




Cost-Causation and Integration Cost Analysis for Variable Generation

Michael Milligan, Erik Ela, Bri-Mathias Hodge, Brendan Kirby (Consultant), and Debra Lew
National Renewable Energy Laboratory

Charlton Clark, Jennifer DeCesaro, and Kevin Lynn
U.S. Department of Energy

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

Technical Report
NREL/TP-5500-51860
June 2011

Contract No. DE-AC36-08GO28308

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Cost-Causation and Integration Cost Analysis for Variable Generation

Michael Milligan, Erik Ela, Bri-Mathias Hodge, Brendan Kirby (Consultant), and Debra Lew
National Renewable Energy Laboratory

Charlton Clark, Jennifer DeCesaro, and Kevin Lynn
U.S. Department of Energy

Prepared under Task No. WE110820

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, Colorado 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-5500-51860
June 2011

Contract No. DE-AC36-08GO28308

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information

P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) PIX 16416, PIX 17423, PIX 16560, PIX 17613, PIX 17436, PIX 17721

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

Cost-Causation and Integration Cost Analysis for Variable Generation

Michael Milligan, Erik Ela, Bri-Mathias Hodge, Brendan Kirby (consultant), Debra Lew
National Renewable Energy Laboratory

Charlton Clark, Jennifer DeCesaro, Kevin Lynn
United States Department of Energy

Introduction

Wind and solar power generation are prized for their environmental benefits, their low and stable operating costs, and their help in reducing fuel imports. Advances in both technologies are reducing capital costs and providing significant control capabilities. Still, the primary energy source for both technologies is variable and uncertain and a power system with significant wind or solar penetration must be operated differently than a power system based exclusively on conventional resources. It is very natural to ask what the additional cost of accommodating wind and solar generation is. However, calculating the integration cost of variable generation turns out to be surprisingly difficult.

Integration cost analysis has progressed significantly over the past ten years. There is also a much better understanding of the cost drivers among the system stakeholders now than there was when wind and solar generation were in their infancy. This report examines how wind and solar integration studies have evolved, what analysis techniques work, what common mistakes are still made, what improvements are likely to be made in the near future, and why calculating integration costs is such a difficult problem and should be undertaken carefully, if at all.

Analysis techniques are now very good at simulating power system operations with time-synchronized load, wind, and solar data. The best studies model security-constrained unit commitment and economic dispatch with hourly (or shorter) time steps covering one year or longer. They account for forecast errors for wind, solar, and load as well as actual output and consumption. Cases with and without wind and solar can be compared. Total system costs with and without renewables can be calculated fairly accurately under a range of system conditions. These cost differences are typically dominated by the fuel cost savings that renewables provide, however. Calculating an "integration cost" that only includes the added cost the power system incurs dealing with the variability and uncertainty of wind and solar, and excludes the fuel cost savings, is much more difficult. The many complex interactions among components of the power system and assumptions regarding the no-wind base case all have important influences on integration cost estimates, and in fact raise questions of whether cost components that are commonly thought to be integration costs can be correctly untangled. We discuss many of these concerns and implications, shedding some light on the difficulties involved in measuring and interpreting integration cost estimates.

Variable Generation Impacts on Balancing Requirements

The variability and uncertainty of wind and solar power generation increase the response requirements from conventional generators and responsive load, but they do not increase the overall capacity requirements of the power system. Peak load with wind and solar is never higher than peak load without wind and solar. Wind and solar generation can only reduce the net load (i.e., system load minus wind and solar generation) which must be served by conventional generation; they can never increase the capacity required to serve load unless the existing generation fleet cannot respond quickly enough ─ in that case, the problem is not that there is insufficient capacity; the problem is that the existing is not flexible enough.

Power system variability during normal operations is commonly separated into three time frames as shown in Figure 1. Regulation deals with the random, minute-to-minute variability of loads and generation. Load following (or more properly simply "following") deals with slower trends that extend from minutes to hours. Aggregate load exhibits a typical daily cycle. Solar exhibits a daily cycle as well. Wind patterns vary with daily cycles in some regions and storm-front driven patterns in other regions. In all cases there is diversity among individuals (wind plants or turbines, solar plants or devices, or loads) in the minutes-to-hours (following) time frame than in the minute-to-minute (regulation) time frame. Consequently, aggregation does more to reduce regulation requirements than it does to reduce following requirements. The scheduling process deals with the longer term, often day ahead, unit-commitment decision time frame driven by the forecasts of expected load, wind, and solar generation.

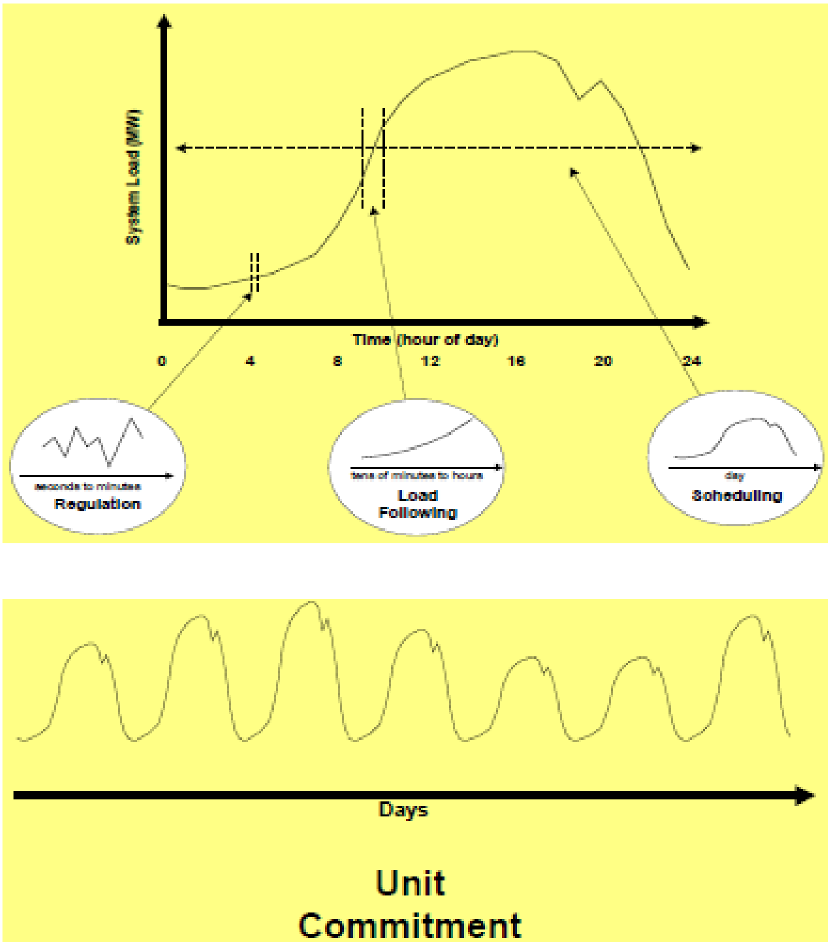

Figure 1 — Time scale for power system operation.

Power systems must also balance generation and load under contingency conditions. No power plant or transmission line is 100% reliable and the power system must be continuously prepared to respond to the sudden failure of any resource. A series of reserves are maintained to provide immediate and sustained response to a contingency, as shown in Figure 2.[1] Wind and solar generators typically have little impact on contingency reserve requirements because individual wind turbines and solar panels or plants are small compared with the largest conventional power plants, and contingency reserves must be constantly maintained so they are ready to deal with the largest credible event. However, large wind and solar ramping events do share some characteristics with conventional generation contingencies; the largest events are relatively rare, with reserve standby costs being more important than response costs. Large wind and solar ramping events differ from conventional contingency events in that they are much slower.

[1] B. Kirby, 2004, *Frequency Regulation Basics and Trends*, ORNL/TM 2004/291, Oak Ridge National Laboratory, December.

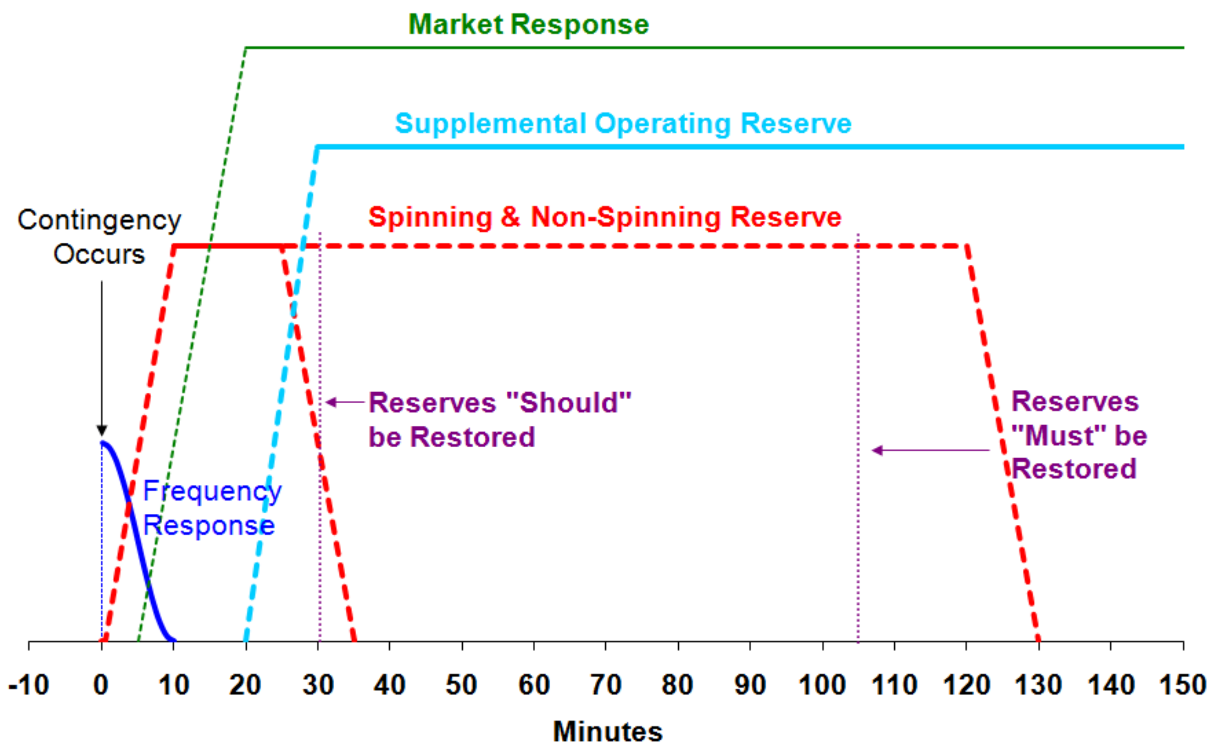

Figure 2 — A series of contingency reserves provide instantaneous and sustained response.

Regulation

The impact of wind power on the regulation time scale is generally well-understood. The impact is relatively easy to calculate when synchronized high-resolution load and wind data are available.[2] There is less operating experience with solar power than with wind, but aggregation benefits for both dispersed resources and regulation impacts should be similar in principle. Because the minute-to-minute variability of individual loads, wind plants, and PV plants are highly uncorrelated, total power system regulation requirements are reduced with larger aggregations of load and generation.

