

# Regions in Energy Market Models

W. Short

**Technical Report**  
**NREL/TP-640-40506**  
**February 2007**

NREL is operated by Midwest Research Institute • Battelle Contract No. DE-AC36-99-GO10337



# Regions in Energy Market Models

*Technical Report*  
NREL/TP-640-40506  
February 2007

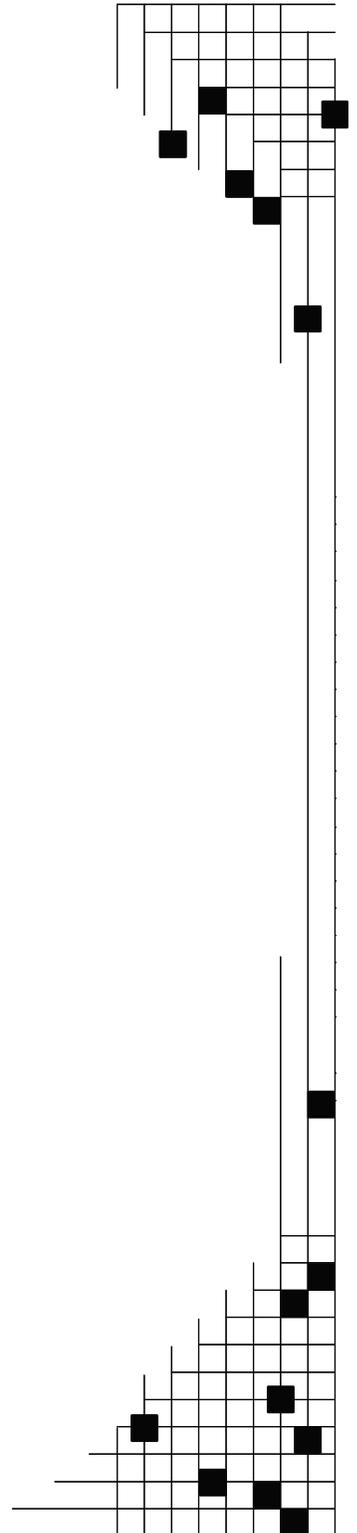
W. Short

Prepared under Task No. ASA6.2023

**National Renewable Energy Laboratory**  
1617 Cole Boulevard, Golden, Colorado 80401-3393  
303-275-3000 • [www.nrel.gov](http://www.nrel.gov)

Operated for the U.S. Department of Energy  
Office of Energy Efficiency and Renewable Energy  
by Midwest Research Institute • Battelle

Contract No. DE-AC36-99-GO10337



## NOTICE

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

Available 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](mailto:orders@ntis.fedworld.gov)  
online ordering: <http://www.ntis.gov/ordering.htm>



## Table of Contents

	<b><u>Page</u></b>
Abstract	1
Introduction/background	1
Reasons for fine spatial resolution	3
Benefits of final spatial resolution	4
Problems with fine spatial resolution	5
Two examples of models with different spatial constraints	8
Modeling with final spatial resolution	17
Modeling with coarse spatial resolution	18
Conclusions	21
References	22

## Abstract

This report explores the different options for spatial resolution of an energy market model – and the advantages and disadvantages of models with fine spatial resolution. It examines different options for capturing spatial variations, considers the tradeoffs between them, and presents a few examples from one particular model that has been run at different levels of spatial resolution.

## Introduction/background

A large number of national energy market models are currently used to do market and policy analysis. The most widely known in the United States is probably the National Energy Modeling System (NEMS), developed by the Department of Energy's (DOE) Energy Information Administration (EIA). NEMS is used primarily for developing the *Annual Energy Outlook* (EIA 2005). However, there are many other models – several of which are listed below (note that this is not a comprehensive list).

- ICF International's Integrated Planning Model (IPM)
- ICF International's North American Natural Gas Analysis System (NANGAS) model,
- Argonne National Laboratory's All-Modular Industry Growth Assessment (AMIGA) model,
- Brookhaven National Laboratory's MARKet ALlocation (MARKAL) model,
- Resource For the Future's Haiku model,
- Altos' North American Regional Gas (NARG) model,
- Energy 2020
- ABB's Grid View transmission and optimal power flow model

As shown in **Figure 1**, one striking fact about these models is that none of them use the same regional structure. NEMS has multiple overlaying sets of regions – North American Electric Reliability Council (NERC) regions for the electric sector, Census regions for the end-use sectors, and various regional structures for fossil fuel supplies. The North American IPM data set has 80 regional definitions, which can be aggregated up based on user needs.<sup>1</sup> Brookhaven's MARKAL has only one electric region for the entire United States.

As we will see in the discussions that follow, the spatial resolution of these models is determined largely by their intended purpose; and, on the other hand, their uses are constrained by their spatial resolution.

Note that in addition to differing regional structures, energy market models also differ in the level of resolution they use for:

- time – e.g., time horizon, length of time periods
- technologies – e.g., types of coal-fired power plants

---

<sup>1</sup> As cited on November 7, 2005, at:  
[http://www.icfconsulting.com/markets/energy/doc\\_files/ipm\\_overview3.pdf](http://www.icfconsulting.com/markets/energy/doc_files/ipm_overview3.pdf)

end uses – e.g., space heating vs. water heating in buildings  
 markets – e.g., distinguishing between commercial and residential buildings, or  
 between NAICS industrial codes/sectors

These other types of resolution are recognized in this report, only from the perspective that they can be a factor in determining the appropriate level of regional resolution.

