Electricity Production and Emissions Simulation

11 December 2005

Yihsu Chen and Benjamin F. Hobbs

Prepared for USEPA

Summary of Findings

The purpose of this portion of the project is to link electricity demands with electricity sector emissions, with a focus on NO_x. Climate change will alter the level and timing of electricity demands, as well as the thermal efficiency of electricity generating units. With a given capital stock of generating plant (short run analysis), this will change their operations and emissions. Even in the presence of seasonal or annual emissions caps, emissions might then increase during the warmest days when ozone episodes are most likely to occur. In the long run, the mix and amounts of various types of generation technologies will adjust, and if climate change occurs, the resulting generation plant will be different than if climate change does not occur. This has further implications for the timing and amount of emissions. This portion of the research project developed and demonstrated methods for obtaining temporal and spatial distributions of NO_x emissions from power plants, and for quantifying the effects of climate change scenarios on those distributions.

In Section 1 of this report, we provide background on the analyses, as well as summarize data sources. In Section 2, we describe the methodology used to assess the effects of climate warming assuming short-run (fixed capital) conditions for the power generation sector; this is an emissions- and transmission-constrained representation of the power market for the mid-Atlantic and Midwest states. Generation plant and demand conditions for the year 2000 were used in that analysis. That section also summarizes the results of the analysis. In Section 3, we turn to the issue of interannual variability of climate, electricity demands, and emissions. Most energy models, when used in climate impact analyses, assume a "typical" or "average" year. However, interannual variability is very important; because of the nonlinear relationships between temperatures, emissions, and their impacts, the average impact on air quality over a number of years may be quite different (and higher) than the impact on air quality in an average year. In Section 4, the procedure used to calculate long run (variable generation plant) scenarios that adapt to changing climate is described, along with selected results. Appendix I summarizes the procedure used to translate the hourly emissions outputs from the electricity generation simulation model into the script files that can be used by the SMOKE module of CMAQ. Appendix II lists publications and conference presentations from this portion of the project, and Appendix III summarizes two journal articles resulting from this project that describe the short run models and some analyses of the interaction of oligopolistic generator behavior and emissions.

The major findings concerning the short- and long-run effects of climate change on pollution emissions from the power sector are summarized as follows:

1. In the short-run, climate change, which is characterized as an increase in ambient temperatures, can affect power system in two ways: increasing power demand (demand effect) and degrading generation heat rates (efficiency effect). According to our analysis, for a 2°F increase in the ambient temperature, the former is greater than the latter by an order of magnitude. For instance, for one ozone episode analyzed, the demand effect contributes a 5% increase in NO_x emissions during ozone episodes, while the efficiency effect accounts for less than 0.1%. These increases occurred even though emissions were assumed to be subject to a seasonal cap. The greatest emission impacts occur in the hours when demand is already high; because those are the hottest days of summer, those are also the days when ozone episodes are most like to occur (Section 2).

2. Besides the short-run demand and efficiency effects, climate change has long run effects on two categories of capital stock: energy-using equipment (e.g., size and numbers of air conditioners) and the mix and amount of generation plant (e.g., the amount of peaking gas-fired capacity versus baseloaded coal capacity). The effects on consumer equipment choices are examined in the GWU portion of this final report. Here, we focus on effects on generation investment, and the implications for emissions. We used a combination of the GWU short-run analyses and long run National Energy Modeling System analyses to analyze the effect of a 415 ^{o}F-day increase in cooling degree-days upon electric load distributions in the year 2025. As a result, we project the highest (peak) hourly demand increases by 11%, compared with just a 4.5% increase in the summer average demand. The CDD increase was obtained by analyzing selected years from the GISS GCM/MM5 scenarios for the 1990s and 2050s climates. The substantial increase in the peak period demand has two implications for the power system and its emissions. First, more peaking units such as combustion turbine need to be installed to meet high demand in the peak hours. Second, an increase in the emission during high-demand hours could worsen regional air quality at precisely those times when ozone episodes are most likely to occur. For a given summer hour, the correlation between temperature and emissions is around 50% to 70%. Our results imply that the aforementioned 15% increase in peak demand translates into a 33% increase in the peak pollutions emissions in the long run, even when there is a seasonal cap on emissions (Section 4).

3. There could be significant inter-year variability associated with NO_x emissions during high-demand hours in the long-run. Based of GISS simulations of two years under 1990s climate conditions and another two years under 2050s conditions, we find significant variation in the distribution of year-to-year peak emissions, even under the restrictive assumption that there is no interannual emissions banking. Higher peak ambient temperatures are associated with higher pollution emissions (Section 3 & 4).

1. Background and Data Sources

1.1 Electricity Market Geographical Coverage

Because of the fine geographic scale required for emissions inputs to CMAQ, our power market simulation model is regional rather than national in scope. Our analysis included generation units in two North American Reliability Council (NERC) regions: MAAC (Mid-Atlantic Area Council, also called the PJM Interconnection)[1] and ECAR (East Central Area Reliability Coordination). MAAC covers four states (Maryland, New Jersey, Delaware and Pennsylvania) and District of Columbia. ECAR includes several states located upwind of MAAC, including Ohio, Kentucky, Indiana, and parts of Michigan. However, state boundaries are rarely coincident with power control areas (PCA) in ECAR, and we only included PCAs that have most of their service territory within these states.

1.2 Generating Characteristics

We replied on several sources for data on generation. In general, power generation simulation models (often called "production costing models") require details on engineering and cost inputs, such as heat rate, fuel cost, capacity, primary mover, and nonfuel variable operation and maintenance costs. The primary source was from Energy Information Agency (EIA) of the US Department of Energy.

The short run analysis represents the generation plant in place in 2000. In summary, the short run generating dataset comprises with 1,453 generating units, of which 731 generating units were located within PJM, and remaining were located in ECAR. The total capacity accounting for the forced outage rate in our model was 114,685MW.

The marginal production cost in our analyses was represented as the sum of fuel cost, SO_2 permit costs, and non-fuel variable operation and maintenance expenses. Fuel costs are exogenous and depend on plant location and type. Since SO_2 allowances trading is national in scope, we treat costs associated with SO_2 allowances as an exogenous component of production cost. In contrast, under the NO_x SIP call, NO_x trading is more regional in nature, and so we explicitly cap NO_x emissions in the region, which makes the shadow price of NO_x allowances endogenous to the solution.[2] To properly account for differences in the reliability of various types of generating units, we included forced outage rates in the model by "derating" capacity by the average outage rate. The value of forced outage

[1] The geographic coverage of PJM Interconnection was the same as MAAC in 1998. But, over the past few years, PJM has grown substantially. In 1998, it only included four states: Maryland, New Jersey, East Pennsylvania, Delaware and District of Columbia in the mid-Atlantic region. Now, its coverage has expanded to comprise 13 states plus DC. As a result, besides the original territory, it now serves a total of 51 million customers with a peakload of 131,330 MW, approximately two and a half times as much load as in 1998.

[2] Of course, the Clean Air Interstate Rule proposed in 2005 by USEPA would make NO_x more of a national market. So as not to overstate the effect of climate change on emissions in the study region, we assume that the region does not import allowances from elsewhere during higher demand (generally warmer) years, which would be consistent with assuming that during those same years, other regions also have higher allowance requirements and so do not have additional allowances to spare.

depends on prime mover and size of plants, and the data was drawn from North American Reliability Council (NERC) (NERC, 2005).

Although the EIA dataset contained information about emission rates for many units, for some units that lack such information we derived emission rate from other sources. In particular, we used data from the USEPA (Environmental Protection Agency) Integrated Planning Model (IPM) (USEPA, 2005d) and the Emission and Generation Resource Integrated Database (eGRID) (USEPA, 2005c). If the information from USEPA IPM and eGRID remained inadequate or inconsistent for a given source, we cross validated them with these in the USEPA Continuous Emission Monitoring Data (CEMS) (USEPA, 2005e).

1.3 Transmission Network

One crucial step in developing the models was to have an adequate representation of the transmission network. The presence of network constraints can greatly affect the total cost and emissions of power generation, as well as their spatial distribution. Including the high voltage (345 kV and above) network allowed us to take into account the effect of network congestion in power generation and emissions.

The main source for transmission network was PowerWorld (PowerWorld, 2005). The PowerWorld data set contains information about network topology and transmission limits. In our model, the entire PJM system (1998 footprint) is represented by a 14-node system with 18 arcs. The ECAR system is only represented as a single node, with two arcs connecting to PJM region. The reason is while there is significant congestion within PJM but less within ECAR region. Furthermore, we are focused more on ozone impacts on the eastern seaboard, so more detail on that region's power system is appropriate. The congestion between PJM and ECAR will be captured by the two connecting arcs between two regions.

In order to correctly represent network congestion, we split some of the PCAs into several nodes so that congestion within PCAs can be represented. To identify the location of each generator and assign each generator to a node in the network, we use information from USEPA eGRID or lat-long information. Figure 1 plots the resulting transmission network in our analysis. In particular, four PCAs have been spilt into more than one node: METED (Metropolitan Edison), JCPL (Jersey Power and Light), PPL (Pennsylvania Power & Light) and BGE (Baltimore Gas & Electric).

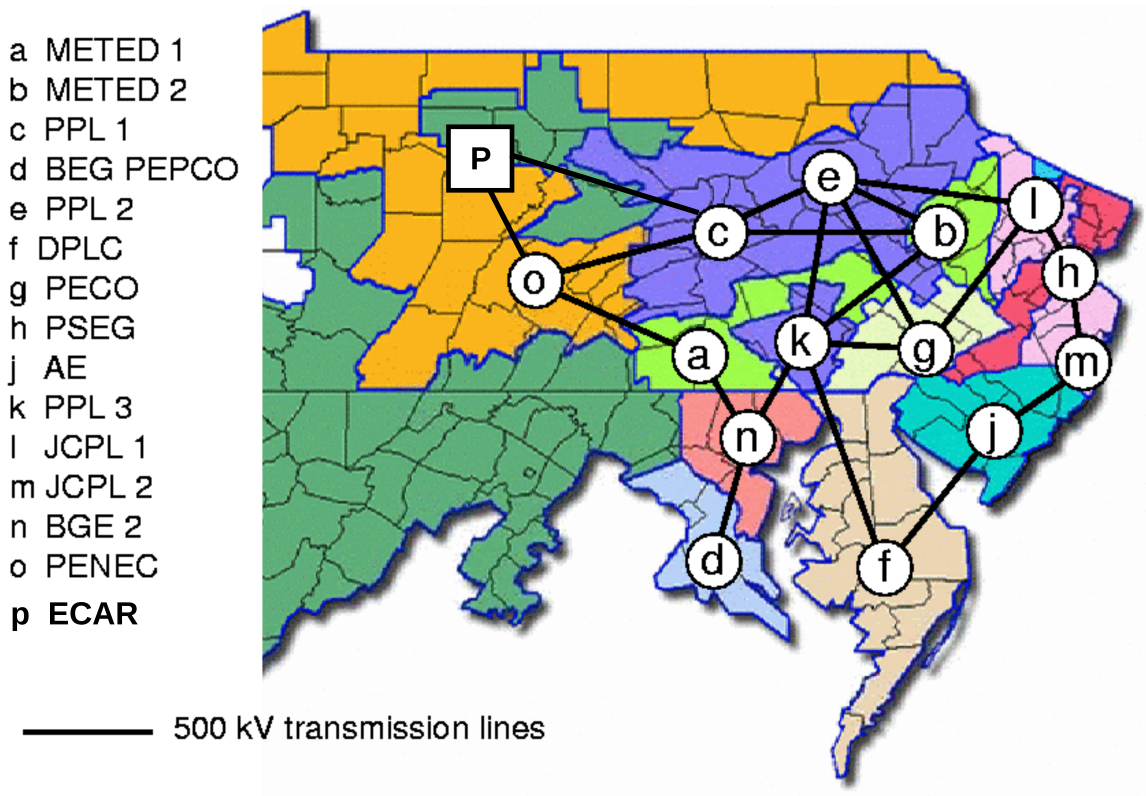

Figure 1: The PJM and ECAR Transmission Network

1.4 Ownership

While ownership is not crucial in our competitive market analysis, it becomes important in our companion analyses, where we look at possible strategic interaction of oligopolistic power producers in the electric and emissions allowances markets within PJM region (Chen and Hobbs, 2005; Chen et al., 2005). The primary source of ownership data is EIA Form 860 (EIA, 2005). For units not in the EIA 860, an internet search or personal contact was used to confirm ownership. To ensure an appropriate representation of the potential of market power under the current ownership, we assume that operational decisions (generation and sale) are controlled by the parent company, replacing any subsidiaries with the corresponding parent company. For the 29 units jointly owned by more than one incumbent company, we treat each as a set of multiple units by splitting capacity in proportion to ownership percentage. Other assumptions, such as assigning control to the owner with majority ownership, could instead be applied to study market power in the presence of partial ownership (Amundsen and Bergman, 2002) .

1.5 Loads

The simulation period was entire year of 8,760 hours, where the ozone season of comprising 3,672 hours was separate from others. The annual load, i.e., load duration curve, was approximated by several block systems. The number of blocks in the model was the balance between the considerations of results' precision and the computational complexity. For instance, in the approximation of the three-day episode analysis, a block could correspond to one hour. Hourly load data for each individual PCA or node were obtained from the PJM and ECAR website, as are boundary conditions (net imports).

1.6 NO_x Allowances

The number of SIP Call allowances allocated to the study region an important issue since it affects operations in the short-run (due to emissions dispatch in response to NO_x prices) and capacity mix in the long-run (as high emission capacity is penalized by the market, and becomes a less attractive investment).