Following

Following imposes a flexibility or ramping requirement on the power system, but not a capacity requirement (i.e., additional generation). The morning load ramp can never ramp above the daily peak load. It is only the required flexibility that increases. Similarly, if wind has been blowing at

[2] M. Milligan, B. Kirby, 2009, *Calculating Wind Integration Costs: Separating Wind Energy Value from Integration Cost Impacts*, NREL/TP-550-46275, July.

night and drops off during the morning load pickup, the morning ramp rate increases but the total system capacity requirement, set by the daily peak net load, does not increase.[3]

Scheduling

Conventional fuel-burning generators typically require preparation time before they can be operated.[4] Large coal fired plants can require a day or longer, depending on their current standby state. Combined cycle plants typically require several hours. Combustion turbines require minutes to an hour or longer. Some hydro and reciprocating engine plants can start and fully load within minutes. Wind and solar never increase the amount of conventional generation that must be scheduled day-ahead (above that which would be required absent the wind and solar), but ignoring the wind and solar forecast can result in excessive amounts of conventional generation being brought on line (above what is actually required with the available wind and solar). This can result in inefficient operations with conventional plants operating well below their optimal outputs. Uncertainty in the wind and solar forecast can result in a change in the optimal scheduling mix, with flexible generation preferred over inflexible.

Contingency Reserves and Large Wind and Solar Ramps

Wind ramping shares at least one important characteristic with conventional generation contingencies; large ramping events are rare. Large solar ramps will likely be found to have similar characteristics once sufficient data is available. Figure 3 compares the maximum one-hour daily ramping requirements for load and wind for the Pacific Northwest with 30% wind energy penetration.[5,6] Load ramping requirements are relatively constant from day to day while large wind ramping requirements are relatively rare. The maximum daily one-hour wind ramp expected every two weeks is only 53% of the peak annual requirement. The average daily one-hour wind ramping requirement is only 24% of the peak annual requirement. This has implications for the type of reserves that are appropriate for wind ramps versus load ramps. This will be discussed further below.

---

[3] B. Kirby and M. Milligan, 2009, *Capacity Requirements to Support Inter-Balancing Area Wind Delivery*, NREL/TP-550-46274, July.

[4] Hydro generators, reciprocating engine driven generators, and some fast-start combustion turbines do not require significant preparation time before they can start operating.

[5] The ramping sign for wind is referenced to load to facilitate comparison. That is, positive ramping is an increase in load or a decrease in wind.

[6] M. Milligan, B. Kirby, J. King, and S. Beuning, 2010, *Potential Reductions in Variability with Alternative Approaches to Balancing Area Cooperation with High Penetrations of Variable Generation*, NREL/MP-550-48427, August.

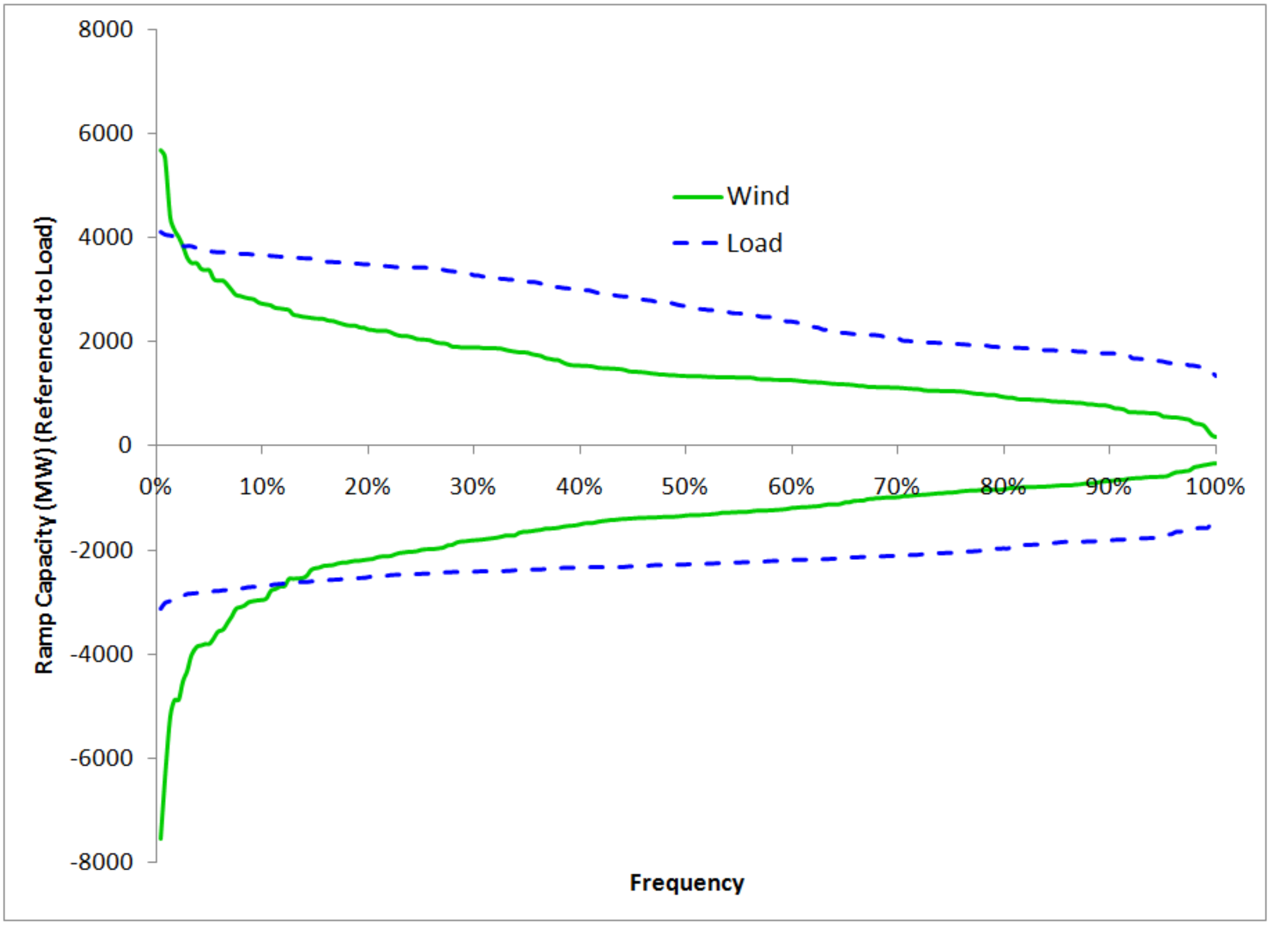

Figure 3 — Large wind ramps are rare compared with relatively consistent load ramping requirements.

Integration Cost

The *concept* of an integration cost for wind and solar generation seems simple and useful. What costs does the power system incur when wind or solar generation are included in the generation mix? Although this appears to be a very simple question, in practice, calculating the integration cost has proven to be very complex. To date it has not been done in a completely satisfactory manner. The complexity does not come from an inability to model the power system or from an inability to calculate system costs under various conditions with and without wind and solar generation. The complexity comes from trying to establish what conditions to compare, and from the complex and numerous interactions between generation resources as their aggregate output is adjusted to maintain load balance. Wind and solar integration costs cannot be measured directly. Instead, total power system costs with and without wind and solar generation need to be compared.

Modeling the *with* wind and solar conditions is now relatively easy; simply decide what penetration of wind and solar generation is desired and then collect the years of data necessary to accurately model the power system. A great deal of progress has been made over the past several years advancing the science of modeling the power system with high penetrations of wind generation. Although some work has been done to create similar solar data sets that can be used in integration analysis, the maturity of this solar modeling needs to evolve to make it comparable to what is done with wind modeling. That progress should be extendable to solar generation as data becomes available.

There are two difficulties in calculating wind and solar integration costs. The first comes from trying to define the *without* wind and solar case. If the power system is simply modeled without the wind and solar generation, the energy that wind and solar would have provided must come from some other resource. The fuel cost (and the value of the energy itself) for that other resource will dominate the difference in costs between the cases. If fuel costs are included, then wind and solar always present a system benefit, not an integration cost; stated differently, integration cost will be a relatively small fraction of the reduction in fuel cost from the displaced generation. The concept of "wind and solar integration cost" is usually used to cover the non-energy costs, greatly adding to the complexity. This leads us to three realizations. First, it is now (relatively) easy to calculate the difference in total power system costs with and without wind and solar generation.[7] Second, a more explicit definition of "integration costs" is required. Third, the *without*-wind-and-solar case must be carefully designed. This discussion also leads to the realization that it may not be possible to rigorously define or calculate integration costs in the first place. The concept of integration cost may be simple. But it may be impossible, or at least difficult, to calculate.

The second difficulty arises because of the complex interactions between resources and loads that make it difficult or impossible to untangle costs and allocate them to the cost-causer.

Wind and solar generation provide valuable energy, they reduce emissions, and they help meet renewable portfolio standards. However, integration cost studies are supposed to determine what *additional* costs the power system incurs as it responds to the variability and uncertainty of wind and solar generation. Additional reserves will be required to maintain reliability. Conventional generators will cycle more as wind and solar output varies. The unit commitment is likely to be less optimal to deal with wind and solar uncertainty. Production cost modeling captures these costs, but it also captures the value of the wind and solar energy itself.

[7] We do not mean to trivialize the difficulty of performing this analysis. As will be discussed below, significant data is required and significant care must be taken to perform the analysis correctly. There are many common errors that some analysts continue to make.

Integration studies that have attempted to calculate integration costs have typically considered the variability component of the net load. Adding VG to the system increases variability and uncertainty, and integration cost studies attempt to capture this on different time-scales (regulation, load following, and scheduling). However, increasing the generator-supplied reserve requirement is only one of the options available to the balancing authority (BA). The WWSIS pointed out that increasing demand response would be a cheaper option than increasing contingency reserve generation capacity to address shortfalls. BA cooperation and intra-hour scheduling between BAs may be a still cheaper way of dealing with the variability than increased generation or load-based reserves.

Two basic schools of thought have emerged concerning integration costs. One school tries to develop a zero-energy-cost-proxy for the without-wind-and-solar base case resource that supplies the wind and solar energy, but that does not have the wind and solar variability and uncertainty. Total system costs can be compared between the base case and the with-wind-and-solar case to determine the added integration cost that results from variability and uncertainty. The other school maintains that it is not possible to develop a suitable proxy and that only total costs with and without renewables can and should be compared.

Proxy Resource

Determining the integration cost associated with the variability and uncertainty of wind and solar generation should be straightforward. Model the power system with and without the renewable resource and compare the total system costs. Use security-constrained unit commitment and economic dispatch production-cost modeling. Base the modeling on time-synchronized load, wind, and solar data with as high a resolution as possible for one or more years.[8] In the likely event that the wind and solar generation does not yet exist, be very careful to create realistic time-series wind and solar generation data through mesoscale modeling of the atmosphere for each renewable generation site. All of this takes considerable effort, but advances in analysis techniques show how to do this effectively. The real difficulty comes with developing an appropriate non-variable and non-uncertain proxy energy resource for the base case.

A proxy resource is needed in the base case to supply the wind and solar energy without the variability and uncertainty. This will remove the energy value when the total costs are compared from the base case and renewables case, leaving only the integration costs caused by the wind and solar variability and uncertainty.

Initial wind integration analysis efforts used a flat block of energy equal to the annual average wind and solar output. For example, 10,000 MW of wind with a 35% capacity factor would be modeled in the base case as a constant 3,500 MW, zero cost energy supply. Variability and uncertainty are removed from the base case. Unfortunately the *value* of the flat block energy is

[8] Historically, 10-minute wind data has been acceptable, but higher resolution may be required for solar variability and for more finite dynamic analysis.

significantly different from the value of the wind and solar energy, so the calculated integration cost is not valid.[9] The proxy flat block likely has more energy in the high-priced summer season than the actual wind. The proxy also likely has more energy during peak hours of the day than wind. The bias for solar is likely opposite, with solar generation having more daily on-peak and summer generation than the flat block proxy. The flat block proxy also has high capacity value which wind and solar may not be able claim. While the cost differences between the flat block proxy and the renewables are very real, they are not a good representation of the costs of variability and uncertainty because they include a significant component based on the difference in the temporal value of the energy itself.