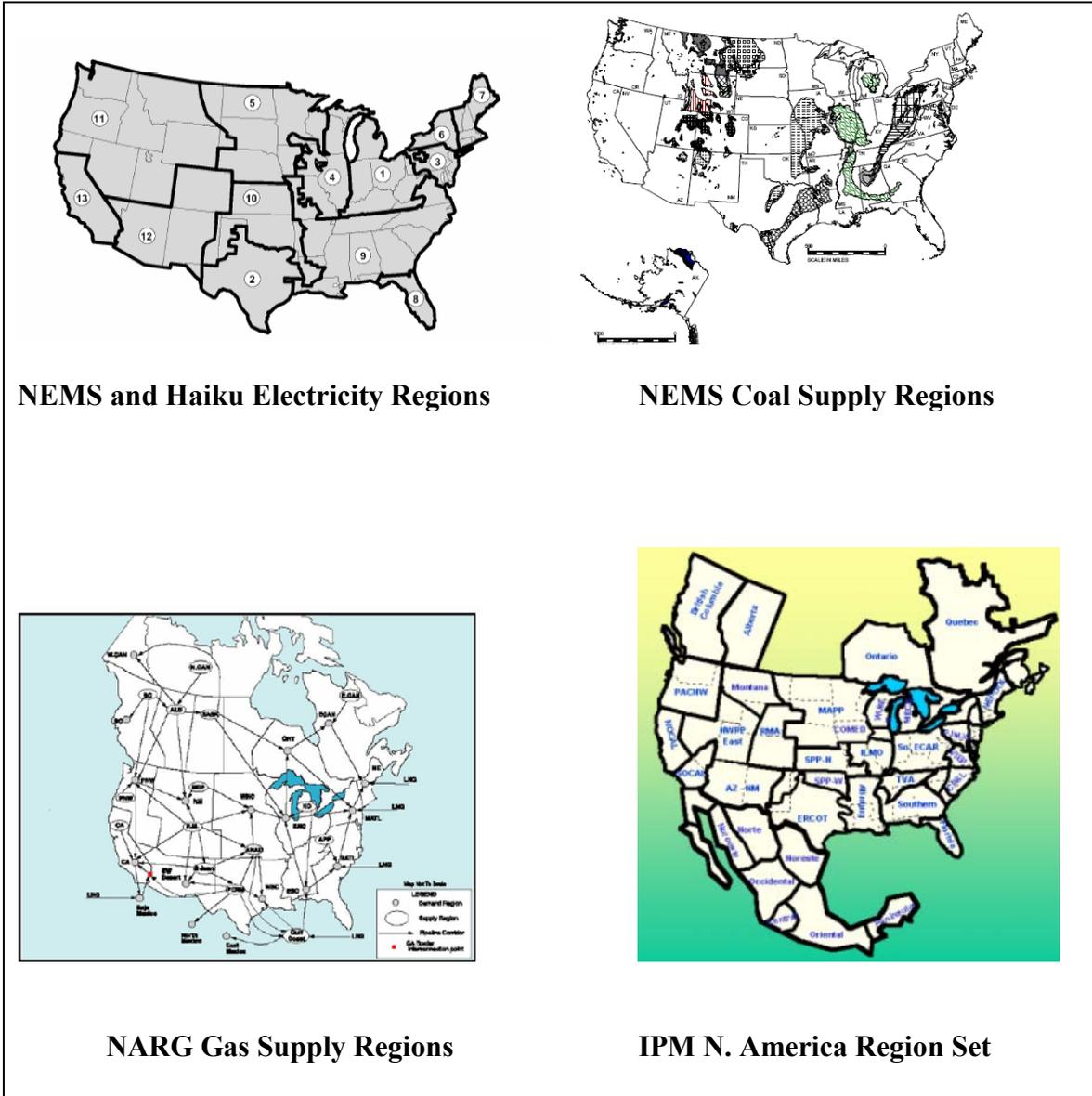


Figure 1. Regional Structures of Major Energy Market Models

## Reasons for fine spatial resolution

There are many good reasons to include as many regions as possible in an energy market model. They are explained below.

**Question to be answered:** The single most important criterion in deciding on the appropriate level of regional resolution is the question – or set of questions – the model is being built to address. For example, in its analysis of the impact of the Comprehensive Electricity Competition Act, the DOE Policy Office had the DOE/EIA’s National Energy Modeling System (NEMS) regional structure modified from 13 NERC region/subregions to 114 power control areas. The revised model, Policy Office Electricity Modeling System (POEMS), was better able to capture more local issues such as stranded assets, environmental policies, and other impacts of utility deregulation.

This modification of the electric-sector regions of NEMS to a finer spatial resolution in the POEMS model is a relatively rare occurrence. Such a spatial resolution modification usually requires excessive time and financial resources. Other approaches include finding an existing model with the appropriate spatial resolution, disaggregating the outputs of a model that is too coarse spatially, or building an entirely new model to meet the need.

**Regional market differences:** An equally compelling reason for fine spatial resolution is that the regions differ with respect to important variables that drive the model. For example, the retail price of electricity to residential consumers can vary by more than a factor of 2, e.g., Washington State vs. California (EIA 2006). California’s prohibition on coal-fired power plants is another example that is based on environmental regulation, not energy prices. California’s geothermal resources also stand in stark contrast to Arizona’s geothermal potential.

Regional differences can create niche markets that would be missed entirely in a model with coarse spatial resolution. The energy contribution in a small regional niche market may not be significant on a national scale. However, the initial deployment of a technology in a niche market can lead – both in the real world and the model – to improvements in the deployed technology that allow it later to penetrate larger, more significant markets.

**Interregional energy transport and markets:** Regional differences in prices, environmental exclusions, resources, and other factors might not be so important if it was possible to ship energy between regions easily and at low cost. For example, California uses coal-fired power generation transmitted into southern California from the “Four-Corners Area” of New Mexico, Arizona, Utah, and Colorado. This transmission cost is not large enough to change the competitive balance between coal and photovoltaic generation, but it could be a significant factor in deciding between coal and natural gas in California.

Thus, the modeling of multiple regions allows one to capture transport bottlenecks and the costs of transporting energy between regions. It also yields estimates of the transport levels and the transportation infrastructure requirements.

**Correlation:** In some cases, it may be possible to retain a coarse regional structure and use average values across each region for the market drivers. However, the use of such averages can substantially mislead when multiple drivers are correlated. For example, U.S. wind resources are prevalent through the middle of the country from North Dakota to northern Texas. This is also one of the least densely populated areas of the country (people do not tend to live in windy regions) with the fewest electric loads. Thus, much of the wind resource may require extensive transmission lines to be useful. Trying to capture such spatial correlation between wind resources and population, with an average wind resource and total electric load across the entire country, omits a key issue in estimating wind energy market penetration.

**Diversity:** In reality, energy prices and quantities vary from one region to the next. A small shift in the real-world price of one energy commodity could make it the cheapest source in one region, but not all regions. Such a price shift usually will have some impact on energy supply and demand (especially in that region in which the prices reverse), but the demand in any particular region is not usually large enough to turn the real-world market upside down with all regions shifting to this new cheaper supply source. However, this is exactly what can happen in an optimizing model with only a single region for all interactions. With no diversity in supply, a slight increase in the price of one source of energy can make it more expensive than the price of an alternative source, sending the entire market away from the first source.

This kind of “knife-edge” result limits the usefulness of such a model, because it can’t appropriately capture the impact of small changes in energy prices. The effect often is not that noticeable in single-region models, because they are diverse in other factors (e.g., multiple time slices and/or technology variants). The effect is also less noticeable in models with market-penetration algorithms that implicitly assume variation around the average prices, e.g., a logit-based market penetration algorithm as compared to the “knife-edge” results that can be incurred with a linear program optimization.

## **Benefits of fine spatial resolution**

The reasons given for using fine spatial resolution in an energy market model could also be considered the benefits of fine spatial resolution. A fine resolution allows for consideration of different issues and market drivers. It allows one to capture the correlation between different factors that vary between regions and allows the analyst to estimate the energy transport between regions and the need for transportation infrastructure. It also reduces the incidence of – and the impact of – “knife-edge” solutions. Finally, and obviously, it provides results at a finer level of regional detail and can present niche market opportunities.

Unfortunately, the magnitude of these benefits often cannot be known without analysis. Without explicitly representing the regions – at a level sufficient to capture the differences – one can't know whether the differences are indeed material.

## **Problems with fine spatial resolution**

In spite of the advantages listed above for including many regions in an energy market model, most models include relatively few regions. For example, NEMS has 13 electric regions and nine end-use sector Census regions. We explore some of the reasons why most models use only a few regions.

**Computer run time:** Generally, the number of variables and constraints in a model is at least proportional to the number of regions – two regions require at least twice as many variables and constraints as a single region. In fact, the number of variables required can be combinatorial with the number of regions, as even more variables are required to capture the transfer of energy among regions.

These additional variables and constraints almost invariably increase the computer time required to solve the model; sometimes the increase is exponential. With the phenomenal increase in computer speeds that has occurred over recent years, energy market models have grown in scope and can delve deeper into market details. Nonetheless, run-time considerations – which can amount to hours or sometimes days – continue to discourage the addition of more detail and finer regional structure.

**Computer memory requirements:** With today's ever-increasing computer memory capabilities, memory requirements are less of a constraint on model size than they used to be. However, market models frequently expand to consume the computer resources available.

**Analyst resources:** In addition to the increase in computer and data resources required with more regions, more analyst time is required to prepare model inputs and to check the outputs at the lowest regional level.

**Interregional transfers:** One reason more analyst time is required with more regions is that the analyst needs to examine not only the regional results, but also the interregional transfers of energy. These transfers can become quite intricate with transfers through regions, multiple routes between two regions, etc. The cost of these transfers also must be modeled, along with the ability to expand the transportation links (e.g., pipelines) between the regions. The actual expansion of transportation links can depend as much on the local topography, land use/ownership, impacts on current energy providers/transporters, regulation, and politics as on the economics of the transportation link development itself – thus, economic modeling can be limited in its ability to capture such regional detail.

One particularly difficult interregional transfer problem to model in an energy market model is the transmission of electricity. These difficulties arise for a number of reasons:

- Real power flows must obey Kirchhoff's circuit laws<sup>2</sup> and, therefore, the energy that can flow across any given line is a function of the flows on and capacities of interconnecting lines, the amounts generated and the locations of the generators, and the amounts demanded and their locations.
- Reactive power requirements can impact real power flows.
- Transmission line costs per MW/mile can vary by an order of magnitude due to economies of scale, siting issues, topography, etc.

Nonetheless, there are approximations that can be made when transmission is critical to a technology or market. For example, NEMS adds a general transmission cost (\$/KW) to all new power plants to represent their average share of new transmission builds, and NEMS adds a distinct transmission cost to new coal plants built outside of California to meet loads in California. Similarly, NREL's WinDS model (Short 2006), explicitly considers the addition of new transmission lines for new wind installations because transmission can be a critical cost component of remote wind farms.

