ELECTRIC POWER RESEARCH INSTITUTE

# Simulating Annual Variation in Load, Wind, and Solar by Representative Hour Selection

**3002008653**

# Simulating Annual Variation in Load, Wind, and Solar by Representative Hour Selection

3002008653

Technical Update, June 2016

EPRI Project Manager

J. Bistline

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

THE FOLLOWING ORGANIZATIONS PREPARED THIS REPORT:


**Electric Power Research Institute**

**Department of Management Science and Engineering**


**This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.**

## NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

# ACKNOWLEDGMENTS

The following organizations prepared this report:

Electric Power Research Institute (EPRI)
3420 Hillview Avenue
Palo Alto, CA 94304

Principal Investigators
G. Blanford
J. Bistline
D. Young

Department of Management Science and Engineering
Stanford University
475 Via Ortega
Stanford, CA 94305

Principal Investigator
J. Merrick

This report describes research sponsored by EPRI.

The authors would like to thank Steve Wan for his research assistance. The views expressed in this paper are those of the individual authors and do not necessarily reflect those of EPRI or its members.

This publication is a corporate document that should be cited in the literature in the following manner:

*Simulating Annual Variation in Load, Wind, and Solar by Representative Hour Selection.* EPRI, Palo Alto, CA: 2016. 3002008653.

# ABSTRACT

The spatial and temporal variability of renewables have important economic implications for investments and system operations. This study describes a method for selecting representative hours to preserve key distributional requirements for regional load, wind, and solar time series with a two-orders-of-magnitude reduction in dimensionality. We describe the implementation of this procedure in the US-REGEN model and compare impacts on energy system decisions with more common approaches. The results demonstrate how power sector modeling and capacity planning decisions are sensitive to representation of intra-annual variation and how our proposed approach significantly outperforms simple heuristic selection procedures with lower resolution. The representative hour approach preserves key properties of the joint underlying hourly distributions, whereas seasonal average approaches over-value wind and solar at higher penetration levels and under-value investment in firm capacity by inaccurately capturing the corresponding residual load duration curves.

**Keywords**

Capacity planning
Intermittent renewables
Market integration
Power system modeling
Representative hours
Variability

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1
# INTRODUCTION

A key research question for energy system modeling is what role wind and solar could or should play in the transition to a low-carbon energy system. To explore this question effectively, a model must capture the strong effect of hourly and spatial variability on the fundamental economics of intermittent renewable energy. While maintaining full hourly resolution in a model is possible in applications with limited scope, more compact alternatives are required for use in energy system or integrated assessment models with national or global coverage and long timeframes. In this paper we propose the "representative hours" method for capturing the essential economic implications of intra-annual variability in a computationally efficient manner, namely with resolution two orders of magnitude lower than hourly.

Prior to the recent emergence of wind and solar power as major potential sources of electricity supply, the capturing of intra-annual variability in energy system models required the capturing of the variability of electricity demand alone. Many existing studies show the relevance of increased temporal resolution to model outputs when wind and solar power are options (Ludig et al., 2011; Nicolosi et al., 2011; Pina et al., 2011) and provide overviews of approaches to incorporate this increased source of variability in models (Merrick, 2016; Nahmmacher et al., 2014). Examples of such approaches are varied (Swider and Weber, 2007; Ueckerdt et al., 2015; Van der Weijde and Hobbs, 2012). Note that in a static setting where wind and solar power capacity is fixed, load can be translated into residual load (demand net of wind and solar production), and the relevant variability can again be reduced to a single dimension (De Sisternes et al., 2015). The modeling challenge arises in a dynamic setting in which investment in renewable energy is a decision variable, so that the approach must be robust to any level of deployment of wind and solar power.

The generalized objective of any such approach is to find the unique hours of the year, not only in terms of electricity demand, but in terms of joint demand and wind and solar availability (Merrick, 2016). This framing suggests the use of clustering methods, which are employed elsewhere to find the number of unique hours (Merrick, 2016; Nahmmacher et al., 2014). The operations research literature on aggregation of linear programs also points to clustering methods and associated guarantees on model accuracy (Rogers et al., 1991). For a sample dataset (Merrick, 2016), a clustering method reduced the number of hours from 8760 to the order of 1000 while maintaining the characteristics of the original problem.

However, this order of magnitude is still intractably high for a detailed electricity model. Currently, most numerical applications in a dynamic investment (i.e., capacity planning) context are only able to reflect intra-annual variability with resolution on the order of 100 hours due to computational limitations. Our novel representative hours method, used by the US-REGEN model (EPRI, 2014; Blanford et al., 2013), consists of strategic selection of particular hours during a calibration year that satisfy simultaneously key distributional requirements for load, wind, and solar series across multiple inter-connected model regions. In particular, to reduce the resolution from thousands to hundreds of hours, we use a priori information about the relevance of different hours to the model solution. A key principle of the representative hours method is that the extreme points of the annual load, wind, and solar distributions must be captured. These hours are important for representing potential capacity shortfalls as well as the potential extent of surplus renewable energy production. The relevant extremes include not only the peaks and minimums of the individual series, but crucially also the *joint extremes*, for example the moments when load is high and both wind and solar are low, represented for each region. However, while extreme hours are essential, by themselves they do not suitably reflect the distribution across the entire year. To supplement the selected extreme hours, we also include several hours identified by a standard clustering algorithm to represent the interior of the distribution.

The selected hours are then weighted to sum to a full year while minimizing the sum of errors between the approximated and hourly duration curves for each regional series. See the Methods section for further details on implementation. The resulting set of "representative hours" is used as the domain for dispatch of electric generation and transmission capacity in the model, where load and wind and solar availability factors are equal to their levels in the actual underlying hour and the duration is equal to the weight. We provide a range of diagnostic tests demonstrating the method's performance, as well as a comparison to a more typical approach of choosing a small number of points calculated as seasonal averages (see Section 2). This "seasonal average" approach uses a limited number of segments traditionally defined to capture the load curve (peak, shoulder, and base in summer, winter, and fall/spring) and assigns wind and solar coefficients to each segment based on average resource availability during the corresponding load period. While such simpler approaches can be effective at reproducing a load duration curve and average wind and solar capacity factors, they poorly represent the distribution and co-variation with load of renewable resources, as well as the co-variation among regions needed to effectively model power transmission.

The most meaningful metric presented here is an ex-post calculation of marginal value curves for wind and solar using a static model with full hourly resolution as well as approximations based on our method and the more typical approach. We show that the "representative hours" method reproduces far more closely the value of wind and solar as measured in the hourly model. A presentation and discussion of our results follow. In the methods section at the end of the paper, we provide a brief literature review of modeling approaches to aggregation of intra-annual hours, a description of the underlying data and US-REGEN modeling context, and an exposition of the three components of the "representative hours" method: (a) selection of "extreme" hours; (b) selection of "cluster" hours; and (c) weighting of hours. We also briefly describe our implementation of the more typical approach based on seasonal averages. Further methodological details are provided in Appendix A.

# 2
# METHODS

The representative hours method has been implemented in the US-REGEN model, a detailed equilibrium model of electricity investment and dispatch with 15 distinct sub-regions of the continental US (see Figure A-1 in Appendix A for regional definitions). Parameterization is based on synchronized hourly data series for each region for load, as provided by FERC at the service territory level (FERC, 2010), and for renewable resource availability. Profiles for potential wind and solar output were developed for the model by AWS Truepower based on detailed simulations at potential sites in each region, as described in (EPRI, 2014). Although the simulations produced a range of profiles in each region corresponding to differentiated quality classes, which are represented in the model, for the purposes of representative hour selection we consider only a single profile for each region averaged over all classes. It is also possible to include multiple classes in certain regions, as discussed in the Appendix A. In this analysis, all profiles are based on data from 2010. The model can be solved as a dynamic optimization through 2050 with five-year time steps, or alternatively as a static equilibrium with dispatch and capacity rental for a single year using full hourly resolution as well as more aggregate configurations.

## 2.1 Representative Hours Method: "Extreme" Hour Selection

The first phase of the hour selection algorithm is to identify a minimal set of hours that adequately covers each relevant extreme in each region. To begin, the algorithm identifies the hours with minimum and maximum values of load, wind, and solar individually, that is, six hours per region. Next, the algorithm identifies the hours at each vertex of the three two-dimensional planes (four each), and of the three-dimensional load-wind-solar space, that is, eight hours corresponding to each possible combination of maximum and minimum in each dimension, again in each region. If there were no overlap among these selected extremes, this would result in 26 points in 15 regions, or 390 candidate hours. However, the two- and three-dimensional extremes often coincide, and one extreme is sometimes represented by the same hour in multiple regions. For our 2010 profiles, overlap between dimensions and regions reduces this total to 223 unique extreme hours.

