Verification of Power Calculations

To validate the wind power calculations, VCE® obtained Security Constrained Economic Dispatch (SCED) data from Electric Reliability Council of Texas (ERCOT) for years 2017 and 2018. Data from 109 wind farms in ERCOT is obtained for the intercomparison. The metadata of these wind farms is obtained from EIA and SARA reports. In addition, the shape of the wind farm and the locations of individual turbines is obtained from United States Geological Survey (USGS) United States Wind Turbine Database (USWTDB). Figure 1 shows the locations of wind farms and wind turbine locations found from the USWTDB dataset. The USWTDB dataset contains all the required metadata information to model the wind farm such as nameplate capacity, wind turbine types in the farm, hub-heights, rotor diameter and wind turbine rated capacity. The following section describes how wind farm and wind turbine information is used to model the power output from each wind farm.

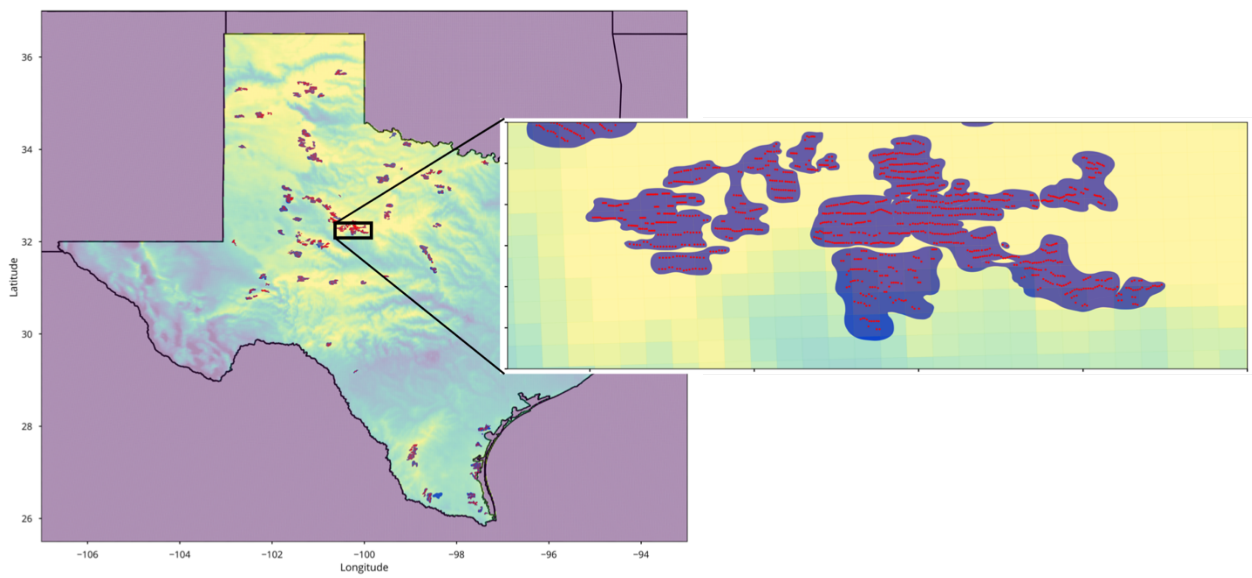

Figure 1. Wind turbine location information from the USWTDB dataset. Inset shows a zoomed in image of the area denoted by the solid black box. The background colors are average wind power capacity factors for 80 m hub-heights for years 2014 to 2016. Brighter colors indicate higher capacity factors.

There are many different open data sources available to find metadata on existing wind farms. Each source can have both large and small differences between them. There is no answer that is 100% accurate from any one data source. Having different sources is pertinent to achieving a more complete picture of what is out there for any generator technology, wind included. In ERCOT, three sources were used for analyzing and setting the metadata for all the utility scale wind farms under that ISO's umbrella. The sources included: Annual EIA 860 generator and plant data, the latest max capacity values from the Seasonal Assessment of Resource Adequacy (SARA) report and finally the max power output observed from the Security Constrained Economic Dispatch (SCED) observation data reported by ERCOT for individual wind farms. Certain processes would be considered such as the following:

1. Fundamentally aligning the actual wind farms reported from SARA, the EIA 860 and the SCED data was the first main address. Each source might have a slightly different count of plants or plants named differently.