Daily flat blocks have been tried as a proxy resource. These correct the seasonal bias but not the on-peak/off-peak bias. Daily blocks also introduce an unrealistic daily step change in the base case. This step change can be partially mitigated by ramping between days. Daily on-peak and off-peak blocks, with ramps, have also been tried as proxies in integration studies. While an improvement, none of the proxies has proven to be entirely satisfactory. Figure 4 shows one week of daily flat-block energy and wind energy (top panel), the difference in market value of the wind energy and the daily block (middle panel), and the difference in the market value of wind energy and a 6-hour fixed energy block (bottom). Use of a daily flat-block proxy energy resource would overstate the wind integration cost by $1.06/MWh based simply on the difference in the energy values.[10]

[9] See Milligan, M.; Kirby, B. (2009). Calculating Wind Integration Costs: Separating Wind Energy Value from Integration Cost Impacts. 28 pp.; NREL Report No. TP-550-46275.
[10] M. Milligan, B. Kirby, 2009, *Calculating Wind Integration Costs: Separating Wind Energy Value from Integration Cost Impacts*, NREL/TP-550-46275, July.

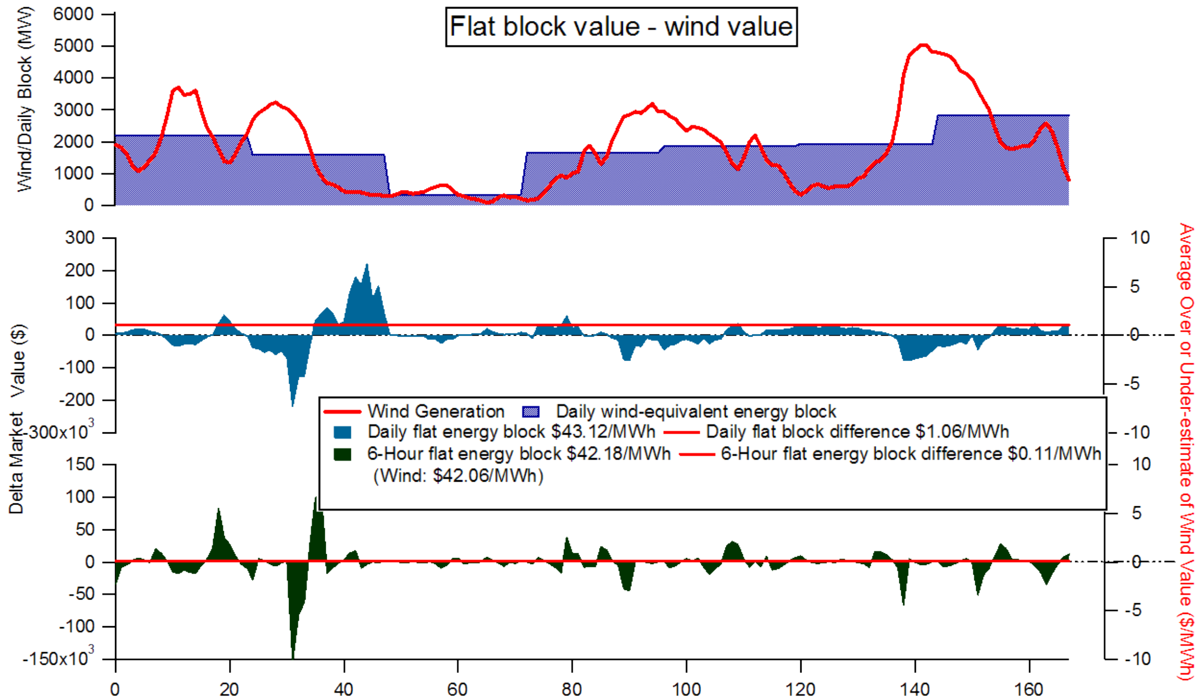

Figure 4 — One week of the wind energy and daily flat energy block (top); difference in market value (24-hour block, middle) and 6-hour block (bottom)

Comparing Total System Costs

An alternative analysis approach does not attempt to directly calculate integration costs. Instead, it focuses on calculating the total system costs and benefits of integrating renewables. The same time series, mesoscale data-based security-constrained unit commitment and economic dispatch modeling is performed but without the proxy resource in the base case. The total change in production costs is calculated including the value of saved fuel, as well as the inefficiencies introduced by the variability and uncertainty of the wind and solar generation.

Adjusting the Conventional Generation Mix

The conventional generation mix presents an additional complication for wind and solar integration studies. If the study is examining the addition of wind and solar generation over the next few years, the composition of the conventional generation fleet will be generally known. If the study is examining a wind and solar build out over the next decade or two, the composition of the conventional generation fleet may not be as certain. The optimal conventional generation mix will likely be different with and without renewables. The best conventional generation mix without renewables will likely include more baseload generation, focusing on the lower production costs from nuclear, coal, and combined cycle plants. The addition of significant amounts of wind and solar will increase the value of flexibility in the conventional generation fleet, while also reducing the capacity factor of many conventional generators. More flexible combustion turbines and reciprocating engine plants with lower capital cost but higher variable costs will likely be included in the optimal generation mix. Total cost can be calculated in both cases, but explicit integration costs are harder to define.

Higher Resolution Modeling

Production cost models typically run with an hourly resolution. Sub-hourly impacts have been effectively captured by adjusting the reserve requirements based on statistical analysis of ten-minute or faster wind, solar, and load data. While these methods have been refined in studies such as the Eastern Wind Integration and Transmission Study (EWITS), WWSIS, and the New England Wind Integration Study (NEWIS), advances in computational power and production cost modeling may be making it practical to reduce the modeling time step to ten minutes or even one minute. This could enable the modeling to directly capture the minute-to-minute variability and uncertainty and avoid the need to separately specify additional reserves.

Other Types of Generation Impose Integration Costs

Integration impacts are not exclusive to wind and solar. Nearly all generators can impose costs on the power system or other generators when they are added to the power system. These impacts are seldom calculated as integration costs and never applied to conventional generators as integration costs. Figure 5 shows an example of this effect by displaying the outputs of two similar coal fired generators. The generator in the top figure is providing regulation while the generator in the lower figure is imposing a regulation burden on the power system.[11]

[11] E. Hirst and B. Kirby 1996, *Ancillary-Service Details: Regulation, Load Following, and Generator Response*, ORNL/CON-433, Oak Ridge National Laboratory, Oak Ridge, TN, September.

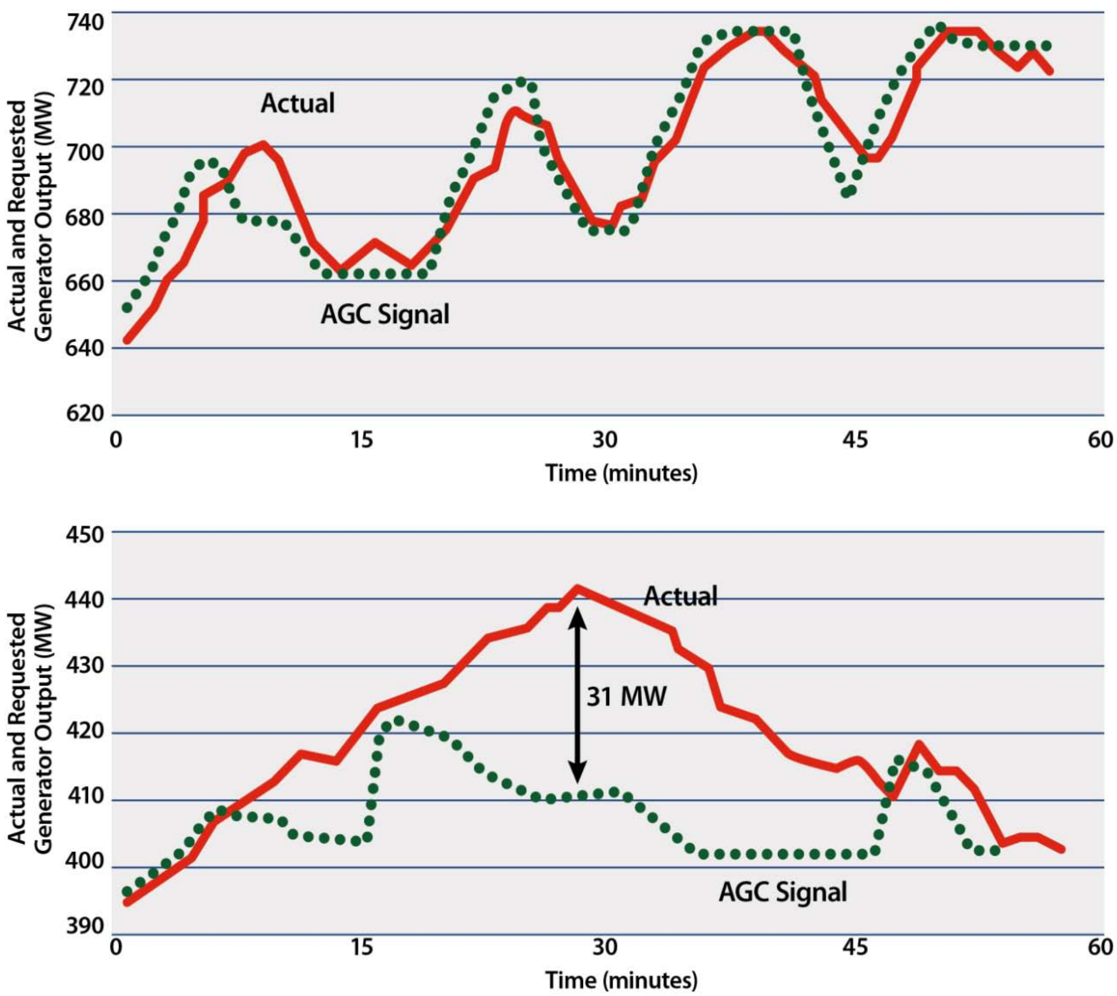

Figure 5 ─ Two similar coal fired generators: the upper generator is providing regulation while the lower generator is imposing a regulation burden on the power system.

Baseload plants can also increase the costs of operating other generators. Figure 6 provides a conceptualized comparison of the addition of wind generation to the addition of a new baseload generator on the cycling requirements for the rest of the generation fleet. The top figure shows coal generation providing flat output for the week. Combined cycle plants and combustion turbines cycle daily and follow the load ramps. The middle figure shows that the addition of wind energy does force coal to cycle somewhat. Wind also increases cycling of the combined cycle plants and reduces the total energy produced by the combustion turbines. The bottom figure shows that adding a nuclear plant (or any lower cost baseload generation) forces coal to cycle and also displaces both the combined cycle and combustion turbine-based generation.