The simplest approach might be simply to proceed using all 223 extreme hours. However, though the extremes themselves are unique, other hours may be quite similar in terms of their joint values in the three dimensions of load, wind, and solar, and moreover may be near extremes in multiple (usually neighboring) regions. Thus it is possible to reduce substantially the number of hours needed to adequately represent the extremes by allowing other "qualifying" hours within a certain radius (based on a Euclidean norm) from the true extreme. Geometrically, one may imagine a "bubble" around each extreme point in each region (in one, two, or three dimensions respectively), as shown in Figure 2-1 for the three-dimensional space in Texas. The selection algorithm is designed to find the minimum number of hours such that all 390 bubbles are populated with at least one hour. The algorithm is implemented as a straightforward integer programming problem in GAMS/CPLEX with a binary decision variable for each hour corresponding to whether it is selected, a constraint that the sum of selected qualifying hours for each extreme is greater than or equal to one, and a minimand equal to the sum of selected hours. If the radius for each bubble is set to zero, the algorithm must choose the extremes themselves, that is, the 223 unique extreme hours. With a non-zero bubble radius, the algorithm can take advantage of hours that are near-extreme in multiple regions or dimensions and choose fewer total hours. Some bubbles may be more important than others and thus may be assigned a smaller radius or tolerance. There is naturally a trade-off here between accuracy and computational tractability of the model, whose solution time is convex in the number of hours. Our goal is to arrive at a configuration on the order of 100 hours, including the cluster hours described in the next section. Table 2-1 summarizes the relationship between radii of the various bubbles and the minimum number of spanning hours. The current analysis uses the configuration shown in bold with 76 extreme hours.

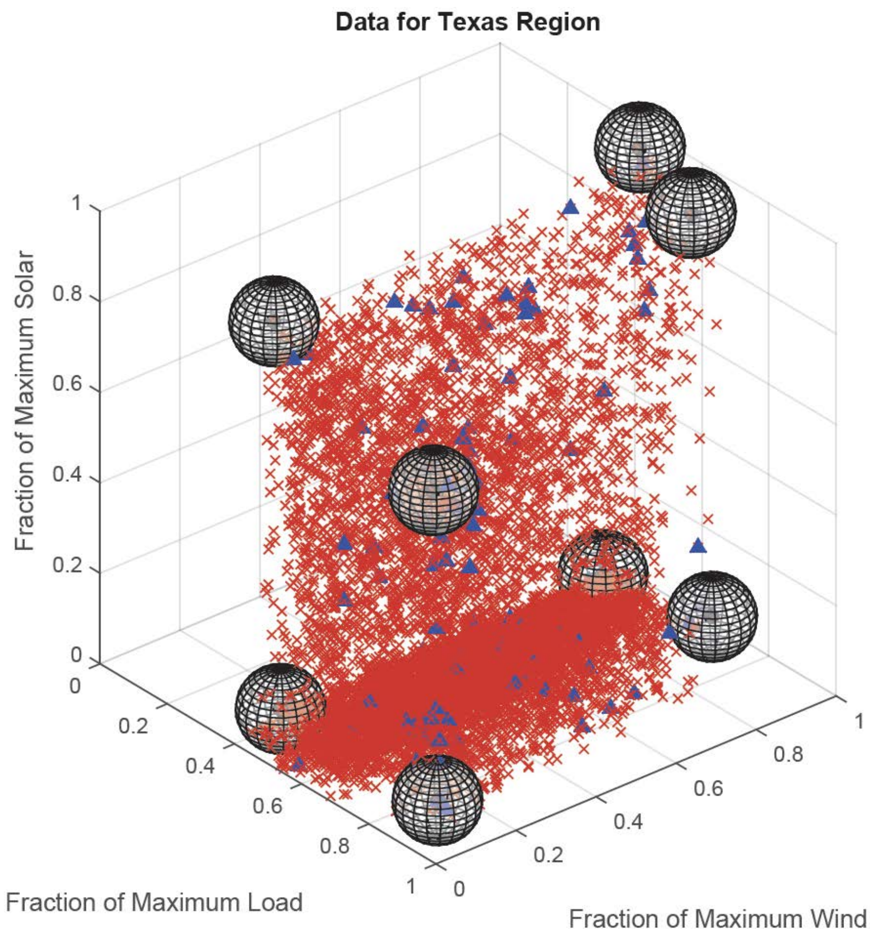



**Figure 2-1**
**Normalized hourly load, wind, and solar data for the Texas region (red), bubbles around corner points (black), and chosen segments from the hour selection procedure (blue).**

**Table 2-1**
**Relationship between bubble tolerance are number of selected extreme hours.**

| Extreme Description [number] | Bubble radius (% difference from true extreme allowed in each dimension) | | | | |
|---|---|---|---|---|---|
| 1D load max/min [2] | 0 | 1 | 1 | **1** | 5 |
| 1D wind/solar max/min [4] | 0 | 1 | 1 | **5** | 5 |
| 2D load max vs. wind/solar min [2] | 0 | 1 | 1 | **1** | 5 |
| 2D load min vs. wind/solar max [2] | 0 | 1 | 1 | **5** | 5 |
| 2D load max vs. wind/solar max [2] | 0 | 1 | 5 | **10** | 10 |
| 2D load min vs. wind/solar min [2] | 0 | 1 | 5 | **10** | 10 |
| 2D wind vs. solar max/min [4] | 0 | 1 | 5 | **10** | 10 |
| 3D load max and wind or solar min [3] | 0 | 1 | 5 | **5** | 10 |
| 3D load min or wind and solar max [5] | 0 | 1 | 5 | **10** | 10 |
| Number of Selected Hours | 223 | 177 | 123 | **83** | 56 |

## 2.2 Representative Hours Method: "Cluster" Hour Selection

By design, the first phase has focused only on characterizing the convex hull of the three-dimensional space in each region. While extreme hours in one region may in fact be interior points that are more centrally located in another region (as illustrated in Figure 2-1), it is clear that the interior of the load-wind-solar space remains under-sampled by the algorithm described above. Thus we add a second phase in which a second set of hours is selected based on a standard clustering algorithm. In contrast to the first phase, such an algorithm will by design identify hours near the center of the joint distribution. Whereas in the first phase, the number of selected hours was an outcome of the algorithm (and the choice of qualifying radii), in a clustering algorithm the number of clusters is an input. With one cluster, the algorithm will attempt to find the centroid of the entire distribution. With multiple clusters, the algorithm will first partition the space into clusters and then identify the centroid of each cluster, with the objective of minimizing the sum of distances from each point to its corresponding centroid. By definition, the hours nearest the selected cluster centroids will have virtually no overlap with the hours identified as extremes, thus providing a complementary subset. Again, there is a trade-off between sampling accuracy and computational tractability. In this analysis we applied the cluster algorithm phase with a target of 20 hours, for a total of 103 representative hours.

## 2.3 Representative Hours Method: Hour Weighting

Finally, the representative hour method must calculate weights for the hours selected by the first two phases such that the sum equals 8760, or a full year. The weights are chosen to minimize the error not only with respect the annual average, but also with respect to the shape of the cumulative distribution function (i.e., sorted duration curve) of each series in each region. The objective of the error minimization is the sum of squared differences between each representative hour's sort position in the full hourly curve and its sort position in the representative weighted curve. An additional subtlety is that errors are more heavily weighted at points where the sorted duration curve is steeper, that is, where errors in the sort position will lead to a greater distortion of the shape of the curve. We note that the residual error after optimal weights are chosen is an indicator of whether the number of selected hours by the first two phases has been sufficient.

## 2.4 Alternative Method ("Seasonal Average" Approach)

Traditionally, modeling the electric sector in a reduced-form context required only a relatively simple representation of the load duration curve, with a small number of segments, typically less than ten or even five, capturing the peak, shoulder, and base load periods. Many models were developed with this structure, which allows a reasonably effective approximation of trade-offs between high fixed-cost / low variable-cost base load generators and low fixed-cost / high variable-cost peaking generators in a conventional power system. However, with the growing importance of intermittent renewable energy, models with this type of structure face challenges in adding variation in wind and solar to the framework. A typical response to these challenges is illustrated by the National Energy Modeling System (NEMS) used and made publicly available by the U.S. Energy Information Administration (EIA). The NEMS model, as described recently in EIA (2014), uses nine segments traditionally defined to capture the load curve (peak, shoulder, and base in summer, winter, and fall/spring), and assigns wind and solar coefficients to each segment based on average resource availability during the corresponding load period. For this analysis, we have reproduced a similar set of coefficients based on the same 2010 hourly data used in our representative hours method.