For the short-run simulations, we relied on EPA annual compliance reports for tradable permit data (USEPA, 2005b). A total of 131,440 permits were available at the end of 2000 for affected facilities in the 1998 PJM region: Maryland, Pennsylvania, New Jersey and Delaware. There were 109,227 permits allocated in 2000 by the NO_x budget program, and 22,163 permits were carried over (banked) from previous years. Only 92,107 permits are assigned to power plants; the remaining permits, which are owned by other industrial sources, are left out of our analysis. Consistent with empirical observation of generator behavior in PJM, we assumed only 80% of available permits are used for compliance purpose, and the remaining 20% are banked for future use during the short run (2000) simulation. Thus, in the short-run analysis, the amount of available allowances for PJM region is 73,686 tons. A similar approach is used to estimate the number of the available NO_x allowances for the ECAR region.

For the long-run analyses, the tightness of the NO_x cap is based on the CAIR (Clean Air Interstate Rule) (USEPA, 2005a). The table below
Table 1

Table 1: NO_x Cap for SIP Call and Clean Air Interstate Rules (CAIR) [tons]

	SIP	CAIR	
ST	2003-2007	2009-2015	2015-
DC	207	112	94
DE	4,306	2,226	1,855
IN	7,088	45,952	39,273
KY	19,654	36,045	30,587
MD	14,519	12,834	10,695
MI	25,689	28,971	24,142
NC	31,212	28,392	23,660
NJ	9,716	6,654	5,545
OH	45,432	45,664	39,945
PA	47,224	42,171	35,143
VA	17,091	15,994	13,328
WV	26,859	26,859	26,525
Total	265,078	312,506	267,985

2. Short-run Transmission-Constrained Emission Simulation

The research objective of this subtask is to translate the short-run climate-sensitive power demands estimated by the research group at GWU into the hourly pollutions emissions. This short-run analysis focuses on Pennsylvania-Jersey-Maryland (PJM) power system, and assessed the impact of climate change on hourly emissions a three-day ozone episode, while ensuring that total emissions during the ozone season satisfy the assumed cap..

2.1 Analysis Procedure

Figure 2 provides a flowchart to describe the procedures that we used for the short-run three-day episode analysis. The details of each step are explained in the next few paragraphs.

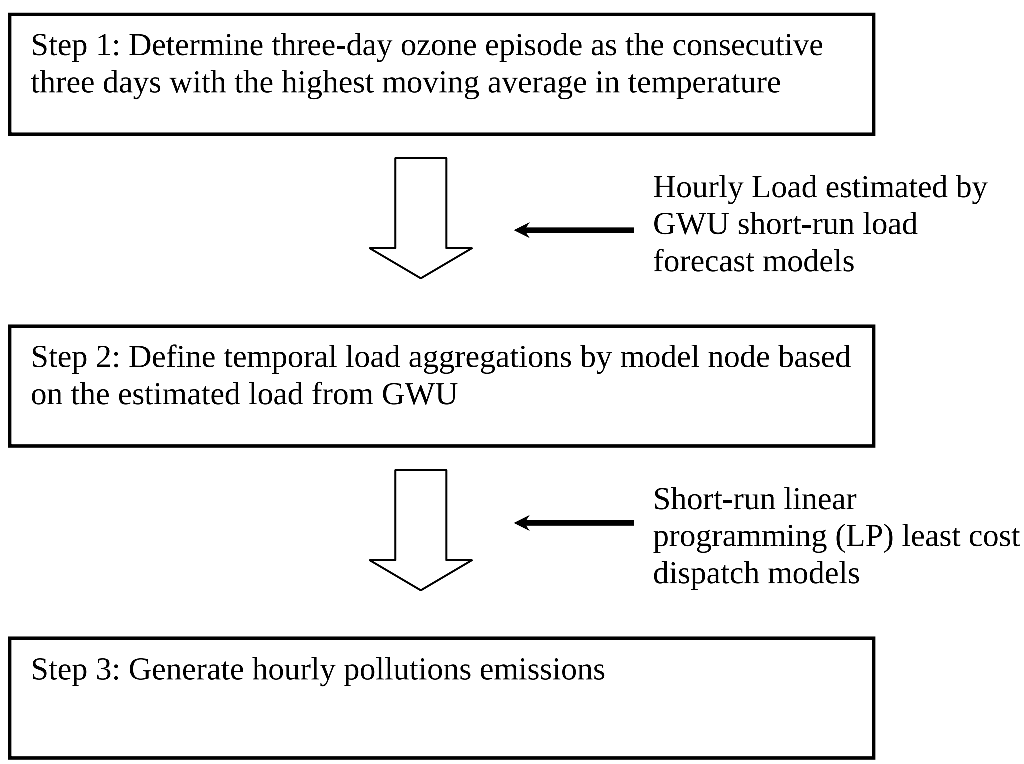

Figure 2: Flowchart for short-run transmission-constrained emissions analysis

The details of each step are summarized as follows:

Step 1: We assessed candidates for ozone episodes by looking at 3-day moving averages of temperature during July and August of 2000 in two cities in the PJM region: Philadelphia and Washington DC. With a moving average of 82˚ F, Aug. 7 to Aug. 9, 2003 is selected as our episode period.

Step 2: In order to include the regional NO_x cap in our model, our analysis considers the ozone season of May 1 to Sep. 30, 2000, a total of 3,672 hours. Excluding the 3-day ozone episode period considered in the first step, the load distribution of remaining 3,600 hours is approximated by five load periods. The peak period has 60 hours, while the other periods have 885 hours each. Together with the 72 periods (one for each hour) for the episode days, our model includes 77 periods. These loads are disaggregated to each of 14 nodes in an aggregated PJM network, based upon the GWU analysis.

Step 3: Using the load blocks generated in the second step, we then applied a short-run least-cost linear programming (LP) dispatch model to simulate the hourly pollutions emissions in the PJM region during five-month ozone period. The assumed sensitivities of generator characteristics (MBTU/kWh heat rate and MW(e) capacity) to temperature changes are that an 0.075% increase in the heat rate and an 0.4% decline in available capacity for per 1 F° in temperature for gas turbines, respectively.[3] To ensure that the market structure that we simulated in the LP models is close to the reality, we also performed analysis using two alternate market structure assumptions: oligopolistic Cournot and Stackelberg.[4]

2.2 Simulation Model

In this subsection, we present the mathematical formulation of the LP short-run least-cost dispatch model that we used to simulate short-run power operation. We begin with the definition of notations, followed by the explanation of objective function, constraints and solution procedure.[5]

2.2.1 Definition of Notation. In the following presentation, the parameters are in capitalized letters; primal variables are in alphabetic letters; and dual variables are in greek letters.

---

[3] Based on F.J. Brooks (Brooks, 2000), Figure 9 shows that the heat rate is 96% of nominal at 0 º F and 105% of nominal at 120 º F. Since the curve is linear, this translates into a sensitivity of 0.075% per º F. Degradation in steam plant efficiencies with increases in ambient temperatures were approximated based on Carnot efficiency calculations. The result is a heat rate degradation of 0.068% per ºF, very close to the turbine values. Capacity effects were estimated by comparing summer and winter capacities from the Energy Information Agency NEMS data set, and assuming a 30 degree average temperature difference between the seasons.

[4] The details of the analyses using alternative market structure are not presented here, but can be found in the two published articles included in Appendices A and B, i.e., (Chen and Hobbs, 2005) and (Chen, Hobbs et al., 2005). In general, we concluded that the PJM market during 2000 was relatively competitive; and the market simulated using LP least-cost dispatch models is a good approximation of actual market conditions. Or, equivalently, the oligopolistic Cournot prices were closest to prices assuming that generators in PJM were heavily forward contracted, which dampens incentives to restrict output and raise prices in the spot energy market. This was indeed the case with PJM generators in 2000, which accounts for the relative competitiveness of that market relative to California in that year.

[5] The use of linear programming to simulate generation system operation goes back to the 1950s and Electricite de France, see (Turvey and Anderson, 1997). The embedding of linearized DC load flow models to represent transmission flows ((Schweppe et al., 1988) , Appendix A) is more recent, e.g., (Ji and Hobbs, 1998).

Table 2: Set and index definitions for short-run transmission-constrained model

Indices/Sets	Description
$t \in T$	Time periods, or blocks
$t^{OTC} \in T$	Periods under ozone cap
$i, j \in I$	Nodes in network
$k \in K$	Constrained transmission interface in network
$h \in H$	Generators
$H(i) \subset H$	Set of generators located in node *i*
$H^{OTC} \subset H$	Set of generators whose emissions are included in NO_x program

Table 3: Parameter definitions for short-run transmission-constrained model

Parameters	Unit	Description
C_h	[$/MWh]	Marginal production cost of generator *h*
X_h	[MW]	Production Capacity of generator *h*
E_h	[tons/MWh]	Emission rate of generator *h*
FOR_{ih}	[]	Forced outage rate of generator *h*
T_k	[MW]	Capacity limit for interface *k*
L_{it}	[MW]	Demand of power at node *i* in period *t*
B_t	[hours]	Duration of period *t* in the load duration curve approximation
$\overline{NO_x}$	[tons]	Seasonal NO_x emission cap
$PTDF_{ki}$	[MW/MW]	Power transfer distribution factor for a unit power injection at an arbitrage hub node and unit withdrawal at node *i* for transmission interface *k*
Z_{it}	[MW]	Power imported from outside the study region to node *i* in period *t*

Table 4: Variable definitions for short-run transmission-constrained model

Variables	Unit	Description
x_{iht}	[MW]	Output level of generator *h* in period *t*
s_{it}	[MW]	Sales of power at node *i* in period *t*
p^N	[$/ton]	NO_x allowances price
p_{it}^E	[$/MWh]	Power price at node *i* in period *t*
ρ_{ht}	[$/MW]	Dual variable of capacity constraint for plant *h* at period *t*
λ_{kt}	[$/MW]	Dual variable of transmission constraint for flowgate *k*
w_{it}	[$/MWh]	Wheeling charge of bringing 1 MW of power from hub to node *i*
y_{it}	[MW]	Power delivered from hub to node *i* in period *t*. This is net power injected at node *i*

2.2.2 Mathematical Model. The short-run transmission-constrained model presented below only covers the summer ozone period (called STM-summer). The model consists of an objective function (cost) to be minimized (equation (1)), a set of variables (defined above) whose values are chosen to minimize that function, and a set of constraints (2)-(7) that define what values of the variables are feasible. Dual variables (shadow prices) for each constraint are shown to the right of the constraint in parentheses.

PROBLEM STM-SUMMER:

$$\underset{x_h}{MIN} \sum_{h,t} B_t C_h x_{ht} \tag{1}$$

subject to:

$$x_{ht} \le X_h(1 - FOR_h), \forall h,t \qquad (\rho_{ht}) \quad (2)$$

$$\sum_{h \in H(i)} x_{ht} - z_{it} - s_{it} = y_{it} \quad \forall i,t \qquad (w_{it}) \quad (3)$$

$$B_t s_{it} \ge B_t L_{it} \ \forall i,t \qquad (p_{it}^E) \quad (4)$$

$$\sum_{h \in H^{OTC},t} B_t E_h x_{ht} \le \overline{NO_x} \qquad (p^N) \quad (5)$$

$$\sum_i y_{it} PTDF_{ki} \le T_k \quad \forall k,t \qquad (\lambda_{kt}) \quad (6)$$

$$s_{it}, x_{iht} \ge 0$$

The objective function in STM-summer is to minimize the overall production cost, which is comprised of fuel, variable operation and maintenance costs. There are five sets of constraints. First, the energy output x_{ht} is less than or equal to the plant's capacity, derated for the forced outage rate. The second constraint states that power flow in each node has to be balanced. That is, the output minus the sum of the import and nodal consumption has to be equal to the net withdrawal/injection at that node. Third, the demand has to be met at each node in every period *t*. The fourth constraint says that the total NO_x emissions need to be no more than the cap. Finally, the power flow in through each transmission flowgate *k* cannot exceed its thermal capacity at each period *t*. The term $PTDF_{ki}$ is called power transfer distributor factors, which are a matrix that governs the flow in the network. It is derived from Kirchhoff's voltage and current laws, assuming there are no transmission losses in the network. The solution from STM-summer gives the output level of each generator and power flow in the network at each period *t*. This solution can then be used to calculate hourly emissions.

2.2.3 Solution Procedure. The STM-summer model is formulated in the GAMS modeling environment (General Algebraic Modeling System). It is then solved using the LP solver CPLEX (Brooke et al., 1998).

2.2.4 Short-Run Example Results. In this section, we summarize the impact of a 2 ˚ F increase in ambient temperature on pollution emissions for the ozone season, assuming that the mix and amount of generation capacity does not change. In particular, we look at the overall impact on entire ozone season as well as the hourly emission change during the hypothesized three-day episode.

The load sensitivity shows that the overall impact on ozone season demand is an average of 4.3% increase of load over 2000 ozone season as a result of a 2 °F increase in ambient temperature, or 2.15% per °F. This increase in demand leads to no change of the total NO_x emissions, since, by definition, the seasonal emissions are capped. However, substantial impacts on fuel cost are observed among scenarios. The increase in fuel cost compared with the base scenario is 21%, 0.4% and 22% as a result of load increase alone, deterioration in generator efficiency due to temperature, and both effects together, respectively for the entire ozone season.

