare the ENS calculation at each site, and choose the site with the lowest ENS (best reliability). The process then simulates the building of 25 MW of wind capacity at the chosen site, and this becomes the new base case. We repeat the process, running Elfin for each site combined with the chosen site from the previous step. Choose 25 MW from the site with the best ENS, and repeat until all 500 MW of wind has been installed. The algorithm then simply counts the number of 25-MW increments of wind plants added at each of the sites, and that is our result.

The optimization using LOLE selects 250 MW at Brewster, 225 MW at Currie, and 25 MW at Luverne. The ENS optimization selects 450 MW at Brewster, 25 MW at Currie, and 25 MW at Luverne. The results from this set of optimizations appears in Table 1. However, these results do not tell the whole story. When we examined the selection part of the optimization, we found that there were often extremely small differences in either LOLE or ENS between the chosen site and the second or third runner-up. This issue is discussed further in the next section.

5.3. Inter-annual Variations and Uncertainty

The wind-speed data used for this study were collected by anemometers mounted on a single tower at each of the six sites we analyzed. Using a power-curve for a modern wind turbine, we calculated hypothetical power output, after accounting for wake effects, and mechanical and electrical losses. If 25-MW clusters of wind turbines were built on any of these sites, however, each turbine would respond to somewhat different winds, depending on the terrain and microscale meteorological events. We are therefore forced to accept the proposition that each time-series of wind-speeds represents one of many possible series. Although it is possible that each of these meteorological towers has been placed in a “representative” location for the overall site, we have no assurances that this is indeed the case. The implication of this is that the precise calculations from our models are based on somewhat imprecise data.

In previous work, we have also been somewhat skeptical of modeling that does not explicitly take inter-annual variations in wind speed into account (Milligan, 1996b). So far, this paper is also subject to that critique. Due to data

<table>
<thead>
<tr>
<th>Table 1. Comparison of the Elfin Optimizations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MW Built by ENS Optimization</strong></td>
</tr>
<tr>
<td>Alberta</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td><strong>MW Built by LOLE Optimization</strong></td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>
constraints we were not able to perform a full analysis of the underlying time-series properties from multiple years of data at each wind site, although that would be our preferred approach. This would allow the use of sequential Monte Carlo runs with Elfin, resulting in probability distributions of the reliability measures of each wind plant and combination of plants (see Milligan and Graham, 1997). Such an analysis would allow for the explicit accounting of the underlying probability distributions, so as to help the decision-maker assess the impact of these variations.

In the absence of a more detailed analysis, we applied a technique borrowed from fuzzy logic (such as Monteiro & Miranda, 1997, and Pereira et al. 1997) to this problem. The use of fuzzy logic allows our modeling to incorporate the uncertainty associated with the issues discussed above. We hypothesized that the LOLE and ENS measures that we obtained from the Elfin optimization are fuzzy values, with variations ranging up to ±0.5% of the calculated value. The choice of this range of values is based on the partial results of our optimization runs. As we examined the reliability values of the best site compared to the runners-up, there appeared to be a clustering of reliability values very close to the optimal values, whereas the least optimal plants' reliability values were significantly worse. In our judgement, the choice of 0.5% is a reasonable one, based on our data. In the absence of specific probability distributions, we hypothesize that the reliability measures are distributed uniformly on this interval, which is similar to other approaches using fuzzy analysis. We then modified the selection decision portion of the optimization algorithm to select not only the best single site, but any site whose optimization parameter is within some small distance of the best choice. Since we have no a priori knowledge of which fuzzy value is best, we performed an analysis using stepwise increments from 0.0% up to 0.5% of the differences in the reliability measure and averaged the cases. This amounts to choosing a 25-MW block of installed wind capacity whenever

\[ c_p(1-\epsilon) \leq c_i \leq c_p(1+\epsilon) \]  

where \( c_p \) is the reliable capacity of the best site, \( \epsilon \) is the fuzzy parameter expressed as a decimal, and \( c_i \) represents the capacity of plant \( i \), \( 1 \leq i, p \leq 6, \text{ and } p \neq i \). The results of this approach appear in Table 2.

Our preferred method is the fuzzy ENS approach because ENS represents the area under the load probability distribution, whereas LOLE represents the height of the tail. Depending on relative costs of purchasing on-peak capacity and energy, a utility could use whichever method is most appropriate.

### Table 2. Comparison of Fuzzy Modeling Results

<table>
<thead>
<tr>
<th></th>
<th>Alberta</th>
<th>Becker</th>
<th>Brewster</th>
<th>Crookston</th>
<th>Currie</th>
<th>Luverne</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW Built by Fuzzy ENS Optimization</td>
<td>38</td>
<td>0</td>
<td>215</td>
<td>0</td>
<td>120</td>
<td>127</td>
</tr>
<tr>
<td>MW Built by Fuzzy LOLE Optimization</td>
<td>65</td>
<td>0</td>
<td>155</td>
<td>0</td>
<td>173</td>
<td>108</td>
</tr>
</tbody>
</table>

5.4. Comparison of Results

A summary of results appears in Figure 8. Each bar in the diagram represents a single method. The first three methods provide similar results: the best combination of sites excludes Becker and Crookston. In each of these cases, the lowest recommended capacity is at Alberta. Brewster and Currie are both recommended in the range of 125 MW
to about 200 MW, and Luverne's share ranges from about 100 MW to about 150 MW. The right side of the graph shows how unstable the results can be when we use a deterministic approach. A small amount of capacity at Luverne is chosen in both cases, but there is clearly a very large difference in the capacity recommendations for Brewster and Currie. The reason for this disparity is because of the extremely close values in reliability that were often found among the runners-up. In this case, Brewster and Currie were very close in both the LOLE and ENS reliability measures, so small differences between these measures altered the relative ranking of the sites. This is one reason for the application of a method that recognizes the role of uncertainty in the modeling. The method of choice, in our judgement, is the fuzzy ENS approach. We believe that ENS provides a more robust measure of reliability, in general, than does LOLE, and is more likely to be stable over short variations in load and generator parameters.

Figure 8. Comparison of methods

6. CONCLUSIONS

Production-cost/reliability models can be applied to the problem of selecting among competing sites for wind generators. However, the use of these models must be tempered with some judgement. The wind sites that we have analyzed exhibit some overall correlation, but also provide some benefit to the overall system reliability because of time lags in hourly generation. We believe that the fuzzy ENS analysis provides the best means of analysis of such problems.

Several additional factors could be introduced into future studies. First, given additional intra-site data, the results would be more accurate. Second, these results are sensitive to the specific load and generator characteristics used by the model. Additional data on wholesale power transactions from the state of Minnesota would improve the accuracy of these results. Finally, a complete analysis of multiple years of hourly wind data at the various sites would provide additional information about the trade-offs that could be expected between sites in future years. Constraints in the transmission system and power flow have not been considered here, but it would be important to analyze these factors before embarking on the installation of a large geographically diverse wind power system.

ACKNOWLEDGMENTS

This paper was prepared as an account of work performed at the National Renewable Energy Laboratory in support
of the U.S. Department of Energy under contract number DE-AC36-83CH10093.

The authors would like to acknowledge support from the FIRST/STEP program at the National Renewable Energy Laboratory, and the Minnesota Department of Public Service.

REFERENCES


