Market Protocols in ERCOT and Their Effect on Wind Generation

Ramteen Sioshansi*

Integrated Systems Engineering Department, The Ohio State University, Columbus, Ohio, United States
Phone: +1-614-292-3932

David Hurlbut

National Renewable Energy Laboratory, Golden, Colorado, United States

Abstract

Integrating wind generation into power systems and wholesale electricity markets presents unique challenges due to the characteristics of wind power, including its limited dispatchability, variability in generation, difficulty in forecasting resource availability, and the geographic location of wind resources. Texas has had to deal with many of these issues beginning in 2002 when it restructured its electricity industry and introduced aggressive renewable portfolio standards that helped spur major investments in wind generation. In this paper we discuss the issues that have arisen in designing market protocols that take account of these special characteristics of wind generation and survey the regulatory and market rules that have been developed in Texas. We discuss the perverse incentives some of the rules gave wind generators to over-schedule generation in order to receive balancing energy payments, and steps that have been taken to mitigate those incentive effects. Finally, we discuss more recent steps taken by the market operator and regulators to ensure transmission capacity is available for new wind generators that are expected to come online in the future.

Key words: Wind integration, reliability, transmission operation

1. Introduction

In 1999 the Texas state legislature passed Senate Bill 7 (see TSL (1999)) which established a framework for restructuring the electricity industry in Texas and put into place a renewable portfolio standard (RPS) for the state. The restructuring efforts expanded the role of the Electric Reliability Council of Texas (ERCOT), the independent system operator that lies entirely within the Texas state boundaries and is not subject to federal regulation for pricing and operations.[1]

ERCOT began operating as a single control area in 2001, covering 85% of the load and 75% of the land area in the state of Texas. Aside from public power entities, ERCOT has no traditional utilities. Investor-owned monopoly utilities were unbundled into separate generation and retail entities in 2002, with new load-serving entities and power generating companies allowed to enter the market to compete. Transmission remained regulated with respect to rate recovery, route approval, and determination of need for upgrades, but the task of operating the grid was transferred to ERCOT (see VTAS (2007)).

Generators and load-serving entities in the ERCOT market do not deal directly with system operators, but are represented instead by qualified scheduling entities (QSEs). These QSEs handle such functions as resource scheduling, control error management, and financial settlement. Each QSE is responsible for providing ERCOT with an hourly schedule matching the total planned output of the QSE's committed units with the total amount of total load it expects in its portfolio. Financial settlements for all grid operations instructed or man-

*Corresponding author
Email address: sioshansi.1@osu.edu (Ramteen Sioshansi)

[1] The ERCOT control system does have some DC connections to neighboring reliability regions, but these do not subject ERCOT to federal regulation. ERCOT also functions as a reliability region , and in this respect is subject to federal standards.

aged by ERCOT are done with QSEs (see ERCOT (2007)).

As the independent system operator, ERCOT procures and deploys balancing energy and ancillary capacity services for the entire region. It is also the regional reliability organization. Neither ERCOT nor the transmission owners whose assets it operates are permitted to own generation facilities; all energy and capacity services that ERCOT requires for operating the grid are procured via auction or in certain cases direct contract (for reliability must-run units, for example).

A market clearing price for energy (MCPE) is calculated for every 15-minute operating interval based on energy offers from QSEs, and the amount of balancing energy ERCOT needs to match real-time load. Most of the energy that ERCOT deploys for ancillary services is paid according to the MCPE.

ERCOT began operating as a zonal market in 2001, with zones defined each year on the basis of major transmission paths with degrees of congestion that were deemed commercially significant. In December 2010, ERCOT is expected to complete its transition to a locational marginal pricing (LMP) model in which each node for generation or load will have its own local energy price based on generation offer prices, all transmission constraints, and demand response from loads.[2]

In addition to the market restructuring provisions, Senate Bill 7 also mandated that 2,880 MW of renewable energy capacity be installed in Texas by 2009,[3] which amounted to a nearly 2,000 MW increase in renewable generating capacity. Senate Bill 20 (see TSL (2005)), which was passed in 2005, mandated further increases in renewable generating capacity. Due to excellent wind resources, especially in western Texas, most of this renewable energy capacity came in the form of wind generation.