Next, in our three-episode analysis, Figure 3 shows the sum over nodes of loads for the three-day episode period. As anticipated, a greater impact occurs during peakload periods when demand is already high (and temperatures are likely to be high), and a lesser impact during the off-peak period. The average impact is around 5.0%, but impact during peak-demand hours could be as high as 5.6%. Table 5 summarizes the emissions and fuel cost impacts. The impacts due to load changes are almost two orders of magnitude greater than the impacts due to changes in generator efficiency. (In the latter analysis, we will focus exclusively on the effect of heat rate changes alone.) Note that for this three day period, NO_x emission impacts are not zero, because emissions are capped only for the entire ozone season, and not for particular days. NO_x emissions have increased during this time, implying that emissions at other times are lower in order to meet the overall cap.

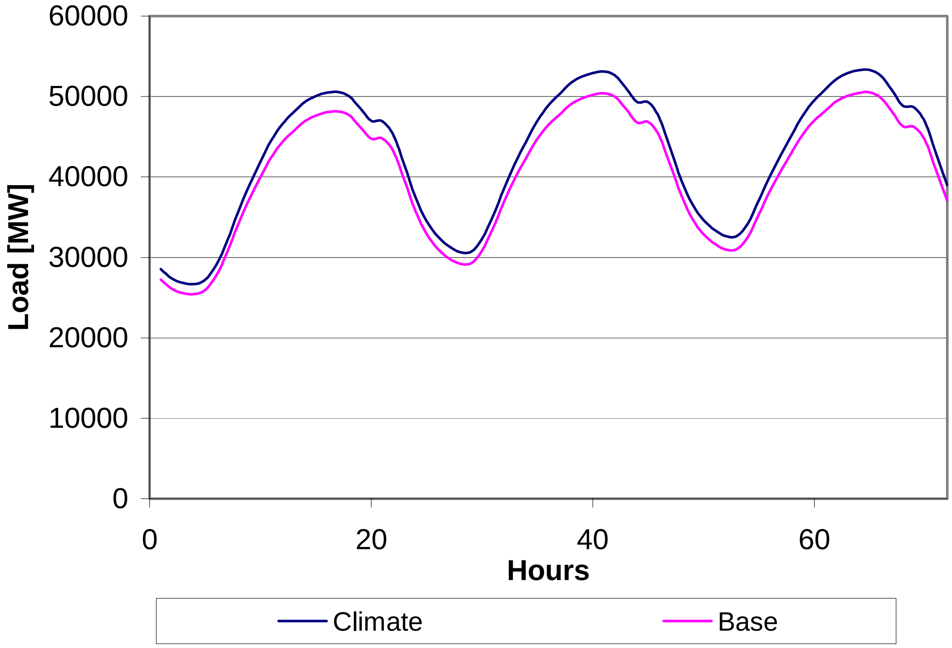

Figure 3: Load Profile over 3-Day Episode Period in PJM (Aug 7 – 9)

Table 5: Electricity Demand and Generators Performance Impact Relative to Base Case for Three Day Ozone Episode, Assuming 2 ^{o}F Temperature Increase

Impact\Category	NO_x Emission	SO_2 Emission	Fuel Cost
Demand Impact	5%	5.5%	19.9%
Generator Efficiency Impact	0.076%	-0.001%	0.25%
Joint Impact	4.9%	5.4%	20.3%

Figure 4 illustrates the hourly emissions impact during our three-day episode period, where the left graph represents the NO_x emission and right one is for SO_2 emission. The climate change scenario is the joint impact of load increase and generator efficiency deterioration due to the ambient temperature increase. The average impact is 4.9% and 5.4% for NO_x and SO_2, respectively. While most NO_x emission increases occur during peakload hours, i.e., day time, the increase of SO_2 emission tends to occurs during both day and night time.   This is because the marginal generators at night are coal plants, which are high SO_2 emitters.

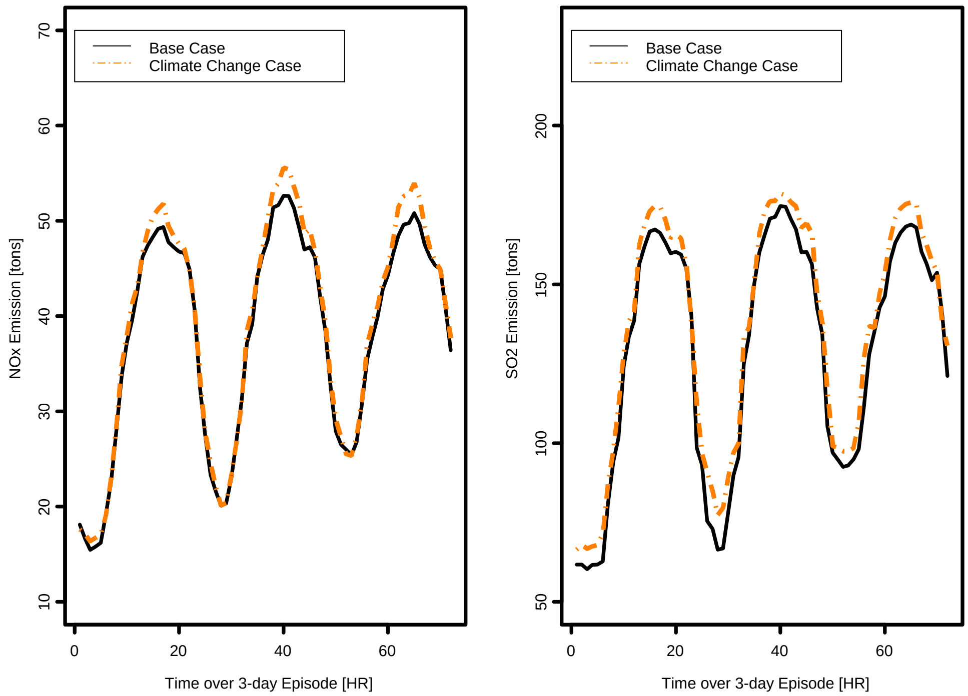

Figure 4: NO_x and SO_2 Emissions Impact during Three-Day Episode Period, PJM

Figure 5 further decomposes the aggregate emissions impact of the three-day episode into state-level results. In terms of NO_x emissions, the greatest tonnage impact occurs in Maryland and Pennsylvania, amounting to 6.3 % and 3.3% increases compared with base case, respectively. The percentage increases for NO_x in Delaware and New Jersey is 10.1% and 3.5%, respectively, but the tonnages involved are much smaller than the other states. Among states, the SO_2 tonnage impact is highest in Pennsylvania and lowest in Maryland.

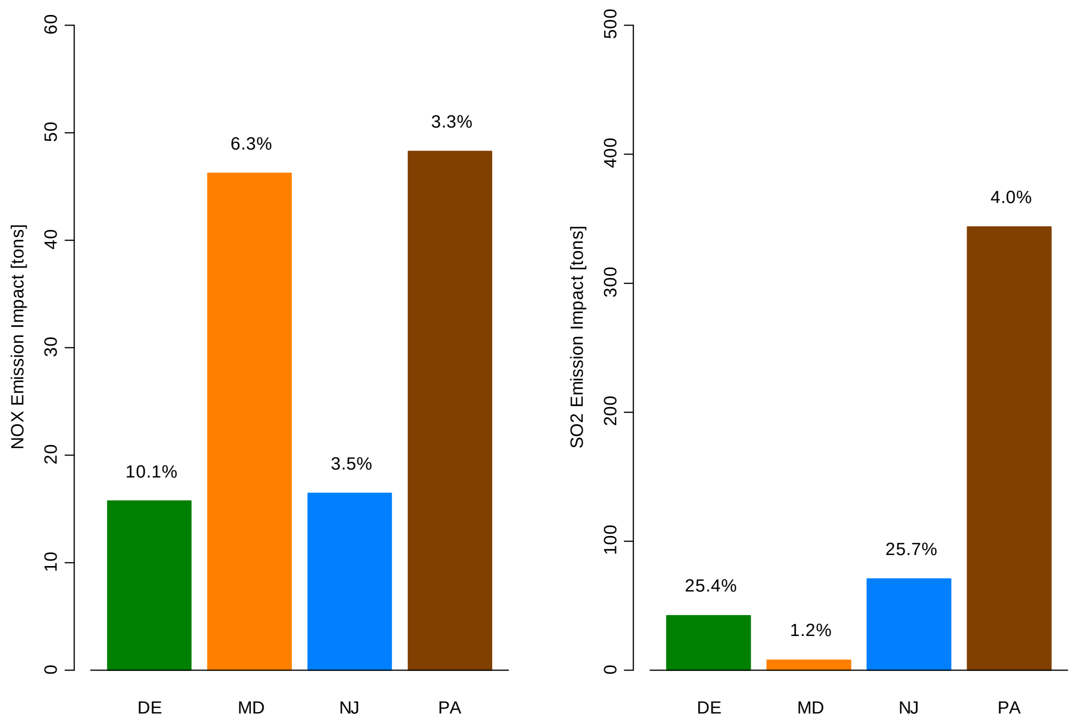

Figure 5: State-Level Impact of 2 ^{o}F Warming during 3-Day Episode

3. Inter-Year Variability of Load and Temperature

3.1 Motivation

The analysis in Section 2 focused on load and temperature data in 2000. One important aspect about climate change is its variability from year to year. To explore the impact of such inter-year variability on the demand and, thus, on emissions, GWU used time-series temperature data from the GISS Global Circulation Model (GCM) from several cities within the study domain for fourteen years as inputs to their demand models . In this analysis, we treated years1991-1998 (90s) as the reference years, assuming no climate change. On the other hand, the six years of 2050-2055 ("50s") are seen as experimental years, which weather conditions are affected by climate change. We assume that there are two distinct distributions that characterize the temperature data for 90s and 50s, respectively. The temperature of a given year, e.g., 1991 and 2050, is simply a realization of a random draw from the respective temperature distribution i.e., 90s and 50s. The temperature data is then used by the GWU research group to simulate hourly loads using short-run load forecast models (for the 1990s) and combining short run noise from their short-run forecast models with long run averages. In the next two sections, we present the analyses with regard to inter-year variability of ground-level temperature and the predicted load.

3.2 Inter-Year Temperature Variability

Figure 6 plots the summer average ground-level temperature for fourteen selected PCAs in the PJM and part of ECAR region during 1991-1998 and 2050-2055. Each PCA is represented by a city within the region. For example, BC (Baltimore Gas & Electric) is represented by the city of Baltimore. The years 1-8 and 9-14 in the X-axis correspond to 1991-1998 and 2050-2055, respectively. The difference in temperature between 90s and 50s varies by PCA, with highest of 3.15 °C for ME (Metropolitan Edison) and lowest of 2.06° C for AE (Atlantic Electricity). Not only is the temperature generally higher among the years in the 50s than among the 90s, but so is the variance. This implies that extremely hot years are considerably more likely in the 50s. The variance in different PCAs is more or less the same, so the variability of temperature among cities of our study is quite homogeneous.

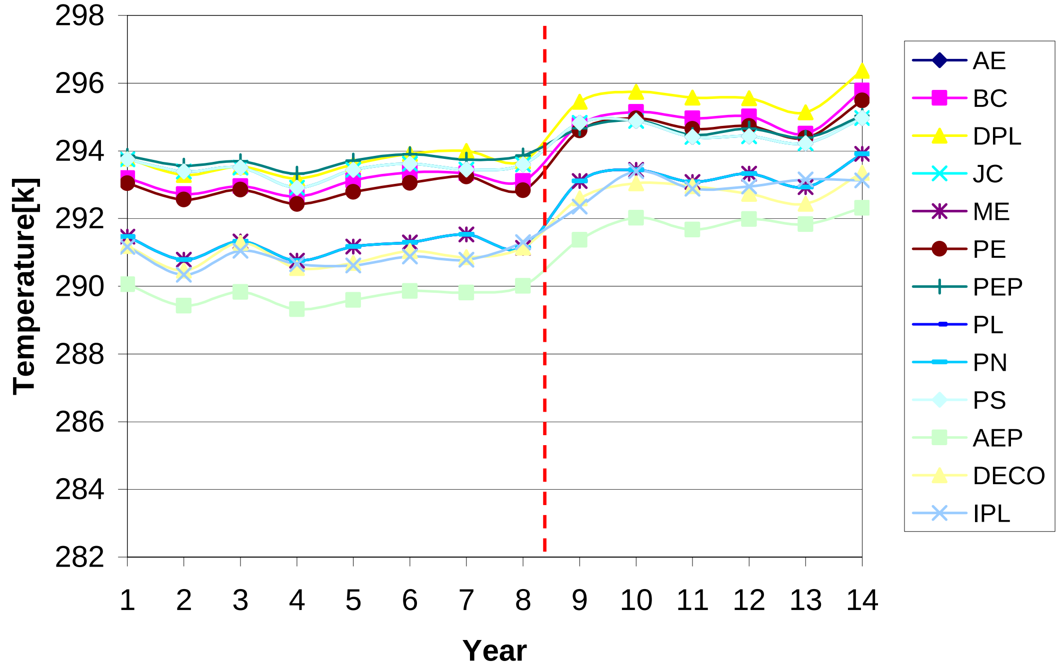

Figure 6: Average Temperature of fourteen PCAs during years 1991-1998 and 2050-2055

3.3 Inter-Year Load Variability

Figure 7 illustrates the inter-year variability of temperature (left) and load (right) for one selected PCA – Baltimore Gas and Electric (BC). This choice of PCA is for illustrative purposes. Similar results apply to other power control areas. The variability within a year is portrayed using temperature- and load-duration curves, respectively. A duration curve shows the number of hours during the summer that the respective quantity exceeds the value given on the y-axis.

For temperature, the respective standard deviation across years is 0.41°C and 0.49 °C for the 90s and 50s. Thus, the variation of temperature across 2050-2055 is only marginally larger than that of 1991-1995.

