# Predicting Avoided Emissions from Policies that Encourage Energy Efficiency and Clean Power

**Prepared by:**
**Geoff Keith and Bruce Biewald**
**Synapse Energy Economics**
**22 Pearl Street, Cambridge, MA 02139**
**www.synapse-energy.com**
**617-661-3248**

**Prepared for:**
**The Ozone Transport Commission**

**June 24, 2002**

# Table of Contents

## List of Acronyms

| | |
|---|---|
| ACEEE | American Council for an Energy Efficient Economy |
| CCAP | Center for Clean Air Policy |
| $CO_2$ | Carbon Dioxide |
| DOE | U.S. Department of Energy |
| EIA | U.S. Energy Information Administration |
| EPS | Emissions Performance Standard |
| ERC | Emission Reduction Credit |
| EPA | U.S. Environmental Protection Agency |
| Hg | mercury |
| ICLEI | International Council of Local Environmental Initiatives |
| IPM | Integrated Planning Model |
| ISO | Independent System Operator |
| ISO NE | New England Independent System Operator |
| kWh | kilowatt-hour |
| mmBtu | million British thermal units |
| MWh | Megawatt-hour |
| NEMS | National Energy Modeling System |
| NERC | North American Electricity Reliability Council |
| $NO_x$ | Oxides of nitrogen |
| NREL | National Renewable Energy Laboratory |
| NY ISO | New York Independent System Operator |
| OTC | Ozone Transport Commission |
| OTR | Ozone Transport Region |
| PJM | Pensylvania/New Jersey/Maryland control area |
| $SO_2$ | Sulfur dioxide |
| STAPPA | State and Territorial Air Pollution Program Administrators |

# Executive Summary

Efforts to integrate energy and environmental regulation more effectively have highlighted the need to understand the emission impacts of energy programs such as subsidies for energy efficiency and renewable generation. However, because the interconnected regional electricity systems in the U.S. operate in a highly complex way, the future emissions impacts of these kinds of programs are not easy to predict.

This report explores what is involved in predicting avoided emissions from energy efficiency and clean generation and reviews four projects underway that focus on one or more aspects of this task.

## *The Task of Predicting Avoided Emissions*

The task of predicting avoided emissions from energy efficiency and clean generation can be viewed as two distinct subtasks:

- describing quantitatively the energy saved or clean energy generated, and
- predicting how the regional electricity system(s) will react to the energy saved or generated.

These two subtasks require different analytic methods and tools. Regarding the first subtask, questions of where and how much energy is saved by energy efficiency and renewable programs are not difficult to answer. Regulators are usually interested in the impacts of a program implemented in a specific region, and the magnitude of the program's effects can be estimated. It can be more difficult to predict what types of renewable generation will result from a program subsidizing renewable energy, and thus assessing the average emission rate of the renewable generation can be difficult. Analysts must make an informed prediction of the emissions associated with renewable generation. In some cases, the rules of the program will provide guidance, as in the case of a Renewable Portfolio Standard for which certain technologies are not eligible. In all cases it is probably wise to explore the implications of different assumptions.

The most complex task in characterizing an energy efficiency or renewable program for analysis is predicting *when* the program will reduce load or generate electricity. To do this, the analyst must obtain data on the "load profile" of the technologies installed, that is, the pattern of operation in each time period of the year. Data on the load profile of renewable generators are available from the National Renewable Energy Laboratory. Publicly available data on the performance of energy efficiency technologies, however, are harder to find. More work needs to be done to identify and review sources of these data.

Only one of the projects we review – the Environmental Protection Agency's (EPA) project – will collect and disseminate information on load profiles. EPA currently plans to provide load profile information on a number of energy efficiency and renewable technologies along with displaced emission rates to allow users to assess avoided emissions from programs subsidizing these technologies.

The second subtask listed above is to predict how the regional electricity system(s) will react to the reduced load or new generation from the policy being assessed. For this task, it is important to distinguish between the short term and the long term, because projecting emission reductions over the short term requires a different analytic approach than projecting reductions over the long term. Here, the "short term" is defined as the period during which the capacity mix will not change due to the program being analyzed. Over this time horizon, the analytic task is to predict how the existing regional electricity system will react to reduced energy use or additional clean generation. Because changes in generation and load in one control area can affect generation in neighboring control areas, it is important to assess net regional changes in generation.

We define the "long term" as the period when old generating units can be retired and new ones added. This will typically be the period starting three or four years from the present. For long-term assessments, the task is to predict how reduced load or additional generation will affect (a) existing resources and (b) plant additions and retirements. The main challenge here is in predicting which generating units will be retired and when, and what kinds of new units will be built and when.

For predicting avoided emissions over the short-term, a system dispatch model is a necessity. A dispatch model is important because regional electricity systems operate in complex, integrated ways. A credible prediction of how a particular system will respond to reduced load or increased generation must be based on a detailed simulation of the system. Dispatch models do this by simulating unit dispatch to meet hourly loads. The most detailed dispatch models simulate dispatch on a chronological basis. That is, they dispatch generating units to meet load in each hour of the year in chronological order. Input data to these models includes highly detailed data on generating units and regional loads. These models cannot, however, assess large geographic regions (such as the entire U.S.) with such detailed data, and they are not designed to predict (endogenously) unit additions and retirements.

There is more uncertainty around assessments of long-term avoided emissions from energy efficiency or renewable generation than short term. This is because long-term assessments are highly sensitive to predictions of unit additions and retirements, and these predictions are uncertain. There are essentially two approaches to predicting unit additions and retirements over the long term. One approach is to use a forecasting model, a model designed to predict plant additions and retirements. These models are much broader in scope than the dispatch models discussed above. They forecast the evolution of all the major energy sectors by simulating the interaction of these sectors in areas such as fuel prices and supply and demand. Forecasting models generally operate iteratively, converging on an optimal solution based on the input assumptions. The second approach is to make a prediction of unit additions and retirements based on a close study of the key indicators within a region, such as the wholesale market, the regulatory climate and key economic and financial indicators. This approach involves informed judgments about the dynamics of capital investments in the region and how the programs being modeled are likely to affect these investments.

The use of a forecasting model to predict additions and retirements is appealing in that it provides an automated result. That is, after the algorithms are set and input data entered, the model operates objectively. Some people view a prediction based on this as more credible than a prediction made by informed judgment. However, the optimization routines used in forecasting models may not effectively account for the full range of dynamics that affect plant additions and retirements over time, and the incorporation of forecasting into the modeling effort requires aggregation of data in the dispatch simulation process that can compromise near-term predictions.

The selection of an approach to long-term forecasting is the most significant aspect of any methodology to predict emission reductions from energy efficiency and clean generation. A key conclusion of this report is that, as the Ozone Transport Commission (OTC) moves forward in its efforts to develop a credible methodology for assessing emission reductions from these programs, the group should carefully consider what level of resources it is willing to allocate to predicting power plant additions and retirements in the Northeast and what approach to making these predictions it prefers.

Regardless of which approach one takes to predicting additions and retirements, this is an inherently uncertain endeavor. For this reason, many analysts adopt a scenario-based approach when forecasting. With this approach, different assumptions representing a range of possible scenarios are explored to get a sense of the range of possible outcomes and the key sensitivities.

### *Specific Projects*

The specific projects we review in this report include: ISO New England's Marginal Emissions Analysis, the Center for Clean Air Policy's (CCAP) work for the Great Lakes Protection Fund, EPA's Average Displaced Emission Rate Work, and the STAPPA/ICLEI Strategic Planning Software.[1] The models we review include: the Integrated Planning Model (IPM®), the National Energy Modeling System (NEMS), Energy 2020, PROSYM and a dispatch model developed by CCAP. Both methods and models are reviewed with an eye to the OTC's stated goals. These goals include developing a methodology for predicting avoided emissions from energy efficiency and clean energy programs with the following characteristics. The system must be able to:

- predict avoided emissions over both the short term and the long term;
- predict reductions of $NO_x$, $SO_2$, mercury and $CO_2$;
- assess programs implemented in any combination of the three northeastern power control areas; and
- assess emission reductions from energy efficiency, clean energy, emissions performance standards (EPSs) and multi-pollutant regulations.

It is important to note that none of the projects we review shares all of these goals with the OTC. The ISO New England and CCAP projects, for example, do not seek to develop a methodology for predicting future emission reductions from energy efficiency and clean generation. The ISO New England work calculates retrospective marginal emission rates, and the CCAP work will rank energy efficiency and clean energy technologies in relative terms with regard to their effectiveness in reducing emissions.

[1] STAPPA is the State and Territorial Air Pollution Program Administrators, and ICLEI is the International Council of Local Environmental Initiatives. A companion paper, *The OTC's Emission Reduction Workbook: Description and User's Manual*, discusses a fifth project, the development of an avoided emission calculation methodology for the OTC.