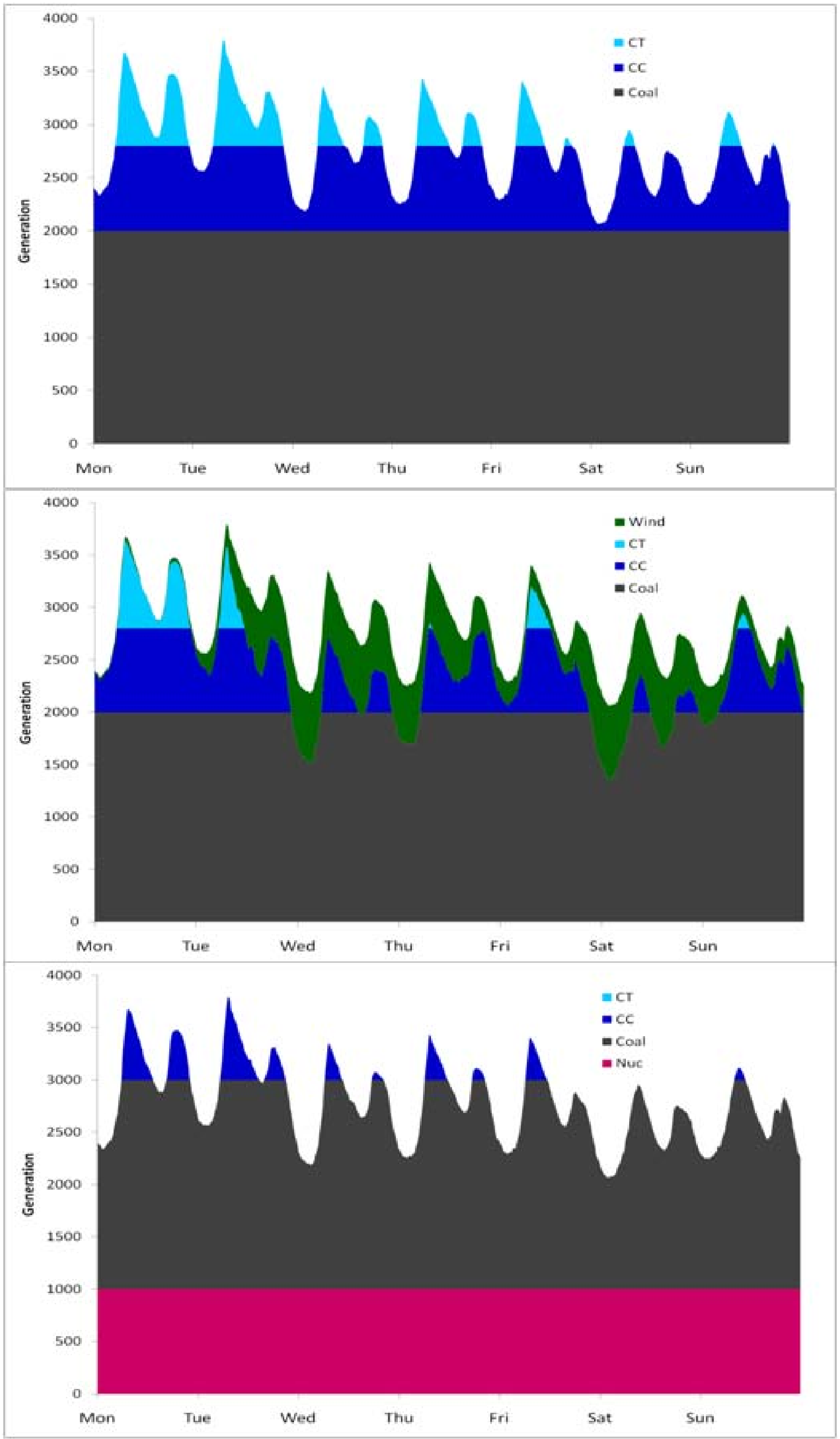

Figure 6 ─ The addition of any generation can impact the cycling requirements of the existing conventional generators.

Contingency reserve requirements result largely because some conventional generators are large. No generator is 100% reliable and the power system must continuously stand ready to respond if a large generator or transmission facility suddenly fails. Exact rules vary from region to region, but contingency reserve requirements are typically based on the size of the largest generator. Each BA or reserve sharing pool must keep enough spinning and non-spinning reserve ready to respond if a generator fails. The cost of maintaining these reserves is not allocated to the generators that cause the need, however. Instead the cost is broadly spread across all users of the transmission system. This has the effect of allocating costs based on capacity or output rather than on size or contribution to contingency reserve requirements. Costs could be allocated based on cost causation, as shown in Figure 7, but they are not.[12] Instead, these costs are socialized, and have been for many years. Current practice has the effect of subsidizing the large generators at the expense of the small generators (or their customers).

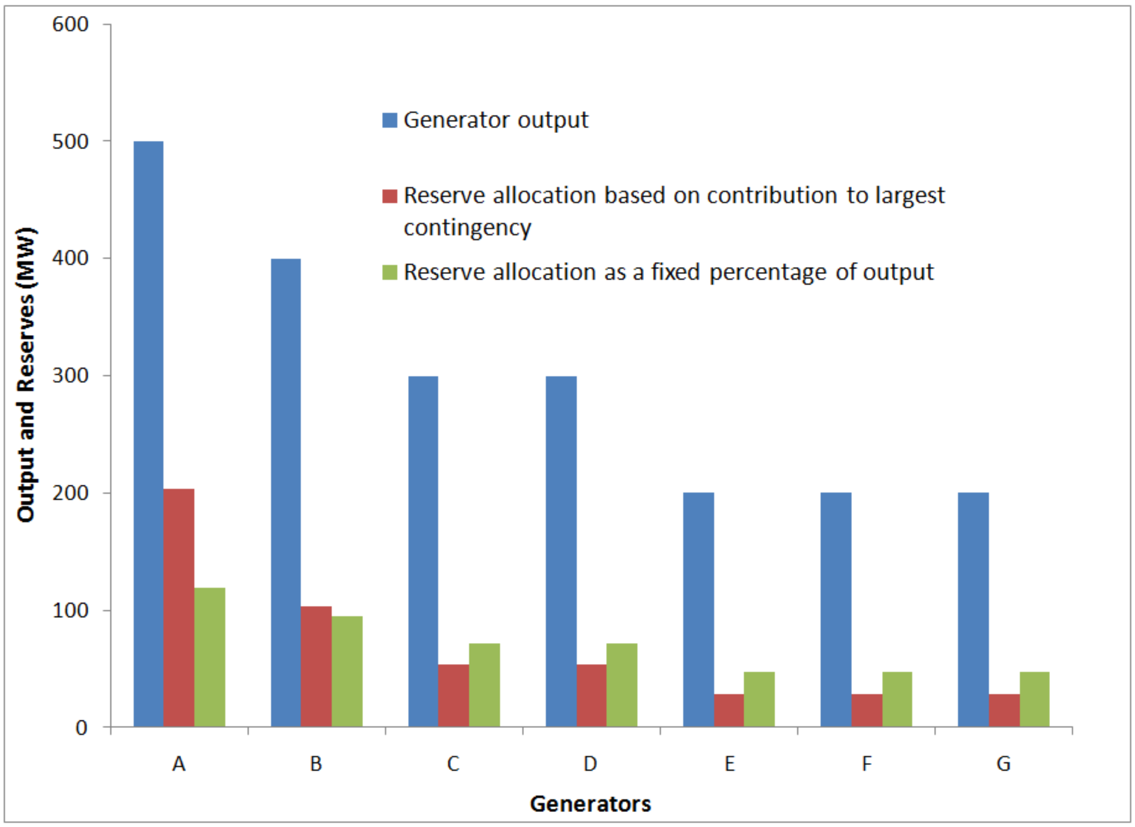

Figure 7 — Contingency reserve costs could be allocated to each generator based on their contribution to the contingency reserve requirements.

[12] E. Hirst, B. Kirby, 2003, *Allocating Costs of Ancillary Services: Contingency Reserves and Regulation*, ORNL/TM 2003/152, Oak Ridge National Laboratory, Oak Ridge TN, June.

Hydro Integration Costs

Run-of-river hydro is variable and somewhat uncertain, similar to wind and solar, though both the short-term variability and uncertainty of hydro are typically much less than for wind or solar generation. The same analysis techniques used to determine wind and solar integration costs are appropriate for run-of-river hydro.

Hydro generation with storage is typically very responsive with low cycling costs. It is a flexible resource that was often built at least partially for its flexibility. There is both seasonal and annual variability and uncertainty in the water availability, which can reduce the long-term capacity value. More recently, environmental restrictions associated with preserving endangered fish have reduced the flexibility available to the power system from many hydro projects, and may impose an integration burden. For example, excess water during times of light power system load may exceed the reservoir storage capability. Historically, the power system would use as much of the water as possible for generation and the rest would be spilled. This was an unfortunate and unavoidable economic lost opportunity, but nothing more. Better understanding of fish biology has led to additional operational restrictions. Spilling water can increase dissolved gasses, however, and now can represent an unacceptable threat to fish. The water cannot be spilled but must be run through the turbine generators. The power system must therefore accept the excess power. This may require uneconomic cycling of thermal power plants or curtailing wind with a loss in production tax credits and renewable certificates. What was previously just a lost economic opportunity (spilled water that did not generate electricity revenue) is now a direct cost (uneconomic cycling of thermal plants and curtailed wind production). This represents a real integration cost of constrained hydro.

While excess generation at times of minimum load can be an operating challenge, a real reliability concern and a significant economic burden, it also represents an opportunity. Fundamentally, this is a surplus of a valuable commodity rather than a shortage. The real solution should be in the use or storage of the excess, rather than its disposal. If the excess is significant and can be characterized, it seems likely that some load processes could be redesigned to make profitable use of the surplus energy. The load would probably not be designed to use only surplus power. Instead, an already viable process might adjust its operations and/or equipment to take advantage of surplus power when it is available. The type of process that could use the energy would depend on the frequency and duration of surplus events as well as the amount of warning or notification that the power system could provide. An initial step towards assessing the viability of this concept would be for the utility to estimate the surplus energy characteristics. How many MWh might be available each year for the next ten years? How long would surplus energy be available for each event? How many hours or minutes warning might the utility be able to provide? Further research could provide an initial assessment of what types of loads might profitably make use of the surplus energy. The initial assessment and the utility estimate of the resource amount could be publicized to solicit industry interest.

This particular power system problem should be an opportunity. At present, however, it does represent an integration cost of environmentally constrained hydro.

Gas Integration Costs

The physical flexibility of natural gas-fired generation depends on the technology. Engine-driven plants are extremely flexible. Some, such as reciprocating engines, can start and fully load in two minutes with zero cycling costs. They have low minimum loads and a broad operating range with high efficiency. These plants ramp quickly and accurately and can provide the full set of generation-based ancillary services.[13] Large gas-fired steam plants represent the other extreme, requiring hours to days to start. They are slow to ramp and are more limited in their operating range. Combined cycle plants and combustion turbines fall in between, although gas plants generally tend to be more flexible than most other generators.

Gas scheduling and contracting can limit the flexibility of gas fired generators significantly below the physical capability. Gas is typically nominated day-ahead, committing the generator to operate or not operate in essentially the same time frame as coal plant commitment. While there is ample physical capability to respond to changing load conditions and changes in the rest of the generation fleet up until the operating hour, the gas scheduling restricts this flexibility. This problem is compounded on weekends and holidays. Gas schedules are typically set on Friday for Saturday, Sunday, and Monday's operations. With a holiday, this stretches to four days and includes Tuesday. Gas scheduling restrictions represent a significant integration cost to gas generators that is not based on limitations in the physical capabilities of the generator.

Gas also presents another integration cost related to the potential for a common mode failure. Extreme weather conditions can result in gas shortages which impact all gas fired generators in a region. This represents a much larger contingency than the power system is designed to survive. System operators are forced to shed firm load to cope with the loss of generation. This occurred in Texas and the Southwest in February 2011. This significant cost is born directly by the loads that are blacked out and represents an integration cost of natural gas, which is not allocated directly to natural gas generation.

Principles of Cost-causation

Wind integration costs can be thought of as a tariff that is assessed to recover the increased cost that wind causes to power system operations; they are a special case of a cost-causation based tariff. Cost-causation based tariffs provide transparent signals to markets and regulators that, if well defined, provide appropriate incentives for efficient investment and behavior. Kirby et al (2006) describe cost-causation based tariffs in the following principles:

---

[13] B. Kirby, 2007, *Ancillary Services: Technical and Commercial Insights*, Wärtsilä North America Inc., www.consultkirby.com, June.