**Accuracy of results:** It is usually possible to increase the accuracy of the results for a region by disaggregating the region to multiple subregions. For example, assume an energy quantity is estimated for a region as  $X$ . The error in the estimate,  $e$ , can be considered to be a random variable with mean 0 (the estimate is unbiased) and standard deviation  $\sigma_e$ . We assume the relative error,  $p$ , can be expressed as  $p = \sigma_e/X$ . Now, assume we can divide the region into nine equal subregions, and that the quantity in each subregion  $i$  is estimated as  $X_i = X/9$  with the same relative error,

$$p_i = p$$

This implies

$$\sigma_{ei} / X_i = \sigma_e / X$$

or since  $X_i = X/9$

$$\sigma_{ei} = \sigma_e / 9$$

Now the question is "What would the error be if we determined  $X$  by adding all the  $X_i$ ?" To estimate this error  $\sigma_{e'}$  in  $X$ , we first assume the errors in each subregion are independent of one another. With this assumption, then we can add their variances  $\sigma_{ei}^2$  to get the variance in the total estimate as

$$\sigma_{e'}^2 = 9 \sigma_{ei}^2$$

or

$$\sigma_{e'} = 3 \sigma_{ei}$$

or

$$\sigma_{e'} = 3 \sigma_e / 9 = \sigma_e / 3$$

So, for this simple example and our assumptions of independent proportional errors in each of the nine subregions, the error  $\sigma_{e'}$  of the estimate in the total derived by adding the estimates from the nine subregions is one-third of the error  $\sigma_e$  of the quantity for the large region estimated directly. In other words, for this example, estimating  $X$  by adding up the

---

<sup>2</sup> In their simplest form, Kirchhoff's circuit laws state that 1) at any point in an electric circuit, the sum of all entering and the sum of all departing currents must be equal, and 2) the sum around a circuit of the electrical potential differences must be zero.

$X_i$  at the subregion level is a more accurate estimate of  $X$ , than making a direct estimate of  $X$ .

The example above assumed the subregion estimates were proportionately as accurate as a direct estimate for the region as a whole. However, the accuracy of the results at the subregion level often is not as good as the accuracy at the original larger region. This is somewhat analogous to estimating the movement of individual stocks compared to the stock market as a whole.

One reason subregion results may not be as accurate is that the drivers may change as one makes the spatial resolution finer. For example, at the NERC region level, the primary driver in the choice between a coal-fired power plant and a gas-fired, combined-cycle power plant is probably the cost of electricity from the two different technologies. However, at the county level, the driver might be the availability in the county of a gas pipeline or a railroad for coal transport.

Such inaccuracies at the subregion level can lead to unfounded questions regarding the validity of the model. There may be value in not reporting results at the finest level of spatial resolution, but rather rolling those results up to the next level.

**Data requirements:** Another reason that subregion results may not be as accurate is that data frequently is available at one regional level, but not another. For example, electricity demand may be available at the utility service territory level or at a power control area, but not at the county level. Surrogate data is frequently manufactured in such cases. For example, electricity demand data at the county level or below could be manufactured by dividing the demand at the utility service territory among counties proportionately to the population of each county within the service territory – or by more sophisticated spatial statistical techniques (Johanesson 2005).

Thus, the benefits of added precision of spatial resolution may, in some cases, be negated by the increased uncertainty that comes from associated data gaps. Similarly, the issue of unknown and unspecified correlations between data is increased with more regional detail. Occasionally, one kind of data is available with one data structure, and another needed data type is available with another regional structure. For example, NEMS uses electric generation capability at the NERC regions and subregions level for the electric sector, but electric load data from the Census regions for the end-use sectors.

Often data may be available at a fine spatial resolution level, but very difficult to collect because it is not always in one place or of a consistent nature. For example, data available from states on commercial-building energy use may be broken down by different types of buildings in different states. Collecting such data from agencies within the 50 states also may be prohibitive in cost and analyst effort.

**Time horizon:** Energy market models differ in the level of detail they include in ways other than just their spatial resolution. Another dimension that is treated differently is time. Models differ regarding the time horizon they simulate, the time steps or periods for

which they produce results, and the variations within a time step that they consider. For example, NEMS covers a 25-year time horizon with annual time periods. Within each one-year NEMS period, 11 time slices are considered in the capacity expansion module of the electric sector of the model (EIA 2005). MARKAL, on the other hand, has a 50-year horizon, but covers it with five-year periods. This report does not focus on the time resolution within models. However, there are certain connections between time and the appropriate level of spatial resolution.

Although there are certainly exceptions, a model with a long time horizon is more likely to be built with fewer, larger regions. Over a longer time horizon, more basic infrastructural shifts are likely to occur that cannot be captured by the model. These shifts may dramatically alter any regional differences that exist at the time the model is developed. In such a case, the extrapolation of regional differences over long time horizons is not only a waste of computer resources and analyst effort, but also has the potential to be less accurate. For example, regional price differences for natural gas can change dramatically over time as pipelines are built from resource-rich regions like the Rocky Mountains to load centers on the East Coast and West Coast. Similarly, existing regulatory differences among regions can change dramatically over time, as has occurred in the past two decades with utility deregulation and the advent of independent system operators, regional transmission organizations, etc. On the other hand, there are regional differences that do not change with time and for which fine spatial resolution may be appropriate, regardless of the model's time horizon. For example, renewable resources vary significantly regionally, but do not vary widely or in any predictable fashion over time.

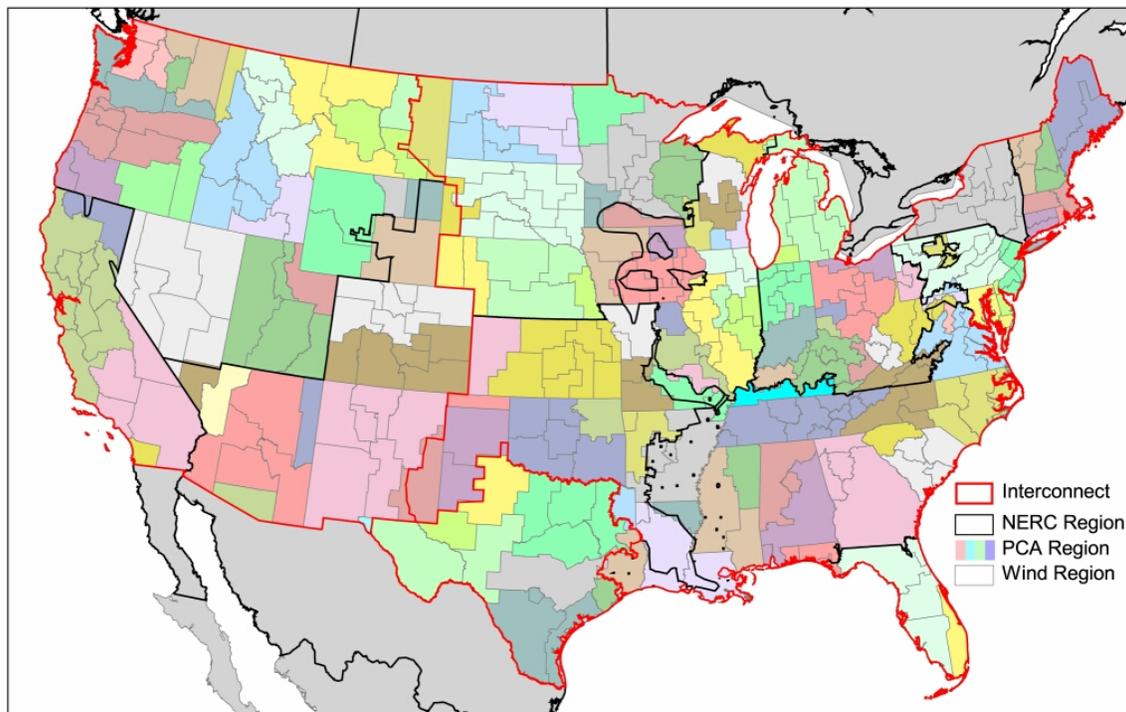
## **Two examples of models with different spatial constructs**

In this section, we describe two models that differ widely in their structural spatial resolution and, consequently, their spatial data requirements and results. NREL recently developed two models of the U.S. electric sector, which are at opposite ends of the spectrum with respect to spatial resolution – the Wind Deployment System (WinDS) model (Short 2006a) and the Stochastic Energy Deployment System (SEDS) model (Short 2006b). It is useful to contrast these two models and the rationale behind the level of spatial resolution used in each.