2. SARA report capacities and maximum outputs from the SCED observation data might reveal that the EIA 860 annual numbers were out-of-date and usually needed to be adjusted upward.
3. SARA report capacities might align well with EIA 860 numbers, but the SCED data might not reveal the same thing. This would be cause for further investigation into what could be going on for this farm.
4. In cases where all three sources aligned, there was much higher confidence in the final metadata decisions.
5. Cases where no resources aligned were rare. However, when they did happen, further investigation occurred into these sites to see if other sources (as an example, the USGS wind turbine location dataset) might provide any insight.

All of this work helped provide a backdrop to the representation of wind farms physically installed in the ERCOT system.

Modeling the Wind Farm Generation:

As described in the previous sections, the wind power capacity factors are available at 3-km horizontal resolution and 5-min time resolution for hub-heights of 80m, 100m, 120m, 140m and 160m. For each wind farm, the wind turbine metadata is used to select the capacity factor profiles at the appropriate hub-height. For older wind farms where hub-heights of lower than 80m are present, a cubic function is used to de-rate the wind power capacity factors from 80m level. If turbines of multiple hub-heights are present in a given farm, currently the farm is modeled using a capacity weighted average hub-height. Future work will model each wind turbine hub-height separately in order to get the most accurate results.

Next, the wind turbine type is used to select an appropriate power curve for that wind turbine type. Using a transfer function, the capacity factors from the IEC-3 power curve used for our calculations is converted the power curve of the turbine type present in the farm. If power curve information is not present then the IEC-3 power curve is used. Finally, it is checked is the sum of all the wind turbine rated capacities add up to the wind farm's nameplate capacity. If this is not so (which can happen due to missing wind turbine location information), then each wind turbine capacity is adjusted equally so that they add up to the wind farm nameplate capacity.

Now the turbine location information is used to retrieve the power capacity factors from HRRR cell the turbine is located in. Appropriate corrections to the capacity factor profiles are made as described above and then this generation information is saved. In a similar manner, all turbine generation profiles are created and finally added together to get the wind farm generation output.

Validation of the Wind Power Generation:

The wind power generation profiles created as described above are validated by comparing against the SCED data obtained for 109 wind farms in ERCOT. First the appropriate settlement point(s) for each wind farm is identified. Data from all the settlement points associated with a wind farm are summed to obtain total SCED generation from that wind farm. Initial filtering of the SCED data is performed to only keep data with "Telemetered Status" was "ON". The SCED data goes through a second, more rigorous quality control where periods of obvious curtailment are removed. It is important to remove periods of curtailment as they cannot be simulated by the power calculation model and will result in over-estimating the model errors. The quality control for curtailed periods is performed as follows for 5-minutely data (some thresholds applied below would change for other granularities):

1. Scanning the data for any sharp spikes. The spike can be in the upward or downward direction. Any given timestamp was flagged if the power changed more than +/-20% of farm capacity and in opposite directions in the both the forward and backward direction in time. As an example, a suspect period would occur when looking back one time step, the power changed more than +20% of maximum capacity and looking forward one time step the power dropped more than -20% of max capacity. This would also be flagged if it had been a spike in the downward direction. This threshold was considered since the data was 5-minutely.
2. When power flatlines or hovers around +/- 1% of a value that is not within a certain range of max capacity or 0% capacity for more than two hours it is flagged as suspect. The offending value has to flatline or hover more than 5% below max capacity or more than 5% above zero.
3. Any sharp drop or rise to or from 0% capacity where the farm was at 0% capacity for more than two hours or the farm dropped to 0% capacity for more than two hours is suspect. For 5-minutely data, a 15% rise or drop in capacity would be flagged.
4. Any drop or rise from any point that was greater than 50% capacity in a 5-minute time frame is flagged as suspect.

Any of the above thresholds can be changed. Loosening the thresholds will release certain curtailment periods from being captured or flagged. Tightening the threshold will remove more periods of data that are actually physical weather events that create sharp ramps. Figure 2 shows a partial time-series of SCED generation that was quality controlled. It is observed that the obviously curtailed periods (dark blue line) are rejected while keep the rest of profile (light blue line) intact.