1. **1.** Because maintaining power system reliability is critical, tariffs should base prices on costs so that the costs of maintaining reliability can be obvious to users of the system and its reliability future.
2. **2.** Tariffs should be based on cost-causation and the cost of providing the service.
   a. Those individuals who cause costs to the system should pay for those costs;
   b. Those individuals who mitigate costs to the system should either incur a lower cost or be paid for helpful actions;
   c. Complex systems like electric grids produce both joint products and joint costs of production that must be allocated among users of the system;[14]
   d. Tariffs should allocate joint production costs on the basis of the use of joint products (the cost allocation principle of "relative use").
3. **3.** Tariffs should not collect revenue if no cost is incurred.
4. **4.** Tariffs should be based on the physical behavior and characteristics of the power system.
   a. Recognize the need to balance aggregate system load and aggregate system generation;
   b. Recognize that balancing individual loads or resources is not necessary, is inconsistent with power system operations and, is very costly.
5. **5.** Tariffs should result in an efficient allocation of resources.

Tariffs can be tested empirically, both with real-world data or detailed modeling. During tariff design, hypothetical cases can be tested by the tariff to ensure underlying principles are upheld.

There are some broader principles that tariffs should also support. The first is horizontal consistency. Horizontal consistency means that if two individuals (loads or generators) each cause equal increases in costs, then the tariff should assess each of them the same amount. A corollary to this principle is that if two individuals impose similar costs, then they should be assessed similar payment amounts. We can extend the principle of horizontal consistency in cases where individuals contribute to cost mitigation. Equal cost mitigations or reductions should be matched by either identical reductions in cost assessment to the individuals, or equal payments to the individuals. If two individuals have similar cost mitigation impacts, then their payments should be similar.

Vertical consistency is the second additional principle. Vertical consistency implies that if individual A imposes a larger cost than individual B, then A should pay more than B. We can extend the concept of vertical consistency to cases where two individuals mitigate costs in a straightforward manner.

---

[14] The classic example is a sheep. A farmer raises a sheep. She sells mutton, hide, and wool. These are joint products. She incurs various costs for raising the sheep: feed, medicine, and a sheperd. These are joint costs of production. The electric system produces joint products: reliability, energy, capacity, convenient system access, ancillary services. The costs for producing these joint products (for fuel, engineers, capital, maintenance) must be allocated to the joint products. The most common allocation principle is relative use; the more you use, the more you pay.

Horizontal and vertical consistency can be empirically tested, either through real-world experience or through detailed modeling of the grid and the individual behaviors in question. Application of the tariff to the individual behaviors can determine whether horizontal and vertical consistency is achieved by the tariff.

However, it is important for regulatory bodies to exercise great care in creating such tariffs lest they elect to only create tariffs that recover integration costs from only some parts of the system while allowing free-riders in other parts of the system. Unfortunately, this is currently the case in some parts of the electric industry where utilities are requesting separate wind or renewables integration tariffs without creating similar tariffs to recover integration costs for other conventional forms of generation. Rather than focusing on technology-specific tariffs, it would be appropriate to focus on performance-specific characteristics. This approach would allow any technology to adapt so that it could supply needed response, and converge to a cost-causation approach that would reduce or elminate deadweight loss.[15]

De-Composition and Re-Composition

The variability of wind and solar is often de-composed into regulation and load following components. We agree with and support this type of analysis because the decomposition allows the variability to be analyzed in the context of normal system operational procedures, as shown in Figure 1. We can illustrate this concept with a simple mathematical formulation:

(1) $V(g) = L(g) + R(g)$

where g represents the variable generator fleet, $V(g)$ is the net load vector, $L(g)$ is the load-following component of the net load accounting for g, and $R(g)$ is the regulation component of the net load accounting for g. Each of the components L and R can be calculated from the original vector:

(2) $L(g) = p[V(g)]$
(3) $R(g) = q[V(g)]$

In a typical integration study, equations (2) and (3) are used to separate the load following and regulation signals for further analysis. However, equation *(1)* allows us to make the simple observation that *the sum of the regulation and load following signals must sum to the original time series.* We call this the 'principle of re-composition.' The power system itself only balances the total net load; $V(g)$. $L(g)$ and $R(g)$ must be defined so that they sum to the actual system requirement $V(g)$. This is important in integration analysis for the simple reason that we have seen this principle violated in numerous studies carried out by utilities. $L(g)$ and $R(g)$ are often

[15] See Kirby, B. and Milligan, M.; Wan, Y., (2006). Cost-Causation-Based Tariffs for Wind Ancillary Service Impacts: Preprint. 26 pp.; NREL Report No. CP-500-40073. Available at http://www.nrel.gov/docs/fy06osti/40073.pdf.

defined independently with the result that the total variability exceeds the actual system balancing requirements. That is, *V(g)* calculated from equation 1 is not the same as the actual power-system total net-load when it is recalculated from the independently calculated components *L(g)* and *R(g)*. This is similar to the problem of measuring the variability of individual loads and generators themselves and failing to account for the aggregation benefits that result from the lack of perfect correlation in individual fluctuations. We discuss this in more detail in a later section of this report.

This concept is not new for the utility industry. The power system only has to meet the system's coincident peak load, not the sum of the peak requirement of each customer or the sum of the peak requirements of each piece of load equipment. If a utility charged each residential customer based on the capital cost of generation multiplied by the sum of the ratings of the water heater, oven, stove, dryer, all lights, TVs, computers, air conditioning, etc., the utility would collect many times the total cost of all generation needed to serve load. Instead, the cost of generation is allocated based on the customer's contribution to coincident peak load, not the sum of the customer's equipment ratings and not even the customer's peak load itself.

The cost to follow system load is similarly much less than the sum of the costs to follow the individual loads that, in aggregate, comprise the system load. This benefit occurs because the individual loads, especially the fluctuations, are generally not correlated with each other.[16]

Testing a Tariff with Thought Experiments

Thought experiments provide a means for testing a tariff to assure that it does what is intended and that it does not have undesired consequences. The behavior of the wind and solar plants, other generators, loads, and power system components are carefully specified to test each tariff attribute of concern. Here we present five thought experiments that can be used to test how a regulation tariff assesses a volatile resource like wind. Each thought experiment is mapped to at least one of our tariff principles.

Thought Experiment #1: Perfect Following of a Volatile or Block Schedule

In formal transactions, both loads and generators forecast their expected behavior and establish a schedule for generation or consumption. Regulation tariffs often impose penalties if a resource does not follow its schedule. Some tariffs are based exclusively on schedule deviations. The reasoning is that the system operator must have a reserve of regulating resources available to immediately compensate for unexpected changes in a generator or load's output or consumption. This is true. But does the regulation resource requirement go away if the resource follows its schedule perfectly? Figure 8 presents a typical system daily load with blocks of generation scheduled to meet that load. If the generation follows its schedule perfectly, is there a regulation burden imposed on the system? What charge does the tariff impose?

---

[16] E. Hirst and B. Kirby 1996, *Electric-Power Ancillary Services*, ORNL/CON-426, Oak Ridge National Laboratory, Oak Ridge, TN, February.

A regulation tariff that is based exclusively on schedule deviations would impose no charge on the block-scheduled generator. Indeed, many feel that scheduled imports and exports impose no regulation burden because the schedule is precisely known, often days in advance, and it is typically adhered to.

The right side of Figure 8 shows that block scheduling imposes severe ramping requirements on the system, adding $2.26 to the cost of each MWh delivered through the block schedule in this example (based upon modeling an example control area). The fact that these requirements always happen at the top of the hour and they are known well in advance does not reduce the amount of fast response capability the system operator needs to have to balance the system and meet CPS 1 & 2 requirements. The tariff needs to assess the individual's impact on total system variability.

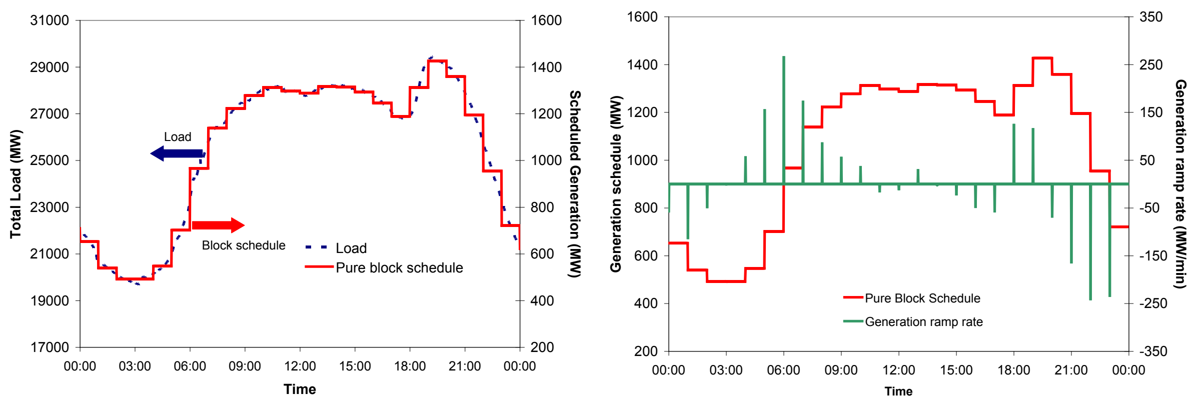

Figure 8 − How does the tariff treat perfect following of a volatile schedule?

This example tariff would violate principle #2 (cost-causation, see Principles of Cost-causation, earlier section) because under the (unlikely) scenario of a perfect wind or solar energy forecast, the tariff would not assess any cost to the wind or solar generator even though there is a cost of moving the regulating units to mitigate variability in the wind output signal. It also violates principle #4, which says that individual movements (or in this case schedule deviations) of individuals do not need to be matched by a responsive unit – only the aggregate variability of the entire system must be compensated. Extrapolating this type of tariff to a case when all schedules and loads are known perfectly in advance, the implication is that there is no cost to the system to manage the total system variability. This is clearly wrong, and would result in distortions in the market.

Thought Experiment #2: Reduced Ramping

It is tempting to design a regulation tariff that simply quantifies the peak-to-peak movements of the generator or load. But this ignores the speed at which the resource moves from one power level to another. If the block schedule used in Thought Experiment #1 (where the schedules changed abruptly at the top of each hour) is provided with 20-minute ramps (where schedules linearly ramp from ten minutes before the hour to ten minutes after the hour), as shown in Figure 9, the regulation costs imposed on the power system drop to $0.20 per MWh (again based on modeling an example control area).[17] Note that the ramp rate scale on the right axis of Figure 9 is one tenth of that in Figure 8.

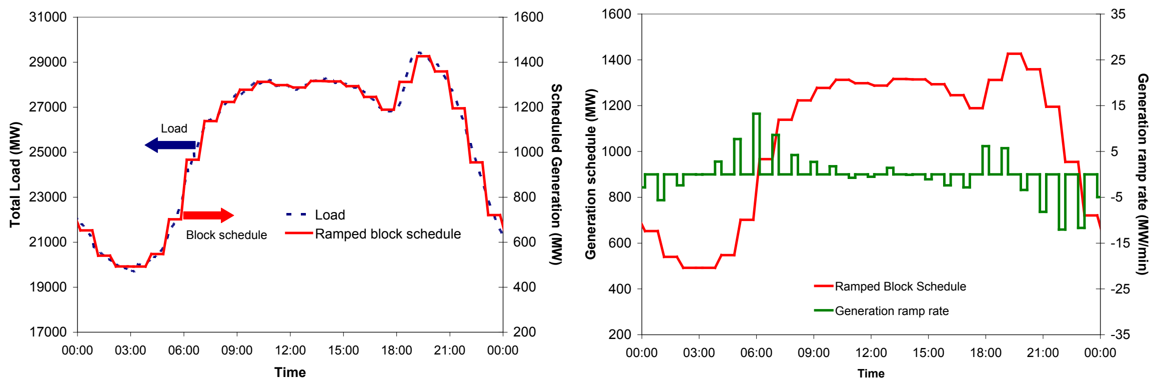

Figure 9 − Ramping the block schedule does not impact the energy delivery or forecast accuracy but reduces regulation requirements.