WinDS is a model of capacity expansion in the U.S. electric sector through 2050. In each two-year period, it optimally selects the type of generation needed (coal, gas combined cycle, nuclear, wind, gas combustion turbines, etc.) in each region to meet electric loads and to maintain system reliability at least cost. WinDS was designed to capture the major market issues related to wind energy – intermittency, transmission, and siting – and to consider these in competing wind energy against conventional electricity generation..

The SEDS model has a totally different purpose from WinDS. SEDS is designed to simulate the U.S. energy system<sup>3</sup> through 2050 by explicitly incorporating uncertainty through the use of probability distributions on uncertain market drivers. Whereas WinDS can estimate the market potential of wind in great detail under a selected scenario, SEDS uses a Monte Carlo approach<sup>4</sup> to examine the market potential of wind and other energy supply, conversion and demand technologies under a full range of possible futures, all in the same run.

WinDS and SEDS are designed to address the future market potential of some of the same technologies in the same sector. Yet WinDS has 358 supply regions (see **Figure 2**) while SEDS has one region for the electric sector. Both models would probably benefit from further spatial resolution. If WinDS had more regions, the areas would be smaller and would provide more accurate estimates of the length and cost of transmission lines for wind to be built from the center of one region to the next. With greater spatial resolution, SEDS would represent more regional nuances in the energy system, and this would potentially lead to a more accurate overall forecast. However, the number of regions has not been increased in either model, largely because of computer run time – but also due to the complexities of modeling at finer spatial resolution.



**Figure 2. WinDS Regions**

<sup>3</sup> As of December 2006, SEDS includes only the electric sector, but it is being expanded by NREL and other national laboratories to include the full U.S. energy system.

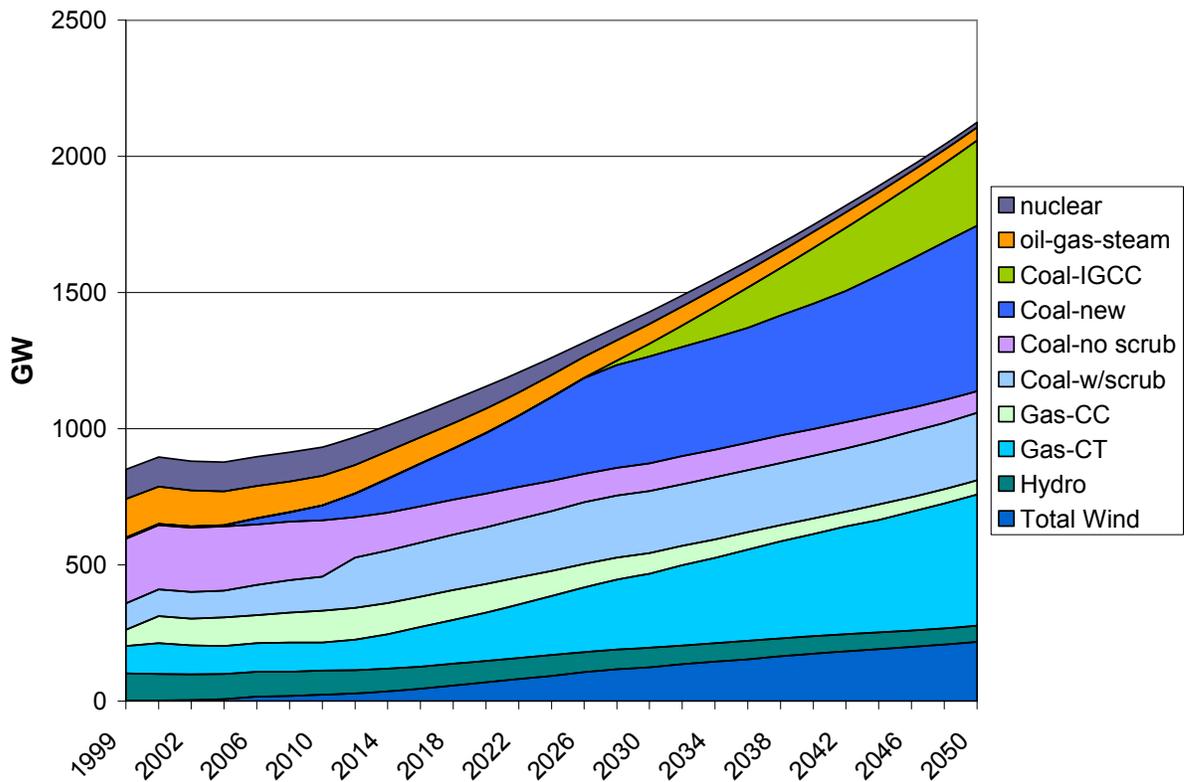
<sup>4</sup> SEDS employs a Latin-Hypercube algorithm that can achieve the same level of statistical accuracy as a Monte Carlo approach, with significantly fewer trials and less run time.

### **Wind Deployment System (WinDS) Model**

The 358 regions in WinDS were built up from counties, no region includes counties from two states, wind resource areas were separated from high load areas, and five regional levels overlay one another – wind supply/demand regions, 136 utility power control areas (PCA), Regional Transmission Organizations (not shown in Figure 2), 13 NERC regions/subregions (as of December 2005), and the three synchronized grid interconnects within the United States. WinDS not only uses many regions but effectively achieves even more spatial resolution through supply curves within each of the 358 wind supply/demand regions. This helps estimate the cost of building transmission lines for each class of wind from each site within the region to the nearest available transmission capacity, or to the nearest load center within the region. So, in some sense, WinDS has a finer spatial resolution than the 358 wind supply/demand regions.

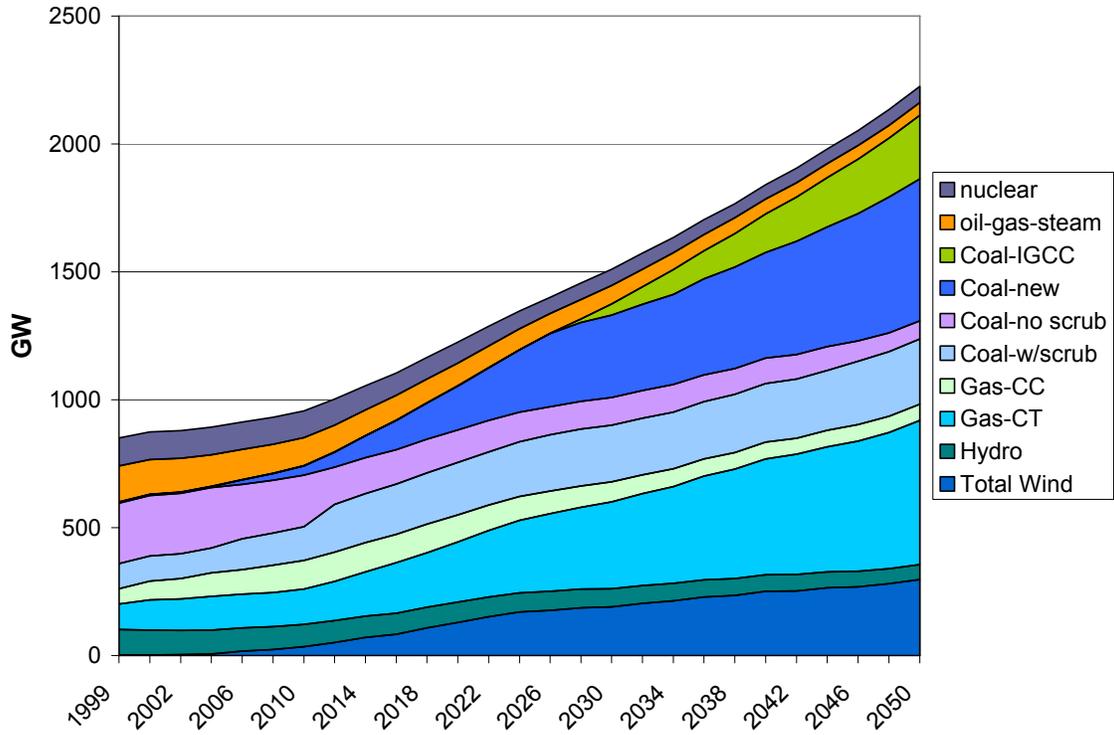
With its 358 regions, WinDS has more than 300,000 variables and 100,000 constraints that are solved in a linear program, which is run for each of 26 two-year periods as the model marches through 50 years from 2000 to 2050. A typical run can require 5-10 hours. The marginal increase in accuracy that might be attained with a finer spatial resolution of WinDS is probably not worth the increase in run time, data requirements, and analyst effort. This cannot be known for sure, even if a second version of WinDS was built because of the difficulty in gauging whether the accuracy of the results has increased.