The difficulty with this type of approach is that it insufficiently describes both the individual distributions of wind and solar resource availability and the joint distribution of wind and solar with load. It also makes no attempt to capture regional correlation, meaning that transmission between model regions during a given segment occurs with non-simultaneous conditions on the two ends of the transaction. Although the NEMS model also employs other ad hoc constraints to account for potential moments of both low and high renewable output, its underlying approach of using a small number of load-based segments with averaged wind and solar coefficients is quite common and can lead to a substantial misrepresentation of the value of renewable technologies.

# 3

# RESULTS

## 3.1 Diagnostic Results

We begin with an assessment of the representative hour (103 segments) and the seasonal average (9 segments) methods with respect to preserving distributional characteristics of the underlying hourly time series for load, wind, and solar. While the peak, minimum, and average of each series are preserved to a high degree of accuracy by construction in the representative hour method, the seasonal average method only ensures accurate capture of the averages and the peak load (see Table A-1 in Appendix A). Figure 3-1 shows duration curves (values sorted in descending order and weighted by segment length) for both the hourly and the approximated series for Texas. Results are qualitatively similar across other model regions, as shown in the Appendix A. The representative hour method successfully captures the shape of the annual distribution for all three individually, but the seasonal average method underestimates variation in wind and solar. Most importantly, we assess the performance of each method with respect to correlation among the series by examining residual load duration curves. Residual load is calculated as demand less the available generation from intermittent renewable resources, which represents load that must be met through dispatchable assets. We illustrate residual load in Texas with a hypothetical introduction of 80 GW of wind and solar respectively (roughly equivalent to peak load in 2010), re-sorted to form a duration curve.

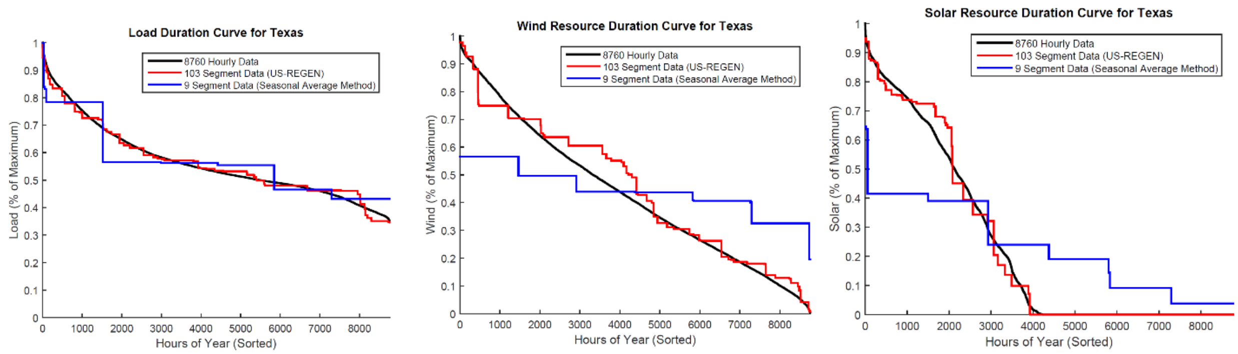



**Figure 3-1**
**Duration curves for load (left panel), wind (middle panel), and solar (right panel) in Texas. The hourly duration curve (black) is approximated with 103 segments (red) and 9 segments (blue).**

As shown in Figure 3-2, the hourly residual load duration curves indicates two important properties of wind and solar: (i) they alone contribute little to capacity needs, as peak residual load is unaffected by large renewable capacity additions; and (ii) they provide energy disproportionately at hours with low residual load. These two properties drive decreasing returns to renewable energy over the long term, as discussed below and elsewhere in the literature (Grubb, 1991; Fripp and Wiser, 2008; Edenhofer et al., 2013). The representative hour method preserves both properties with a limited downward shift in residual on the left side of the curve and a much larger downward shift on the right side. By contrast, the seasonal average method results in a larger contribution to the residual peak and a more limited contribution at the low end of the residual load curve. These results indicate that the representative hour method is much more likely to preserve key economic properties of intermittent renewable technologies in a reduced-form model than the more typical seasonal average approach.

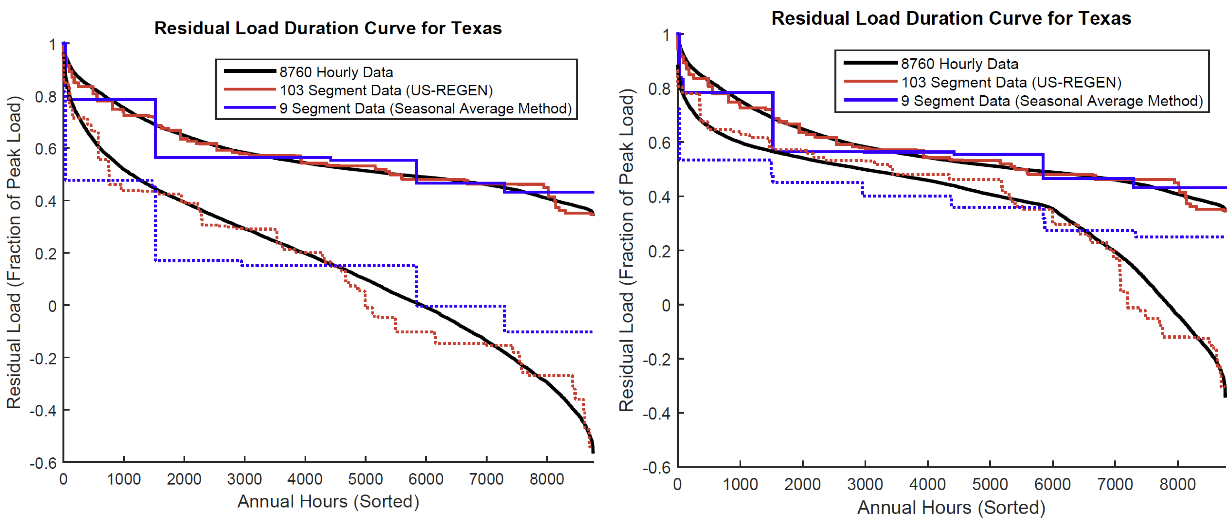



**Figure 3-2**
**Residual Load Duration curves with 80 GW wind (left panel) and 80 GW solar (right panel). The hourly duration curve (black) is approximated with 103 segments (red) and 9 segments (blue). Solid lines at the top of each panel show curves without renewable deployment, and dotted lines show wind and solar penetration scenarios.**

We have also examined correlations between multiple time series (e.g., between load and resources, between different resource types, and across regions). In particular, models should reflect the joint distribution of load, wind, and solar characteristics to reflect the economics of intermittent renewable technologies. As shown in Figure A-2, the representative hour method captures correlation coefficients between load and renewable output well, but the simplified seasonal average approach does not sufficiently represent these characteristics. In addition to the interdependence of load and renewables, correlations between different regions can be an important dynamic in trade outcomes and large-scale system balancing. Figure A-3 shows how the representative hours capture the cross-correlations in the underlying data better than the heuristic approach, which understates the heterogeneity across regions. Additional correlation statistics are examined in the Appendix A.

We now turn to a comparison of model outputs. We begin with an illustrative analysis using the static version of the model, described above, in which the full hourly resolution can be used (as well as the two approximation methods). The experiment, based on EPRI (2015), consists of a series of static model simulations where the cost of wind (resp. solar) is systematically varied from current levels down to zero. As the cost decreases (and all other system parameters remain unchanged), total wind (resp. solar) deployed in the U.S. increases, thus revealing a marginal value curve for each resource. This measurement of value is a strong indicator for the role of renewable technologies in a dynamic simulation where costs and other components of the system evolve over time. Figure 3-3 shows the results of this analysis using hourly resolution as well as the reduced-form methods. Again, the representative hour method closely matches the results from the hourly simulation, while the seasonal average method, which fails to account for key distributional impacts on the value of renewable energy profiles, significantly overestimates the marginal value curve for this dataset.