In Figure 7 (right), we plot load duration curves for the fourteen years for the same PCA. The average standard deviation across years is 47 MW and 76 MW for 1991-1998 and 2050-2055, respectively. The implication of such inter-year variability in load and temperature on pollutions emissions will be explored and discussed in the latter sections of this report. Finally, Figure 8 plots the average summer load against average summer temperature for BC. As anticipated, there is an apparently upward trend, showing that higher average loads are associated with higher temperatures. Two clusters are apparent, which correspond to data point for the 90s and 50s, respectively. The great spread for 50s also validates that the temperature variance/standard deviation is larger for years in that decade.

Figure 7 shows that a load response of approximately a 2.5% change in demand for each degree F (4.4%/^{o}C), consistent with Section 2's results. However, the load model underlying Figure 7 is based on just the short-run response; the long-run response will be greater, as discussed later in this report.

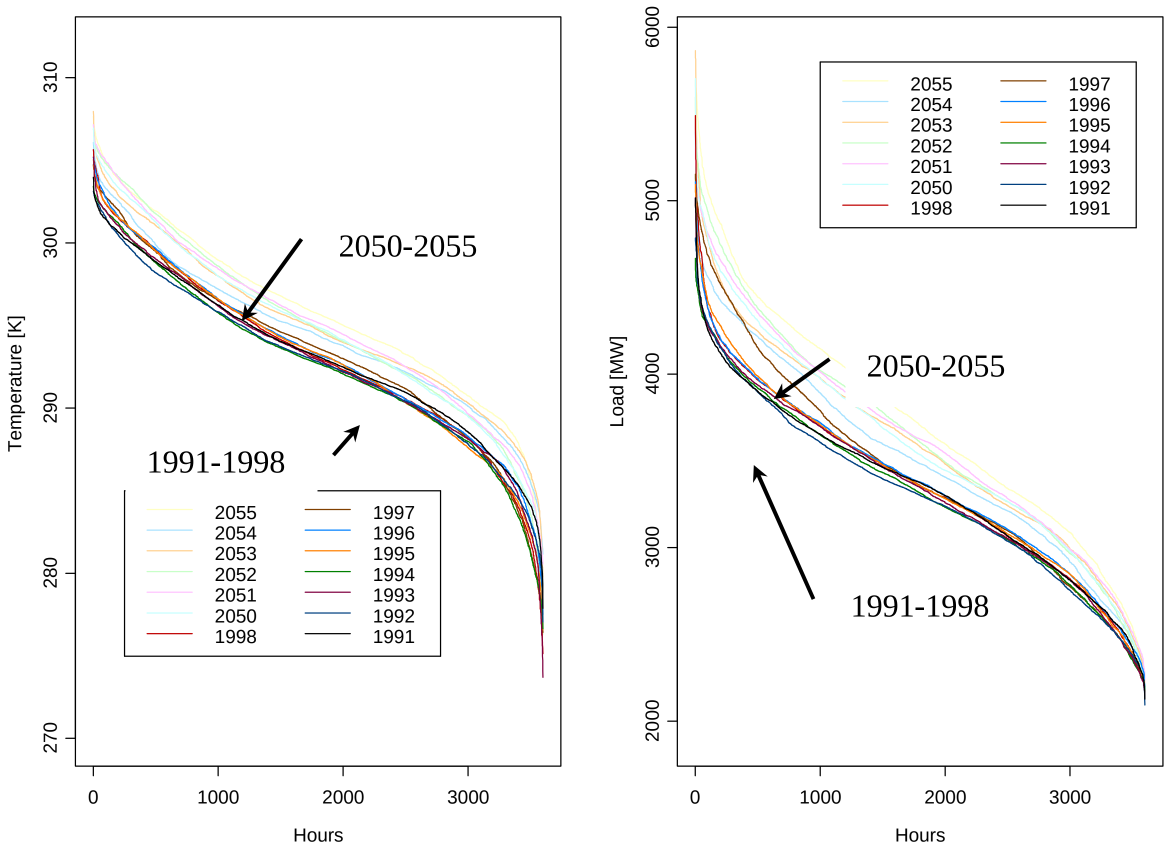

Figure 7: Inter-year temperature (left) and load (right) variability of Baltimore Gas and Electricity

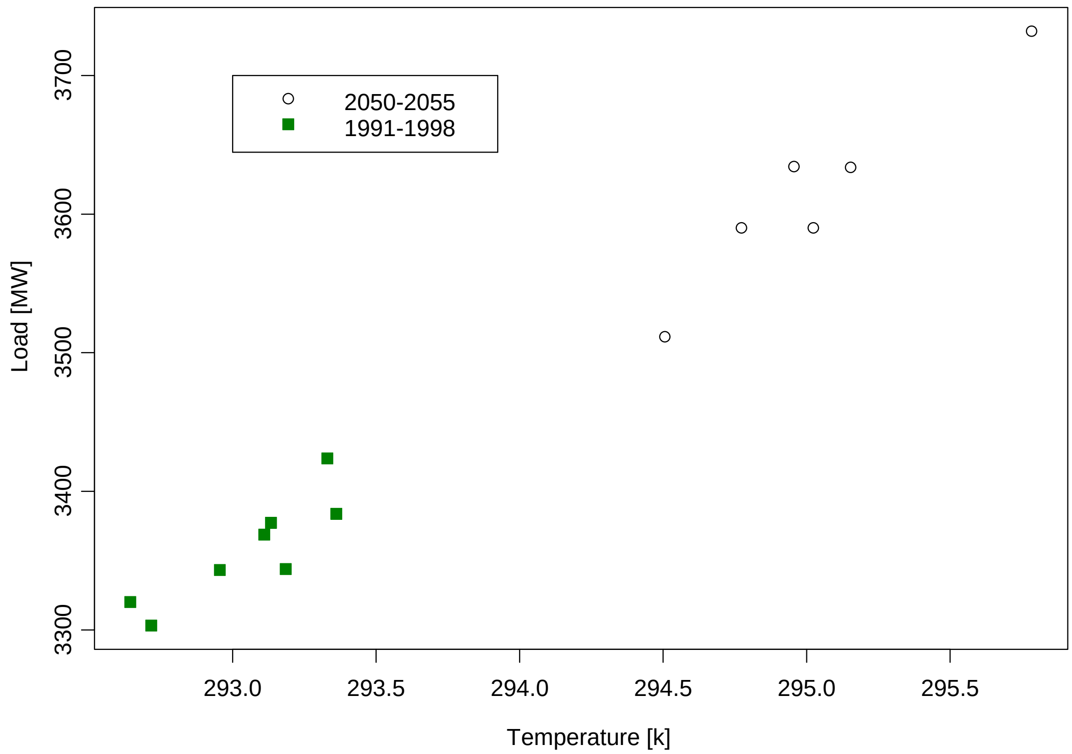

Figure 8: Plot of summer average temperature against load for Baltimore Gas & Electric

4. Long-run emission simulations

In contrast to the short-run analysis of Section 2, this section explores the long-run effect of climate change on NO_x emissions. In particular, the geographic coverage of the analyses in this section is expanded from the mid-Atlantic region represented by PJM to include upwind Midwestern and Southern states in ECAR.

Figure 9 provides a flowchart of the research tasks of the long-run analysis. First, in the long-run, the load duration curve has to reflect not only demand growth due to macroeconomic factors but also the changes in load shape because of the installation of climate-sensitive energy-using equipment by consumers. We developed a procedure in Section 4.1 that constructs such a curve under alternative climate scenarios. This procedure takes into account both the short-run load variability induced by weather variations as well as long-run load shape changes in response to climate change; the procedure does so by imposing short-run temperature induced fluctuations (resulting from changes in consumer utilization of energy using equipment) upon long-run average load distributions (reflecting changes in equipment stock due to average climate). Second, to estimate future mixes of generation capacity based on the predicted future load duration curve developed in Step 1, we deploy two approaches in Section 4.2: (a) graphical screening curve analysis and (b) least-cost LP analysis. The third and final step is to simulate hourly NO_x emissions. Although the emissions are capped throughout the season, the increase in demand under warmer climate conditions will have a redistributive effect on hourly emissions; that is, larger impact during peak periods than that during off-peak periods. Each step will be discussed in the below subsections.

Step 1: Construct long-run NEMS adjusted load duration curve and hourly loads consistent with hourly temperature scenarios and long run climate averages (Section 4.1)

↓

Step 2: Analyze future capacity mixture using screening curve approach and transmission-unconstrained long run LP model (Section 4.2)

↓

Step 3: Generate year-specific hourly NO_x emissions using a long-run capacity expansion models with fixed capacity mix from Step 2 (Section 4.3)

Figure 9: Flowchart of long-run analysis

4.1 Step 1: Adjust Long-Run Demand Based on NEMS Load Shape

4.1.1 Procedures. For the long-run emissions scenarios, we use 2025 loads as forecast by NEMS (National Energy Modeling System).[6] Loads in that year are used because that is the last year simulated in NEMS system.[7] There are five substeps in the procedure that adjusts long-run load shapes for the five month ozone season based upon the output

---

[6] The National Energy Modeling System is an integrated system of models of individual energy sectors in the US economy that is used by the US Energy Information Agency in policy analyses and its Annual Energy Outlook (Gabriel et al., 2001). This model is used because it is in the public domain, and does represent linkages between heating- and cooling degree-days and electricity consumption over time in the residential and commercial sectors.

[7] There is an inconsistency in using 2025 electricity loads and 2050 climate. The use of 2025 NEMS results means that we can also a set of generation cost and emissions assumptions for that year from the NEMS database, allowing us to construct a consistent set of load and generation assumptions for a scenario year well into the future. This is sufficient for our purpose, which is to illustrate the use of the emissions modeling system to project emissions under climate change scenarios.

of short-run load simulation models to produce a time series of hourly loads that is consistent with both:

- a given series of hourly temperatures and
- an assumed average climate.

These steps include (1a) constructing a load duration curve consisting of several load blocks, each with a constant load, using NEMS definitions; (1b) analyzing average cooling degree-day forecasts (CDD) for the 2050 decade based upon temperature time series from the Goddard Institute of Space Sciences GCM; (1c) calculating hourly loads under assumption of a stated increase in CDD in a short run forecast model (done by GWU); (1d) using hourly temperatures from the GISS GCM to calculate hourly-specific adjustment ratios to allow superposition of those short run fluctuations upon an arbitrary block load duration curve; and (1e) finally estimating the hourly loads that is consistent with the hourly temperatures and average climate assumed.

We first present the notation (variables and parameters) used by this procedure in Table 6, followed by a flowchart (Figure 10) and description of each step.

Table 6. Variables, parameters and set used in long-run demand analysis

Variables	Definition
$h \in H(b)$	Set of hours *h* that belong to the NEMS load block *b*.
$b \in B^s(b)(b \in B^w(b))$	Subset of blocks belong to the five months of the summer ozone season (winter time)
$y \in Y^{90s}$	Set of years *y* that belong to years in the GISS GCM 1990s simulations considered here (1991-1998) .
$y \in Y^{50s}$	Set of year *y* that belong to years in the GISS GCM 2050s simulations considered here (2050-2055)
$\vert H(b) \vert$	Number of hours in block b
$D_{h,y}^{GW}$	Hourly demand (Load, in MW) estimated by the GWU short-run (artificial neural network) models calibrated to recent load data for the study region, using as an input GCM temperature data in hour *h* and year *y*. This reflects short-run responses to temperature variations in year *y* under a GCM scenario, and are not adjusted for long run changes in consumer capital stock. $y \in \{1991-1998, 2050-2055\}$
$L_{b,y}^{GW}$	Average $D_{h,y}^{GW}$ over $h \in H(b)$.
$L_{b,90s}^{GW}$ ($L_{b,50s}^{GW}$)	Average $D_{h,y}^{GW}$ over $h \in H(b)$ over years 1991-1998 (2050-2055)
$L_{b,2025}^{NEMS}$	Load for block b in year 2025 under NEMS load shape, base case (no change in climate). These loads are available by NEMS regions (for example, one region corresponds to PJM).
$L_{b,2025}^{NEMS}(X)$	$L_{b,2025}^{NEMS}$ under an assumed increase in cooling degree-days of $\Delta CDD = X$ in year 2025. This was accomplished by running the commercial and residential electricity demand modules of NEMS (GWU).
$D_{h,y}^{NEMS}$	Hourly demand (load) adjusted to be consistent with the NEMS load duration curve (i.e., the average of $D_{h,y}^{NEMS}$ in each block equals the NEMS projection).
ΔCDD	Assumed change in the CDD (cooling degree days) under the climate change scenario. For our simulations, based on Philadelphia and Washington, DC region temperatures from the GISS GCM simulations, this was calculated to be 414 F° on average between the 1990s and 2050s simulations.
$\lambda_{b,y}$	NEMS block adjustment factor for block *b* in year *y*