The WinDS model can estimate the capacity of wind and other technologies that might be installed by 2050 (**Figure 3**). The results of Figure 3 reflect our base case, which uses fuel prices; electricity demands; and conventional generation costs, performance, and other assumptions from the DOE Energy Information Administrations *Annual Energy Outlook 2005* (EIA 2005). Wind turbine cost and performance values in this base case are based on the DOE Wind Program goals. This case also reflects many costs and constraints associated with wind siting, transmission, and intermittency. In particular, these results assume all the regional structure of WinDS described above with the reserve requirements for wind calculated at the RTO level (or power control area level where no RTO exists). This case shows a wind penetration of 218 GW by 2050.



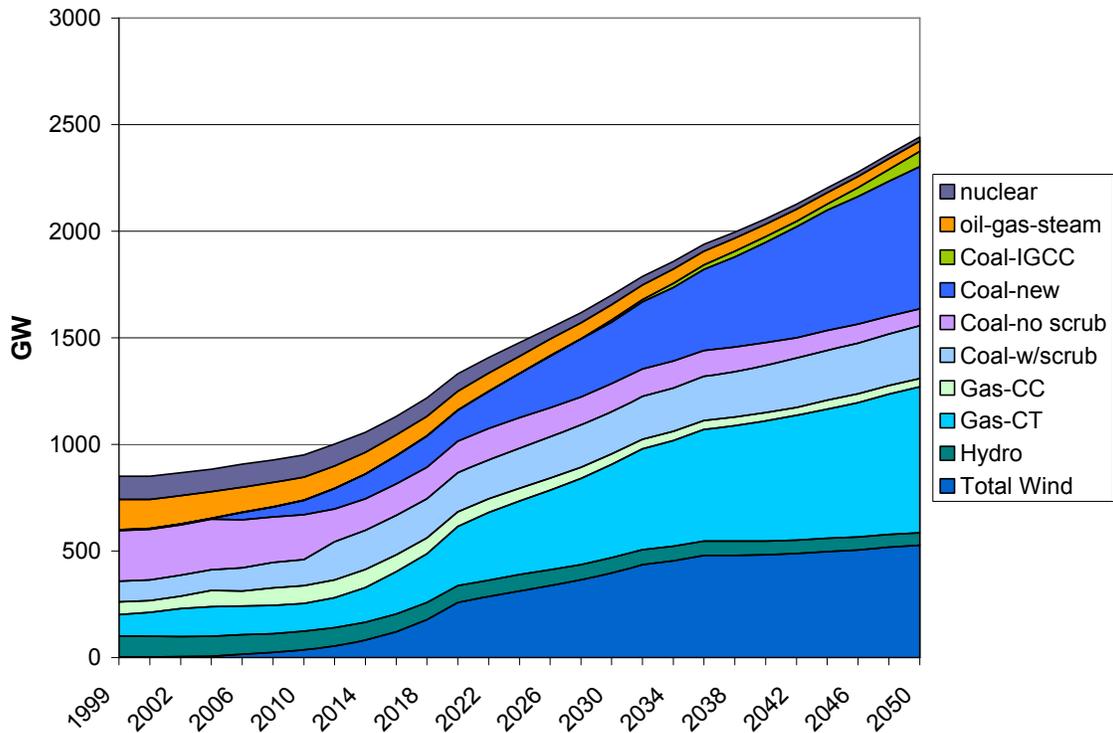
**Figure 3. WinDS Base Case Capacity Results**

To assess the impact of WinDS' fine spatial resolution, we conducted a second run of the WinDS model without the fine spatial resolution. In this second run, we aggregated all the regions up to the level of the 13 NERC regions/subregions, similar to the regional structure used in the electric sector of the NEMS and Haiku models. **Figure 4** shows the capacity expansion results for this second case with the market penetration of wind increased by 36% in 2050 from 218 GW to 297 GW. This increase in wind penetration is a direct result of the aggregation of the regional structure in WinDS up to the NERC-region level. With the larger regions, the cost of transmitting wind power (and conventional generation as well) is ignored within each of the 13 NERC regions. In effect, all wind resources within each NERC region are assumed to be collocated with all the loads in the NERC region. The wind penetration might have been even greater were there not a second factor associated with this coarser spatial resolution that disadvantages wind. With all the wind within a NERC region assumed to be at essentially the same point, the WinDS model assumes the wind resource availability is the same for all the wind within the NERC region. This dramatically reduces the planning reserve contribution of wind and increases the operating reserve requirements due to wind.



**Figure 4. WinDS Base Case Results From the Single-Region Mode of Operation**

We examined one more extreme case in WinDS wherein we assumed only one national region. In this case, the wind penetration increased yet again to 527 GW by 2050, or a 242% increase over our base case. As can be seen by comparing Figures 3 and 5, this increase in wind capacity reduced the penetration of coal integrated gasification combined-cycle capacity (IGCC), gas combined cycle, and other conventional plants.



**Figure 5. WinDS Base Case Results From the Single-Region Mode of Operation**

The results of **Figure 5** are clearly inaccurate and would not be presented by any reasonable modeler using a “single-region” model. Normally, some form of adjustment or supply curve is used to prevent wind from dominating as it does in Figure 5. The problem such models face is determining how large the adjustment to wind power costs and performance should be. To address this problem, one might use a fine-spatial-resolution model like WinDS to develop a supply curve that captures the additional (beyond bus-bar generation) costs of wind associated with siting, transmission, and intermittency. Such a curve could be used in other models lacking the same level of spatial resolution to implicitly capture these extra-generational costs.

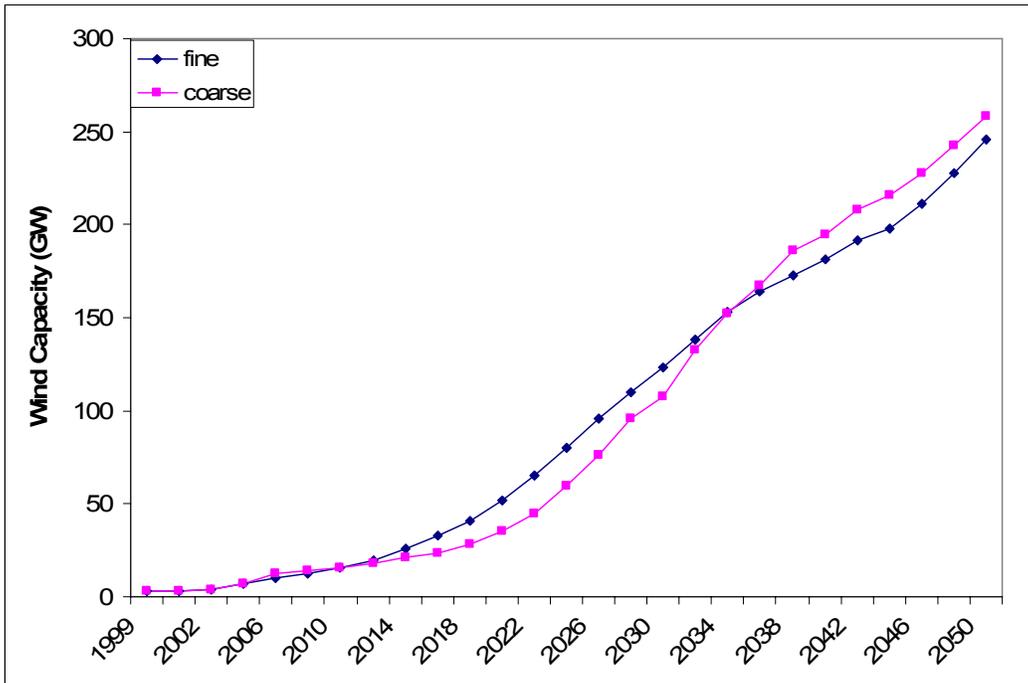
Finally, we conducted one more spatial resolution sensitivity with the WinDS model to examine the impact of different data scales. In this case, we aggregated the resource regions used to develop the supply curves for wind resource that are input for each class of wind in each of the 358 regions. These supply curves are built using NREL’s Geographic Information System (GIS). In the GIS, the wind-resource data is available at a resolution as small as cells 200 meters by 200 meters. The wind resource in each of these cells is estimated to be all of a single class (i.e., class 3, 4, 5, 6, or 7).