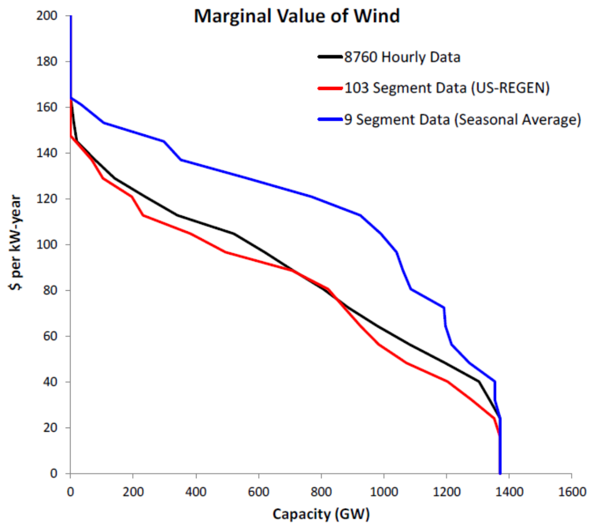



**Figure 3-3**
**Marginal value curves for wind (left panel) and solar (right panel) using the full hourly data (black), representative hour approach (red), and seasonal average approach (blue).**

## 3.2 Dynamic Model Results

To characterize the electric sector transition to a long-run equilibrium and to evaluate the impacts of model representations of annual variation on these outcomes, we conduct a dynamic analysis in US-REGEN through 2050 using the representative hour and season average approaches described above. These two variability specifications are run under three policy scenarios:
1. Reference (no additional climate policies); 2. Carbon tax of \$25/t-$CO_2$ beginning in 2025;
3. Carbon tax of \$50/t-$CO_2$. Scenario assumptions are detailed in the Appendix A.

These experiments demonstrate how model results are sensitive to the segment selection procedure and its accuracy in approximating temporal and spatial distributions of load, wind, and solar. Model recommendations for capacity investments (Figure 3-4) as well as other dispatch, emissions, and cost metrics (Appendix A) are sensitive to these specifications across a range of market settings but are most responsive under scenarios that incentivize low-carbon technologies. The seasonal average approach described above does not account for the correlations between intermittent resources and load, which *ceteris paribus* incorrectly values these resources and operational flexibility of other assets. Cumulative investments through 2050 in solar (wind) are 113 GW (35 GW) larger with the seasonal average approach under baseline policy conditions and 217 GW (156 GW) larger under a \$50/t-$CO_2$ tax compared with the representative hour approach. Moreover, investments in conventional capacity, in particular gas combined cycle and gas turbines, are considerably lower in the seasonal average approach. An assessment of capacity adequacy of the dynamic model solution against the true underlying hourly distribution shows that the seasonal average approach leads to a capacity shortfall of as much as 200 GW nationally, while the representative hour approach, which explicitly accounts for extreme hours driving capacity needs, falls short by a maximum of only 34 GW. Although the specifications above lead the seasonal average approach to overestimate the value of renewables, the direction of the bias depends strongly on the implementation details and data of different approaches (Merrick, 2016).

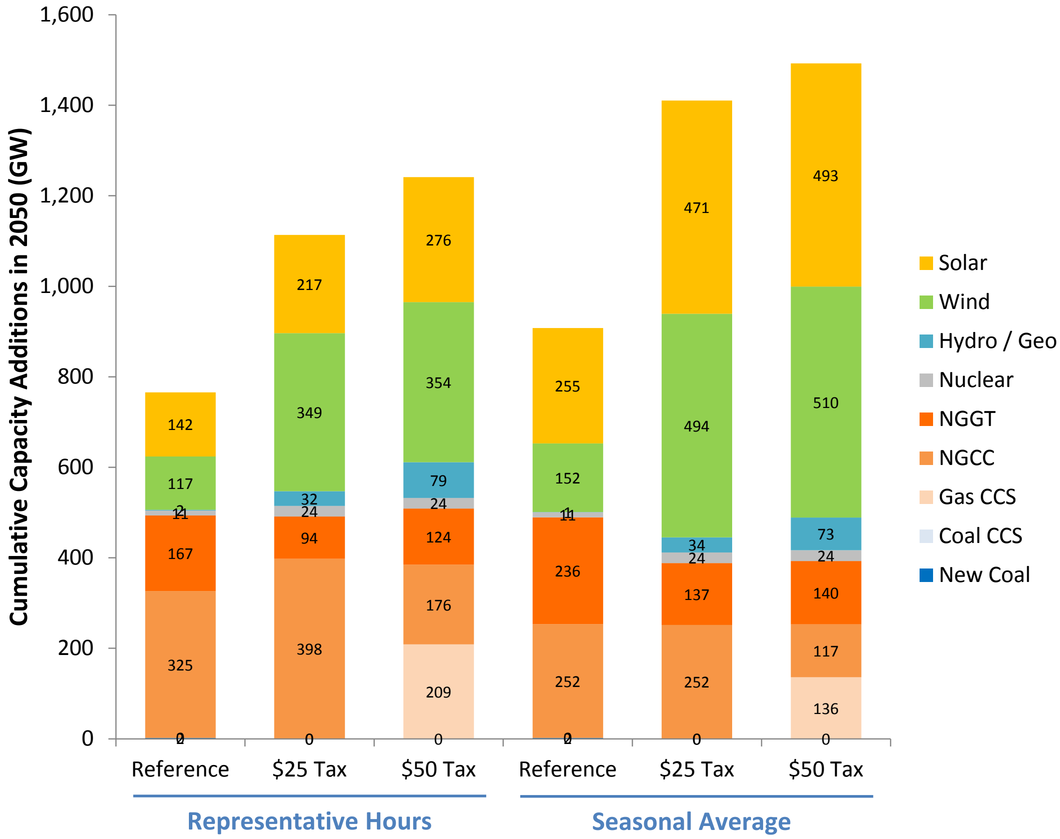



**Figure 3-4**
**Cumulative electric sector capacity investments through 2050 (GW) for three carbon policy scenarios under the representative hour approach (left) and seasonal average approach (right).**

# 4

# DISCUSSION AND CONCLUSION

The results described here demonstrate how power sector modeling and capacity planning decisions are sensitive to the representation of intra-annual variation and how our proposed approach significantly outperforms simple heuristic selection procedures while maintaining computational tractability. In particular, the value of intermittent renewable energy from wind and solar is inextricably linked with the timing of their production relative to load. Clustering methods have been shown to guarantee that a model with aggregated temporal resolution will reproduce the outputs of a model with full hourly resolution (Merrick, 2016). Since such a resolution is still too great for a dynamic, national-level model, we have reduced the resolution further by drawing upon knowledge of the influence of extreme hours on electric sector investments. The goal of our representative hour approach is to model dynamic investment decisions that reflect as accurately as possible (within computational feasibility) the true economic implications of intermittency.

We first demonstrate that key properties of the joint underlying distributions are well-preserved by the representative hour approach for our comprehensive U.S. dataset. The changing shape of the residual load duration curve with increasing penetration of wind and solar capacity is an essential attribute of renewable value, and the approximated curve with weighted representative hours closely matches that based on hourly data. By contrast, the seasonal average approach yields a much poorer fit for residual load. We note that Ueckerdt et al. (2015) has focused on parameterizing directly the relationship between wind and solar capacity and the shape of the residual load duration curve, which is a compact and useful approach for modeling a single region. However, because our method is based on a sampling of simultaneous hourly data across multiple interconnected regions, we are able to reflect spatial variation of resources, including spatial variation in temporal correlations, as well as transmission capacity requirements. In this regard, the representative hour approach is a significantly more powerful (though computationally intensive) aggregation method, and one more suitable for detailed regional electricity dispatch and investment models, than the approach of Ueckerdt et al. (2015). Moreover, in the conventional seasonal average approach, each region’s averages are derived separately, so that only seasonal correlation between regions is retained, making it very difficult to relate transmission flows in a given model segment to actual capacity requirements.

We next demonstrate that model results using the representative hour approach for aggregation align closely with results from an otherwise identical hourly model. The marginal value curve, which describes how the value of wind and solar additions at the margin falls with cumulative national capacity, can be calculated in a rental context with a single year at hourly resolution, making it an ideal point of comparison for alternative aggregation approaches. Again, the seasonal average approach misses the mark, over-valuing wind and solar at higher levels of penetration as a result of not accurately capturing the corresponding residual load duration curves. Finally, we show results for a dynamic simulation with alternative carbon policies. Consistent with the marginal value curve results, the seasonal average approach leads to a greater deployment of wind and solar and lower investment in firm capacity than the representative hour approach, particularly at higher carbon prices. Although the hourly benchmark is not available in this setting, the capacity results from the aggregate model can be compared to the hourly data, which reveals much larger shortfalls in the seasonal average approach due to its lack of attention to extreme moments. Overall, these experiments illustrate how a representative hour approach can provide more reliable energy modeling insights and accurate asset valuation relative to simplified approaches that appear in many similarly oriented models.