Integrating wind generating resources (WGRs) into power systems can present unique challenges due to the limited dispatchability of wind generation, errors in forecasting real-time wind availability, and other design limitations of wind turbines (see DeMeo et al. (2005, 2007) and Smith et al. (2007)). In other countries, wholesale electricity markets with high wind penetration levels have often had to adjust their treatment of WGRs under their market protocols to accommodate these unique characteristics of wind and not unduly penalize wind generators for characteristics that are outside of the wind operators' control. This typically includes a more lax treatment of uninstructed deviations, fewer penalties for over- or undergeneration, and less stringent reactive power requirements.

ERCOT implemented many of these types of provisions in its original zonal market protocols, including allowances for reactive power requirements and uninstructed deviations from a unit's scheduled output. While these market rules were intended to accommodate WGRs, they gave perverse incentives for wind generators to overschedule generation in order to receive decremental energy payments. With input from the staff of the Public Utility Commission of Texas (PUCT), wind generators, and other stakeholders, ERCOT revised the market protocols in 2003 to eliminate payments to overscheduled wind generators, however this has been viewed as an *ad hoc* approach. With the recent move towards adopting a nodal market design, market protocols are again being developed to accommodate WGRs.

In addition to accommodating WGRs in system operations and settlement, ERCOT has also had to deal with transmission issues in integrating wind generation into its system. This is due to the fact that the most abundant wind resource is in western Texas, whereas most of the load is in the east. The limited transmission capacity out of western Texas has been a bottleneck for wind generators, and has required massive wind generation curtailments in some cases. The PUCT has recently begun taking a proactive approach to dealing with this issue by identifying regions of the state that would provide the most cost-effective wind generation, and establishing procedures to ensure there is sufficient transmission capacity installed in those areas in anticipation of wind capacity being added.

This paper surveys the design of the original zonal market in ERCOT and the new nodal market proposals, as they relate to WGRs. In sections 2 and 3

[2] ERCOT originally proposed for the nodal market to go live December 1st, 2008. Setbacks in getting the nodal system into operation, however, delayed this start date, and the nodal market is now expected to go live in 2010 (see Hinsley (2008)).

[3] It also includes incremental renewable requirements before 2009, with 1280 MW required by 2003, 1730 MW by 2005, and 2280 MW by 2007.

actions taken by the PUCT to ensure transmission capacity is available for WGRs in the future, and section 5 concludes.

2. Zonal Market Protocols

The early ERCOT protocols allowed very few major exceptions for WGRs with respect to operations and settlement.[4] Perhaps the most problematic allowance concerned uninstructed deviations from a unit's scheduled output. This protocol gave qualified wind power generators a much more forgiving standard than was required of conventional units with respect to unscheduled variations in a plant's real-time power output (see Robinson (2006)). However, this allowance had a secondary effect on other important protocols that treated wind resources the same as other resources.

2.1. Uninstructed Deviations

Under the zonal protocols, QSEs would normally combine a number of generating units into one portfolio, and present output schedules for the portfolio rather than for individual units. Resource plans indicate the specific units the QSE plans to commit. These day-ahead schedules and resource plans submitted by QSEs were crucial to ERCOT's ability to manage the transmission system. They provided a picture of which generators would be available at any given moment, and how much additional energy ERCOT might need to procure in each 15-minute balancing energy market in order to match total system generation with forecasted load. Unanticipated changes from scheduled generation complicated grid management, and increased the chance that ERCOT would have to use more of its operating reserves.

"Uninstructed deviation" is the difference between the total real-time *metered* units, it became possible for wind units to receive energy payments even when the wind was not blowing.