To examine the significance of this fine spatial resolution, we ran a case with coarser estimates of wind resources. Specifically, we aggregated the wind-resource cells to 30 km on a side, or 900 km<sup>2</sup> squares. Then, we found the average wind class within each of these 900 km<sup>2</sup> cells and rounded to the nearest class, i.e., class 5 as opposed to class 5.2 or 4.8. These larger cells were then used to rebuild the supply curves for each of the 358

WinDS supply regions. The WinDS model was then run with these aggregated supply curves to see the impact of this coarse spatial resolution on wind market penetration.

**Figure 6** shows that the coarse wind-resource supply curves have a small but complex impact on the model’s penetration of wind. In the early years of 2006 to 2010, wind penetration with the coarse supply curves slightly exceeds the fine curves of the base case. This occurs because there are a few 900 km<sup>2</sup> cells where class 6 and class 7 wind resources dominate the cell, and all wind in the cell is assigned the high-class label. However, such high-class wind areas are limited, and the coarse-wind penetration curve crosses below the fine curve by 2012 and stays below it until 2034. During these years, the predominant wind resource cells used are those that have mostly class 5 and 4 wind. In these cells, the remaining class 6 and 7 resources are essentially reclassified by the coarse grid as class 5 and 4. As such, the wind is not as attractive, and the market penetration of wind falls below that of the fine cell case. After 2034, the reverse happens. The 900 km<sup>2</sup> coarse cells are still predominantly class 4 and 5, but have a lot of class 3 in them (which is classified as class 4 or 5 along with the rest of the 900 km<sup>2</sup> cell). Thus, there is effectively much more class 4 and 5 resource in the coarse case, which leads to greater penetration in the late years.

The coarse curve of Figure 6 appears pretty much as expected. Essentially, most of the wind is classified with the mid-quality class 4 and 5 resource, and the market penetration occurs in a more “knife-edge” fashion with little entering the marketplace until class 4 and 5 are economic – at this point, wind penetrates very quickly. So, a coarser spatial resolution does bias the results, but the direction of the bias is dependent principally on the point in time that one is examining.



**Figure 6. Impact of Wind Resource Spatial Resolution on Wind Capacity Installed**

## Stochastic Energy Deployment System (SEDS)

Unlike WinDS, SEDS has only a single region for the electric sector and can be run deterministically in a fraction of a second. When run stochastically, the run time is proportional to the number of stochastic trials that are undertaken – a run of only 200 trials requires 20-30 seconds. Eventually, we may use a finer spatial resolution in SEDS. However, we are reluctant to increase the resolution and run time now, because we are expanding SEDS to include the transportation sector and will eventually expand it to all the end-use sectors and a macroeconomic feedback component. There is one more consideration: SEDS is being developed for use by other analysts. With that in mind, we want to keep it as quick and transparent as possible.

The question is whether the answers provided by SEDS with its single-region electric sector are sufficiently accurate, or are they unacceptably misleading due to the model's regional simplicity? The answer is not entirely clear cut. **Figure 7** shows that, historically, there have been major swings in the types of generators selected by electric utilities, driven by factors outside the scope of most energy market models. In the 1950s and '60s, a mix of coal, gas, and oil entered the market. However, as the availability has increased – and the cost declined for fuel transport (rail lines for coal, pipelines for gas) – the more recent shifts from one technology to another seem to have pervaded the whole country.

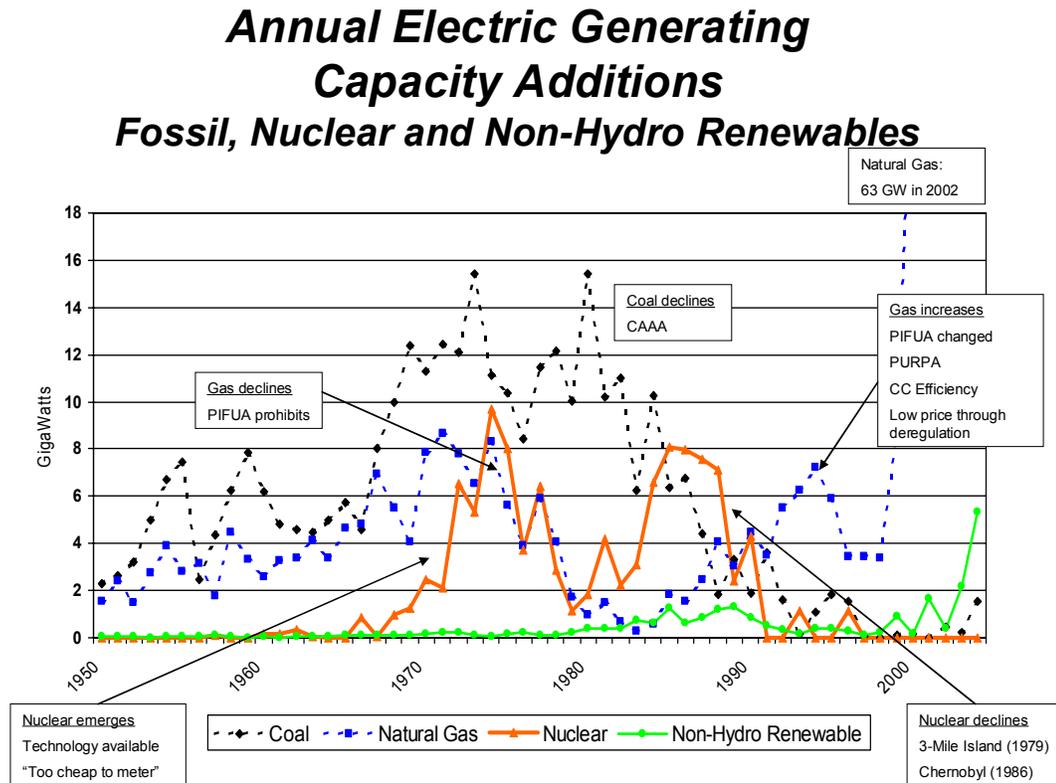


Figure 7. U.S. Historical Electric Capacity Additions by Fuel Type

For example, new U.S. capacity was dominated by nuclear and coal power plant additions in the 1970s and 1980s. When nuclear additions fell in the mid-1980s with high interest rates, increasing cost of regulation, and public safety concerns, they fell across the country – no new plants have been ordered anywhere in the United States since that time. Similarly, in the late 1970s, natural gas plant additions came to an abrupt halt across the country with the approval of the Power and Industrial Fuel Use Act (PIFUA), which prohibited new large industrial and electric boilers from using natural gas as a fuel. Then, with the rescission of the PIFUA provisions in the mid-1980s and the approval of the Public Utility Regulatory Policy Act (PURPA) legislation, independent power producers led a switch back to natural gas and almost entirely away from coal plant additions in the 1990s.

Utility deregulation, low gas prices, and combined-cycle technology advances contributed to the change in the late 1990s with a flurry of capacity expansion centered around natural gas, which exceeded by several times the historical annual additions of any single fuel's generation capacity. This flurry ended just as suddenly in 2004 with the unexpected increase in natural gas prices. And, today, another shift is occurring with more than 120 coal plants planned across the country (DOE 2005).