One important limitation of the representative hour approach, and indeed of the simpler seasonal average and any similar approach, is that the chronology of hours is not preserved. Thus, it is not possible to explicitly model electricity storage (unless one makes an unrealistic assumption of unlimited reservoir size) or operational constraints on ramp rates and start-up/shut-down cycles. While it is possible to study these issues with full hourly resolution, and in the case of operational constraints with a unit commitment formulation, it is not currently possible to conduct national-level dynamic investment analysis in such a context. One idea is to represent the year with a small number of full weeks, possibly at slightly less than hourly resolution, and examine the deployment of electricity storage, subject to reservoir constraints, within each of these "representative weeks." However, it is unclear whether such an approach sufficiently captures annual variation at the hourly level, nor whether a weekly horizon is sufficient to capture the potential value of storage. A more conservative approach is to complement dynamic investment modeling using representative hours with separate but harmonized supplementary analysis using hourly resolution for a single year. For example, EPRI (2015) shows that unless costs of electricity storage become far lower than they are today, bulk storage does not significantly change the optimal mix of generation technology in a static equilibrium for a range of policy scenarios. A key research question for future work is to what extent opportunities for electricity storage and operational constraints fundamentally change the value of generation investments and the trade-offs between intermittent renewable and dispatchable technologies during the transformation to a low-carbon energy system.

# 5
# REFERENCES


Blanford, G.J., Merrick, J.H., Young, D. (2013). A clean energy standard analysis with the US-REGEN model. *Energy Journal* 35:137–164.

De Sisternes, F.J., Webster, M.D., Perez-Arriaga, I.J. (2015). The impact of bidding rules on electricity markets with intermittent renewables. *IEEE Transactions on Power Systems* 30(3):1603–1613.

Edenhofer, O., et al. (2013). On the economics of renewable energy sources. *Energy Economics* 40:S12–S23.

Electric Power Research Institute (2015). *Decreasing Returns to Renewable Energy*, EPRI Technical Update #3002003946 (EPRI, Palo Alto, 2015).

Electric Power Research Institute (2014). *US-REGEN Model Documentation 2014*, EPRI Technical Update #3002004693 (EPRI, Palo Alto, CA).

Energy Information Administration (2014). *The Electricity Market Module of the National Energy Modeling System: Model Documentation 2014* (EIA, Washington, DC).

Federal Energy Regulatory Commission (2010). *FERC Form 714 Part III Schedule 2* (FERC, Washington, DC).

Fripp, M., Wiser, R.H. (2008). Effects of temporal wind patterns on the value of wind-generated electricity in California and the Northwest. *IEEE Transactions on Power Systems* 23:477–485.

Grubb, M.J. (1991). Value of variable sources on power systems. *Generation, Transmission and Distribution IEE Proceedings C* 138:149–165.

Ludig, S., Haller, M., Schmid, E., Bauer, N. (2011). Fluctuating renewables in a long-term climate change mitigation strategy. *Energy* 36(11):6674–6685.

Merrick, J. (2016). On representation of temporal variability in electricity capacity planning models. Forthcoming (http://stanford.edu/~jmerrick/temporal_resolution.html).

Nicolosi, M., Mills, A., Wiser, R. (2011). *The Importance of High Temporal Resolution in Modeling Renewable Energy Penetration Scenarios*, LBNL-4197E (LBNL, Berkeley, CA).

Pina, A., Silva, C., Ferrão, P. (2011). Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy* 39(9):4692–4702.

Nahmmacher, P., Schmid, E., Hirth, L., Knopf, B. (2014). *Carpe Diem: A Novel Approach to Select Representative Days for Long-Term Power System Models with High Shares of Renewable Energy Sources*, USAEE Working Paper Series No. 14-194.

Rogers, D.F., Plante, R.D., Wong, R.T., Evans, J.R. (1991). Aggregation and disaggregation techniques and methodology in optimization. *Operations Research* 39(4):553–582.

Swider, D.J., Weber, C. (2007). The costs of wind's intermittency in Germany: Application of a stochastic electricity market model. *European Transactions on Electrical Power* 17:151–172.

Ueckerdt, F., et al (2015). Representing power sector variability and the integration of variable renewables in long-term energy-economy models using residual load duration curves. *Energy* 90(2):1799–1814.

Van der Weijde, A.H., Hobbs, B.F. (2012). The economics of planning electricity transmission to accommodate renewables: Using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty. *Energy Economics* 34(6):2089–2101.

# A
# MODEL AND DIAGNOSTIC RESULTS

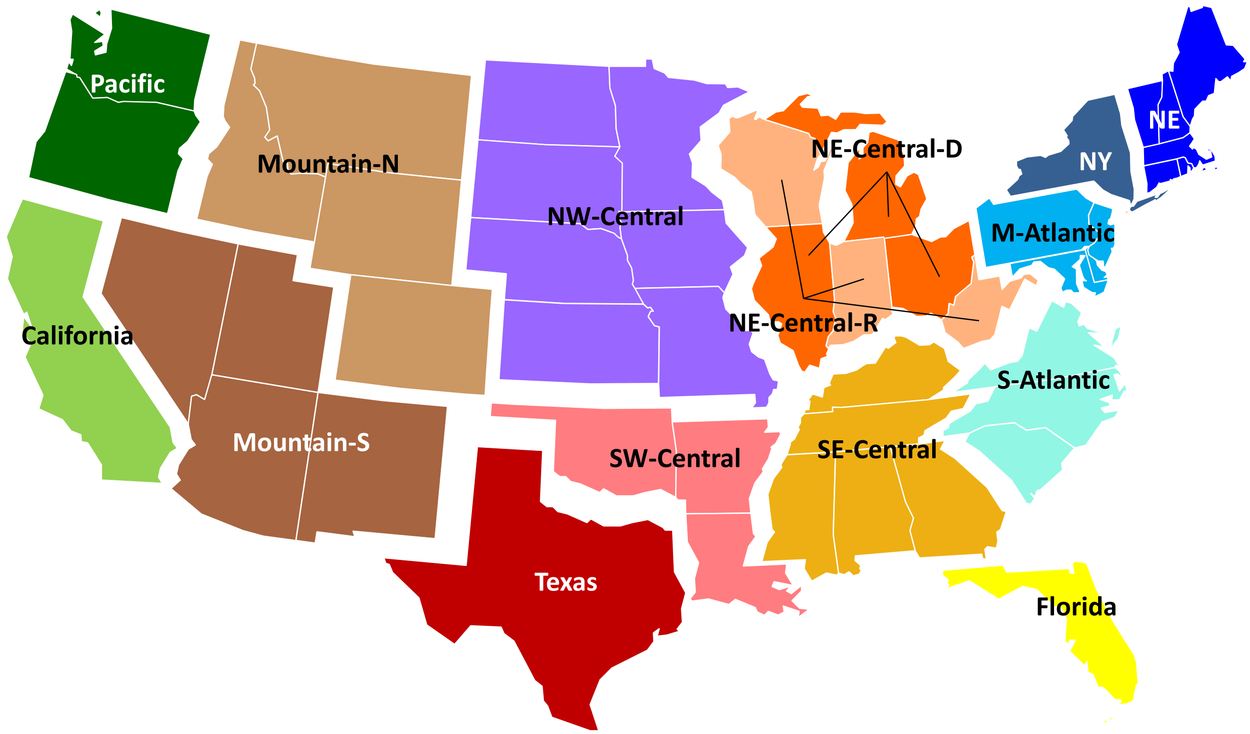



**Figure A-1**
**Regional structure of the US-REGEN model.**

The U.S. Regional Economy, Greenhouse Gas, and Energy (US-REGEN) model and analysis in this paper is organized into 15 state-based regions, as shown in Figure S1. For two regions, South-Atlantic (VA, NC, and SC) and Florida, a wind series was not simulated by AWS Truepower because of the low quality of wind resources in those regions. Otherwise, a single series for wind and for solar is assigned in each region based on a weighted average of identified potential sites across a range of quality classes. These classes are represented individually in the model, with the same hours and weights derived in the methodology described in this paper based on the weighted average regional series. There is typically a strong correlation within a region between different quality classes, and we have not observed large errors with this approach. However, it is relatively straightforward to explicitly include multiple quality classes within a region, for example a high-quality wind class in a region that is expected to be important in the model solution. Rather than add an additional wind dimension for this class, we have found it more practical to add a duplicate region with a new wind series and unchanged values for the load and solar series.