In the zonal market, ERCOT operators use out-of-merit energy (OOME) to relieve congestion on lines within a zone that may remain due to the flow of scheduled power and the deployment of balancing energy. ERCOT procures balancing energy market-wide on a merit basis, *i.e.* by awarding the procurements to the QSEs that have offered the energy at the lowest cost. In addition, balancing energy awards and instructions go to the QSE's *portfolio* of resources, not to a specific generating unit, without considering the feasibility of the resulting intrazonal network flows. Consequently, the units that a QSE uses to provide balancing energy are selected without regard to the feasibility of the resulting intrazonal network flows and may be located in such a way that their deployment would cause certain transmission lines to become overloaded. To solve these problems, ERCOT procures OOME from units that can relieve congestion on the line by generating more or less energy at specific points on the network.

Unit curtailments to relieve transmission congestion come as OOME-Down instructions from ERCOT operators. Because they are instructed to deviate from their schedules, OOME-Down does not count towards a QSE's uninstructed deviation penalty. OOME-Down during any given 15-minute operating interval is settled at the current MCPE, so that the QSE is made whole for the difference between its scheduled output, and the lower output instructed by ERCOT (see ERCOT (2002)).

A higher scheduled operating level increases the payment a QSE receives for the same OOME-Down instruction. For conventional units this was simply a mathematical artifact of the protocols, and der the new market protocols are a nodal, as opposed to zonal, congestion management system and more centralized coordination of the market. In addition to the general changes to market operations, the nodal protocols have a number of provisions that are meant to better integrate WGRs into the market by speci cally taking account of their unique properties. This includes (i) an assessment of WGRs in long-term resource planning; (ii) a consistent and more accurate forecasting methodology, which is conducted by ERCOT, to determine potential wind generation for day- and hour-ahead scheduling; and (iii) reduction of imbalance and deviation penalties for WGRs in real-time.

3.1. Long-Term Resource Assessment

One of the provisions of the nodal pr-tioesource which model regional weather patterns; (ii) statistical models, which use historical and real-time data to estimate wind generation; and (iii) telemetered real-time data provided by wind generators. These models are combined to produce hourly probability distributions of generation from each WGR for the following 48 hours. These hourly probability distributions are then used to produce hourly resource forecasts, called the Wind Generation Resource Production Potential (WGRPP), which is a generation level with an 80% probability of exceedance (*i.e.* if we let x denote the WGRPP, x is chosen such that there is an 80% probability that the actual wind generation is greater than x) (see ERCOT (2008c) for a description of these forecasting activities).

The WGRPP forecast is used primarily in the RUCs, and bounds the amount of energy from wind generators that can count towards resource requirements. Furthermore, under the current nodal protocols, WGRs are not qualified to provide ancillary services. This use of the WGRPP in the RUCs is to ensure an exceedingly low loss of load probability, which could occur if an overly optimistic wind forecast is used and insufficient generating capacity is committed day- and hour-ahead. Although the WGRPP bounds wind generation in the RUCs, wind generators are permitted to submit bids into the DAM or bilaterally trade more energy than the WGRPP, since bilateral and DAM transactions are treated as being purely financial transactions. This provision allows wind generators to realize the value of their generation assets by selling their expected wind generation, while being subject to the financial obligation to replace any energy shortfall in the real-time market.

Although the use of this conservative WGRPP in the RUC will reduce loss of load probabilities, the use of a point estimate of wind (as opposed to explicitly modeling the random nature of wind in a stochastic unit commitment framework) will tend to be inefficient. One shortcoming of this approach is that the RUCs may overcommit generation day- and hour-ahead due to an overly conservative estimate contrast, are subject to deviation penalties if they deviate by more than 5% from their dispatched output.