This thought experiment violates principle #2, the principle of cost causation. Recognition of only the peak-to-peak ramp does not distinguish between the two behaviors illustrated here that have significantly different cost impacts. This also violates the principle of vertical consistency because there is a significant difference in imposed cost that would not be picked up in the tariff.

Thought Experiment #3: Ramp Rate or First Derivative Metrics

Another tempting regulation tariff simplification is to measure average ramp rate or the average first derivative of the minute-to-minute energy consumption. This can also be characterized as a "distance traveled" metric referring to the amount of "movement." This attempts to quantify the amount of ramping or changing of output that a generator has to provide. The flaw in this simplification is that behaviors with very different system impacts can result in the same measured performance, as shown in Figure 10.

Figure 10 compares the behavior of three hypothetical individuals (loads, wind, or solar generators, or balancing areas). The minute-to-minute change ("line slope"), integrated over the hour, is the same for all three; 60-MW-minutes. Clearly, however, the regulation burdens imposed by the three are radically different. In this very simple example, the solid red entity requires 1 MW of regulation compensation. The dashed green entity requires 5 MW. The dotted blue entity requires a total of 60 MW, but not of regulation. A sustained ramp is a following requirement that can be, should be, and is (in most locations) supplied by moving the baseload and intermediate generators. There is no regulation burden imposed by the dotted blue ramp.

[17] Ten-minute ramps for interchange scheduling are standard in the eastern interconnection. The western interconnection benefits in reduced regulation costs from the use of twenty-minute ramps.

Metrics based on average rate of change of an individual violate principle #2 (cost causation) and principle #4 (failure to recognize aggregation benefits).

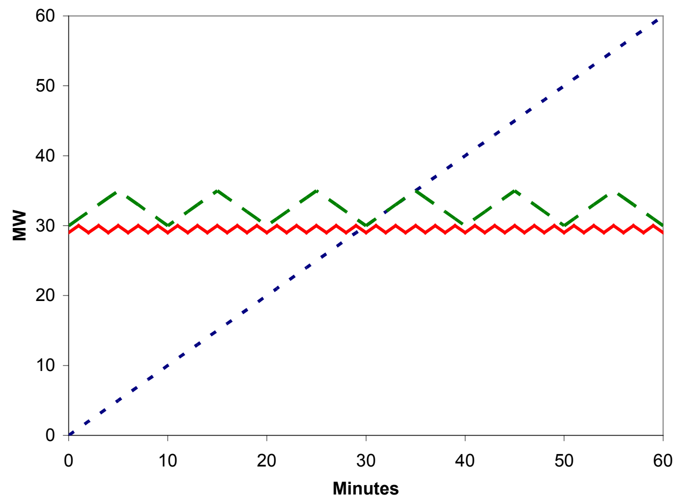

Figure 10 − These three individuals impose radically different regulation requirements but have the same minute-to-minute-change metric performance.

Thought Experiment #4: Equal but Opposite Behavior

One very powerful feature of thought experiments is that they can be carefully tailored to examine specific behavior characteristics. They do not have to be realistic to be useful in determining if a tariff will produce desired results. Unrealistic examples can be useful in understanding the pieces of complex behavior that are often buried in the intricacies of actual operations.

When designing a regulation tariff, it is tempting to assess the generator's or load's variability in isolation. This ignores the fact that the underlying reliability requirement to balance generation and load is imposed on the BA (hence its name) rather than on the individuals. Figure 11 shows two mirror-image wind plants and a total system load. If the wind plants were assessed for their variability in isolation of each other and the total system load they would both receive an identical regulation variability assessment. Together they present an absolutely constant output with no regulation burden.

This thought experiment is completely unrealistic but it illustrates an important point. A tariff that can not recognize complete compensation of one plant for another will not recognize more subtle interactions or uncorrelated behavior that, consequently, does not add linearly.

A tariff that does not recognize the impact of equal but opposite behavior would collect payment from both of these hypothetical wind plants. However, because their impacts net to zero, there would be no cost to the system. This type of tariff would therefore violate principle #3 (the principle that if no cost is incurred, the tariff should not collect revenue) and principle #4 (the recognition that only the aggregate system variability must be compensated for).

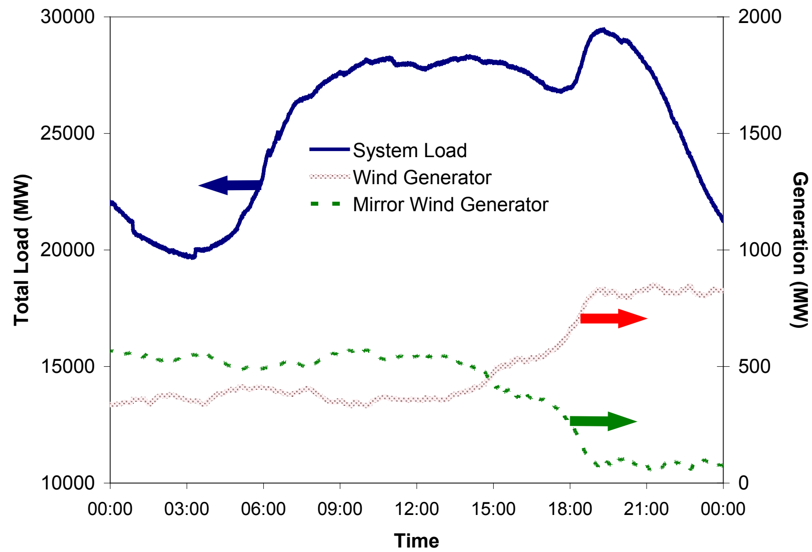

Figure 11 — How is equal but opposite behavior treated?

Thought Experiment # 5: Beneficial Movement

The last thought experiment asks how the tariff treats movement that is beneficial. Regulation tariffs that only assess variability (total range, ramp rate, or adherence to a schedule) can penalize a resource that is actually helping reduce the total system aggregate variability. Figure 12 presents the measured variability of a number of generators and a total system. A tariff that simply charged for variability would penalize the automatic generator control (AGC) generator that is deliberately balancing the system. Presumably the tariff would not be applied to this generator but the principal remains the same. A generator that inherently has favorable response characteristics for whatever reason should not be penalized.

A tariff that assesses a cost based on an individual's variability in isolation of what the system needs would discourage helpful behavior. Because this type of tariff would impose a cost on the generator when in fact the resource is providing a system benefit and reducing system costs by helping to mitigate system variability, this kind of cost in a tariff clearly violates principle #2 (cost causation), principle #3 (imposing a cost instead of paying the generator), and principle #4 (does not recognize system balance).

Common Errors in Integration Analyses

In our experience participating in technical review activities for most major wind integration studies in the United States, we have seen honest mistakes made in the technical analysis. Because it is not our intent to single out entities that have committed these errors, we do not identify them, but attempt to extract the issues so that they can be identified and subsequently contribute to more accurate analyses. We also point out that the studies we have been involved with have occurred over the past decade, and there has been considerable evolution in methods and data. However, in spite of significant progress, there are many entities that are apparently not engaged in this evolution and repeat some of the errors that have been previously identified, and for the most part, corrected with the evolution of studies.

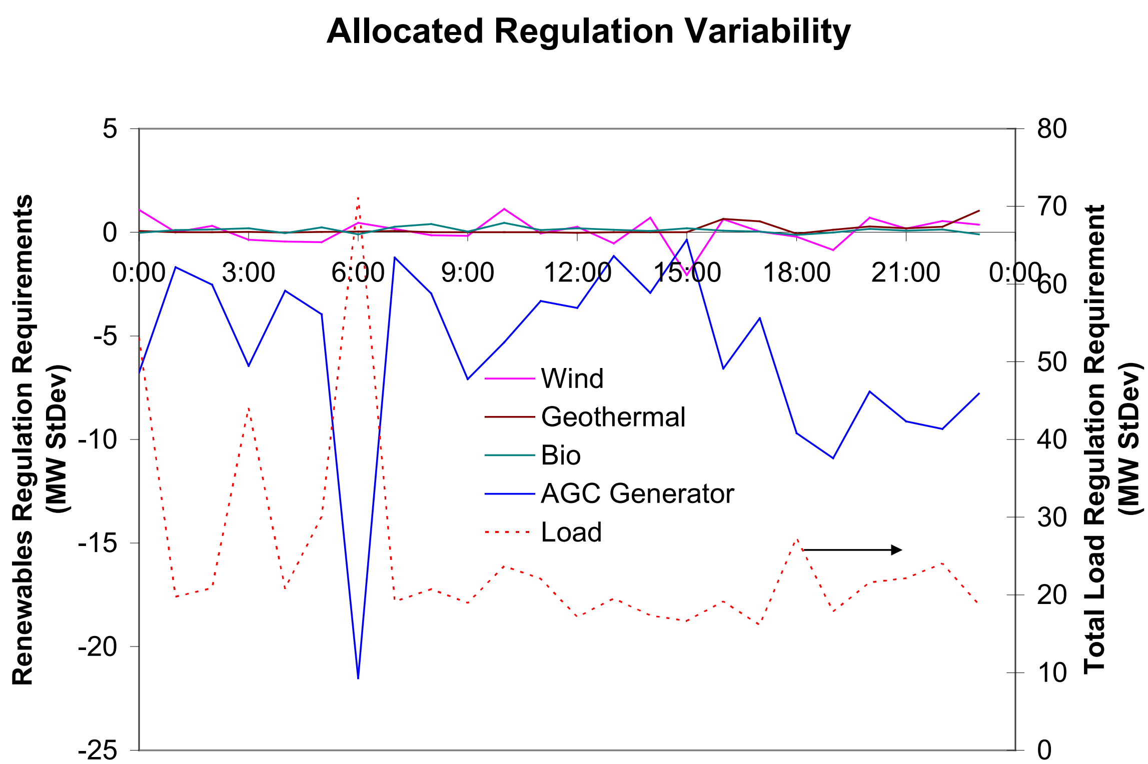

Figure 12 — Tariffs should recognize and reward favorable behavior.

Double Counting

Double counting in one form or another is probably the most common error made in integration studies. This usually results from failing to account for aggregation benefits, either among wind facilities and/or between wind and load.

Double counting can also result from including the same variability or uncertainty in multiple services. There are several ways in which this error can be manifest in the integration studies. The most common is violating the principle of re-composition. For example, a rolling average or other suitable filter is applied to the net load to separate regulation. Load following is then estimated by applying heuristics or rules of thumb based on utility scheduling practice. When these rules of thumb include the entire wind or solar output, which naturally contains both following and regulation components, double counting occurs. The sum of the reserves required for regulation and following should not exceed the total system balancing requirements. Similarly, wind, solar, and load balancing requirements are often calculated separately, which is only valid if these parameters are perfectly correlated (which is not plausible). The sum of the balancing requirements allocated to individual wind and solar plants, and loads should not exceed the total system balancing requirements.

Forecasting errors are often another source of double counting. Production simulation runs may be performed assuming perfect foresight: the load, wind, and solar are forecast perfectly and system costs are calculated. Another simulation is run, this time using wind and solar forecasts for the unit commitment process, and using "actual" wind and solar data for the economic dispatch. The total costs of the two simulation runs are compared and the difference is the integration cost. However, the impact of load forecast error has not been removed nor accounted for. Therefore, part of what remains in the erroneously calculated wind/solar integration cost is in fact the integration cost of load forecast errors. Similarly, the variability of the variable generation (VG) resource and the relevant forecasts are not perfectly correlated, and are thus not additive.