EPA's work and the STAPP/ICLEI work have broader goals. Both of these projects seek to develop tools with which users can perform their own calculations and both seek to predict avoided emissions over a long period. Each of these projects uses a single computer model to simulate system dispatch and predict plant retirements and additions over time. EPA's work is using the IPM® model, and the STAPPA/ICLEI project is using NEMS.

Below is a summary of the key aspects of each project we review, with a focus on project strengths and limitations vis-à-vis the OTC's current goals.

ISO New England's *Marginal Emissions Analysis* does not meet the goals established by the OTC for avoided emission analysis for the following reasons.

- It focuses only on New England.
- It does not assess mercury emissions.
- It does not predict future marginal emission rates, and
- It and does not attempt to assess the emissions impacts of specific policies.

However, the ISO's calculation of retrospective marginal and average emission rates using a system dispatch model is methodologically sound. Further, because the ISO calculates system emission rates retrospectively, there is less uncertainty around its numbers than there is around predictions of future system emission rates.

The methodology employed in the CCAP project would be not be effective in meeting the OTC's current goals for the following reasons.

- Its geographic focus is New York (and potentially areas of the Midwest).
- It does not predict future marginal emission rates, and
- It does not attempt to quantify the emissions impacts of specific policies.

The CCAP methodology appears to be sound in the context of the projects goals. Moreover, this project is one of the first attempts to simulate the operation of DG within regional power systems.

The methodology used by EPA in its Average Displaced Emission Rate (ADER) work comes closer to meeting the OTC's goals than the two methodologies discussed above. Ways in which EPA's ADER methodology is well suited to meet the OTC's goals include the following.

- The work will result in a tool with which users can predict the emissions impacts of energy efficiency and clean generation.
- It will allow users to assess energy efficiency and clean generation using 11 different time blocks each year. This will allow users with high-resolution load profile data to "describe" programs with ADER parameters with considerable precision.
- ADER parameters are being developed for $NO_x$, $SO_2$, $CO_2$ and mercury.
- The model runs used to develop the ADER parameters take into account transmission constraints and interregional effects.

- If EPA releases load profiles for energy efficiency programs and renewable technologies, consistent with the 11 ADER time periods, this will be a significant strength of the work. If they do not, this will be a limitation, as the ADER parameters will be less useful without load profile data.

The following are ways in which the EPA's ADER work is not well suited to meet the OTC's goals.

- The ADER method only provides one displaced emission rate per pollutant for the entire Northeast region. Differences between control areas are not captured.
- IPM® simulates unit dispatch using aggregated data on generating units and system loads. Ideally, modeling focused on a particular region should dispatch units chronologically, based unit-specific information and hourly load data.
- The ADER approach provides displaced emission rates for selected future years rather than for each year of a given study period, and the first year for which ADER parameters will be available is 2005.
- The ADER parameters will not be useful in assessing emission reductions from multi-pollutant regulations and EPSs.[2]

In addition, a key aspect of the EPA methodology is the use of the IPM® model to predict plant additions and retirements over the long term. When the ADER parameters are released, it will be important to evaluate in detail the predicted additions and retirements that underlie them.

Like EPA's work, the methodology employed in the STAPPA/ICLEI project has strengths and weaknesses vis-à-vis the OTC's goals for assessing emission reductions from energy efficiency and clean energy programs. Ways in which the STAPPA/ICLEI methodology is well suited to meet the OTC's goals include the following.

- It develops separate displaced rates for each of the three northeastern control areas.
- It models unit dispatch with data specific to existing generating units, and
- It will result in a tool with which users can predict the emissions impacts of energy efficiency and clean generation.[3]

The following are ways in which the STAPPA/ICLEI methodology is not well suited to meet the OTC's goals.

- It only provides one displaced emission rate per year (i.e., it does not reflect the difference in marginal emission rates during different seasons or times of day).
- Annual load data has been aggregated into 11 load levels for dispatch modeling in NEMS and dispatch is not chronological.
- It does not account for changes in emissions in neighboring regions (due to load reductions in a given region), and

[2] In separate projects, EPA has used the IPM® model to assess the impacts of multi-pollutant bills.
[3] The STAPPA/ICEI software is not reviewed in this report. We review only the development of avoided emission rates for that software.

- It does not assess reductions in mercury emissions.

In addition, a key aspect of the STAPPA/ICLEI methodology is the use of the NEMS model to predict plant additions and retirements over the long term. When the STAPPA/ICLEI software tool is released, it will be important to evaluate in detail the predicted additions and retirements that underlie them.

# 1. Introduction

Efforts to integrate energy and environmental regulation more effectively have highlighted the need to understand the emission impacts of energy programs such as funding for energy efficiency and renewable generation. Because the interconnected regional electricity systems in the U.S. operate in a highly complex way, the future emissions impacts of these kinds of programs are not easy to predict.

One challenging aspect of this task is that projecting emission reductions over the short term requires a different analytic approach than projecting reductions over the long term. Here, the "short term" is defined as the period during which few, if any, generating assets will added or retired. Over this time horizon, the analytic task is to predict how the existing regional electricity system will react to reduced energy use or additional clean generation. We define the "long term" as the period when old generating units can be retired and new ones added – this is the period starting three to four years from the present. Here, the task is to predict how reduced load or additional generation will affect (a) existing resources and (b) plant additions and retirements. The major challenge is in predicting which generating units will be retired and when, and what kinds of new units will be built and when.

A second analytic challenge lies in the question of geographic scale. The most accurate and credible results are likely to come from a detailed analysis of a small region. However, this type of analysis risks missing important interactions and secondary effects in neighboring regions. Broadening the scope too much, however, often means sacrificing detail in the primary region of interest.

This report explores the challenges of predicting future emission reductions from energy efficiency and clean generation, and it reviews several existing and emerging methods for doing so. The various computer models analysts are using to assist in this task are an important focus of the report. Section 2 explores the task of predicting emission reductions and the models being used. Later sections explore specific efforts to calculate avoided emission rates and/or predict emission reductions from specific programs. Models and methods are reviewed here with an eye to the Ozone Transport Commission's (OTC) stated goals. These goals include the development of an analytic tool that can credibly and cost effectively predict emission reductions from a wide range of program types in the three northeastern power pools.

# 2. Calculating Reduced Emissions from Energy Efficiency and Clean Generation

To date, a number of different methods have been used to calculate emission reductions from energy efficiency and clean generation, and efforts are ongoing to improve on these methods. The less complex of these methods have simply calculated or estimated a marginal or average emission rate representative of the regional generating system and multiplied saved energy or clean generation by this emission rate. More ambitious efforts go beyond this approach, using higher resolution data and electricity system models that account for the additional complexities of today's regional power pools. At the core of these more ambitious efforts are two key tasks, shown in Table 1, below.

**Table 1. Key Tasks in Evaluating Emission Reductions from Energy Efficiency and Clean Generation**

| | |
|---|---|
| Task 1 | Develop assumptions about the energy saved or clean energy generated. This includes assumptions about:<br>(a) where the energy is saved or generated,<br>(b) how much energy is saved or generated,<br>(c) when the energy is saved or generated (i.e., how it is distributed across seasons and days), and<br>(d) the emissions characteristics of the energy, if it is generation. |
| Task 2 | Develop assumptions about how the regional electricity system will react to the energy saved or generated. For short-term analyses, the critical task is discerning which generating unit or units will be affected by the savings or generation during each time period. For long-term analyses the key task is to predict what types of generating units will be added and retired. |

These two tasks require different analytic methods and different computer models. The following two sections explore each of these two tasks in turn.

## *2.1 Predicting the Operation of Energy Efficiency or Clean Generation*

The first question listed under Task 1 above – where is the energy program implemented – is relatively simple to answer. Policy makers usually have state, regional or national implementation in mind when considering a program to support energy efficiency or clean generation. Modeling the implementation of a program at each of these levels is possible, provided the appropriate model is chosen. One must be sure, however, to choose a model that can simulate the effects of the policy at the appropriate scale. For example, a state energy policy may have regional emissions impacts that are important to consider. A regional policy may have interregional impacts. Compromises often have to be made between geographic scope and the level of detail in the model inputs and results.

The second two questions – how much energy is saved or generated and when – are more difficult to answer than the question of where. Each energy policy under consideration will generate energy or reduce load at different rates during different hours of the day and different seasons of the year. That is, each policy will have a unique "load profile." Data on program load profile are an *input* to the modeling tool used to predict the emissions effects of the program. These data will be from a different source and possibly developed using a different modeling tool.[4]

The most detailed data on program load profile is hourly data. An hourly data set consists of 8,760 hours of output (or energy savings) data for the technology or program. To assess emission reductions from most program types, hourly data is not necessary; data for several annual time periods is sufficient. However, it is important that the data for these periods is built up from high-quality hourly data to ensure that it accurately

[4] For example, the load profiles EPA is using in developing its ADER parameters are from the Department of Energy's DOE-2 model of building performance.

reflects the performance of the technology or program. Ultimately, predictions about emission reductions will only be as accurate as the load profile data used to describe the program.