This more tolerant treatment of WGR deviations reflects the fact that WGRs have much less control over real-time output than conventional generators, but helps put sufficient incentives in place to ensure performance by WGRs. This is achieved by subjecting WGRs to purchase replacement energy for over- and under-generation from the dispatch instruction, as well as any financial obligations entered into bilaterally or in the DAM. Although undergeneration is not subjected to any deviation penalties, the fact that wind generation has a zero marginal generation cost will give WGRs an incentive to generate up to its dispatch quantity and receive LMP payments. Overgeneration is similarly penalized by the real-time price of decremental energy, as well as deviation penalties if the WGR is more than 10

4. Transmission Access for Wind Generators

One of the largest impediments to integrating wind generators into the ERCOT market is the geography of the state and access to transmission capacity. The most abundant wind resource in ERCOT is in the western end and panhandle region of the state, whereas most of the population and load centers are in the east. This geography and the limited transmission capacity out of western Texas has proven to be a challenge to integrating wind generators into the ERCOT power system. As Baldick and Niu (2005) note, this issue is exacerbated by the fact that ERCOT's interconnection policy allows wind generators to connect to the power system even without sufficient transmission capacity to carry the power. Moreover, the cost of any upgrades or additions to the high-voltage transmission grid that may be necessitated from generator interconnection are assigned to loads (as opposed to generators) using a postage-stamp tariff.

For example, in 2002 758 MW of wind generators were interconnected in the McCamey area in western Texas, despite there only being 400 MW of transmission capacity in the substation. LCRA (2003) estimates that this resulted in about 380 GWh of wind generation, with an estimated market value of more than $21.4 million, being curtailed until mid-2003 when the substation was upgraded.

In order to address this issue, the Texas legislature passed Senate bill 20 in 2005 (see TSL (2005)), which mandated that the PUCT take steps to ensure transmission infrastructure improvements are undertaken for wind generators.[6] The new law directed the PUC to (i) designate regions within Texas that would deliver the most beneficial and cost-effective wind resource, (ii) develop a plan to build transmission capacity into those zones, and (iii) take into account financial commitments of WGR developers in determining the competitiveness of a potential zone. The purpose of this legislation is both to encourage investment in wind right of way. Kirby et al. (2002) and Weigt et al. (2009) discuss and analyze these and other benefits of using HVDC connections to directly deliver wind to load centers.

In the case of Texas, however, HVDC and 765 kV AC present several disadvantages that have ultimately made their use uneconomic. Integrating HVDC connections into an AC transmission network requires costly high-voltage DC-AC converters at the two ends of the HVDC connection. This in turn requires the HVDC connection to cover a long distance and carry high capacities to exploit economies of scale, allowing for only a limited number of HVDC connections to be built (the cases ERCOT examined only had two HVDC lines). This presents two challenges for wind integration in ERCOT. One is that significant transmission upgrades are required within the CREZs to deliver energy to the limited number of HVDC connections. Secondly, in order to ensure the transmission system is compliant with $N-1$ reliability requirements, a number of electrically parallel AC transmission lines would have to be upgraded, which reduces the cost advantage of HVDC relative to upgrading the existing infrastructure. 765 kV AC connections also suffered from these issues — exploiting the reduced right of way cost with high-voltage AC connections would require upgrades to transmission infrastructure within the CREZs, and $N-1$ reliability requirements would also require upgrades to the existing infrastructure. In the end the lower cost of upgrading the existing infrastructure, along with reasonable system operations costs and wind curtailment levels with this type of transmission topology, led the PUCT to approve upgrades to the existing infrastructure without use of HVDC or 765 kV AC connections. Figure 1 shows the CREZ that have been designated by the PUCT, and the conceptual transmission scenario. Current estimates place the cost of development at nearly \$5 billion, with an estimated ratepayer impact of around \$4 per customer per month (see PUCT (2008a) and ERCOT (2008a)).

di cult to predict in advance. The PUCT and ERCOT have developed an innovative means of ensuring transmission capacity is available for future wind projects. The aim of these e orts is not only to ensure there is su cient transmission capacity available, but also to proactively \direct" wind investment in parts of the state that have been identi ed as having the best wind resources. One of the major challenges confronting ERCOT in the future will be the determination of wind power's true ca-