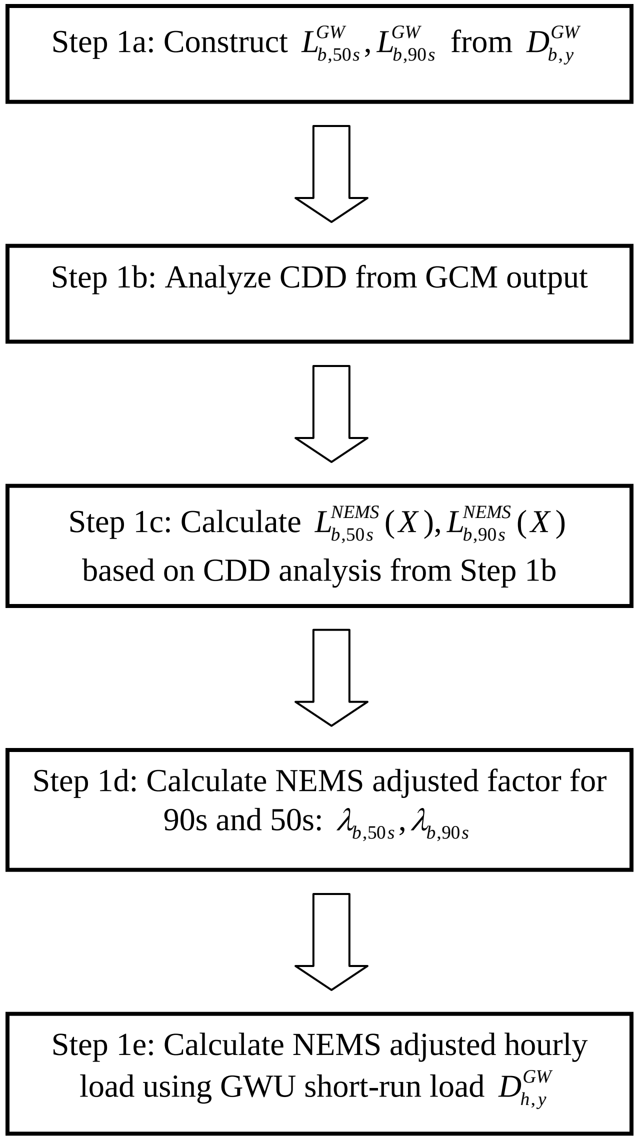

Figure 10: Flowchart for generating NEMS adjusted load duration curves

The five steps associated with Figure 10 is described in below in more detail:

Step 1a: The first step is to construct an initial block-based load-duration curve for block *b* in year *y*, i.e., $L^{GW}_{b,y}$, based on GWU hourly load data ($D^{GW}_{h,y}$). The load block is defined by load group and load segment. In particular, there are three NEMS load groups, depending on the time of a given day: mid day (9:00 -16:00), morning evening (6:00–8:00 and 17:00-24:00) and night (1:00-5:00); and three levels of load segments in each load group: highest 1%, next highest 33% and the lowest 66%. The following equation shows that the block height of a given block is the average hourly load in that block.

$$L_{b,y}^{GW} = \frac{\sum_{h \in H(b)} D_{h,y}^{GW}}{|H(b)|}, \text{ for } \forall b, y \qquad (7)$$

Step 1b: The second step calculates the block-based load assuming an X °F-day increase in the CDD ($L_{b,y}^{NEMS}(X)$). The year 2025 output from the NEMS analyses by GWU simulates two specific conditions: an 0 F°-day and 150 F°-day increase in the CDD (base case and climate change case, respectively). That is, NEMS gives us $L_{b,y}^{NEMS}(0)$ and $L_{b,y}^{NEMS}(150)$. We assume $L_{b,y}^{NEMS}(X)$ is approximately linear in X. Thus, the term $L_{b,y}^{NEMS}(X)$ then can be calculated by extrapolating a linear function that passes through $L_{b,2025}^{NEMS}(0)$ and $L_{b,2025}^{NEMS}(150)$. More specifically, it can be estimated by the following:

$$L_{b,2025}^{NEMS}(X) = (1 - \frac{X}{150})L_{b,2025}^{NEMS}(0) + (\frac{X}{150})L_{b,2025}^{NEMS}(150) \qquad (8)$$

Step 1c: We analyzed the results of the GISS GCM (as downscaled using MM5) by considering the average increase in the CDD in the MM5 cells (ground level) corresponding to two cities (Baltimore and Philadelphia).[8] The average difference in CDDs between the 1990s and 2050s scenarios is 414°F-day. The below table shows the input data for this analysis.

[8] The GISS GCM scenario were downscaled using MM5, a regional scale meteorological model that is driven by the coarser scale GCM. A 36 km grid scale was used. (Hugh? Is that right?)

Table 7: Ground-level May-September temperatures for 36 km MM5 cells containing Baltimore and Philadelphia from GISS GCM/MM5 scenarios for 1991-1998 and 2050-2055

Scenario	Baltimore		Philadelphia	
Year	CDD	^{o}F (mean)	CDD	^{o}F (mean)
1991	583.0	68.1	548.3	67.8
1992	496.8	67.2	478.7	66.9
1993	603.5	67.7	601.6	67.5
1994	561.5	67.1	514.7	66.7
1995	638.0	68.0	585.3	67.4
1996	631.7	68.4	571.8	67.8
1997	731.5	68.3	718.7	68.2
1998	633.3	67.9	570.2	67.4
2050	984.7	70.9	936.0	70.6
2051	1046.1	71.6	997.7	71.3
2052	1048.0	71.3	992.6	70.7
2053	1014.5	71.4	952.0	70.8
2054	874.2	70.5	853.2	70.2
2055	1223.7	72.7	1149.4	72.2

This CDD data means that we need to obtain $L_{b,2025}^{NEMS}(414)$ to simulate the change in average demands (by block) in an average year under a 2050s climate. This corresponds to an average of 3.5°F increase in ground-level temperature for the ozone season.[9]

Step 1d: The next step is to calculate NEMS block-adjustment factor $\lambda_{b,y\in Y^{90s}}$ ($\lambda_{b,y\in Y^{50s}}$) for years in 1990s (2050s), where the 1990s (or "90s") represent the no climate change scenario and the 2050s ("50s") are the climate change scenario. These factors are used to adjust the GWU short-run hourly simulations so that their averages are consistent with the NEMS load blocks. Those factors are calculated in the following two equations.

$$\lambda_{b,y\in Y^{90s}} = L_{b,2025}^{NEMS}(X=0)/L_{b,y^{90s}}^{GW}\text{, for } \forall b, y \in Y^{90s} \qquad (9)$$

$$\lambda_{b,y\in Y^{50s}} = L_{b,2025}^{NEMS}(X=414)/L_{b,y^{50s}}^{GW} \text{ for } \forall b, y\in Y^{50s} \qquad (10)$$

Step 1e: Calculate NEMS adjusted hourly load using GW hourly load.

[9] The CDD metric for potential cooling demand is calculated by taking the positive difference between average daily temperature and 72^{o}, and summing this over all days in the cooling season. For instance, two days, one with an average temperature of 75^{o} and the other with an average of 80^{o}, results in 11 CDDs.

$$D^{NEMS}_{h\in H(b), y\in Y^{90s}} = D^{GW}_{h,y} \lambda_{b,y^{90s}} \text{, for } \forall h, y \in Y^{90s} \tag{11}$$

$$D^{NEMS}_{h\in H(b), y\in Y^{50s}} = D^{GW}_{h,y} \lambda_{b,y^{50s}} \text{, for } \forall h, y \in Y^{50s} \tag{12}$$

The above procedure produces hourly loads consistent with hourly meteorology and long run climate change (in the 2050s scenarios). Meanwhile, the construction of the nonozone season ("winter") hourly load is rather straightforward: we use the output from NEMS analysis directly. Hourly loads are not necessary because we do not do air quality simulations for the winter months. For example, if there are N hours in block b during winter with a load of $L^{NEMS}_{b,y}$ based on NEMS output, we then create N number of hours of such load in the load duration curve.

$$D^{NEMS}_{h\in H(b), b\in H^{w}(b), y} = L^{NEMS}_{b\in B^{W}(b), y} \tag{13}$$

Therefore, in this illustrative analysis, we do not consider how long run climate change affects winter loads. This can be an important omission, if changes in winter loads (due to decreases in heating degree-days) translates into changes in the generation capacity mix.

4.1.2 Results. The result of the above process is an estimated distribution of hourly loads in each year considered, representing a total of 14 years of GISS GCM climate simulations under conditions for the 1990s and 2050s (1991-1998, 2050-2055). The hourly load distributions are shown as load duration curves in Figure 11 and Figure 12, respectively.