Data on the generation profiles of renewable technologies are available in the public domain. One source of good data on renewables performance is the National Renewable Energy Laboratory (NREL). Publicly available information on the load profiles of energy efficiency technologies, however, is much harder to find. For this project, we sought publicly available data from the following sources:

- The U.S. Department of Energy (DOE),
- The Electric Power Research Institute (EPRI),
- The American Council for an Energy Efficient Economy (ACEEE), and
- Several private sector firms.

We found no detailed (i.e., hourly) data on the load profiles of energy efficiency technologies available to the public.[5] To support efforts to predict emission reductions from energy efficiency programs, more work needs to be done to locate and review publicly available sources of detailed load profiles for energy efficiency technologies and programs.

Finally, when assessing a program designed to incentivize clean generation, one must develop assumptions about the average emission rate of the resulting generation. Clean generating technologies like biomass, landfill gas and fuel cell generation have pollutant emissions, and these will often have to be taken into account when assessing net reductions. (One or all of these technologies are eligible to meet most of the Renewable Portfolio Standards (RPSs) in the country today and to receive funding from many subsidy funds.) For programs such as these, assumptions must be made about (a) the emission rates of these technologies and (b) the percentage of total program generation that these technologies will represent. It is unlikely that there will be a high level of certainly around these assumptions, and they can affect predictions of avoided emissions significantly. Thus, a sensitivity analysis approach may be best. That is, two assessments with bounding assumptions (high and low assumptions) about program emissions may be more useful in a policy-making context than one analysis based on a best guess. Often such sensitivity analyses can illuminate changes that could be made in policy design to make policies more effective.

### *2.2 Predicting Emission Reductions Over the Short Term*

As discussed above, predicting reduced emissions over the short term requires a different analytic approach than predicting over the long term. We first address approaches to predicting short-term events, and next we turn to predicting over longer time frames.

[5] One notable source of efficiency load profile data is Regional Economic Research (RER), a research and consulting company based in California. RER sells hourly load profile data on residential, commercial and industrial equipment for each U.S. state. RER develops these data using its SitePro model, a model based on analysis of many hourly load profile data sets. For more information see: www.rer.com.

To make an informed prediction about how a new efficiency program or generating unit will affect an *existing* regional electricity system, a system dispatch model is a necessity.[6] A dispatch model is important because regional electricity systems operate in highly complex, integrated ways, with fluctuations in demand and randomly occurring outages of equipment. A credible prediction of how a system will respond to reduced load or increased generation must be based on the ability to simulate these changes in a way that reasonably represents the manner in which the real power system is operated. Detailed dispatch models do this by simulating unit dispatch in a system to meet hourly loads on a chronological basis.

When using a dispatch model, detailed information is entered about the generating units within the region, as well as the regional transmission system and regional electricity loads. This information includes, for each generating unit: unit size, fuels used, emission rates, efficiency (typically in the form of a heat rate curve or heat rates for different load levels), operating limitations (such as start-up ramp times and minimum up and down times), operating costs and fuel costs. Operating costs include the costs of environmental controls and emission allowances needed to operate units affected by allowance programs. Detailed information on the regional transmission system is also entered, such as transmission constraints within each control area and interconnections between control areas and their capacities.

For models that simulate dispatch chronologically, hourly load data is entered for each control area being simulated. In other words, 8,760 system load levels are entered, one for each hour of the year. By simulating dispatch chronologically, these models capture the full implications of unit-specific operating constraints, maintenance outages and other time-sensitive events.[7] Unplanned unit outages are represented in chronological dispatch models probabilistically. That is, the models algorithms recognize and reflect the randomly occurring nature of forced outages. This randomness can be an important aspect of system operations.

Dispatch models can accommodate highly detailed data inputs, because they are not usually used to simulate dispatch across large geographic areas such as the entire U.S. This is an important trade-off in electric system modeling. When large areas are modeled, more aggregated data on loads and generators must be used as well as simplifying assumptions for dynamics like forced outages. When smaller areas are simulated (say, five control areas or fewer), detailed input data can be accommodated.

In a world of unlimited resources, one would simulate the near-term effects of each policy of interest with a regional dispatch model to determine near-term emissions impacts. This approach would ensure that the analysis captured the complex interactions

[6] These models, usually called dispatch models, have more recently been modified to represent bid-based electricity markets, and may be called "electricity market simulation models" and similar terms. We will use the term dispatch model in this report, since the key feature of interest is the manner in which generators are dispatched by the system operator, whether it is a traditional regulated utility control area operator, or a newer entity such as an "Independent System Operator" (ISO) or "Regional Transmission Organization" (RTO).

[7] For example, an idle generating unit may not be available at a specific time because of a "minimum down time" limitation – when the plant is taken off line, it cannot be started again for a certain number of hours. Another unit might not be available because it cannot be started quickly enough. These constraints on system operation can have significant emissions implications, and they can only be captured by chronological simulation.

of all transmission and generation resources in the region. However, if there is interest in many policies, this approach would be time consuming and expensive. This fact has led policy makers to seek a method for assessing the emission impacts of a range of programs without running a dispatch model for each scenario of interest. One approach to developing such a method is the use of a dispatch model to generate generalized emission factors that can then be applied to energy efficiency and clean power programs in a simpler setting. The challenge here is to ensure that the factors developed using the model capture enough detail in the regional system to provide credible results when applied in a more generalized way to load reductions or clean generation.[8]

Models commonly used for regional dispatch analysis include PROSYM, GE MAPPS ELFIN and PROMOD. PROSYM and GE MAPPS simulate dispatch chronologically, while ELFIN and PROMOD do not. These models can be used to assess long planning horizons – and they often are – however, they are not designed to predict plant additions or retirements over time. When a dispatch model is used to simulate a multi-year period, the user directs the model to add specific generating units (or unit types) in specific years or to maintain a certain capacity reserve margin by a adding a specific type of plant as needed. In contrast, when analysts want *the model* to predict the capacity mix in the future, they turn to a forecasting model. These models are discussed in the following section.[9]

## *2.3 Predicting Emission Reductions Over the Long Term*

Over the long term, decisions made by power plant owners and new plant developers will take into account many of the changes in the regional system that took place during the near term. Demand forecasts made in 2007, for example, will take into account many of the conservation and load management programs implemented in the period 2002 through 2006 as well as new generators installed in this period. Thus, at first, energy efficiency and new renewables will displace energy from existing resources, but over time, many of them will displace energy from a mix of existing resources and potential new resources – and they will affect plant retirement decisions. Therefore, the question of what kind of generating units will be added and retired is extremely important to predicting emission reductions over the long term. Unfortunately, predicting what kind of units will be added and retired – and when it will happen – is difficult.

Predicting plant additions and retirements is difficult, because it is not simply a question of costs. Many factors – regulatory, political, economic and financial – influence the decision to build a new unit or retire an existing one, and plant developers and owners do not always behave like the rational market participants assumed in economics books. In addition to basic economics (such as relative fuel prices and technology costs), some of the important factors affecting plant additions are as follows.

[8] Efforts to develop emission factors to be used to assess energy program impacts include: the Environmental Protection Agency's "ADER" project, discussed in Section 4; the software planning tool being developed by the State and Territorial Air Pollution Program Administrators (STAPPA), discussed in Section 6; and the OTC's Emission Reduction Workbook, discussed in the OTC report The OTC Emission Reduction Workbook..

[9] There are computer models that focus exclusively on the demand side, and forecast load growth (e.g., peak hour demand, annual energy requirements, load shapes). These are not what we mean here by the term "forecasting model." Rather, we use the term "forecasting model" to refer to models that develop a forecast *internally* of how the system capacity mix will change over time.

- *The prevailing attitudes of capital markets toward the power generation sector.* The current state of affairs regarding new plant financing illustrates well the fickle attitudes of these markets. Lenders are currently retreating from the power generation sector to a degree that does not appear warranted by market fundamentals.
- *Energy policies designed to support the construction of certain unit types and discourage the construction of others*. Subsidies, portfolio standards, tax incentives and other policies can cause power projects with marginal economics to move forward. Predicting the effectiveness and longevity of these programs is difficult.
- *Irrational behavior in the project development process*. For many of the players involved in a power project, compensation is directly linked to the project's success. If the project is scrapped, people like project developers and lawyers make much less money than if the project succeeds. Thus, there are strong incentives for these players to push a project forward even if the economic outlook for the project becomes weaker.[10]

Decisions regarding plant retirements are in many ways harder to predict than decisions about new units. Many different costs and benefits factor into unit retirement decisions, and these costs and benefits are very difficult to quantify. Key aspects of plant retirement decisions include the following.