These movements might suggest that regional differences are no longer as strong a driver in the decisions regarding what form of conventional generation should be installed to meet growing electric demand. However, there are counter examples. For example, no new coal plants are planned in California due to environmental restrictions imposed by state planners, and California is moving to restrict the import of power from coal plants located in other states. Whether this becomes a national trend remains to be seen. This type of uncertainty argues for a stochastic SEDS-type approach.

While fossil-fueled power plant additions may follow national trends, renewable-generation capacity additions have definitely exhibited regional variations. While gas plant additions dominated the first five years of the 21<sup>st</sup> century, there were small, but growing additions of wind capacity in selected areas of the country with favorable wind resources, transmission access, and state-level mandates and incentives. These cannot be captured directly in a single-region national model – this is where SEDS and WinDS join forces.

With only a single region, SEDS cannot explicitly capture wind-transmission issues. However, the authors hope to develop a supply curve from WinDS model runs to explicitly capture regional wind transmission and intermittency variation in SEDS. In this way, the coarse-spatial-resolution SEDS model will be able to model the heterogeneity of wind with only a single region, saving run time and allowing hundreds of possible futures to be examined in a single Monte Carlo run.

## Modeling with fine spatial resolution

Once the use of a fine spatial resolution has been decided, questions still remain regarding the number of regions required. How should their boundaries be formed? Should there be different regions for different purposes? Should there be multiple levels of regions?

The question the model is trying to address is a strong determinant of the number of regions a model should have. For example, a model designed to evaluate the transmission requirements for wind energy will need to have small enough regions to estimate the distance to either existing transmission lines or load centers.

Once the general level of spatial resolution is decided, the specific boundaries of individual regions must be determined. Probably the most important determinant of the individual boundaries is the availability of data. Continuing with the wind example, there are several necessary data elements. These include wind resources and loads within each region for which transmission generally must be provided. Wind resource data is available as finely as grid squares of no more than 200 meters on a side. On the other hand, electric load data are available only at the utility service territory or power control area level. In spite of this limitation, it is possible to use smaller regions through geospatial statistical techniques (Johannesson 2005). For example, electric loads could be approximated at the county level.

It is also desirable to have some representation of utility service territories, because this is the regional level at which investment decisions in the different generation technologies (e.g., wind) are made. Utility service territories are largely coincident with the power control areas modeled in WinDS. However, utility service territory boundaries usually do not coincide with county boundaries, often cross state boundaries, and can even include islands. Fortunately, in many states, the counties are small enough that their aggregation into wind supply and electric demand regions can be made to conform reasonably well to utility service territories.

An additional advantage in this wind example, with spatial resolution built up from the county level, is that the county level resolution is consistent with aggregation back up to the state level. State-level regions are useful for imposing state-level regulation and taxes. For example, many states have renewable portfolio standards (RPS) that require a certain amount of electric generation or capacity from renewable energy. These can best be imposed if the wind model has states broken out spatially.

Another consideration in choosing the wind supply and electric demand regions in our example is ensuring that regions with significant wind resources and significant electric demands are well-defined. For example, the counties containing a major metropolitan area might be separated into their own region. Similarly, counties containing significant wind resources might be formed into a single region, if it is not too large an area. This separation of wind regions and loads ensures that the transmission distance estimates made in the model are more accurate.

A possible disadvantage in this wind example of spatial resolution down to the county level is that the number of variables and constraints in a national model would probably be overwhelming because there are more than 3,000 counties in the United States. For example, Texas has 254 counties. These small counties could be aggregated back up to regions that would still be relatively small (i.e., allow a relatively accurate approximation of the transmission distance required for wind), fit within a power control area, and are all within the same state.

Our wind model example illustrates that the creation of regions and their boundaries is more of an art than an exact science. However, it is an art where certain considerations such as data availability, separation of drivers, political constraints, etc. are important.

## **Modeling with coarse spatial resolution**

Several reasons for using a coarse spatial resolution were presented above – computer run time, analyst resources, data availability, accuracy of results at the finest spatial resolution level, and long time horizons. If these concerns dominate, and a model with coarse spatial resolution is desired, there are several techniques one can use to better approximate reality using a coarse spatial resolution.

The typical approach is to use an average value for any parameter that varies across a region. For example, the price of electricity to a building owner at a given point in time might be represented by a single average value across a state. This would be fine if all the relationships among market drivers were linear. But this is almost never the case in reality, or even within a model. For example, even a linear programming optimization model will decide on whether to purchase electricity from the electric grid or from a photovoltaic system, based implicitly on a highly nonlinear, “winner-take-all” scheme. If different electricity rate schedules and prices were modeled, there might be opportunities for photovoltaics (PV) to capture a portion of the market where rates are particularly high (e.g., congested areas), net metering is permitted, and/or a distribution feeder is overloaded or must be extended.

There are alternatives to average values that come closer to capturing actual variations without requiring further spatial resolution. One commonly used approach in energy market models is the use of a market-share algorithm that spreads market share out in such a way that the least-expensive energy source receives the bulk of the market share, but not all of it. One of the simpler and more commonly used market-share algorithms that spreads out the shares is the logit. The EIA’s NEMS model uses logit market share algorithms to capture diversity in transportation, electric, and other market sectors within its regions. The logit is stated in its most fundamental form as:

$$MS_i = \exp(U_i) / \sum_j \exp(U_j)$$

Where

MS<sub>i</sub> is the market share of technology i,

$U_i$  is the utility value of technology  $i$

The utility  $U_i$  can be expressed as the sum of a value function  $V_i$  and a random disturbance  $r_i$  that captures both variability in market conditions and investor preferences, and uncertainty in future values of  $V_i$  or the independent variables on which  $V_i$  depends.

$$U_i = V_i + r_i$$

Using price,  $p_i$ , as the only independent variable, and by setting  $V_i = \ln(p_i)$ , the logit can be reduced to:

$$MS_i = p_i^{-\alpha} / \sum_j p_j^{-\alpha}$$

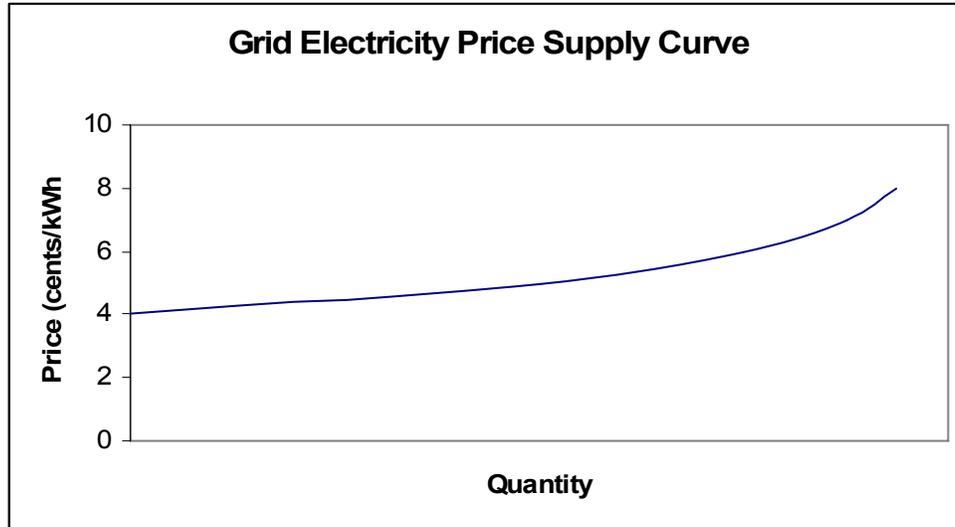
It can be shown that this single-nominal logit is equivalent to assuming that the market share of energy supplier  $i$  is the same as the probability that the price of energy from supplier  $i$  is less than the price from all other suppliers. It can also be shown that, in this formulation, the random component  $r_i$  in the utility of each competing energy supplier is assumed to be independent and identically distributed according to a type I extreme value distribution in standard form (Cramer 1991), i.e., with mean 0 and standard deviation of 1.

This assumption of identical distributions in the random component of the utility provided by each competitor is one of the basic limitations of the logit market share calculation. It is particularly egregious when various generation supplies whose prices may be relatively homogenous within a region (e.g., natural gas) compete with other supplies that may be very heterogeneous in their prices within a region (e.g., wind energy that varies from site-to-site with resource quality, road access, transmission access, terrain slope, etc.). One would prefer a market share algorithm that allows the random component to reflect these different distributions in the independent variables that determine the value of a supply. There are other market-share algorithms that distribute the shares based on relative prices, but the author is not aware of any closed-form algorithms that allow one to easily capture different distributions in supply prices.