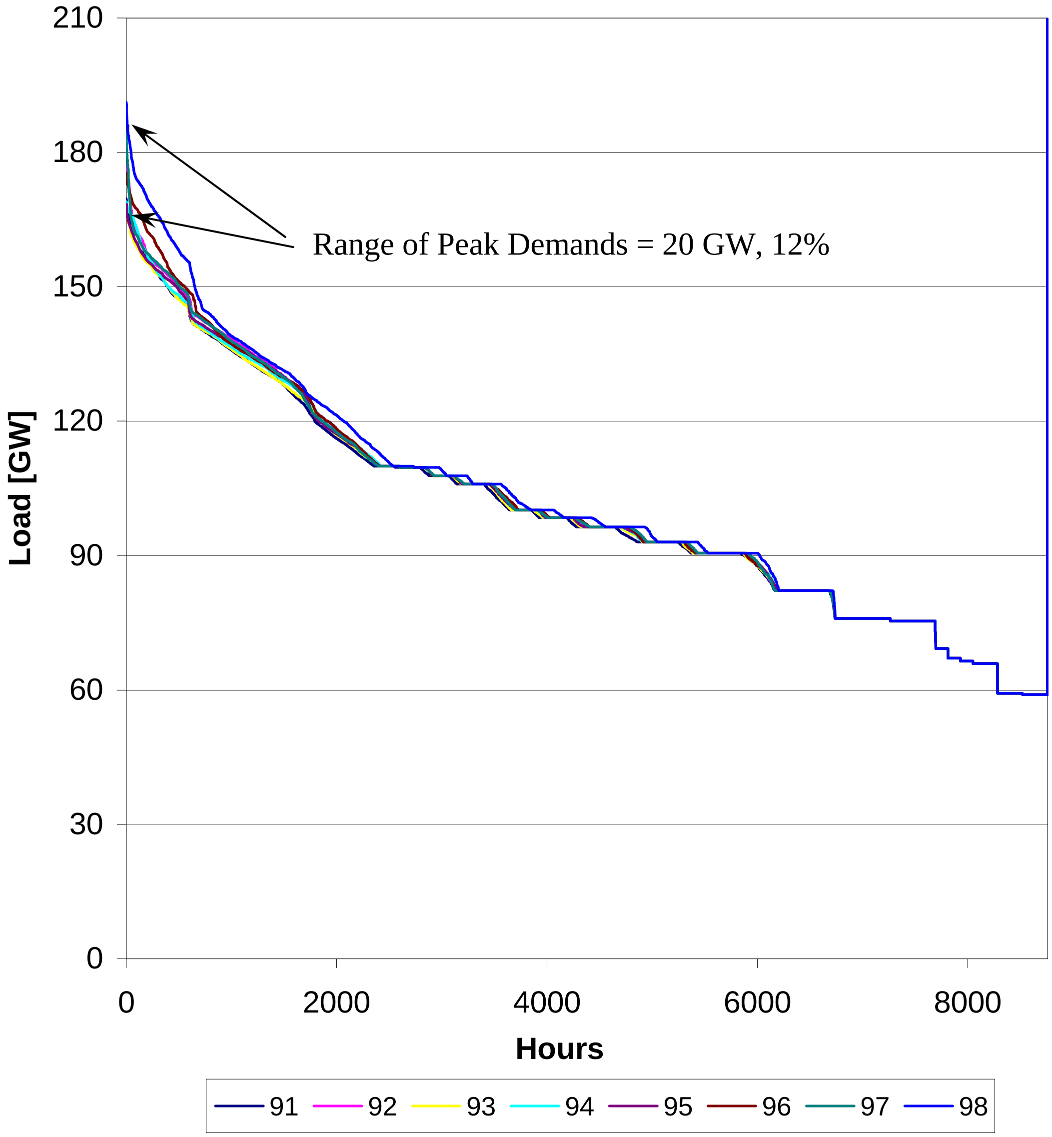

Figure 11: NEMS Adjusted Load Duration Curves for Years 1991-1998

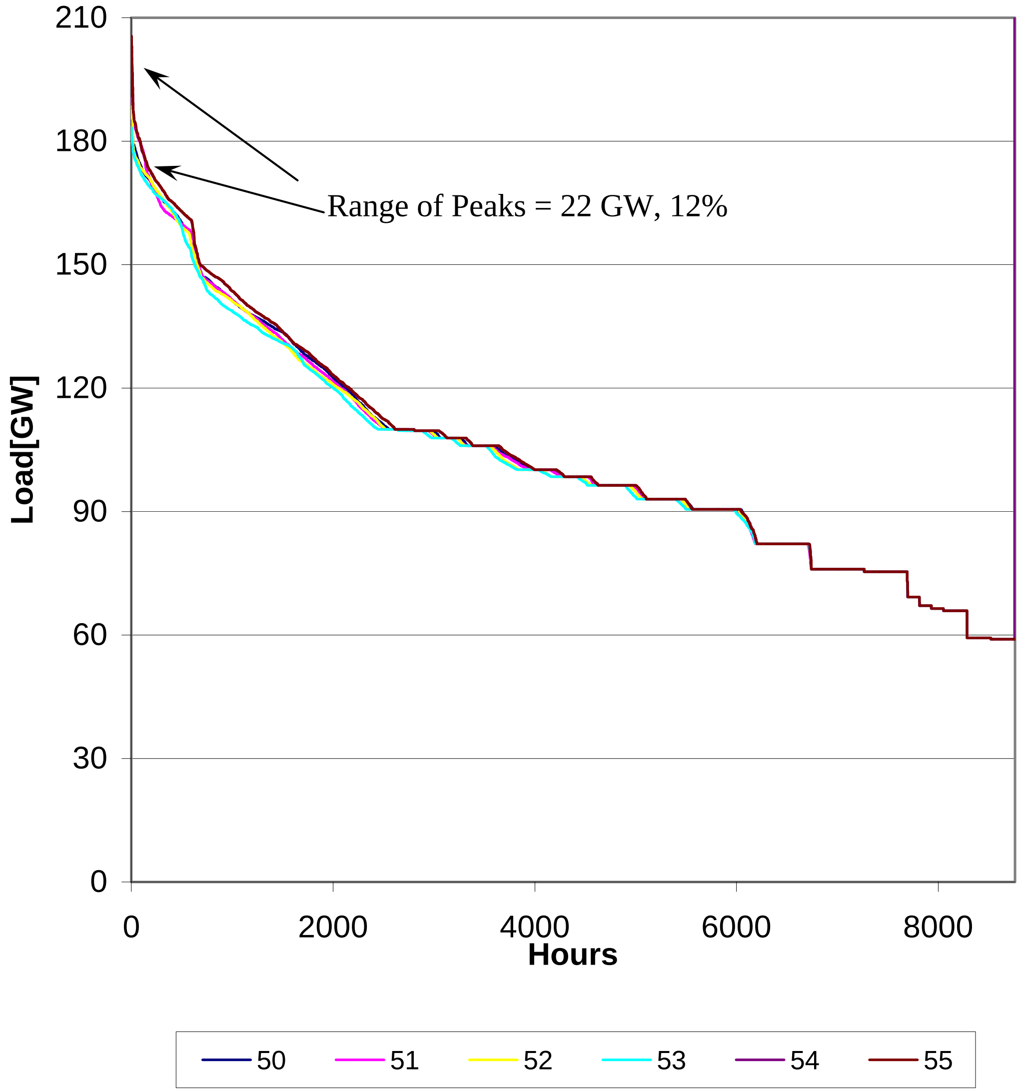

Figure 12: NEMS Adjusted Load Duration Curves for Years 2050-2055

Table 8 presents the descriptive statistics for NEMS adjusted load duration curves for 1991-1998 and 2050-2055, respectively for the PJM and (partial) ECAR region considered in this study. In particular, we report peak load, total energy consumption, average load and its standard deviation. Given a 414 degree-day increment in the CDD for the 2050s series compared to 1990s series, the average peak load for the 2050s series is 193.9 GW, which is 11% higher than that the 174.6 GW average peak for the 1990s. The effect of CDD on the average load is much less, however. The average ozone season load in 2050s is only 4.4% higher that that in 1990s.[10] The 414 degree-day increase results from an average 4°F increase (approximately) in the summer time temperature.

Table 8: Descriptive statistics for NEMS adjusted load duration curve for years 1991-1998 and 2050-2055

Year	Peak Demand [GW]	Total Energy [TWh]	Mean [GW][a]	Standard Deviation [GW]
1991	169.7	876.7	100.1	25.3
1992	177.5	883.7	100.9	26.1
1993	166.5	878.3	100.3	25.3
1994	169.0	881.1	100.6	25.6
1995	168.5	881.3	100.6	25.6
1996	175.4	885.5	101.1	26.5
1997	184.2	883.2	100.8	26.0
1998	185.9	901.6	102.9	28.5
2050	191.3	897.0	102.4	28.1
2051	190.4	899.9	102.7	28.5
2052	188.4	897.8	102.5	28.2
2053	183.3	893.7	102.0	27.8
2054	204.5	906.7	103.5	29.3
2055	205.5	907.0	103.5	29.4

a. The seven winter months were assumed to have no change; thus, all changes are in the five summer months

Thus, the average temperature sensitivity is roughly 2.5% per °F for the peak period, but only 1.1% per °F on average for the ozone season. This can be compared to the short run sensitivity described in Section 2, which was 2.15% per °F on average. There are at least two possible reasons for this difference:

1. In the long run, investment in more efficient energy using equipment (since greater utilization would more easily justify such investment) dampens the energy use response to temperature changes.
2. Modeling limitations in NEMS that result in understating of temperature response. For instance, choices from among possible energy-use technologies may be restricted.

The results in the table also show that the inter-year variability measured by the range of the peak demands is 11.6% and 12.1%, respectively, for the 1990s and 2050s series (relative to the lowest peak in the period). For example, the range of peak-hour load for 1990s series is 19.4 GW, which is 11.6% of the lowest peak load (166.5 GW) among the years 1990-1998. These results show that year-to-year variation in weather can result in

[10] The total demand over the entire year is only 2% in the 2050s than 1990s, according to the Table. However, all the load increase in the 2050s is assumed to take place in the summer months.

significant changes in loads. Because hotter, higher load years will also be more likely to have conditions favorable to ozone episodes, there is likely to be great variation in year-to-year average ozone levels as well as numbers of severe ozone episodes. More importantly, from the point of view of impact analysis, consideration of emissions from an average load year will likely understate the ambient ozone impact. That is, the average ozone concentration (or average days of ozone NAAQS exceedences per year) will be less for an average load year than when calculated over a sample of years reflecting year-to-year temperature variations.[11]

The corresponding range for average summer-time loads is 6.7% for the 1990s and 3.5% for the 2050s. The relatively high value for the 1990s arises from one very high year (1998), and so the difference in ranges could be due to sample error.

4.2. Step 2. Long-Run Generation Mix Analysis

The purpose of this step is to estimate the mix of generation capacity that a competitive electric power market will provide in response to a given distribution of electricity loads.[12] The output form long-run screening curve analysis will give the amount of new capacity in MW by types of technologies. In this project, we deploy two approaches: a graphical approach ("screening curves"), and LP transmission-unconstrained model. The screening curve approach is excellent for communicating the intuition behind different solutions. The LP, although more complex, has the advantage of being able to account for side constraints such as reserve margin requirements, limited or intermittent energy output (hydro, wind), emissions limitations, and limits on annual capacity factors (ratios of actual output to capacity). For this reason, screening curve results are presented only to provide some insights, and it is the result of the LP model that is used to generate emission scenarios.

Table 9 lists the generation and emission characteristics of the seven most attractive generation technologies we considered (a number of others were also analyzed, but were inferior to these). The primary data source is EIA (EIA, 2004), supplemented with information about emission rates from USEPA IPM (USEPA, 2005d).

[11] This results from Jensen's inequality: the expected value of a convex function E(f(X)) is at least equal to, and generally greater than f(E(X)), the function evaluated at the expected value of the input. In this context, X is annual weather conditions, and f(X) is an air quality index such as frequency of ozone NAAQS exceedences. For instance, in a cool year, there may be 2 exceedences, while an average year might have 5 exceedences. But a hot year might have 23 exceedences. If those three possibilities are equiprobable, then $E(f(X)) = (2+5+23)/3 = 10 > f(E(X)) = 5$. It is a well accepted principle of risk analysis that if a system is nonlinear, a full distribution of inputs (such as meteorology) should be considered, not just average or typical conditions.

[12] A competitive response is equivalent to a least-cost optimum, if demand is perfectly inelastic, as assumed here. Therefore, our model can be viewed as simulating a competitive market, or as simulating the result of an efficient regulatory process that induces power companies to make least-cost investments.

Table 9: Data for Seven Selected Generation Technologies (Source: NEMS)

	Levelized Cost [$/kWy]	Fuel Cost [$/ MBTU]	Variable Cost [$/ MWh]	Fixed O&M [$/ MW-y]	Heat Rate [BTU/ kWh]	Cap Factor []	Forced Outage Rate []	SO_2 Rate [tons/ MWh]	NO_x Rate [lbs/ mmBTU]
Scrubbed Coal	149.5	1.23	3.10	24.81	9000	0.84	0.0470	0.01015	0.11
IGCC	174.59	4.88	2.07	34.11	8000	0.80	0.0181	0	0.11
Adv. Nuclear	230.65	0.00	0.43	59.17	10400	0.90	0.0541	0	0
Conventional CC	69.727	4.88	2.07	12.40	7444	0.90	0.0181	0	0.02
Advanced CC	76.889	4.88	2.07	10.34	6268	0.90	0.0181	0	0.02
Conventional CT	53.536	4.88	4.14	10.34	10878	0.92	0.0235	0	0.08
Advanced CT	60.33	4.88	3.10	8.27	9289	0.92	0.0235	0	0.08

4.2.1. Screening Curve Approach. This graphic approach relies on plotting total cost of generation from 1 MW of a power plant against the capacity factor (CF) of the plant.[13] This results in a linear function, with the y-intercept being the fixed (annualized investment) cost and the slope being the marginal operating cost per MW-year. For any given CF, the plant type whose line is the lowest is the most economic means of meeting the demand which is exceeded that fraction of the year. As a result, we can then superimpose those curves on the load duration curve, which allows us to identify the ranges of loads that should be served by each generation type. This assumes that generation is perfectly reliable and that there are no constraints other than that total generation has to meet demand in each hour.

Figure 13 is a screening curve plot that uses the PJM and ECAR load duration curve based on 1990s and 2050s temperature data derived in Step 1. Let capital P donates the complete set of future technologies and p is a subset of P. The relationship between load (L) and capacity factor (z) is described by a function K, i.e., $K(z)\rightarrow L$, where z is the capacity factor. A set of 18 possible technologies (P) are included in Figure 13. For a given capacity factor, a technology is said to be the dominating technology if no technology has a lower total cost. In other words, the key step in screening curve analysis is to find the set of technologies defining the lower envelope of the cost curve. Figure 13 implies that the seven technologies (p=1,..,7) that constitute the lower envelope are conventional combustion turbine, advanced combustion turbine, conventional combined cycle, advanced combined cycle, coal plants with scrubbers, IGCC and advanced nuclear. The generation characteristics and cost information about seven technologies in the lower envelope are displayed in Table 9, above.

[13] Turvey and Anderson, op. cit.; S. Stoft, *Power System Economics, Designing Markets for Electricity*, IEEE, Piscataway, NJ, 2002.

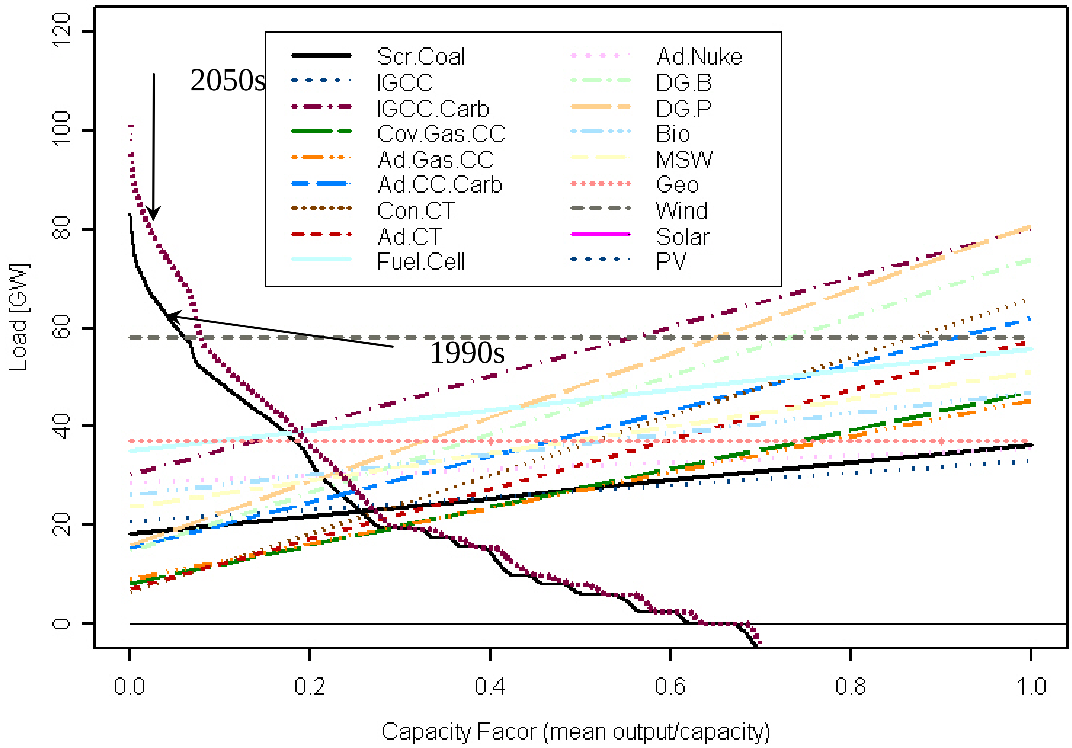

Figure 13: Traditional Graphical Screening Curve Analysis with All 18 Generation Technologies

To estimate the amount of the new capacity of each type ($\boldsymbol{u_p}$), we first calculate the set of the intersection points $\boldsymbol{z_p}$, which can be obtained by solving a series of the system of two linear equations. For example, the intersection for the first two technologies, i.e., conventional combustion turbine and advanced combustion turbine, yields a capacity factor or x-axis equal to 0.0846($\boldsymbol{z_1}$). Other intersection points can be estimated by applying the same procedure: 0.09565 ($\boldsymbol{z_2}$), 0.3115 ($\boldsymbol{z_3}$), 0.5117 ($\boldsymbol{z_4}$) and 0.5803 ($\boldsymbol{z_5}$). Except for the first and last technology, the estimated capacity for p technology can be calculated by $K(z_p)$-$K(z_{p+1})$. For p=1, $K(z_0)$=Max(L); and for p=7, $K(z_8)$=Min(L). In addition, we assume that fossil-fueled plants built prior to 1965 will be retired by 2025, but that post-1964 fossil plants and all nuclear plants will still be in service. The latter are subtracted from demand, so that the load curve in the screening curve figures is only the load net of existing capacity.

Figure 14 plots the screening curve of seven generation technologies associated with the lower envelope. The results of the screening curve analysis are summarized in the Table 10. They show that the higher demands associated with the 2050s scenario results in more capacity being required, primarily peaking capacity. In particular, significantly more peaking generators are added in the 2050s scenario, i.e., 13,000 MW of combustion turbines, because of the large increase in peak period demands. In addition, a moderate amount more (about 4000 MW) of baseload generators, i.e., IGCC, and combined cycle generators are constructed for the 2050s than for the 1990s climate..