- *The true operating costs of the unit*. As a generating unit ages, assumptions about operating costs based on unit type become less reliable, because different companies maintain and retrofit units differently. The portion of operating costs that are fixed versus variable is also important to the retirement decision and difficult to know without discussions with plant operators.
- *The hedging value of avoiding "retired" status*. The possibility of future electricity shortages and exorbitant wholesale prices makes it very attractive *not to* fully retire older generators but to minimize their fixed costs and keep them available, such as by "mothballing" them. Once a unit has been officially retired, the regulatory process of bringing it back on line is arduous. Such units must apply for new permits and are subject to the environmental standards applicable to new units.
- *The capacity value of the unit*. In most regions, generating units receive payments for both energy and "capacity" – for being available to operate. The value of being operable (as opposed to being mothballed or retired) is highly dependent on rules established by the local ISO.
- *Case-specific negotiations*. In many cases, old units become the subject of negotiations between their owners and environmental regulators. Often regulators push for retirement of older units in the context of permitting new units. Companies tend to prefer adopting operating limitations at older units rather than retiring them, for the reasons stated

[10] While this is rational behavior at the individual level, it can lead to irrational outcomes at the market level.

in the previous bullet. The outcome of these negotiations is difficult to predict.

In general, the factors cited above tend to favor keeping old generating units operational rather than retiring them, and indeed this trend has been pronounced in the industry during the past decade. Many companies have engaged in "life extension" programs at fossil-fired units, and in the Northeast and California, companies who chose to keep older units available have been rewarded handsomely during periods of high wholesale prices.[11]

Thus, predicting future unit additions and retirements is an inherently uncertain endeavor, however, it is an endeavor crucial to predicting long-term emission reductions from energy efficiency and clean generation. There are essentially two approaches to predicting unit additions and retirements. One approach is to use a forecasting model designed to make such predictions, and the other is for the policy maker or analyst to make predictions for a specific region, based on key indicators in that region.

Energy forecasting models are much broader in scope that the dispatch models described above. The major forecasting models have modules that focus on each energy sector (e.g., transportation, industrial fuel use, residential fuel use, etc.). They forecast the evolution of these energy sectors by simulating the interaction of the sectors in areas such as fuel prices and supply and demand in each sector. For example, as new gas-fired power plants are added in the electric sector, the impact of these units' fuel use is factored into natural gas supply and pricing across all energy sectors. If gas prices are predicted to rise, the viability of additional new gas-fired power plants is reduced. Forecasting models generally use a mathematical optimization technique.[12] Some operate iteratively, converging on an optimal solution (or at least a stable one) after a number of runs, while others use techniques such as linear programming to find system expansion plans that best satisfy an objective function (e.g., least total cost) subject to constraints.

The electricity modules of multi-sector energy models perform a dispatch function, but they usually do so at a more aggregated scale than dispatch models. For example, forecasting models often dispatch generating unit types rather than specific units. That is, data is entered into the model on generating unit types rather than on the specific generating units in a region. Units are grouped together on the basis of their fuel type, age, efficiency, and other factors. The total capacity of different unit types in each region is stored in the model, and this capacity is dispatched to meet load.

The load information used by forecasting models is also aggregated and simplified. That is, rather than representing hourly chronological loads, the loads are grouped together into seasons or time periods, and then step functions or a load duration curves for representative time periods are entered into the model. For example, rather than simulating unit dispatch on each day of a future summer, the model dispatches units to

[11] Note that many of these life extension programs have become controversial vis-à-vis the New Source Review provisions of the Clean Air Act.
[12] For this reason, these models are sometimes called "optimization models."

meet several types of summer day, and then takes the predicted unit operation for those hours and extrapolates it to the entire summer.

Forecasting models also usually simplify the representation of forced outages of power plants. Rather than representing forced outages probabilistically, these models usually represent them as "deratings" (i.e., reductions) to the capacity of the generator. While this results in a reasonable estimate of the amount of annual generation from baseload plants, the result for intermediate and peaking units may be inaccurate.

Examples of forecasting models include the Integrated Planning Model (IPM®) used often by the U.S. Environmental Protection Agency (EPA), the DOE's National Energy Modeling System (NEMS), and the Energy 2020 model. (These models are different in important ways, and they are discussed in Sections 4, 6 and 7, respectively.) Importantly, all of these models are highly flexible in their operation, and users can input a variety of different types of data. For example, when simulating U.S. energy markets for its Annual Energy Outlook, the DOE uses highly aggregated data in the electricity module of NEMS. Other analysts focus NEMS more specifically on the electric industry, and enter more detailed data into the model than does the DOE. However, forecasting models are designed to be able to focus on large geographic areas, and because of this they cannot accommodate the kind of detail found in dispatch models.

The use of forecasting models to predict capacity additions is appealing in that it provides an automated result. In other words, once the input assumptions and algorithms have been set up, the model makes objective decisions, seeking the optimal future solution based on the inputs. This can lend a credibility to the model's prediction that may be desirable in a policy-making setting. An important limitation of forecasting models is that their algorithms may not effectively represent the complex dynamics that affect the decisions to built new plants and retire old ones. Regulators and participants in a given regional energy market may be able to make an informed prediction that captures these complexities as well or better than a computer model of that market.

Regardless of whether a model or a person predicts unit additions and retirements for an assessment of emission reductions from energy efficiency or clean energy, two things are clear. First, this prediction, and the assumptions that underlie it, should be clearly stated, because results will be highly sensitive to it. Second, policy makers should assess future emission reductions under several assumptions about long-term additions and retirements in order to understand the range of possible outcomes and the important sensitivities. One group of analysts, who use a large forecasting model, describe this type of scenario analysis as follows.

> Utility planners possess no crystal balls and are unable to discern all the necessary information about the present and even less about the future. A utility, however, can be optimally prepared for the future without knowing what that future might be. A utility so prepared will have developed options that are robust under a variety of potential futures. Instead of concentrating on cost-minimization for a "baseline" future that ***never*** comes to pass, the utility builds a portfolio of optimal strategies that maximize benefits under all plausible conditions, emphasizing flexibility and cost-effectiveness.[13]

[13] Quoted from the Energy 2020 website, produced by Systematic Solutions, Inc. and Policy Assessment Corp, at:

### *2.4 Predicting Emission Reductions of Tradable Pollutants*

When predicting emission reductions from energy efficiency and clean generation, it is important to consider the role of allowance trading programs. The major allowance programs in effect are the Title IV $SO_2$ program, the OTC $NO_x$ Budget program in the Northeast, and several regional allowance programs in Texas and California.[14] An important future allowance program to consider is the SIP Call $NO_x$ allowance program, going into effect midway through 2004 in the eastern half of the U.S.

The key issue with allowance programs is that, when an electricity generator subject to an allowance program operates less, the owner may have extra emission allowances to sell to other sources either in the local area or outside the local area. The result is that emissions reduced at one generating unit (due to, for example, an efficiency program) may well be emitted on another day, in the same region or another, by a source that has purchased allowances from the generating unit. Thus, energy efficiency and new clean generation may simply move emissions around in time (and space) rather than reducing total annual or seasonal emissions.

Most electricity models – both dispatch and forecasting models – include allowance costs in the operating costs of generating units. Thus, a unit with a lower emission rate will operate more than units with a higher emission rate, all other things being equal. Over the long term, if the model is well calibrated and the allowance price is accurate, total emissions should not exceed the cap. However, it is impossible to predict what a generating unit owner will do with allowances freed up by lower than anticipated generation. The allowances may be sold and used in the local area. They might be sold and used in a distant area. Or they might be simply banked for future use. Whatever the decision, it will not be made immediately, but over the course of the compliance period (either the ozone season or the year). So the effects of the allowance program are likely to lag behind actual generation reductions in time.

Given the increasing prevalence of allowance programs, regulators will need to consider carefully how and when to credit energy efficiency and clean generation with emission reductions. Rewarding efficiency and clean generation for its air-quality benefits is an effective way to incentivize its development. However, regulators must find a way to ensure that the emission reductions they credit to these programs are true reductions.

The following Sections describe four different methods to predict emission reductions from policies that support energy efficiency and clean generation. The focus of the discussion is on the key assumptions and techniques that affect the results and differentiate the methods from each other. The models employed in each method are also described in detail. Each section includes a short discussion of the strengths and weaknesses of that particular approach vis-à-vis to OTC's current goals.

www.energy2020.com (June 19, 2002).

[14] These allowance program are distinct from credit trading programs in that allowances are allocated to affected sources. In credit trading programs, such as Emission Reduction Credit (ERC) trading in the Northeast, sources apply for credits for emission reductions below current emission limits.

# 3. ISO New England Marginal Emission Rate Analysis

The New England Independent System Operator (ISO New England) uses the PROSYM dispatch model to calculate marginal and average emission rates each year. This practice began in 1994, when the NEPOOL Environmental Planning Committee began an effort to analyze the impact that demand-side management (DSM) programs had on $NO_x$ emissions in the power pool. The results of this work were used to support applications to obtain emission reduction credits (ERCs) from DSM program activities.[15] Later in 1994, NEPOOL released an analysis of marginal $NO_x$, $SO_2$ and $CO_2$ emission rates to complement the initial $NO_x$ effort. Since 1994, NEPOOL (later ISO New England) has published a marginal emissions analysis every year.[16]

## *3.1 The PROSYM Model*

The PROSYM system, developed by Henwood Energy Services, Inc., is a complete regional power pool analysis and accounting system. It is designed for performing planning and operational studies, and as a result of its chronological structure, it accommodates detailed hour-by-hour investigation of the operations of power control areas.