**Table A-1**
**Comparison of the average, maximum, and minimum hourly values for the annual time-series data for load, wind, and solar. Each section compares the underlying hourly data (8760), the representative hour approach (103), and the seasonal average method (9). Columns represent the 15 model regions.**

| | NE | NY | MA | SA | FL | NEC -D | NEC -R | SEC | NWC | SWC | TX | MTN-S | MTN-N | PAC | CA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Load-Avg | | | | | | | | | | | | | | | |
| 8760 | 0.552 | 0.558 | 0.563 | 0.616 | 0.531 | 0.640 | 0.595 | 0.622 | 0.638 | 0.592 | 0.564 | 0.688 | 0.570 | 0.616 | 0.541 |
| 103 | 0.552 | 0.557 | 0.563 | 0.616 | 0.532 | 0.639 | 0.596 | 0.622 | 0.638 | 0.592 | 0.564 | 0.687 | 0.570 | 0.616 | 0.541 |
| 9 | 0.555 | 0.558 | 0.563 | 0.616 | 0.531 | 0.640 | 0.595 | 0.622 | 0.639 | 0.592 | 0.564 | 0.688 | 0.570 | 0.616 | 0.541 |
| Load-Max | | | | | | | | | | | | | | | |
| 8760 | 1.000 | 1.000 | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 103 | 0.999 | 0.999 | 0.992 | 0.993 | 1.000 | 0.996 | 1.000 | 0.997 | 0.990 | 0.996 | 1.000 | 0.996 | 1.000 | 1.000 | 0.998 |
| 9 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Load-Min | | | | | | | | | | | | | | | |
| 8760 | 0.338 | 0.355 | 0.342 | 0.350 | 0.282 | 0.408 | 0.376 | 0.380 | 0.438 | 0.378 | 0.347 | 0.499 | 0.397 | 0.411 | 0.378 |
| 103 | 0.346 | 0.361 | 0.352 | 0.359 | 0.288 | 0.412 | 0.381 | 0.388 | 0.444 | 0.383 | 0.347 | 0.505 | 0.400 | 0.421 | 0.388 |
| 9 | 0.431 | 0.432 | 0.430 | 0.456 | 0.385 | 0.512 | 0.468 | 0.474 | 0.515 | 0.460 | 0.432 | 0.588 | 0.462 | 0.510 | 0.452 |
| Wind-Avg | | | | | | | | | | | | | | | |
| 8760 | 0.303 | 0.293 | 0.271 | 0.266 | 0.000 | 0.289 | 0.280 | 0.206 | 0.368 | 0.300 | 0.406 | 0.350 | 0.265 | 0.245 | 0.241 |
| 103 | 0.342 | 0.301 | 0.289 | 0.295 | 0.000 | 0.269 | 0.287 | 0.217 | 0.361 | 0.320 | 0.428 | 0.314 | 0.250 | 0.225 | 0.236 |
| 9 | 0.323 | 0.308 | 0.297 | 0.000 | 0.000 | 0.293 | 0.335 | 0.241 | 0.359 | 0.403 | 0.420 | 0.413 | 0.369 | 0.285 | 0.308 |
| Wind-Max | | | | | | | | | | | | | | | |
| 8760 | 0.955 | 0.930 | 0.964 | 0.938 | 0.000 | 0.938 | 0.942 | 0.936 | 0.916 | 0.913 | 0.944 | 0.887 | 0.845 | 0.925 | 0.873 |
| 103 | 0.930 | 0.883 | 0.914 | 0.906 | 0.000 | 0.911 | 0.902 | 0.887 | 0.909 | 0.870 | 0.922 | 0.847 | 0.800 | 0.891 | 0.873 |
| 9 | 0.392 | 0.404 | 0.429 | 0.000 | 0.000 | 0.364 | 0.408 | 0.347 | 0.435 | 0.509 | 0.534 | 0.481 | 0.431 | 0.325 | 0.367 |

**Table A-1 (continued)**
**Comparison of the average, maximum, and minimum hourly values for the annual time-series data for load, wind, and solar. Each section compares the underlying hourly data (8760), the representative hour approach (103), and the seasonal average method (9). Columns represent the 15 model regions.**

| | NE | NY | MA | SA | FL | NEC -D | NEC -R | SEC | NWC | SWC | TX | MTN-S | MTN-N | PAC | CA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wind-Min | | | | | | | | | | | | | | | |
| 8760 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.012 | 0.003 | 0.004 | 0.013 | 0.009 | 0.001 | 0.002 |
| 103 | 0.005 | 0.002 | 0.000 | 0.020 | 0.000 | 0.010 | 0.006 | 0.001 | 0.051 | 0.014 | 0.008 | 0.029 | 0.026 | 0.004 | 0.028 |
| 9 | 0.233 | 0.218 | 0.151 | 0.000 | 0.000 | 0.145 | 0.189 | 0.057 | 0.251 | 0.267 | 0.185 | 0.321 | 0.292 | 0.248 | 0.240 |
| Solar-Avg | | | | | | | | | | | | | | | |
| 8760 | 0.137 | 0.135 | 0.139 | 0.150 | 0.153 | 0.138 | 0.140 | 0.150 | 0.157 | 0.160 | 0.186 | 0.158 | 0.184 | 0.154 | 0.191 |
| 103 | 0.142 | 0.131 | 0.134 | 0.150 | 0.161 | 0.133 | 0.133 | 0.154 | 0.156 | 0.158 | 0.188 | 0.156 | 0.187 | 0.155 | 0.200 |
| 9 | 0.131 | 0.132 | 0.131 | 0.146 | 0.150 | 0.132 | 0.141 | 0.148 | 0.159 | 0.173 | 0.184 | 0.159 | 0.180 | 0.156 | 0.189 |
| Solar-Max | | | | | | | | | | | | | | | |
| 8760 | 0.728 | 0.726 | 0.785 | 0.752 | 0.691 | 0.754 | 0.739 | 0.710 | 0.769 | 0.752 | 0.803 | 0.787 | 0.780 | 0.787 | 0.806 |
| 103 | 0.691 | 0.682 | 0.748 | 0.735 | 0.677 | 0.747 | 0.739 | 0.692 | 0.769 | 0.728 | 0.763 | 0.752 | 0.765 | 0.760 | 0.790 |
| 9 | 0.324 | 0.353 | 0.373 | 0.399 | 0.373 | 0.372 | 0.415 | 0.389 | 0.440 | 0.516 | 0.518 | 0.438 | 0.398 | 0.373 | 0.364 |
| Solar-Min | | | | | | | | | | | | | | | |
| 8760 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 103 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 9 | 0.000 | 0.002 | 0.061 | 0.045 | 0.041 | 0.054 | 0.054 | 0.040 | 0.062 | 0.035 | 0.030 | 0.016 | 0.067 | 0.056 | 0.000 |

Table A-1 compares summary statistics for the load, wind, and solar time-series data across the three approaches on a regional basis. The underlying hourly data (8760) is compared with the results from the representative hour approach (103) and seasonal average method (9) for the average, maximum, and minimum values in each time series. The peaks, minimums, and averages are preserved to a high degree of accuracy by construction in the representative hour method; however, the seasonal average method only ensures that the averages and peak load are accurately characterized. The other extreme values are notably not high enough (for the maximum values) and too high (for the minimum values).

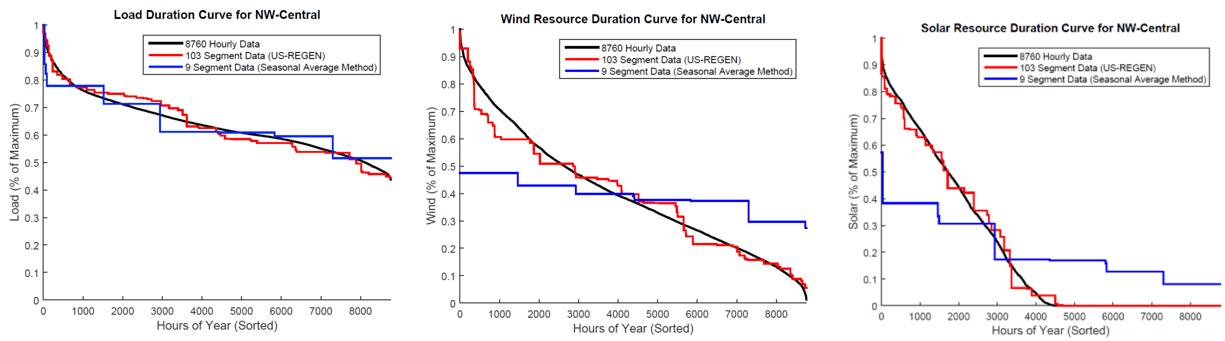



**Figure A-2**
**Duration curves for load (left panel), wind (middle panel), and solar (right panel) in the Northwest Central region. The hourly duration curve (black) is approximated with 103 segments (red) using the representative hour approach and 9 segments (blue) using the seasonal average method.**