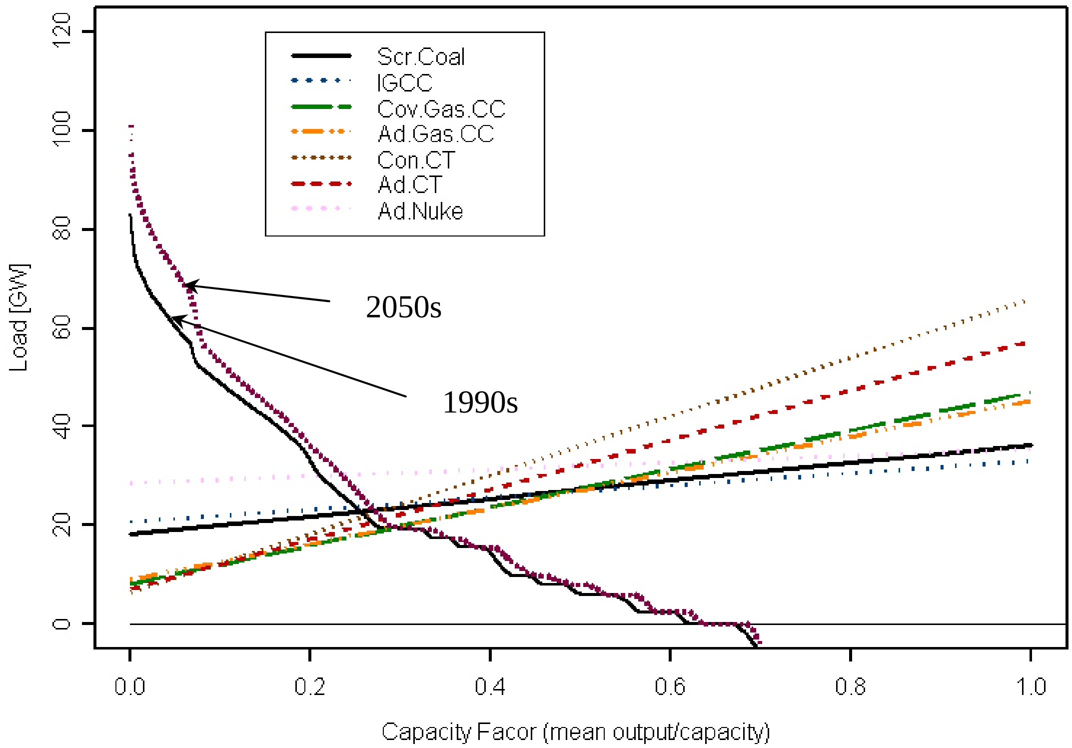

Figure 14: Traditional graphic screening curve analysis with technologies associated with lower envelope

Table 10: Results from screening curve analyses using load duration curve for 1990s (2050s) (Note: LDC adjusted downwards for existing capacity)

Technology	Capacity [MW]
Conventional C.T.	31,990 (45,071)
Advanced C.T.	1,597 (2,083)
Conventional C.C.	30,164 (34,275)
Advanced C.C.	13,236 (11,568)
IGCC	6,023 (7,856)
Scrubber Coal	0 (0)
Total	83,011 (100,854)

The traditional screening analysis is rather intuitive, but also subject to two limitations. First, emissions cap constraints cannot be explicitly considered, therefore, the result could be a bias in favor of technologies with higher emission rates. Also, other side constraints cannot be considered. Finally, retirement decisions for old generators need to be exogenously determined in the screening curve analysis. In the next section, we present an alternate approach that allows these two limitations to be corrected.

4.2.2. Linear Programming Transmission-Unconstrained Least-Cost Model. The second approach to generating future capacity mixes is a least-cost LP capacity expansion model without transmission constraint.[14] The motivation of this approach is that the screening curve method presented in the Section 4.2.1 is only an approximate approach, where other reliability and environmental constraints such as reserve margins and emissions cap are not included in the analyses. As a result, high-polluting technologies may be overbuilt under the screening curve approach, and too little capacity might be constructed. In contrast to a complete long-run least cost capacity expansion model, this model disregards the network. In other words, all generators are assumed to produce and sell power at the same market. In addition, for coal generators built after 1965, the retirement decision is endogenously determining by associating them with an annual cost of 5000 $/kW if the owners decide to maintain them online. The load used in the LP transmission-unconstrained leas-cost analyses is the average of 1991-1998 and 2050-2055, respectively for the PJM and ECAR regions included in the analysis. (Note that this is a larger region than considered in the screening curve analysis, so the results are not strictly comparable.)

Long-run capacity expansion decisions are assumed to be based on average load distributions over a decade; that is, the average over 1991-1998 and 2050-2055 for the no climate change and climate change scenarios, respectively. The specific load associated with a year, e.g., 1991, is simply a realization of a random draw from that underlying load distribution. In other words, we simulate two load distributions – the 1990s and 2050s – and they differ by 414°F-days in terms of CDD. The capacity decisions are assumed to be based on an average distribution of load, calculated by averaging the load duration curves across years within a decade, as the next figure shows. A 20 load block approximation of these curves is used (Figure 15), consistent with the NEMS load block definitions.

[14] See (Turvey and Anderson, 1997) for a review of linear programming methods for capacity mix selection, or (Hobbs, 1995).

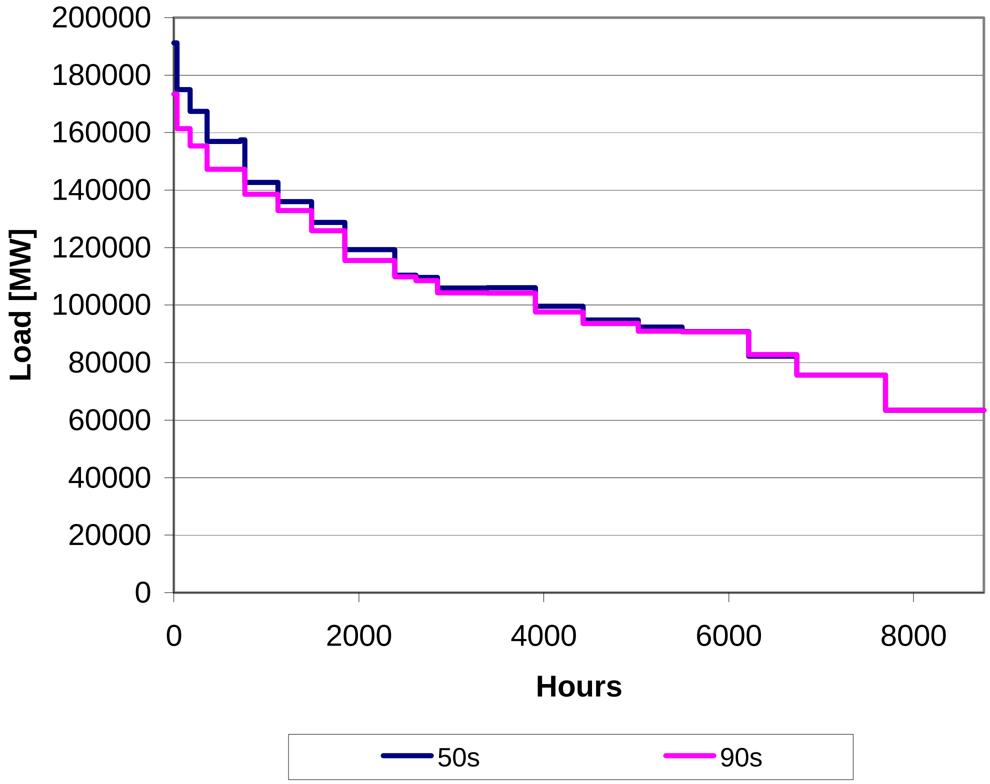

Figure 15: Average approximated load blocks used in LP generation mix analyses for 1990s and 2050s

Table 11 summarizes the definition of the variables, parameters and set that are used in the model. The mathematical equations for the model are presented afterward. Table 9 above contains the data for the seven generation technologies considered in the LP analysis.

One of strengths of the LP approach is its ability to explicitly model the NO_x seasonal constraint. However, the tightness of the NO_x cap depends on which policy that a model intends to simulate. For the purpose of the current project, we will use NO_x cap under the Clear Air Interstate Rule (CAIR); for our study region, we estimated a 2025 limit of 129,110 tons per ozone season.

Table 11: Definitions of Variables, Parameters and Sets used in LP Generation Mix Analysis

Variables/Parameters/Set	Definition
h	index of existing plants
p	Type of new capacity
t	Period (load block)
$t \in oz(t)$	Set of periods subject to ozone season
$h \in H^{OTC}$ or $p \in P^{OTC}$	Set of existing or new sources subject to NO_x cap
x_{ht}	Output level of plant *h* in the *t* period [MW]
u_p	Amount of new installed capacity of *p* type[MW]
v_{pt}	Output level by new plant of type *p* in period *t* [MW]
$\rho_{ht}(\rho_{pt}^N)$	Dual variables for capacity constraint for generator *h* (*p*) in period *t* [$/MW]
$\kappa_h(\kappa_p^N)$	Dual variable for capacity factor constraint for generator *h* (*p*) [$/MWh]
p_t	Power price in period t [$/MWh]
p^{NOX}	NO_x allowances price [$/ton]
p^{RES}	Dual of capacity reserve constraint [$/MW]
B_t	Period width in period (load block) *t* [Hours]
FC_h	Fuel cost of plant h (including SO_2 allowances) [$/MWh]
CAP_h	Capacity of plant *h* [MW]
$FOR_h(FOR_p^N)$	Forced outage rate for plant *h* (*p*)[%]
NFC_p	Fuel cost of new plant *p* [$/MWh]
$E_h(E_p^N)$	NO_x emission rate for plant *h* (*p)* [tons/MWh]
LC_p	Levelized capital cost of new plant *p* [$/MW/yr]
$CF_h(CF_p^N)$	Maximum capacity factor of plant type *h* (*p)* [%]
$NFOR_p$	Forced outage rate for new type of plant *p* [%]
$L_t(L_{peak})$	Load in the period *t* (peak) [MW]
RM	Reserve margin [%]
TC	Total cost [$]

The equations (14)-(21) define a long-run LP transmission-constrained capacity expansion model (LRexp). Problem LRexp can be viewed as a variant of the short-run LP dispatch model, i.e., STM-summer (Section 2, above). Both models attempt to minimize the overall cost. However, in contrast to the production cost objective in the STM-summer, the objective function in the LRexp includes operating cost for newly installed capacity as well as its construction cost. One year's time is considered, so capital costs are expressed in levelized (or annualized) terms. In addition to capacity constraints in the equation (14)-(15), the demand constraint in equation (19) and NOx cap constraint in equation (21), there are two sets of new constraints. First, a capacity factor constraint in equation (17) and (18) limits the annual total energy can be produced by a generator. Next, equation (20) is a capacity reserve margin constraint imposed by the regulator to ensure that adequate capacity is built for the system, including a reserve margin in order to deal with load or generator outage contingencies.

$$MINTC = \sum_{h,t} B_t x_{ht} C_h + \sum_{p,t} B_t v_{pt} NC_p + \sum_{p} LC_p u_p \qquad (14)$$

Subject to:

$$x_{ht} \le (1 - FOR_h) CAP_h, \forall h,t \qquad (\rho_{ht}) \quad (15)$$

$$v_{pt} \le (1 - FOR_p^N) u_p, \forall p,t \qquad (\rho_{pt}^N) \quad (16)$$

$$\sum_t B_t x_{ht} \le 8760(1 - FOR_h) CAP_h CF_h, \forall h \qquad (\kappa_h) \quad (17)$$

$$\sum_t B_t y_{ht} \le 8760(1 - FOR_p) u_p CF_p^N, \forall p \qquad (\kappa_p^N) \quad (18)$$

$$\sum_h B_t x_{ht} + \sum_p B_t v_{pt} \ge B_t L_t, \forall t \qquad (p_t) \quad (19)$$

$$\sum_h CAP_h + \sum_p u_p \ge L_{peak} RM \qquad (p^{RES}) \quad (20)$$

$$\sum_{h \in H^{OTC}, t} B_t E_h x_h + \sum_{p \in P^{OTC}, t} B_t E_p^N v_p \le \overline{NO_x} \qquad (p^{NOX}) \quad (21)$$

$$x_{ht}, y_{pt}, u_p \ge 0$$

We now consider the capacity mix solution constructed from the average, across years, of the 1990s load duration curves constructed in Step 1 of the long-run analysis for the portions of the PJM and ECAR regions we considered. First, there is no endogenous retirement since 5000\$/MW per year going forward cost for existing generators is much less than the levelized cost for constructing new generators. In other words, plant owners would rather keep old generators online than build a new set of generators, even if new generators are more efficient. Next, the model chooses several new technologies in addition to the existing plants: 66,424 MW of IGCC, 4,975 MW for conventional CC, and 37,434 MW for conventional CT. The estimated NO_x allowances price is 5,025 \$/ton. We now consider the results from the LP analysis using 2050s load data. Three technologies are chosen for investment by the LP model: IGCC, conventional CC, and conventional CT. The amount of new capacity is 67,247 MW of IGCC, 9,697 for conventional CC and 52,409 MW of conventional CT, respectively. Thus, it is CT capacity that is most strongly affected by climate warming (almost 15 GW more is added, consistent with the screening curve analysis). The NO_x allowances price increases from 5,025 to 7,581 \$/ton, reflecting the effect of increase in demand which increases pressure on the NO_x market. Overall, the total new installed capacity is larger using LP transmission-unconstrained approach. This is mainly due to the reserve margin requirement, i.e., equation (20).