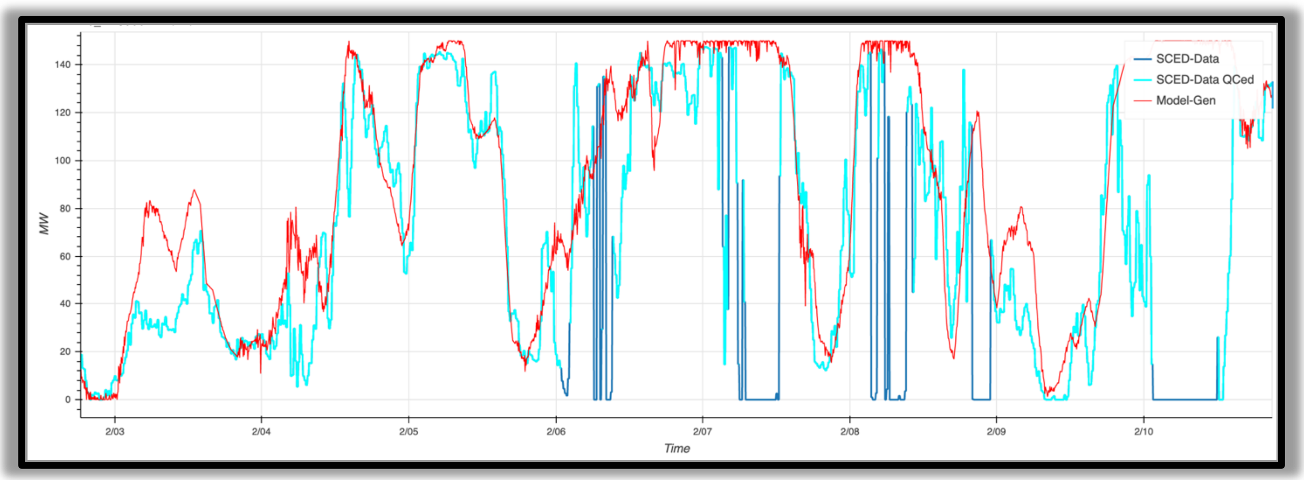

Figure 2: An example wind farm in ERCOT where model output, SCED Data and quality-controlled SCED data is shown. When dark blue can be seen is where data was flagged as potential curtailment through VCE quality control processes.

The metrics used to estimate the quality of the modeled wind power profiles are bias, root mean square error, correlation coefficient and coefficient of determination (R^2). Figure 3 shows an example time-series of the VCE® model wind generation output compared to the SCED data for the Amazon Wind Farm in ERCOT. The timeseries show that the model is able to capture the temporal variability in the power production accurately. This is due to a combination of accurate physical model of wind turbine operation, accurate wind forecasts from the HRRR and finally accurately determining the wind farm and wind turbine metadata to ensure that the wind generation model is working off correct details.