Load and wind forecast errors typically do not add linearly and consequently benefit from aggregation. The sum of the forecast error reserves allocated to wind and load should not exceed the total system forecast error reserves. If individual wind plant forecasts are considered, the problem is amplified. Diversity among the individual wind plant forecasts should be considered, with total reserves reduced accordingly.

Another form of double-counting is the overestimation of reserves that are needed to balance wind and solar energy. Some amount of reserve is naturally provided as a function of economic operation of the system; as mentioned earlier, integration studies that attempt to calculate integration costs typically compare the increased variability of the net load of the case with wind/solar versus the case without wind/solar. For example, increased load-following reserve requirements could be estimated by examining the distribution of 10-minute net load deltas. On average in the WWSIS, the load following reserve requirement increased by a factor of two. However, it was only by running the actual production simulation analysis and looking at the amount of load following reserves online during each hour that it was found that the system *naturally provided* these extra load following reserves, because many of the thermal units were backed down, as opposed to being decommitted. As a result, WWSIS recommended that no additional load following reserves needed to be deployed because they would naturally be available as wind/solar came online. Therefore, no additional cost for committing extra load-following reserves was incurred.

Fixed Schedules and Fixed Resources

Fixing transaction schedules, often hydro schedules, based upon the without-wind case optimization and holding those schedules for the with-wind case typically results in seriously sub-optimal resource scheduling and significantly higher balancing costs. Typically, system operators would not schedule hydro or conventional generation while ignoring the presence of significant amounts of wind generation. Accounting for wind (and load) forecasting errors is appropriate, but fixing hydro schedules based on assuming no wind is not.

A related error is the assumption that only a subset of generation is available for balancing response. This error is typically made when studying a region where response has historically been obtained from only one resource or only one type of resource. A BA with a significant amount of energy-limited hydro generation, for example, modeled integration of large amounts of wind and calculated high integration costs. The modeling showed that the maneuvering capability of the hydro unit was exhausted. The analysts did not allow the production cost software to utilize the response capability of the conventional generators simply because that had not been the historic practice. It had not been historic practice because there was no need, not because there was any actual limitation. Another case results from the assumptions made about hydro response. Even though hydro generation is subject to various constraints, it should not be treated as a constant resource in integration analyses. Models are often capable of accurately representing ramp constraints, over various time steps, and yet these constraints are sometimes substituted by constant constraints on the hydro performance. These will nearly always result in different commitment and dispatch scenarios.

It is appropriate to hold the conventional installed generation mix relatively constant when studying renewable integration for the next year or two. It is not appropriate to fix the resource mix for studies examining conditions over decades, or very large changes in penetration rates of wind or solar generation. There will be generation retirement and new installations with or without wind and solar additions. The selection of the optimal generation fleet will likely be strongly influenced by the expected presence of large amounts of wind and solar. Similarly, demand response will likely be developed if the flexibility it offers is valuable and properly valued.

Scheduling practices are changing within the power industry, with or without wind and solar. Half of the load in the country is now located in regions that have five-minute energy scheduling. The fraction is continuing to grow. Some studies still calculate excessively high balancing cost based on an assumption that only hourly scheduling will be allowed in the future. This restricts access to the response capability that physically exists in the conventional generation mix.

A final scheduling error involves bilateral contracts. While the host utility that is conducting the integration analysis may have no control over some of the bilateral contracts within its BA, it is more reasonable to assume that they will change to reflect economic opportunities rather than assuming they will remain fixed for decades.

Balancing Individual Wind Plants or the Wind Fleet

Power system balance requires the aggregate load to be equal to the aggregate generation (ignoring imports and exports for this discussion). Therefore, not every generation movement in a wind or solar plant must be matched one-for-one with a movement in other generation. If wind generation increases at the same time as load increases, this reduces or eliminates the need for other generation to follow the load increase. Similarly, if solar generation is decreasing when load decreases, there is no need to increase other generation to fully compensate for the decline in solar generation. The concept of balancing the net load with conventional generation is well-understood in the integration literature and power system operations. In fact, the NERC Area Control Error (ACE), Control Performance Standards (CPS1&2) standards, Disturbance Control Standard (DCS), and balancing requirements are based upon it. However, within the past year we have seen two integration analyses that have attempted to balance wind and solar in isolation from the remaining load. This means that when wind/solar and load are both increasing, a conventional generator must decrease output to hold the wind and solar constant, but at the same time, generation must increase to meet the increasing load. This does not reflect how power systems are operated and greatly overstates the balancing costs of wind and solar.

Scaling

Wind and solar integration studies typically study future conditions when there is expected to be a larger amount of wind and solar generation present in the power system. Almost by definition, there is no actual wind and solar data available to study. A common error is to scale the output of an existing generator to represent the expected output of a larger fleet. This greatly overstates the variability of wind and likely overstates the variability of solar.[18] Figure 13 shows the output from a set of 300 wind turbines. Clearly, linearly scaling the output of a single turbine (left curve) would dramatically overstate the variability of the plant (right curve). Unlike conventional power plants where a larger individual generator can be installed, larger amounts of wind and solar require installation of more individual wind turbines and solar collectors. There is inherent geographic diversity, even within a single facility. This reduces the correlated variability.

It is similarly inappropriate to simulate a new wind plant simply by time delaying or advancing the output of an existing plant based on prevailing wind speed and direction. Wind does not remain coherent over inter-plant distances, so the resulting simulation will have too much correlation and too much variability. Mesoscale modeling is currently the best way to generate the required time-synchronized wind and load data needed for valid integration studies. This also

---

[18] See Wan, Y. H. (2004), Wind Power Plant Behaviors: Analyses of Long-Term Wind Power Data, 66 pp.; NREL Report No. TP-500-36551 or Wan, Y. H. (2005), Primer on Wind Power for Utility Applications, 45 pp.; NREL Report No. TP-500-36230.

points out the importance of expanding the database of high-resolution resource (wind, solar) data that is updated annually.

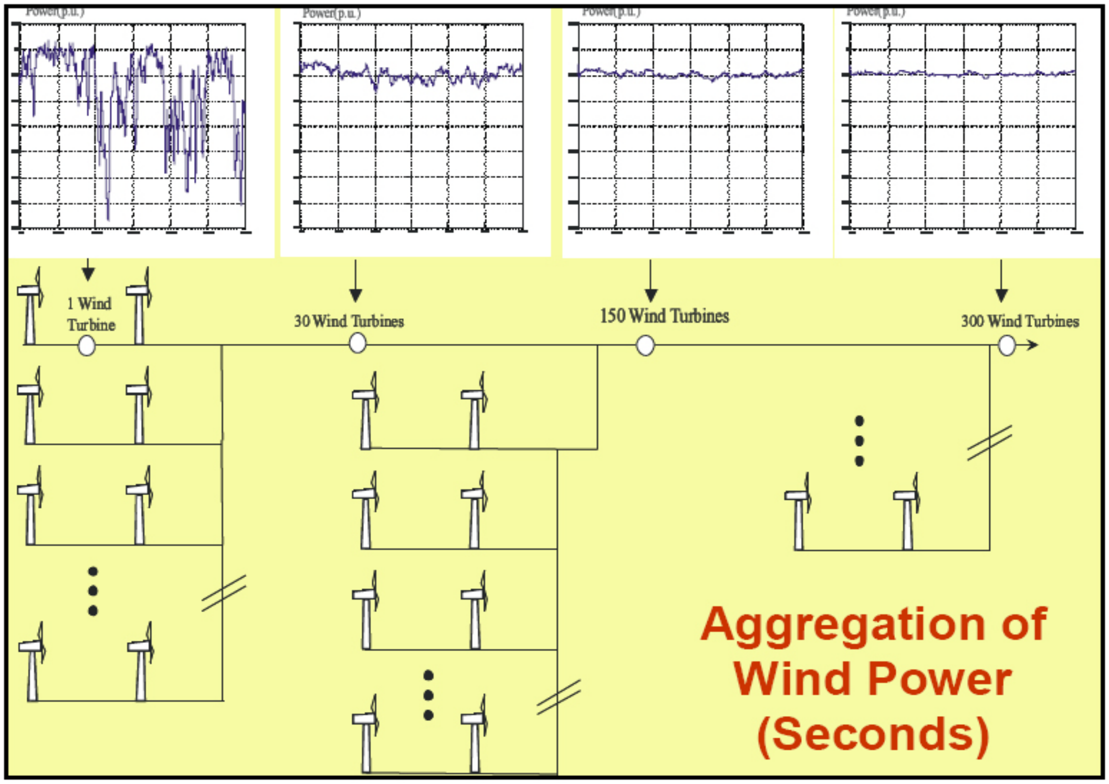

Figure 13 — Linear scaling is not appropriate for wind or solar generation.

Synchronized Load and Mesoscale Wind / Solar Data vs. Statistical Data Synthesis

Wind and solar integration studies typically model wind and solar plants that do not exist yet. Wind and solar output data must be generated for the study. It is tempting to create data with appropriate statistical characteristics based on a shaped series of random numbers. The temptation should be resisted. Not enough is known about all of the interactions between wind and load. Both are driven by atmospheric conditions, but the relationships are not simple or straightforward. The large geographic areas with varying conditions further complicate the relationships. Inter-temporal relationships for wind, solar, and load are complex and not well understood. Results from integration studies performed with statistically generated wind data do compare well with actual data taken after the wind generation was installed.

The best current technology for generating wind data for integration studies is to use mesoscale atmospheric modeling to calculate wind speed at hub height at locations of expected wind plants. Wind speeds are typically generated at ten-minute time steps or less with a geographic grid spacing of 2 kM or smaller for a historic time period of a year or longer. The mesoscale wind data must be time-synchronized with actual load data. The load data can be scaled to future conditions. Linear scaling of load data does introduce some error, but that error is typically reasonably small because the scaling and variability are also reasonably small and well-understood.

Forecast data

Similar to the wind and solar dataset, the wind and solar forecast datasets must be time-synchronized to historical weather patterns. If the forecasts are assumed to be generated by a single provider for a large region, then the forecasts should show similar spatial correlation to the wind and solar datasets, over varying distances. That is, the wind forecast for a plant is more likely to be accurate if the wind forecast for a nearby plant is also accurate. If forecasts are assumed to be generated individually for each power plant by different providers, they may show less spatial correlation over larger regions. Temporal correlation must also be preserved. There is a much higher probability of significantly missing the forecast in the current hour if the provider missed the forecast in the previous hour. Forecast error distributions are not normal distributions. Missing the tails of the forecast error distributions can underestimate the uncertainty impacts of wind and solar.

Excessive or Unknown CPS Performance

It is important to maintain consistent reliability between the base case and the high wind and solar penetration cases. This assures that integration costs are not either subsidized by reduced reliability or charged for increased reliability. Maintaining an excessively high CPS score can be inappropriate as well. If, for example, a BA has abundant reserves and is able to maintain a CPS2 score of 98% at little cost with no wind, it is not necessary or appropriate to require a high penetration case to also have a 98% CPS2. Instead, CPS2 requirements should be relaxed to 92%-93% for both the base case and the high penetration case. This is reasonable because it reflects how operations at that BA would likely evolve as reserves became scarce and therefore more accurately reflects true integration costs.