The basic time unit used in PROSYM is one hour (a half-hour version is available for use in certain control areas). In each hour of a study period, PROSYM considers a complex set of operating constraints to simulate the least-cost operation of the system. This hour-by hour simulation, respecting chronological, operational, and other constraints in the case of cost based dispatch, and relevant pool or independent system operator (ISO) rules in the case of bid based dispatch, is the essence of the model.[17]

PROSYM is a general-purpose simulation model capable of representing most electric load and resource situations. To perform simulations, the PROSYM system requires: at least one basic set of annual hourly loads; projections of peak loads and energies on a weekly, monthly, seasonal or annual basis for the study of any future period; and data representing the physical and economic operating characteristics (the resource mix) of the control area and any relevant pool or ISO rules. ISO rules, such as day-ahead unit commitment rules spinning reserve requirements are important constraints on the operation of generating units. When using PROSYM to assess system operations, ISO New England uses detailed, unit-specific information on the power plants in the region that reflect actual operating constraints and outage patterns. Plant efficiency data takes the form of a heat rate curve or multiple heat rates for different generating load levels. Forced outages of generating equipment are represented as randomly occurring (using a convergent Monte Carlo technique) rather than as unit deratings.

[15] The ERC program allows large sources in Massachusetts to earn tradable $NO_x$ and VOC credits by reducing emissions below regulatory requirements. Sources may apply for credits associated with emission reductions achieved by DSM activities begun after January 1, 2002.

[16] ISO New England's Marginal Emission Rate analyses are available at: http://www.iso-ne.com/Planning_Reports/Emissions.

[17] PROSYM users have the option of simulating regions with competitive wholesale markets using bid-based unit dispatch or simulating regulated regions with a traditional cost-based dispatch.

PROSYM can be used to analyze one control area or a group of contiguous control area, however, due to the large amount of detailed input data necessary. PROSYM is rarely run at a national scale.

### *3.2 The ISO New England Methodology*

Before looking at the assumptions ISO New England uses to develop its system marginal emissions analysis, it is important to note that this analysis is not intended to assess the impact of any specific energy or environmental policy. In addition, ISO New England publishes marginal emission rates retrospectively; it does not attempt to predict future marginal emission rates.

The ISO publishes retrospective system marginal emission rates for each of four distinct time periods. These periods are: ozone season on-peak hours, ozone season off-peak hours, non-ozone season on-peak hours and non-ozone season off-peak hours. The ozone season is defined as the period between May 1 and September 30, and the "peak" period is 8:00 am through 8:59 pm.[18]

ISO New England derives these system emission rates by evaluating two consecutive PROSYM runs. First, the New England system is dispatched to meet the actual loads recorded on each day of the year being assessed. Actual constraints experienced by generators are included (such as start-up ramp times and minimum up and down times) as are transmission dynamics between neighboring control areas. The results of this run become the baseline. Second, the system is dispatched in a scenario in which all hourly loads are increased by 500 MW. This is called an "increment run." To calculate marginal emission rates, the *additional* emissions in the increment run are summed across each of the four time periods. In other words, total $NO_x$ emissions from the baseline run are subtracted from emissions in the increment run for each time period. The incremental emissions in each time period are then divided by the incremental generation in the corresponding period to derive a marginal emission rate in terms of lbs/MWh.[19]

ISO New England has the advantage of hindsight when calculating its marginal emission rates. Thus, it can include actual plant outages and transmission constraints that were experienced in the year being assessed. This lends a particular credibility to the ISO's calculations, but it provides little help in predicting future system emission rates. Another noteworthy aspect of this approach is that, while transmission flows between New England and neighboring regions are modeled, the ISO's marginal emission rates do not take into account changes in generation in neighboring regions. That is, if the additional load modeled in New England would have caused generators in New York or the Pennsylvania/New Jersey/Maryland Interconnection (PJM) to operate more, these emissions are not included in the ISO's marginal emission rates. Table 2 summarizes ISO New England's approach to calculating marginal emission rates.

[18] While ISO New England uses the term "peak" to describe this period, we prefer the term "weekday." In general, people associate the word "peak" with the highest electricity loads, and these loads tend to occur during the afternoon hours, a period which is a subset of the "weekday" period. This distinction is important when thinking about system marginal emission rates.

[19] The ISO also divides total incremental emission in each period by total incremental heat input to derive marginal emission rates in lb/mmBtu.

**Table 2. Summary of ISO New England Methodology**

| Issue | Treatment/Method |
|---|---|
| Regions Assessed | ISO New England |
| Time Frame Assessed | One year – retrospective |
| Pollutants Assessed | $NO_x$, $SO_2$ and $CO_2$ |
| Model Used | PROSYM |
| Model type | Chronological dispatch |
| Data on generating units in the model | Highly detailed unit-specific data |
| Load data in the model | Hourly load data |
| Method of simulation analysis | Increment run |
| Transmission constraints modeled | Yes |
| Interregional effects included | No |
| Number of Annual Emission Rates Produced | Four (for each pollutant) |
| Plant Additions | N/A |
| Plant Retirements | N/A |
| Policies/Programs Assessed | None |
| Data Source for Program Load Shapes | N/A |

### *3.3 Strengths and Weaknesses for OTC*

ISO New England's *Marginal Emissions Analysis* in itself does not meet the goals established by the OTC for emission reduction analysis, because it focuses only on New England, does not predict future marginal emission rates, does not assess mercury emissions and does not attempt to assess the emissions impacts of policies that support energy efficiency and clean power. However, the ISO's development of marginal and average emission rates is methodologically sound and has several strengths. Strengths of the approach include the following.

- It is a retrospective look at system operation, and thus it is subject to less uncertainty than prospective modeling.
- It focuses in detail on an area within the Ozone Transport Region (OTR).
- The area – New England – is modeled using a highly detailed (chronological) dispatch model.
- Energy transfers with neighboring control areas are accounted for, and no data aggregation has been necessary to accommodate a large geographic study area.

Given that ISO New England's approach does not seek to meet the OTCs goals, the important question becomes: how accurate would estimates of emission reductions be if they were based on ISO New England's marginal emission rates? First, they would not be credible as predictions of long-term emission reductions. As discussed, over the long term energy efficiency programs and subsidized renewables are likely to displace other potential new units and speed the retirement of existing units in addition to affecting the operation of existing units. ISO New England's methodology does not seek to simulate these dynamics or make predictions in these areas.

Second, because they only rely on four time periods, ISO New England's marginal emission rates may not be appropriate for assessing some types of policies. In particular, these emission rates would probably not be appropriate for modeling energy programs that operate only during the highest-load hours of the year, such as load management programs. As noted, ISO New England uses the ozone season and non-ozone season and peak and off-peak time periods. The peak period covers the hours 8:00 am through 8:59 pm. Regional marginal $NO_x$ rates, for example, probably fluctuate considerably during this period. To determine exactly which energy programs these four time periods are appropriate for, more research is necessary on the range over which marginal emission rate fluctuates in New England.

# 4. EPA's "ADER" Project

The U.S. EPA does its national and regional power system modeling using the Integrated Planning Model (IPM®), a model developed by ICF Consulting. IPM® is a large multi-sector energy forecasting model. In the past, applications of IPM® have included capacity planning, environmental policy and compliance planning, wholesale price forecasting, and asset valuation. Staff at EPA and ICF are currently using IPM® to develop emission reduction estimates for policies that support energy efficiency and clean generation. The description of the IPM® model below is based on literature distributed by the U.S. EPA.

## *4.1 The IPM® Model*

IPM® is a large-scale model able to simulate plant dispatch at different levels of resolution for all regions of the U.S. It is a linear programming model that has a foresight feature that considers what the cheapest way to operate the power system is over a specified period subject to any specified constraints (e.g. pollutant caps, or transmission limitations) that are placed on power generation units.

IPM® performs both dispatch and forecasting functions. In order to accommodate the large geographic scopes IPM® often simulates (the entire nation) and the model's optimization functions, IPM® uses aggregated and simplified data in its unit dispatch function. The model uses data on plant types rather than specific generating units.[20] Regarding load data, IPM® typically represents load in a limited number of "segments" (e.g., 10 load segments in each of four seasons). Generating unit types are not dispatched chronologically, unit types are dispatched to meet these load segments, and the resulting unit operation is extrapolated to the annual level. IPM® simulates forced outages as capacity deratings. That is, rather than simulate actual random outages of generating units during dispatch, the model reduces the capacity of each generating unit to approximate the effect of forced outages. IPM® predicts plant additions and retirements by selecting from a list of potential resource additions, such that total system cost is minimized.