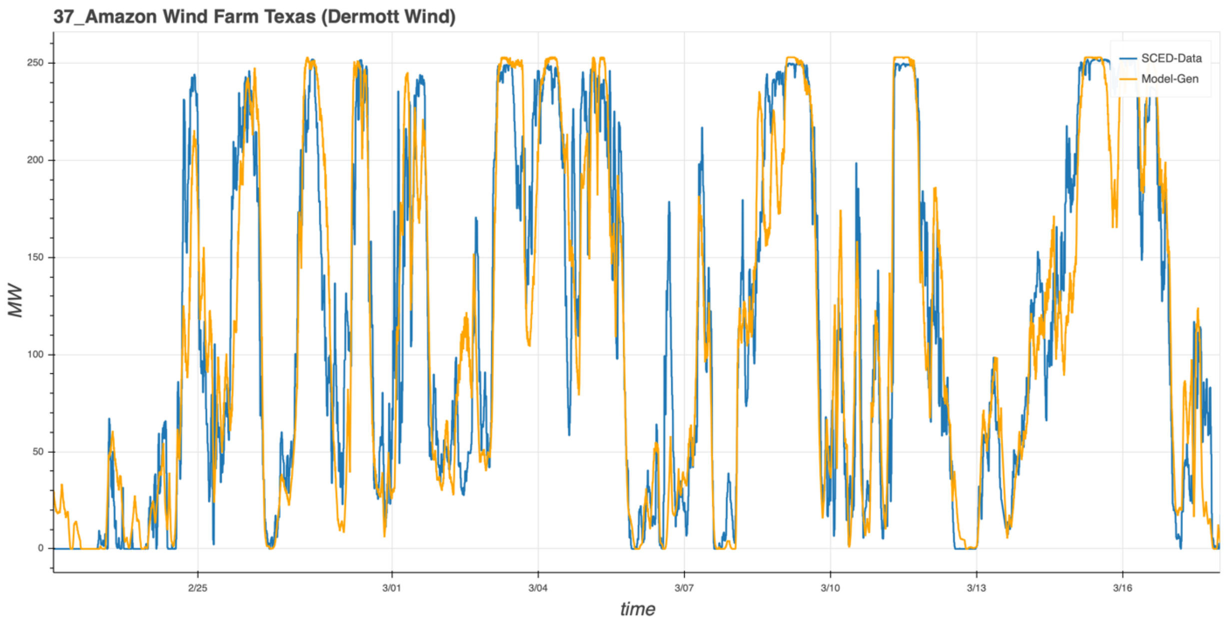

Figure 3. Comparison of VCE® model wind generation output (orange) with SCED data (blue) for the Amazon Wind Farm in ERCOT.

It is observed that of the 109 wind farms for which SCED data is available, 100 wind farms have correlation coefficients of greater than 0.80, while the remaining nine have correlation coefficients of greater than 0.70. Figure 4 summarizes the intercomparison statistics from the wind farm validation. As seen from Figure 4, almost all wind farms show correlation coefficients greater than 0.80. It is also observed that many wind farms are clustered around ratio of standard deviation of 1, which indicates that the model and the SCED data show similar level of temporal variability.

The main challenge in validation of wind farm generation output is removing periods of curtailment as these cannot be modeled by a wind farm generation model. Although it was attempted to remove as many periods of obvious curtailment as possible, many curtailed generation periods still remain. In addition, wind farm output is modified due to turbines being down for maintenance or disbanded (true in cases of many older wind farms). Since this information on wind farm maintenance schedules and turbine status is unavailable, differences in the modeled generation and SCED data arise. These differences are evident in Figure 4 from the centered RMS difference in many wind farms being on the order of 40% of the standard deviation. In addition, the curtailed periods artificially reduce the standard deviation of the SCED data and hence it appears that the model forecasts are over-predicting variability which is probably not the case as the model output is expected to the smoother than the actual generation.