Replacement Power Cost Assumptions

An early wind integration study calculated high integration costs based on an assumed differential in up and down balancing costs. Balancing power required to compensate for a wind power shortfall was assumed to come from quick-start combustion turbines because, it was assumed, no excess coal capacity would be committed day-ahead. Conversely, any excess wind power would be credited with the fuel saving from backing down coal since, it was assumed, only coal-fired generation would be running absent wind. The result was that wind was charged $70/MWh for shortfalls and credited $20/MWh for excess, creating a default $50/MWh imbalance charge that was characterized as an operating cost, not as a penalty. While the described situation could happen during some hours, it will not be the norm during most hours. Imbalance costs should be calculated through economic dispatch and will typically be nearly equal for up and down reserves during most hours.

Constant Reserves

Numerous integration studies continue to assume fixed reserve amounts for wind and solar integration. Clearly, additional up reserves are not required if wind or solar are operating at full capacity. Similarly, additional down reserves are not required if wind or solar are at zero output. Reserves should be adjusted at hourly intervals to reflect the expected operating conditions, with an allowance made for forecast error. Holding fixed reserves overstates integration costs and does not reflect good operating practice. Integration studies performed during the past several years have recognized the dynamic nature of this reserve, and methods have evolved and are expected to continue to improve as more advanced methods are developed and tested.

Failure to Release Reserves

Wind and solar ramp events are similar to contingency events for conventional generators, except slower. While a large thermal generator can trip off instantaneously, it typically takes hours for a similar sized wind ramp. The events are similar, however, in that they are relatively rare. Non-spinning and supplemental operating reserves are often appropriate since the standby costs are more important than the deployment costs. There is an important modeling difference with wind and solar integration, however. Conventional contingency reserve requirements are modeled as reserves that are held throughout the analysis. This is because the contingencies themselves are not actually modeled. Instead, the analysis appropriately assumes that reserves must be available at all times, ready to respond to a random failure. Appropriate reserves must be included in the unit commitment time fame for wind and solar analysis, based on the forecast wind and solar conditions. Wind and solar reserves must be released and made available for response in the economic dispatch time. This is because the wind and solar ramps are modeled in the integration analysis. This is an important distinction between analysis of variable renewables and conventional generators. If the reserves are not released, then the model double counts the reserve requirements because it has to deal with the actual event while simultaneously holding additional reserves.

Other Assumptions that Drive Results (significant sensitivities)

Many assumptions drive an integration analyses. We briefly discuss these with the intent to illustrate that, even if one accepts that integration costs can be calculated accurately in the first place, that comparing them is fraught with difficulties because of the widely varying assumptions that can significantly contribute to the results.

1. **Mix of generation.** The so-called “minimum generation problem” occurs when there is high wind output during a time of low load, and the remaining generation fleet cannot back down far enough. This results in low or negative prices and an overgeneration condition that can be solved by curtailing wind, increasing exports, or deploying responsive loads that can absorb the surplus generation. However, a different generation mix that generally comprises less baseload, or baseload units with lower minimums, may

alleviate the problem. As discussed earlier in this report, integration studies that look out over decades and hypothesize high VG penetrations, yet keep the same generation mix may find problems that can be alleviated with an alternative generation mix. Over longer periods of time, this is a plausible, if not likely, evolution of the generation mix.

2. **Institutional constraints and operating practice does not change even with high penetrations of VG.** Changes in operational practice or institutional constraints are difficult to forecast. However, holding on to ineffective and uneconomic practices one or more decades in the future with high penetrations of VG is not likely nor plausible. The structure of markets and contracts will likely change if there is significant contribution from VG to the overall generation mix. Flexibility will be valued more highly than it is today; because there are no markets in the United States that value flexibility, it is likely that new market products will be available to induce services that have high value but are currently not paid for. At the time of this writing, the Western Interconnection is undertaking analyses of an energy imbalance market (EIM). Although the EIM would not include coordinated unit commitment, the implementation of this market, if undertaken, would have a potentially profound impact on the ability of the interconnection to efficiently integrate large quantities of VG. In addition, an imbalance market without coordinated unit commitment will likely result in systematic overcommitment, and thus incur unnecessarily high operational costs that may induce informal coordinated unit commitment among neighboring BAs, or may even influence the development of a more formal unit-commitment coordination or market to reduce excess costs. Integration study results are clearly sensitive to such developments and the associated assumptions.

3. **The structure of contracts may change in the future.** Bi-lateral contracts with fixed hourly energy delivery may inhibit the ability of the system operator to manage the increased variability and uncertainty of high levels of wind and solar. In the relatively short term, such contracts may not change, but over longer terms, it is possible that they will.

4. **Scheduling intervals may change.** Because wind and solar forecasts become less accurate for longer time frames, late gate-closing allows the system operator to take account of the latest information and more accurate forecasts. At the time of this writing, there is significant effort going into the development of better underlying weather forecast models, such as the High Resolution Rapid Refresh (HRRR) model that can run hourly instead of every six hours, taking advantage of recent weather data to provide more accurate inputs to wind and solar forecast models. At the same time, increasing adoption of computer and communication technology will make it possible to incorporate these forecasts closer to real-time. Thus the "lock-down" period for generator notification and movement to subsequent dispatch levels could be shortened, resulting in more

accurate positioning of the hydro and thermal generation fleet and a corresponding reduction in expensive regulation, which must pick up the dispatch errors.

5. **More frequent dispatch in the West.** Until and unless the EIM (or other similar process) is implemented in the West, schedules and much of the economic dispatch is hourly. An hourly scheduling process strands the physically-available flexibility that is inherent in the generation fleet, calling on the economic dispatch stack once an hour to move to a new operating point. All changes within the hour must then be managed by regulating units, which comprise a relatively small proportion of the online generation. Moving to sub-hourly markets running 5-minute dispatch and schedule changes for imports and exports will substantially increase the ability of the system to manage higher penetrations of VG.

6. **Operating footprint.** Larger operating footprints (balancing areas) can do more with less. The ramping capability of the generation fleet adds linearly with expansion; the ramping needs of the power system add less than linearly. Compliance costs and operational inefficiencies may drive small BAs to coordinate or merge with neighboring systems. EWITS showed that adding a transmission overlay decreases loss of load probability (LOLP), thus decreasing or delaying the need for new generation. Milligan, Kirby, and Beuning[19] have shown that there are considerable efficiencies that accrue in both operating and planning (long-term costs) for coordinated or combined planning and operations.

7. **Methods to develop simulated wind/solar forecasts for integration studies are relatively primitive, and yet have a potentially large impact on integration efficiency.** Representing a future build-out of large-scale wind/solar requires time-synchronized wind/solar data at short time steps of no longer than 10 minutes for at least 3 years. These datasets are derived from numerical weather prediction models that are run for the desired time period, and wind/solar data are extracted at high geographic and temporal resolution. To adequately simulate power system operations, it is necessary to also have forecasts over several time horizons that can inform the unit commitment and dispatch processes; however, there is substantial disagreement within the wind forecasting community (solar forecasting is not yet well-developed) regarding methods to do this. Yet wind forecast distributions and timing can have a significant impact on integration.

In spite of the considerable progress that has been made in integration modeling and analysis, the discipline is still maturing and there are significant questions still remaining. As we have shown, integration analyses are sometimes still subject to error, and even if performed correctly, there

---

[19] "Combining Balancing Areas' Variability: Impacts on Wind Integration in the Western Interconnection", Milligan, M.; Kirby, B.; Beuning, S. Available at http://www.nrel.gov/docs/fy10osti/48249.pdf.

are variations in methods that make comparisons difficult or impossible. Integration studies have grown in scope, complexity, and sophistication. Although we are aware of attempts to develop simplified integration tools, we do not believe that this field of study has achieved sufficient maturity to allow the simplifications and generalizations that would be necessary in the development of such a tool.

Conclusions

While VG integration studies have progressed significantly in the past few years, there is still no universally agreed upon method to calculate the integration costs associated with variability and uncertainty of these resources. Progress in wind and solar integration analysis has been spurred on by the increasing amounts of wind and solar power being deployed in systems around the world. State-of-the-art wind and solar integration analysis now uses the same security-constrained unit commitment and economic dispatch software that is used to operate the power system (at least in regions where the power system is operated efficiently). Mesoscale modeling is used to generate wind and solar time-series data that is time- synchronized with actual load data. Modeling is done for multiple years with ten-minute or faster resolution. Wind and solar forecasts are included for unit commitment. A base case without wind and solar is compared with one or more high penetration wind and solar cases to determine the impact of wind and solar on fuel and operating costs, reserve requirements, and the operations of the conventional generators. Total system costs with and without VG can be calculated with reasonably high confidence.

There is less ability to explicitly calculate *integration costs* for wind and solar. Fuel savings naturally dominate any comparison of wind and solar with conventional generation. Finding an appropriate zero-fuel-cost proxy resource for the base case has proven to be more difficult than expected. Flat blocks have higher or lower on-peak energy content, re-introducing an energy value component that degrades the results. Multiple blocks can try to better match the energy value, but they introduce ramping burdens between blocks. Calculated “integration costs” can be as much an assessment of the characteristics of the proxy resource as they are of the variable renewable. To date no really satisfactory proxy resource has been found. The current status of VG integration modeling is:

- There is no universal agreement on methods for calculating renewables’ integration costs, and even when there is agreement on methods, they are not consistently applied or are applied with errors;
- There are many potential base-cases (no-wind-or-solar) that may be relevant for comparison;
- High penetrations of wind and solar impact the optimal mix of conventional generation, further complicating the base-case selection;
- There is general agreement that wind has an impact on operations, but there is substantial disagreement about whether/how integration costs can be measured.

While there are technical difficulties with calculating wind and solar integration costs, there are also public policy and regulatory questions concerning what to do with renewables integration costs even if they can be accurately calculated. Other generation technologies impose integration costs which are not allocated to those technologies. Large generators impose contingency reserve requirements, block schedules increase regulation requirements, gas scheduling restrictions impose costs on other generators, nuclear plants increase cycling of other baseload generation, and hydro generators with dissolved gas limitations create minimum load reliability problems and increased costs for other generators. None of these costs are allocated to the generators that impose them on the power system. Any policy that assigns integration costs to wind and solar needs to be thought through very carefully to assure that it is not discriminatory.

Generation integration costs are typically broadly shared because the benefits are also broadly shared. Contingency reserves are shared within a large reserve sharing pool because physical aggregation genuinely reduces the physical reserve requirement and therefore reduces everyone's costs. Variable renewables bring fuel diversity, price stability, energy security, and environmental benefits that accrue widely to all users of the power system so it is reasonable that integration costs should likewise be broadly shared.

NREL's WWSIS shows that the current practice of hourly scheduling in most of the West has a larger impact on the cost of system regulation than does operating the system with 35% wind and solar energy installed.[20] The benefits of intra-hour scheduling, however, extend far beyond integration of renewables. Intra-hour scheduling will enable greater utilization of existing transmission facilities, greater flexibility in the provision of ancillary services, sharing of operating reserves, and a reduction in the amount of reserves required to be carried on individual systems. Shorter schedules will unlock access to the flexibility inherent in the existing generation fleets installed in different BAs to provide system balancing. This will reduce costs for all electric service customers on the systems implementing intra-hour scheduling. Although such scheduling changes may require an upfront investment in operations software and hardware, the changes will provide substantial net cost savings over time. With such broad and intertwined benefits, the focus should be on capturing those benefits, which can be accurately quantified, rather than on allocating costs to individuals, which can not be accurately quantified.

Acknowledgments

We thank Paul Denholm and Charlie Smith for their helpful discussion and comments.

[20] Comments of the National Renewable Energy Laboratory's Transmission and Grid Integration Group in Response to FERC's January 21, 2010, Notice of Inquiry on the Integration of Variable Energy Resources, Scheduling Flexibility and Scheduling Incentives, Question 1, p. 5; April 12, 2010.