Another approach often used to capture variability in a parameter within a region is a supply curve. The WinDS model described above uses supply curves to capture transmission cost variability within each of its regions, even though it has relatively fine spatial resolution with 358 regions. A supply curve allows one to replace an average price with a range of prices that increase as more energy is supplied by the source.

Continuing our grid electricity and photovoltaic example from above, a supply curve like that shown in **Figure 8** could be used to capture different grid electricity prices within a state. The first point on the curve would be the lowest price paid by any building owners and the quantity of electricity consumed in the time period by that set of building owners. Similarly, the last point on the curve would be the highest price paid by any building owners and the quantity of electricity consumed in the time period by that set of owners.

One obvious problem with using such a grid-electricity supply curve to allocate market share is that PV may not be applicable to all building types (e.g., those with limited roofspace), yet will compete with the highest cost points on the grid-electricity supply curve, regardless of the buildings that make up that point.



**Figure 8. Representative Supply Curve**

In this example, a second supply curve might also be used to capture the cost of photovoltaics on buildings. If there are significant variations in annual insolation between different parts of a region, a supply curve could be built with the first point representing the cost of electricity from PV with the best resource quality within the region. The quantity associated with this cost could be the PV generation possible from all the rooftops in that part of the region. Or, if excess power cannot be sold back to the grid and no electricity storage is available, the quantity might be the minimum of the possible PV rooftop generation or the electricity consumed by those buildings during daylight hours.

While the above logit and supply curve approaches capture variation within a region, they do not capture correlation between the variables or different cohorts. If the high-insolation PV areas of the previous example are the same areas as the high-cost grid electricity, then real-world PV might possibly capture some market share that the supply curve approach might miss entirely – and the logit might underestimate. Obviously, a model with finer spatial resolution could largely resolve this particular problem.

One viable alternative might be to construct a combined supply curve, in which the cost is the difference between the cost of grid power and the cost of PV generation. In this case, the quantities associated with the first points on the supply curve would be the useful PV generation possible from buildings that are in an area with high grid power costs and low PV power costs (i.e., high insolation). A slightly more complicated approach might be to expand from the single independent variable of a supply curve to multiple independent variables in a “reduced form” curve that captures the results of a more detailed model.

Another alternative is to use Monte Carlo techniques to explicitly capture correlations between market attributes (e.g., grid electricity price and insolation) through the probability distributions on input parameters. The SEDS model described above uses such correlations.

Cohort analysis also can be used to capture correlations. In cohort analysis, markets can be grouped by their similarity in attributes rather than being specified regionally, e.g., all the regional markets with high grid electricity prices and high insolation might be grouped into one cohort. The HYTRANS model developed at Oak Ridge National Laboratory (Greene 2005) is an example of a model with market cohorts that have the same city size.

In some cases, results from models with coarse spatial resolution have been disaggregated (post-processed) to results for finer geographic regions. The validity of such an approach depends on the question being asked, the correlation between market drivers within the initial coarse region, and other factors.

## **Conclusions**

There is no single “correct” level of spatial resolution that is appropriate for all models, and there may not even be a level that is "correct" for a given model. The question the model is attempting to answer is the principal determinant of the necessary level of spatial resolution. Finer spatial resolution requires additional resources – run time, data, analyst effort, modeling detail, etc. It is possible to improve the accuracy of a model without greater spatial resolution by using supply curves, logit market share algorithms, and other techniques. However, supply curves may have to be developed from another model with finer spatial resolution.

Fine spatial resolution can significantly increase the accuracy of some energy models, especially where the options to transport energy are limited or expensive. For example, our analysis shows that ignoring transmission costs associated with wind can result in significant overestimates of the market potential of wind, i.e., fine spatial resolution is required. The recent environmentally driven prohibitions against new coal-fired generation in California introduce another rationale for fine spatial resolution. However, should such environmentally driven restrictions become national in character, then for certain questions, the modeling of conventional energy forms might be reasonably accomplished with more coarse spatial resolution models.

## References

Cramer, J.S. 1991. *The Logit Model: an introduction for economists*, Edward Arnold, London, England.

Department of Energy (DOE). 2005. "Tracking New Coal-Fired Power Plants: Coal's Resurgence in Electric Power Generation," National Energy Technology Laboratory, July 2005,  
[www.fe.doe.gov/programs/powersystems/publications/General\\_Interest/New\\_Coal\\_Plants\\_072505.pdf](http://www.fe.doe.gov/programs/powersystems/publications/General_Interest/New_Coal_Plants_072505.pdf)

Energy Information Administration (EIA). 2005. *Annual Energy Outlook 2005, Assumptions*, DOE/EIA-0383, U.S. Department of Energy, February 2005.

EIA. 2006. [www.eia.doe.gov/cneaf/electricity/epm/table5\\_6\\_a.html](http://www.eia.doe.gov/cneaf/electricity/epm/table5_6_a.html)

Greene, D. 2005, *HyTrans Model Development*, Oak Ridge National Laboratory, Oak Ridge, Tennessee. [http://www.hydrogen.energy.gov/pdfs/progress05/iii\\_3\\_greene.pdf](http://www.hydrogen.energy.gov/pdfs/progress05/iii_3_greene.pdf)

Johanesson, G., J. Stewart, and C. Barr. 2005. *Spatial Statistical Procedures to Validate Input Data in Energy Models*, draft, Lawrence Livermore National Laboratory, Livermore, Calif.

Short, W. 2006. *Wind Deployment System Model: Qualitative Description*, National Renewable Energy Laboratory, Golden, Colorado  
<http://www.nrel.gov/analysis/winds/qualitative.html>

Short, W., M. Leifman, and T. Ferguson. 2006, *Modeling of Uncertainties in Major Market Drivers in U.S. Electricity Markets*, 26<sup>th</sup> USAEE/IAEE North American Conference proceedings, September 2006. Ann Arbor, Michigan.

# REPORT DOCUMENTATION PAGE

*Form Approved*  
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> February 2007		<b>2. REPORT TYPE</b> Technical Report		<b>3. DATES COVERED (From - To)</b>		
<b>4. TITLE AND SUBTITLE</b> Regions in Energy Market Models			<b>5a. CONTRACT NUMBER</b> DE-AC36-99-GO10337			
			<b>5b. GRANT NUMBER</b>			
			<b>5c. PROGRAM ELEMENT NUMBER</b>			
<b>6. AUTHOR(S)</b> Walter Short			<b>5d. PROJECT NUMBER</b> NREL/TP-640-40506			
			<b>5e. TASK NUMBER</b> ASA6.2023			
			<b>5f. WORK UNIT NUMBER</b>			
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> National Renewable Energy Laboratory 1617 Cole Blvd. Golden, CO 80401-3393				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b> NREL/TP-640-40506		
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> NREL		
				<b>11. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>		
<b>12. DISTRIBUTION AVAILABILITY STATEMENT</b> National Technical Information Service U.S. Department of Commerce 5285 Port Royal Road Springfield, VA 22161						
<b>13. SUPPLEMENTARY NOTES</b>						
<b>14. ABSTRACT (Maximum 200 Words)</b> This report explores the different options for spatial resolution of an energy market model – and the advantages and disadvantages of models with fine spatial resolution. It examines different options for capturing spatial variations, considers the tradeoffs between them, and presents a few examples from one particular model that has been run at different levels of spatial resolution.						
<b>15. SUBJECT TERMS</b> National Energy Modeling System; NEMS; Energy Information Administration; EIA; spatial resolution; energy market models; Wind Deployment System model; WinDS; Stochastic Energy Deployment System model; SEDS; Walter Short						
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b> UL	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>	
<b>a. REPORT</b> Unclassified	<b>b. ABSTRACT</b> Unclassified	<b>c. THIS PAGE</b> Unclassified			<b>19b. TELEPHONE NUMBER (Include area code)</b>	

**Standard Form 298** (Rev. 8/98)  
Prescribed by ANSI Std. Z39.18