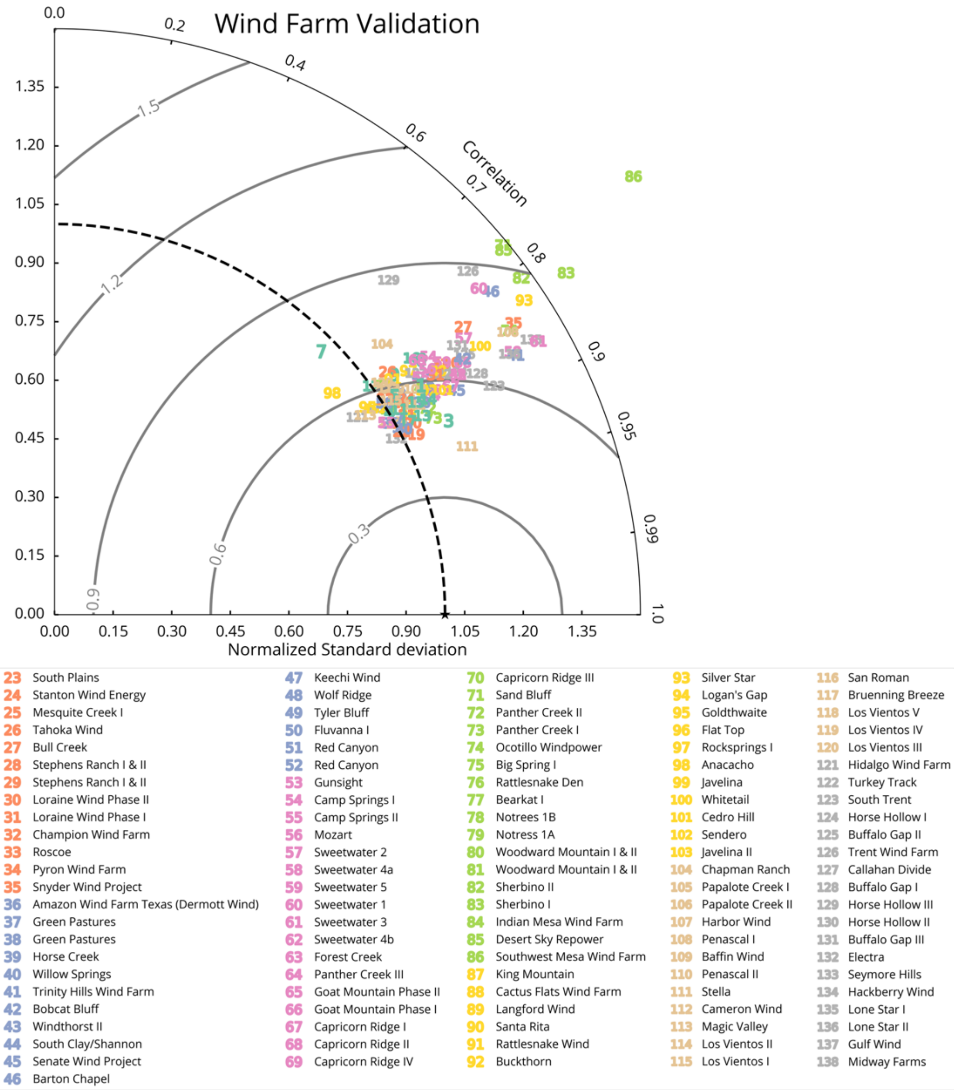

Figure 4. Taylor diagram summarizing the intercomparison metrics between VCE® model predicted wind power generation and SCED data.

To better understand how curtailment, turbine maintenance or disbandment can affect wind farm output, the model forecasted generation is compared against the SCED data for the Southwest Mesa Wind farm in ERCOT in Figure 5. As seen from Figure 5, the SCED data and the model forecasted generation are well correlated, but the SCED data never reaches the maximum power forecasted by the model. The main reason for this is that this wind farm is quite old and from visual inspection of Google Earth®, many turbines are broken down or feathered for maintenance. As a result, although the model is correctly simulating the wind generation from the farm, it shows a high RMS difference but with a high correlation coefficient as seen in Figure 4.

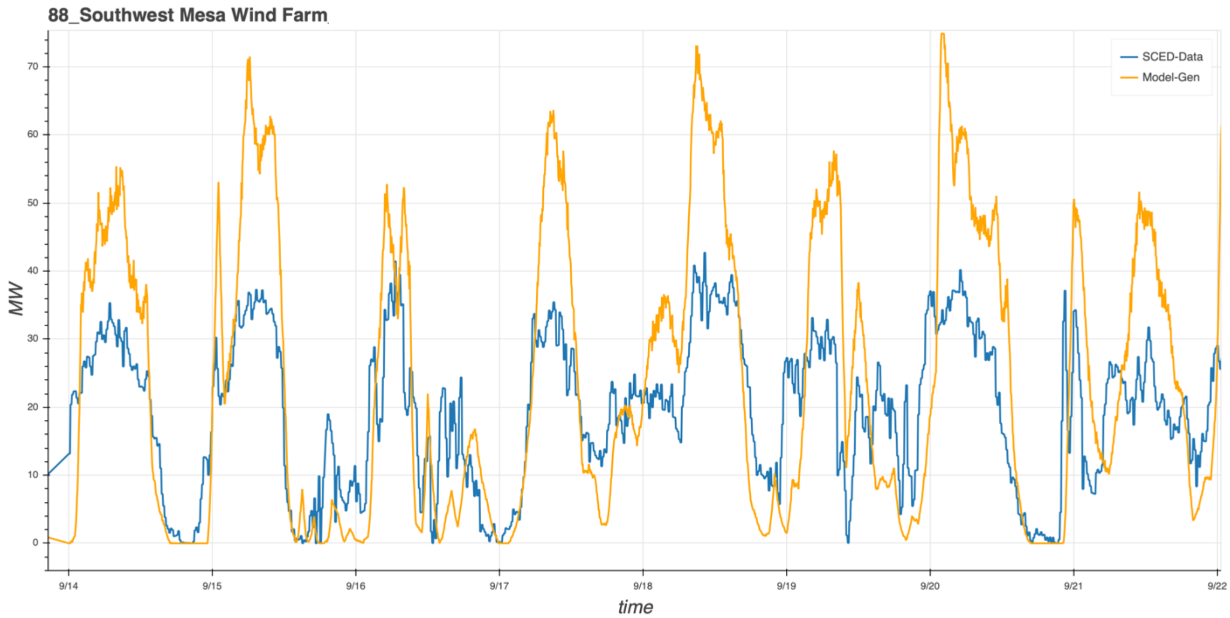

Figure 5. Comparison of VCE® model wind generation output (orange) with SCED data (blue) for the Southwest Mesa Wind Farm in ERCOT.