The version of IPM® used by EPA represents the U.S. electric power market in 26 regions in the contiguous U.S. These regions correspond in most cases to the regions and

[20] In IPM for existing fossil-fired units in 2001, 1,127 coal steam boilers are aggregated into 489 unit types; 284 combined-cycle units are aggregated into 54 unit types; 700 oil/gas steam boilers are aggregated into 91 unit types, and 1,783 combustion turbines are aggregated into 81 unit types. For new units, IPM uses 28 different types of new combined cycle unit, 27 types of new coal unit, and 28 types of new combustion turbine. For existing non-fossil fired units in 2001, 3,548 hydro units are aggregated into 30 unit types, 102 nuclear units are aggregated into 20 unit types, 143 pump storage units are aggregated into 15 unit types, and 243 other units are aggregated into 29 unit types.

sub-regions used by the North American Electric Reliability Council (NERC). IPM® models the electric demand, generation, transmission, and distribution within each region as well as the transmission grid that connects the regions.

The model provides estimates of air emission changes, incremental electric power system costs, changes in fuel use and prices, and other impacts that different approaches to air pollution control in the electric power industry will have. It also provides basic information necessary to consider the consumer price impacts, employment changes, and other types of economic impacts.

Currently, EPA and ICF staff are using the IPM® model to address the question of avoided emissions from energy efficiency and renewable generation. In March 2002, EPA released a draft paper describing the *Average Displaced Emission Rate* (ADER) approach. EPA plans to incorporate the results of the ADER analysis into its forthcoming on-line "Emissions Profile Tool," which will enable individuals and businesses to compute the environmental impacts of their electricity consumption systems. Results of ADER analysis are not currently available, but modeling is underway. EPA expects to make the results of this ADER analysis available in the latter part of 2002.

## *4.2 EPA's "ADER" Method*

The ADER method under development will use the IPM® model to develop generalized parameters with which users will be able estimate emission reductions from energy efficiency and renewable energy programs around the country. The method is based on the development of avoided emission rates, called "ADER parameters," corresponding to specific "hour types" in different regions of the country. ADER parameters are being developed for $NO_x$, $SO_2$, $CO_2$, and mercury, for five different regions of the country and for the nation as a whole. ADER parameters are being developed for the years 2005, 2010, 2015 and 2020.

As an example of how the ADER parameter approach will work, one parameter developed might be for (a) weekday mornings, (b) in the year 2005, (c) in the southeastern U.S. This parameter would represent the rate at which emissions would be reduced during this hour type in the Southeast when load is reduced or new generation added. The parameter would be developed using IPM® model runs for the Southeast, and thus would take into account the mix of existing plant types in this region and neighboring regions, transmission constraints and other region-specific factors. A user of this ADER parameter would apply it, for example, to kWhs saved by an efficiency program on weekday mornings in the Southeast.

To assess the full effects of a given program over a given year, users of the ADER parameters will have to know the load profile of the program. The user would select the ADER parameters for the appropriate hour types using their own data on program load profile. EPA is currently considering providing generic load profile information along with the ADER parameters, so users without load profile data will be able to assess common energy efficiency programs and renewable resources. While EPA has not yet identified the specific load profiles it will provide, the agency expects to provide load profiles developed with the Department of Energy's DOE-2 model of building performance.

ADER parameters are being developed for 11 different hour types. That is, each year is broken into 11 different time blocks such as the one envisioned above (summer mornings). Users with hourly load profile data might use a large number of ADER parameters to assess a given program. Users with more aggregated load profile information will allocate energy savings or production across fewer ADER time blocks.

EPA is developing ADER parameters to be used nationwide (i.e., when assessing policies implemented nationally). It is also developing parameters specific to five U.S. regions. These five regions are aggregations of the 26 NERC regions and sub-regions on which the IPM® model is geographically based. The Northeast region includes all three northeastern power control areas. The model simulates the effects of transmission constraints between these three control areas, but it does not simulate constraints within these areas. The model also simulates transmission between the Northeast and the control areas contiguous to it, including those in Canada.

Because the five regions for which ADER parameters are being developed are so large, the specific plants in each region are aggregated into plant types, or "model plants."

ICF will use a decrement run approach to developing the ADER parameters. For each region, IPM® will be run once to establish a baseline, and then run a second time with load levels in each segment reduced. For each hour type, the difference in total emissions and total generation will be assessed to develop the ADER parameter. Table 3 summarizes the ADER methodology.

**Table 3. Summary of the ADER Methodology**

| Issue | Treatment/Method |
|---|---|
| Regions Assessed | Entire U.S. and five sub regions |
| Time Frame Assessed | 2005, 2010, 2015 and 2020 |
| Pollutants Assessed | $NO_x$, $SO_2$, $CO_2$ and Hg |
| Model Used | IPM® |
| Model type | Dispatch/forecasting |
| Data on generating units in the model | Aggregated unit types |
| Data on loads in the model | Aggregated load segments |
| Method of simulation analysis | Decrement run |
| Transmission constraints modeled | Yes |
| Interregional effects included | Yes |
| Number of Annual Emission Rates Produced | 11 (for each pollutant) |
| Plant Additions | Endogenously determined (by IPM®) |
| Plant Retirements | Endogenously determined (by IPM®) |
| Policies/Programs Assessed | Displacement programs (energy efficiency and clean generation) |
| Data Source for Program Load Shapes | Probably NREL for renewables, U.S. DOE (DOE-2 buildings model) for efficiency |

## 4.3 Strengths and Weaknesses for OTC

Below is a list of ways in which the ADER methodology is well suited to meet the OTC's goals for assessing emission reductions from energy programs.

- The work will result in a tool with which users can predict the emissions impacts of energy efficiency and clean generation.
- It will allow users to assess energy efficiency and clean generation using 11 different time blocks each year. This will allow users with high resolution load profile data to "describe" programs with ADER parameters with considerable precision.
- The model runs used to develop ADER parameters take into account transmission constraints and interregional effects. If load reductions in a given area are predicted to affect generation in neighboring regions, this is accounted for in the ADER parameters.
- ADER parameters are being developed for $NO_x$, $SO_2$, $CO_2$ and Hg, and future regulations (i.e., the SIP Call) are factored into their development.
- If EPA releases load profiles for a number of energy efficiency programs and renewable technologies, consistent with the 11 ADER time periods, this will be a significant strength of the approach. If they do not, it will be a weakness, as the ADER parameters will be less useful without load profile data.

The following are ways in which the ADER approach is not well suited to meet the OTC's goals:

- It only provides one displaced emission rate per pollutant for the entire Northeast region. Marginal and average emission rates differ significantly between PJM, ISO New York and ISO New England, and this difference cannot be explored with the ADER parameters.
- IPM® simulates unit dispatch using aggregated data on generating units and system loads. Modeling focused specifically on the Northeast should dispatch units chronologically, based unit-specific information and hourly load data.
- The ADER approach provides displaced emission rates for selected future years rather than for each year of a given study period. Users of the parameters will have to interpolate to assess a period of years.
- The ADER parameters will not be useful in assessing emission reductions from multi-pollutant regulations and EPSs.[21]

Aside from these limitations, stemming from different project goals, the other key aspect of the ADER work to consider is the use of IPM® to predict unit additions and retirements in future years. As discussed in Section 2.3 above, decisions about unit additions and retirements are highly complex decisions and they are difficult to predict. When the ADER parameters are released, it will be important to review closely the inputs and algorithms that drive unit additions and retirements in IPM®.

21 In separate projects, EPA has used the IPM® model to assess the impacts of multi-pollutant bills.

# 5. The CCAP Work for the Great Lakes Protection Fund

## 5.1 Project Summary

The Center for Clean Air Policy (CCAP) is currently involved in work to assess the air pollution impacts of different energy efficiency and renewable energy technologies. The project is funded by the Great Lakes Protection Fund. The goal of the work is to rank energy efficiency and renewable energy technologies relative to each other in terms of their effectiveness at reducing one or more pollutants. The pollutants assessed include $NO_x$, $SO_2$, $CO_2$ and a number of air toxics. The geographic scope of the CCAP project has not yet been finalized, but CCAP staff expect it to include New York State and selected areas in the northern Midwest.

The work of the CCAP differs from the other projects reviewed here in that the goal is not to develop a tool that can be used to quantify emission reductions. The goal is generate information – a report – on the relative abilities of different technologies to reduce air emissions. This report will be designed to inform policy decisions relating to energy efficiency and renewables. It is unclear at this point whether the report will include quantifications of probable emission reductions from technologies operating within the current resource mixes or whether it will rank technologies relative to each other.

CCAP is assessing energy efficiency and renewable technologies using a system dispatch model developed in house. The model runs on FORTRAN and simulates unit dispatch on an hourly basis using unit-specific data collected from EIA and other publicly available sources. These data are not quite as detailed as the data in other dispatch models such as PROSYM and GE MAPPS. (For example, a single heat rate is used for each unit rather than a representation of the unit's heat rate curve over different loadings.) As in IPM®, forced outages are represented as unit deratings.

The model does not perform any forecasting functions, and CCAP does not plan to make predictions about emission reductions from technologies over a period of future years. However, they may do some limited scenario analysis to predict potential reductions under specific assumptions about the future.