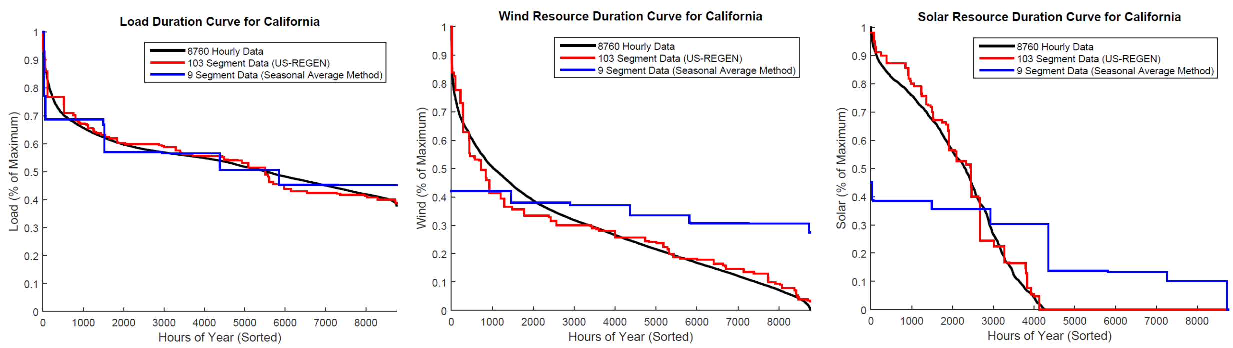



**Figure A-3**
**Duration curves for load (left), wind (middle), and solar (right) in California.**

Figures A-2 and A-3 demonstrate how the duration curves for these data fit these same trends for California and the Northwest Central region. For both regions, the US-REGEN representative hour values (red) closely approximate the distributions for the 8760 data (black). The seasonal average values (blue) only capture the average characteristics.

We have also examined correlations between multiple time series (e.g., between load and resources, between different resource types, and across regions). In particular, models should reflect the joint distribution of load, wind, and solar characteristics to reflect the economics of intermittent renewable technologies. Lamont (2008) shows how the marginal value of a generating technology depends both on a generator's average capacity factor and also on the covariance between marginal system costs and output from that generator. Thus, this theoretical insight motivates the importance of examining the correlation coefficients between load and renewable availability as metrics for systematically testing the quality of variability approximations.

As shown in Figure A-2, the simplified load-targeting approach captures load crudely but does not sufficiently represent the full variability spectrum for wind and solar availability let alone its interdependence. Extreme hours in wind and solar availability are poorly captured with the load-targeting approach due to its focus on averaging renewable characteristics during applicable seasons and load segments, which dampens the resource fluctuations.

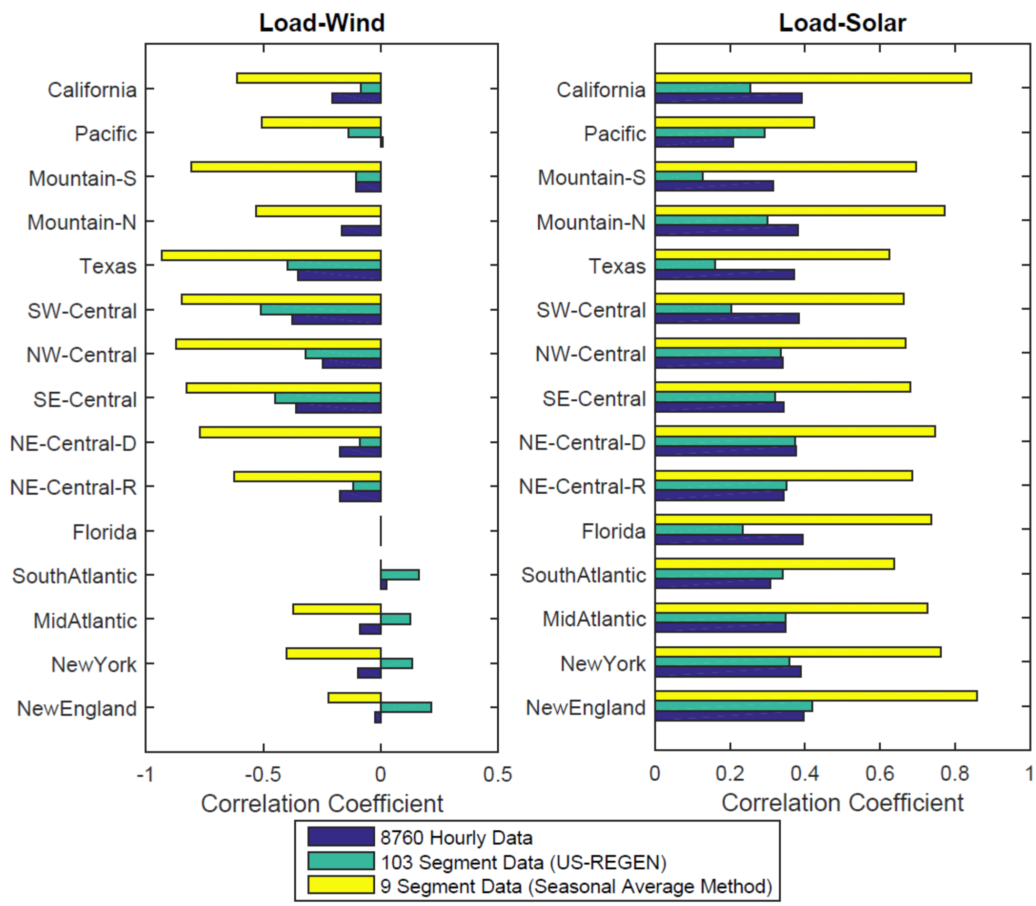



**Figure A-4**
**Correlation coefficient comparison for load and wind (left panel) and load and solar (right panel) across all 15 model regions for all 8760 hours (blue), the 103-segment approximation (green), and 9-segment approximation (yellow).**

Demand, wind speeds, and solar radiation not only vary within a region but also between model regions. Correlations between different regions can be an important dynamic in trade outcomes and large-scale system balancing. Figure S3 shows the interregional correlation coefficients across US-REGEN model regions for the actual 2010 data (left panel) and two approximations (middle and right panels).[1] The 86-segement data capture the cross-correlations in the complete data better than the 9-segment approach, which understates the heterogeneity across regions.

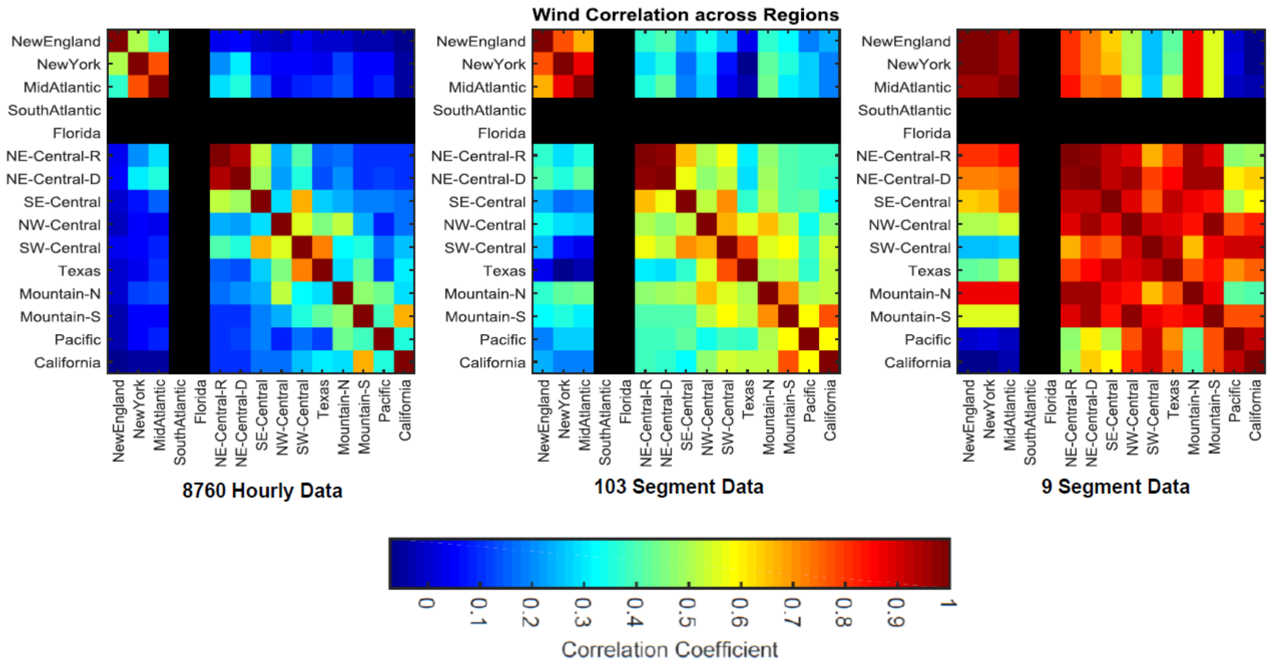



**Figure A-5**
**Interregional correlation coefficients for existing wind resources.**

[1] Data were not available for the South Atlantic and Florida model regions, since no wind capacity was installed in 2010.