One notable feature of the CCAP model is that it is specifically designed to assess the operation of distributed generation (DG), using a simple optimization algorithm to predict the operation of DG at different market-clearing prices. This represents one of the first efforts to simulate the operation of DG and large power plants with a dispatch model.

**Table 4. Summary of the CCAP Methodology**

| Issue | Treatment/Method |
|---|---|
| Regions Assessed | New York and selected areas of the upper Midwest |
| Time Frame Assessed | The present and possibly one future year |
| Pollutants Assessed | $NO_x$, $SO_2$, $CO_2$ and multiple toxics |
| Model Used | CCAP model |
| Model type | Dispatch |
| Data on generating units in the model | Unit-specific data |
| Load data in the model | Hourly loads |
| Method of simulation analysis | Decrement run |
| Transmission constraints modeled | No |
| Interregional effects included | No |
| Number of Annual Emission Rates Produced | Displaced emission rates not provided |
| Plant Additions | Analysis based on current resource mix, no forecasting |
| Plant Retirements | Analysis based on current resource mix, no forecasting |
| Policies/Programs Assessed | Energy efficiency and clean generation |
| Data Source for Program Load Shapes | NREL for renewables, CCAP for energy efficiency |

### *5.2 Strengths and Weaknesses for OTC*

The CCAP methodology would be not be effective in meeting the OTC's current goals. There are three primary differences in the two projects' goals. First, the OTC's region of interest is the OTR, while the CCAP is focused on New York and areas of the Midwest. Second, the OTC seeks to be able to assess quantitatively future emission reductions from efficiency and clean energy technologies. The CCAP will not assess emission reductions quantitatively and will not assess future reductions. Third, the OTC is focused on policies that displace emissions, multi-pollutant regulations and EPSs, while the CCAP is focused only on programs that displace emissions.

Strengths of the CCAP approach vis-à-vis the OTC's goals include the following.

- It focuses in detail on a northeastern control area.
- It focuses on $NO_x$, $SO_2$, $CO_2$ and multiple air toxics.
- The CCAP dispatch model is one of the first attempts to simulate the operation of DG within regional power systems.

# 6. The STAPPA/ICLEI Strategic Planning Software

During 2000 and 2001 the State and Territorial Air Pollution Program Administrators (STAPPA) and International Council of Local Environmental Initiatives (ICLEI) commissioned work to develop a software planning tool for local communities to use in assessing different emission reduction strategies. One aspect of this work was the development of avoided emissions factors for different regions of the country, which the software will use to calculate emission reductions from strategies under consideration. STAPPA and ICLEI hired Tellus Institute to develop these displaced emission factors using the NEMS model. Section 6.1 below provides a brief overview of the NEMS model, and Section 6.2 describes the methodology used by Tellus to develop displaced emission factors.

### *6.1 The National Energy Modeling System (NEMS)*

NEMS is a forecasting model used to predict prices, usage levels and market penetration of energy demand and supply technologies. The Energy Information Administration of the U.S. DOE developed and maintains NEMS to provide projections of domestic energy-economy markets and perform policy analyses.[22] NEMS models energy markets by simulating the economic activity involved in producing and consuming energy products. The time horizon of NEMS is 20 years.

The NEMS model works by balancing the energy supply and demand for each fuel-consuming sector, accounting for the economic competition between the various energy technologies, fuels and sources. The model is organized and implemented as a modular system, with modules representing each of the fuel supply markets, conversion sectors and end-use consumption sectors of the nation's energy system. NEMS also includes macroeconomic and international modules. The primary flows of information between these modules are the delivered prices of each energy commodity to the end user (including to electricity generators and from generators to users) and the quantities consumed by product, region and sector. The delivered prices of fuel include all the activities necessary to produce, import and transport fuels to the end user.

The fuel supply (coal, oil and gas), electricity generation and end–use demand modules reflect technology performance and costs, and supply prices and demands and mutually determined through feedbacks in the model. The integrating module of NEMS controls the execution of each of the component modules. NEMS "solves" future scenarios with an iterative approach. Each supply, conversion and end-use demand module is executed in sequence until the delivered prices of energy and the quantities demanded have converged to within tolerance. Solution is reached annually through the entire planning horizon. Other variables are also evaluated for convergence such as petroleum product imports, crude oil imports and several macroeconomic indicators. Each NEMS component also represents the impact of legislation and environmental regulations (including $SO_2$ and NOx trading systems) that affect that sector and reports emissions of $NO_x$, $SO_2$ and $CO_2$. NEMS reflects current legislation and the cost of compliance with all applicable regulations.

NEMS supports regional modeling and analysis in order to represent the regional differences in energy markets, to provide policy impacts at the regional level and to portray inter-regional flows. The level of regional detail is sector specific. For end-use demand analysis the model uses the nine U.S. Census divisions. Other regional structures include energy production and consumption regions specific to oil, gas and coal supply and distribution, 13 NERC regions and sub-regions for electricity and Petroleum Administration for Defense districts for refineries. National results are presented in the Annual Energy Outlook. Regional and other detailed results are available on CD-ROM and on the EIA website.

---

[22] This description is based on information about NEMS found in the data supplements to the *Annual Energy Outlook*, published each year by the U.S. Energy Information Administration (EIA) and on the Documentation Report, *The Electricity Market Module of the National Energy Modeling System*, published by the EIA in April, 2002 (DOE/EIA-M068(2002)).

The electricity module of NEMS can be used without the other modules to forecast capacity additions and simulate annual system dispatch. The NEMS electricity module includes both an electricity dispatch submodule and a capacity planning module. However, its dispatch function is performed in less detail than in a dedicated dispatch model. The DOE performs dispatch for the 13 NERC regions (including transfers between them) at a highly aggregated level for its annual energy outlook reports. In the DOE's model runs, generating unit types are dispatched based on heat rates and operating costs representative of the unit type. Unit types are not dispatched chronologically, rather they are dispatched to meet aggregated load shapes representative of different time periods.

Load data in NEMS are represented as 11 blocks in each year, corresponding to 11 different time periods. Each block has one load level (MW) specified to represent the demand in that time period.[23] That is, 8,760 hours of load data are aggregated into 11 load levels for different time periods during the year, and generating unit types are dispatched to meet these load levels.

NEMS uses load curves, transmissions costs, and limited foresight to add new units to each region such that reliability is met (using minimum reserve margins in regulated regions and balancing marginal costs of supply with the consumers willingness to pay for new capacity in deregulated regions). The model chooses the set of plants that minimize the total cost across the nation of meeting the regulatory requirements and providing sufficient capacity. Retirement is based on either planned retirements or an algorithm that simulates economic decisions (whether it is more expensive to run an existing plant or build and run a new plant).

## *6.2 The STAPPA/ICLEI "Average Marginal" Emission Factors*

Tellus Institute used the electricity module of NEMS to develop avoided emission factors for the STAPPA/ICLEI software. Avoided emissions derive directly from avoided generation, in this case the changes in generation caused by a decrease in demand. These are also referred to as average marginal emission factors; "marginal" in the sense that they reflect a decrement and thus avoid the highest cost plant additions or dispatch, subject to constraints, and "average" in that they represent changes for an entire year. Tellus derived annual avoided emissions factors for $NO_x$, $SO_2$, $CO_2$ and $PM_{10}$ for 2003 to 2020 for each of the 13 NERC regions (including separate rates for ISO New England, the New York ISO and the MidAtlantic region). Tellus' analysis produced avoided emission factors for each year 2005 through 2020, however the STAPPA/ICLEI software will only use the factors for 2005, 2010, 2015 and 2020. One avoided emission factor was developed for each pollutant for each year. That is, Tellus did not develop separate emission factors for assessing daytime and nighttime load reductions or summer and winter load reductions.

To obtain the avoided emissions factors, Tellus used NEMS to estimate changes in plant retirements and additions over the period 2005 through 2020 due to small changes in

[23]Specifically, the 11 load segments are: Summer Day, Winter Morning/Evening, Winter Day, Summer Day (2), Winter Morning/Evening (2), Spring/Fall Day, Summer Morning/Evening, Spring/Fall Morning/Evening, Winter Night, Summer Night and Spring/Fall Night.

electricity demand. The use of NEMS as a forecasting tool is an important aspect of this methodology. As discussed above, predictions of emissions displaced over the long term are heavily dependent on the model's reflection of planned and economic plant additions and retirements and the way in which demand reductions affect them.

For the dispatch function of NEMS, Tellus used unit-specific data on the existing generating units in the three northeastern regions, from EIA's Form 860 database. Data from this file includes: unit capacity, an average heat rate, fixed and variable O&M costs, and emission rates. Unit dispatch took into account transmission constraints among control areas but not constraints within control areas. Dispatch also accounted for all applicable environmental regulations, such as the Title IV $SO_2$ program, the OTC $NO_x$ Budget program and the (future) $NO_x$ SIP Call allowance program.

To develop displaced emission rates, Tellus performed a model run for a base-case (based on the EIA's *Annual Energy Outlook, 2002*) and a series of runs with a set of annual demand decrements, with the decrement model run reducing load by one percent in each year. For the series of decrement runs, Tellus decremented emissions in each of the NEMS regions individually to estimate avoided emissions resulting from individual community, rather than national-level, reduction programs. For $NO_x$, $SO_2$ and $CO_2$, NEMS provides emissions for both runs, taking account of emission control technology in electricity plants and types of coal. Tellus calculated $PM_{10}$ emissions based on fuel consumption provided by NEMS and emission factors by fuel. The decrement-run emissions were subtracted from base-case emissions to get incremental emission reductions per MWh.

For each regional analysis (i.e., for each decrement run) Tellus performed a national dispatch analysis. However, the displaced emission rates derived only account for changes in emissions in the region where load was reduced. That is, if a load reduction caused a change in generation in a neighboring region, the emissions increase or reduction associated with that change was not included in the avoided emission factor. However, the analysis did capture the change in emissions within a region resulting from changes to electricity imports and exports, as a result of the decrement in electricity demand.

The methodology underlying the development of the STAPPA/ICLEI displaced avoided emission rates is summarized in Table 5.

**Table 5. Summary of the STAPPA/ICLEI Methodology**

| Issue | Treatment/Method |
|---|---|
| Regions Assessed | 13 U.S. NERC regions (including separate analyses of New England, New York and PJM) |
| Time Frame Assessed | 2005 through 2020 |
| Pollutants Assessed | $NO_x$, $SO_2$, $CO_2$ and $PM_{10}$ (also $PM_{2.5}$, CO, NMVOC, lead, nitrous oxide and methane) |
| Model Used | NEMS electricity module |
| Model type | Dispatch/forecasting |
| Data on generating units in model | Unit-specific data (from EIA 860) |
| Load data in model | 11 aggregated load shapes |
| Method of simulation analysis | Decrement run |
| Transmission constraints modeled | Yes |
| Interregional effects included | No* |
| Number of Annual Time Periods | One |
| Plant Additions | Planned additions included; others predicted by NEMS |
| Plant Retirements | Planned retirements included; other predicted by NEMS |
| Policies/Programs Assessed | None – development of avoided emission factors only |
| Data Source for Program Load Shapes | None – development of avoided emission factors only |

** Emission factors only capture changes in emissions in the region where load was reduced. But these in-region changes in emissions do include changes in imports and exports between regions that result from the load reduction.*

## *6.3 Strengths and Weaknesses for OTC*

Below is a list of ways in which the STAPPA/ICLEI methodology is well suited to meet the OTC's goals for assessing emission reductions from energy programs.

- It develops displaced rates for each of the three northeastern control areas.
- It models unit dispatch with data specific to existing generating units, and
- It will result in a tool with which users can predict the emissions impacts of energy efficiency and clean generation.

The following are ways in which the STAPPA/ICLEI methodology is not well suited to meet the OTC's goals.

- It only provides one displaced emission rate per year – it does not reflect the difference in marginal emission rates during different seasons or times of day.
- Annual load data has been aggregated into 11 load levels for dispatch modeling in NEMS and dispatch is not chronological.
- Does not account for changes in emissions in neighboring regions (due to load reductions in a given region), and
- The methodology does not assess reductions in mercury emissions.

Aside from these limitations, stemming from different project goals, the other key aspect of the STAPPA/ICLEI work to consider is the use of NEMS to predict unit additions and retirements in future years. As discussed in Section 2.3 above, decisions about unit additions and retirements are highly complex decisions and they are difficult to predict. A thorough assessment of the avoided emission factors included in the STAPPA/ICLEI software should include a close review of the NEMS-based assumptions about unit retirements and additions.

# 7. Energy 2020

The ENERGY 2020 model is based on the FOSSIL2/IDEAS model developed for the U.S. DOE and is used for many national energy policy analyses. It is maintained and operated by the Systematic Solutions, Inc., an energy consulting group located in Fairbon, Ohio.[24] Energy 2020 is a forward-looking policy assessment model. The analysts who operate Energy 2020 are reluctant to use the term "forecasting model," to caution against the assumption that Energy 2020 provides precise and accurate predictions of the future. Energy 2020 is designed for scenario analysis to help companies and policy makers develop energy policies and strategies that will be robust in a range of future outcomes.

Energy 2020 is more like NEMS than it is like a detailed dispatch model, although it differs from NEMS in important ways. The model is similar to NEMS in that it quantifies future production, consumption and emissions levels consistent with input assumptions using an iterative process that accounts for multi-sector supply, demand and price feedbacks. The supply portion of Energy 2020 includes simulation of electricity capacity expansion and construction, regulated rates and market prices, financial aspects of markets, load shape variation due to weather, and changes in regulation.

The model is different from NEMS in that it is not designed to converge on an optimal solution (i.e., an optimal allocation of resources). Rather, it is designed to simulate the way that energy markets actually work – i.e., to predict the behavior of market participants in deregulated and transitioning markets. Because it focuses on market imperfections such as the exercise of market power, Energy 2020 has been used to assess mergers and acquisitions and the value of energy assets.

Energy 2020 can simulate markets at the national level (including the U.S. and Canada), the state level and the company level. Electric generating units are represented as unit types based primarily on fuel type. There are nine unit types in the model and seven types of purchased power.[25] The model can be configured to make additional distinctions regarding generating units – for example between old and new coal-fired units. Emission rates of $NO_x$, $SO_2$, $CO_2$ and $PM_{10}$ for each plant type are included in the model. Plant types are dispatched based on average heat rates and operating costs of the plant type. The use of unit-specific data is not possible in Energy 2020 due to the large geographic area encompassed by the model.

Load shapes in Energy 2020 are built up by end use sector and summed to produce local market, seasonal, load duration curves. These curves are then sampled over

[24] The following description is based on a model overview provided by Systematic Solutions, Inc., discussions with Policy Assessment Corporation staff (a co-developer of the model), and information from the Energy 2020 website, at: www.energy2020.com. Text quoted in this section is from the model overview provided by the Policy Assessment Corporation.

[25] The electricity plant types in the basic model include: Oil/Gas Combustion turbine, Oil/Gas Combined Cycle, Oil/Gas Steam Turbine, Coal Steam Turbine, Advanced Coal, Nuclear, Baseload Hydro, Peaking Hydro, Renewables, Baseload Purchase Power Contracts, Baseload Spot Market, Intermediate Purchase Power Contracts, Intermediate Spot Market, Peaking PP Contracts, Peaking Spot Market, and Emergency Purchases.

representative hours to produce the dispatch, and then scaled to determine annual production.

The model includes aggregated representations of the U.S. gas and electric transmission systems and the electric generating system. Gas transmission data are provided by the Canadian Energy Research Institute and electric transmission data, by Resource Data, International via the National Electric Reliability Council. The standard version of the model includes 60 transmission interconnections nationwide, although the model can be reconfigured to simulate activity in a given region of the country (e.g., a single control area or group of control areas). To do this, the model is joined with an AC load flow model, Power World, which includes detailed information on interconnections and transmission constraints. The use of Power World data increases the cost of using Energy 2020.

The Energy 2020 model is unique in that it simulates the probable behavior of market players rather than converging on the optimal resource allocation consistent with inputs. Each energy provider is represented in the model by four business units: distribution, transmission, marketing, and generation. The first two remain regulated but the last two can be deregulated to any degree. All market participants use the rules to their best self-interest.

New market entrants, asset sales and purchases, mergers, acquisitions, takeovers, and bankruptcy are explicitly modeled. This allows players to attempt different strategies that, while inconsistent with long-term stability, are successful and therefore economically efficient in the local sense.

Finally, the model includes confidence and validity testing software that places uncertainty bounds on simulation results, quantifies confidence intervals, and ranks the contributions to uncertainty in future conditions. This feature can be used to limit data efforts to information important to the analysis and to determine those strategies and tactics that will most likely result in the desired conditions.

Table 6 below summarizes the Energy 2020 model. Note that “N/A” (not applicable) appears in a number of rows, because we are not reviewing a specific methodology for assessing avoided emissions from energy efficiency and renewable generation.

## Table 6. Summary of the Energy 2020 Model

| Issue | Treatment/Method |
|---|---|
| Regions Assessed | 13 U.S. NERC regions |
| Time Frame Assessed | N/A |
| Pollutants Assessed | $NO_x$, $SO_2$, $CO_2$ and $PM_{10}$ |
| Model Used | Energy 2020 |
| Model type | Dispatch/forecasting |
| Data on generating units in the model | Nine unit types |
| Load Data | Aggregated based on end-use sector |
| Method of simulation analysis | N/A |
| Transmission constraints modeled | N/A |
| Interregional effects included | N/A |
| Number of Annual Time Periods | N/A |
| Plant Additions | Determined endogenously in Energy 2020 |
| Plant Retirements | Determined endogenously in Energy 2020 |
| Policies/Programs Assessed | N/A |
| Data Source for Program Load Shapes | N/A |