## Scenario Details for the Static Analysis

**Table A-2**
**Cost assumptions for the marginal value curve analysis.**

| | Investment Cost ($/kW) | Lifetime (years) | Annualized @ 7% ($/kW-yr)* | Fixed O&M ($/kW-yr) | Variable O&M ($/MWh) | Heat Rate (th.btu/ MWh) | Fuel price ($/mmbtu) |
|---|---|---|---|---|---|---|---|
| **Nuclear** | $6,000 | 70 | $423 | $80 | $2 | 10,000 | $0.5 |
| **Coal** | $2,500 | 50 | $181 | $40 | $3 | 10,000 | $2 |
| **NGCC** | $1,200 | 50 | $87 | $20 | $3 | 7,000 | $5 |
| **GT** | $800 | 30 | $64 | $20 | $4 | 11,000 | $5 |

*Annualized investment (rental) cost is equal to total investment cost multiplied by a capital charge rate.

The charge rate *c* is based on discount rate *r* and technology lifetime *n*: $c = r/(1-(1+r)^{-n})$

| | Charge Capacity ($/kW) | Storage Capacity ($/kWh) | Charge Penalty | |
|---|---|---|---|---|
| **Storage** | $800 | $120 | 25% | Default assumption, varied in analysis |

Transmission Capacity costs $3.85M per mile for a 6.4 GW line = $270/kW between CA and Mtn-S

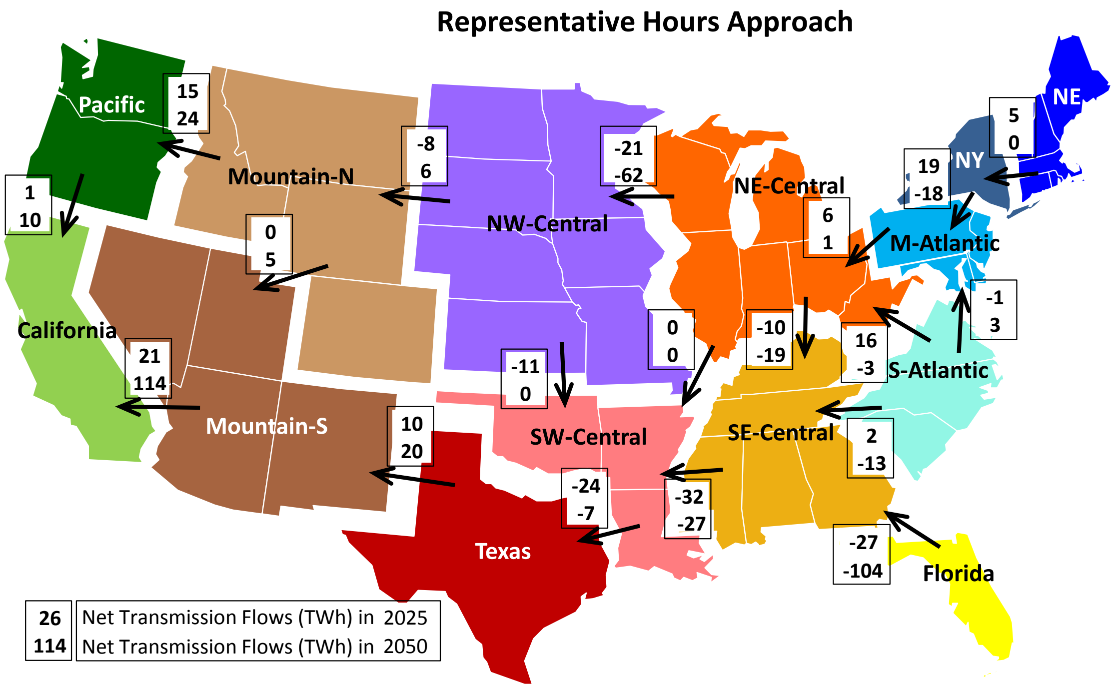
## Scenario Details for the Dynamic Analysis

The scenario specification for the dynamic analysis uses many common assumptions across the reference (i.e., no policy) case and the two carbon tax cases. For this study, we use the electric sector version of the US-REGEN model. Additional documentation about the model structure and assumptions can be found in EPRI (2014). For each of the three scenarios in these numerical experiments, all model features are held constant save for the representation of temporal variability.

All scenarios use fuel prices from the 2015 Annual Energy Outlook (EIA, 2015). Technology cost and performance assumptions come from the most recent EPRI Integrated Generation Technology Options report. In line with AEO 2015 assumptions, there are no forced retirements for existing coal units in the reference case, though retirements for economic reasons are possible in any period. Limitations on new transmission and nuclear capacity additions are based on EPRI expertise and historical experience.

Policies in all scenarios include existing state RPS requirements (as of January 2013), MATS, CWA § 316(b), RCRA CCR, and CAA § 111(b) performance standards for fossil units. Performance standards under CAA § 111(d) for existing sources (i.e., the Clean Power Plan) are not included. Other state and regional policies include RGGI and California's AB32. Extensions of the ITC and PTC are not assumed. Rooftop solar is modeled as a "behind-the-meter" technology and, as per current regulatory practice, receives the retail rate for generated electricity instead of the wholesale price. The carbon tax scenarios apply rates of the \$25/t-$CO_2$ and \$25/t-$CO_2$ to power sector $CO_2$ emissions beginning in 2025.

Figure A-6 provides an additional comparison between the representative hour and seasonal average approaches for regional trade flows.



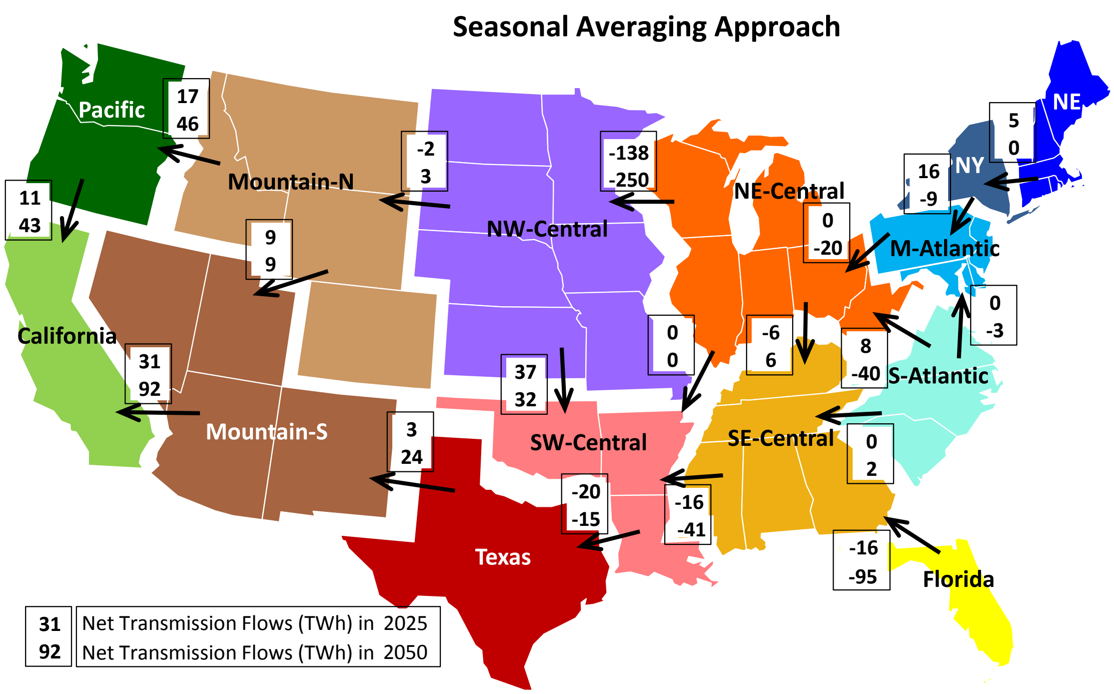




**Figure A-6**
**Annual regional trade (TWh) under the $50/t-$CO_2$ carbon tax scenario under the representative hour (top) and seasonal average (bottom) approaches. Negative values indicate that flows move in the opposite direction of the arrow in a given period.**

3002008653

**Electric Power Research Institute**
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com