4.3. Simulation of NO_x Hourly Emissions

The first two steps of the long-run analysis generate load distributions and mixes of generation capacity. The last step dispatches that capacity against hourly load distributions that are consistent with the hourly meteorology used in CMAQ so that, for instance, emissions are higher when the weather is warmer (and thus more conducive to ozone episodes, at least in the summer).

Although the total NO_x emissions are capped throughout the ozone season, an increase in energy demand under warmer climate conditions, i.e., a climate change effect, will have a redistributive effect on hourly NO_x emissions. That is, emissions during a particular hour incur an opportunity cost of not being able to emit in other hours. In particular, if the incremental NO_x emission occurs during peak demand hours when ozone concentrations are already high, the result could be more frequent occurrences of extreme ozone episodes.

4.3.1 Methodology. In Section 4.2, the LP generation mix model is solved based on a 20-block system – 10 blocks apiece for the ozone and non-ozone periods – using average load duration curves from the 1990s and 2050s series. The dispatch model of this section is formulated using just the ozone season's 10 load blocks, and also includes the transmission grid constraints, as in Section 2.

To illustrate the effect of climate change on NO_x hourly emissions, we simulate power dispatch for just two years: 1991 and 2055. The new capacity is kept at 1990s and 2050s levels (from Section 4.2) for 1991 and 2055, respectively. That is, the variables u_p are fixed at the 1990s and 2050s level for 1991 and 2055 simulations, respectively. These two years are chosen because they are the warmest (2055) and coolest year (1991) among these 14 years. We report the simulated results in terms of "emission duration curves" in the next subsection. However, the hourly time series of emissions by location is meant to be used in the SMOKE module of CMAQ.

4.3.2 Simulated Results for hourly NO_x emissions. Figure 16 plots the NO_x hourly emission duration curve for the two selected years (i.e., 1991 and 2055) over the summer period, which includes 150 days or 3600 hours. The width of each step corresponds to the number of the hours in that block. For example, the peak block contains only 35 hours.

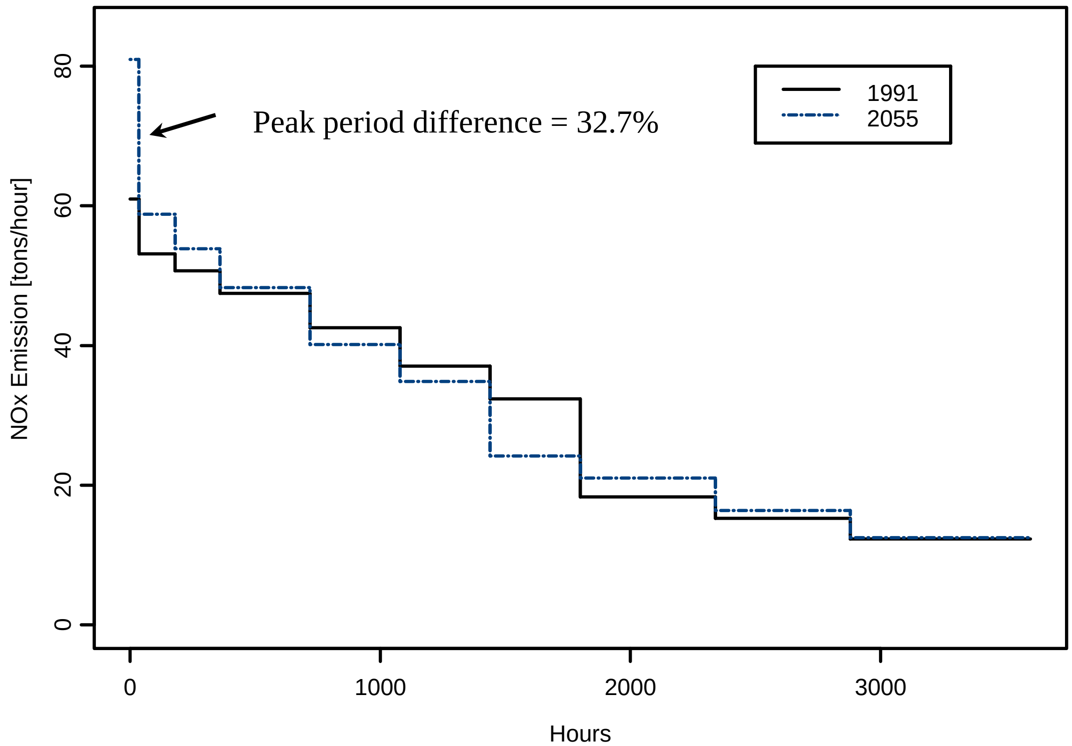

Figure 16: Simulated NO_x duration curves of two selected years: 1991 and 2055

Notice that since total NO_x emissions are regulated under cap, the area under the two curves is equivalent. However, under a warmer climate condition (i.e., 2055) the emission redistributive effect increases peak period NO_x emissions by 32.7% compared to the peak emissions under 1991's weather. Such an increase in the NO_x emissions also occurs for the next three, i.e., 2-4, as well as for blocks 8-10, but to a lesser extent. For instance, the increase in the fourth block is only marginal, roughly 1.7%. The increased NO_x emissions for 2055 in blocks 1-4 and 8-10 are then made up by the less emissions in the blocks 5-7.

In doing the dispatch analyses, we had assumed that the annual cap was firm, in the sense that no banking or borrowing from a bank of allowances is allowed. In reality, a particularly warm year with high generation will have higher emissions than cooler years because generation owners will take advantage of the bank. As a result, the high emissions we calculated during the peak period of 2055 would be even higher if this was allowed.[15]

[15] In future analyses we can simulate this by making more allowances available at a price equal to the shadow price of allowances from the long run analysis.

Finally, Figure 17 plots the hourly NO_x emission against hourly temperature of Baltimore for the hour 14:00 in all days 1991 and 2055. The trend is clearly upward with a correlation of 0.51 and 0.62 for two series, indicating that higher temperatures are associated with higher NO_x emissions. This, together with findings in Figure 16, implies that a significant increase in the NO_x emissions would occur under climate warming during precisely the times when ozone episodes are most likely.

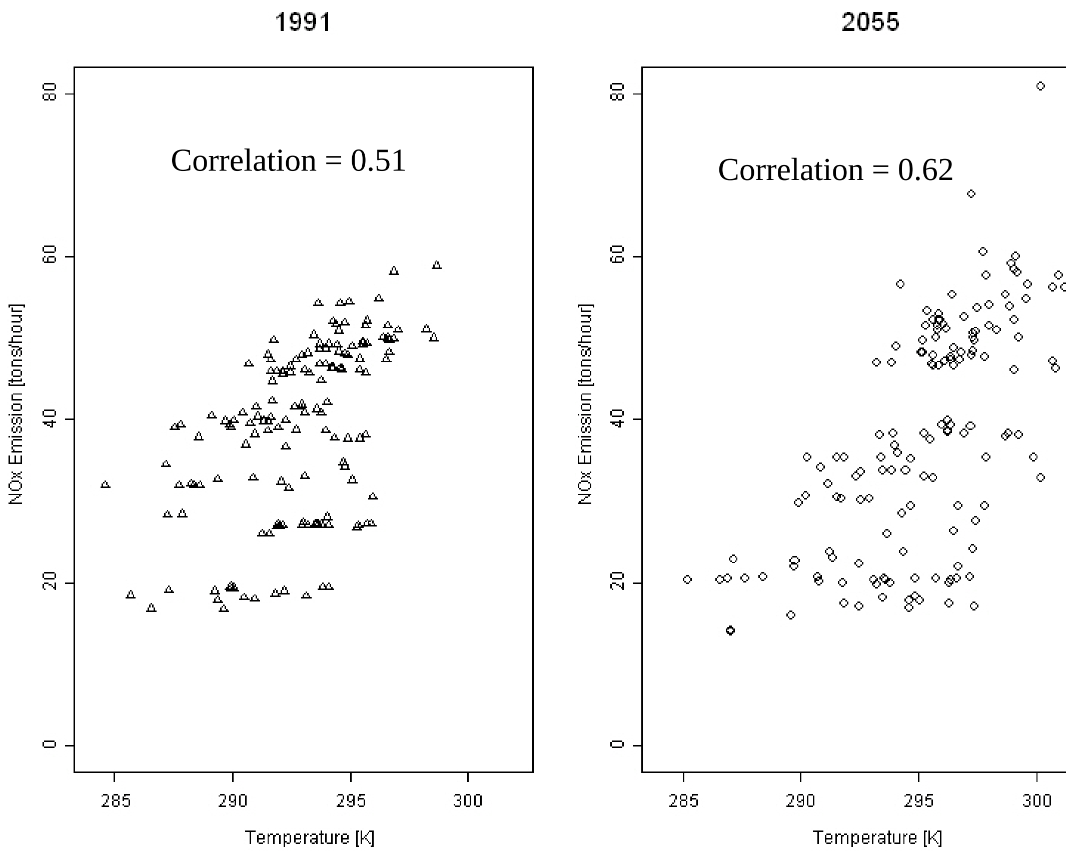

Figure 17: Plot of hourly temperature (Baltimore) vs NO_x emission for hour 14:00 of 1991 (left) and 2055 (right)

Appendix I:
Creation of Script Files of Emissions from Generation Model for use in SMOKE Module of CMAQ

A crucial step in the analysis is to link the effect of climate change on hourly NO_x emissions (the outputs of the electric power sector models) with changes in the ambient ozone concentrations. This relies on creating a script file called PTHOUR, which replaces the existing emission scenarios for power plants in the region of interest in the SMOKE module of CMAQ. Figure 18 details the procedures that we developed to generate the PTHOUR file. PTHOUR file is a text file that specifies the hourly emissions of each stationary emission source at a given hour and date. Each step in the figure is described below.

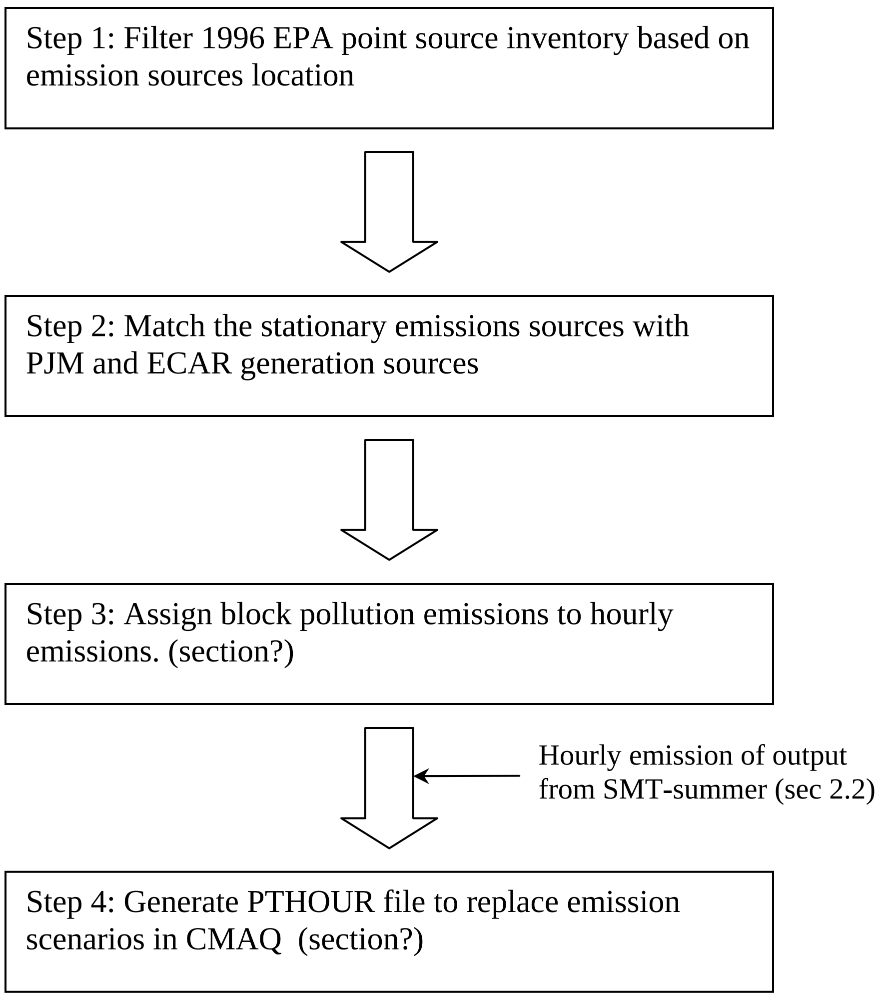

Figure 18: Flowchart of generating PTHOUR file

Step 1: The 1996 US EPA point source inventory contains more than 395,000 point emission sources. They are classified based on a number of fields. For our purpose, we first selected a subset of sources using following fields: state, and city ID. We kept the relevant information such as plant name, plant ID, unit ID, capacity, SCC and annual NO_x emissions for future reference. This helped us narrow down to a total of 25,249 sources. Then, we used SIC codes and locational information, and further identified a total of 3,283 units as our emissions sources of interest (power plants within our study area).

Step 2: In the second step, we then match each source by plant ID and unit ID or plant ID alone. For remaining units that we cannot match with plant ID and unit ID, we then match with latitude and longitude. As a result, excluding non-fossil sources, we were able to account for more than 98.7% of total capacity in the PJM and ECAR area.

Step 3: The output from short-run analysis included hourly emission of each generating unit by load block. In this step, we assigned the membership of each hour to a specific block based on its order in the load duration curve. For example, 14:00 of 7/30/2000 was assigned to be first block (peak block) since its load was among the first 35 hours in the load duration curve. Thus, together with the output from SMT-summer (Section 2.2), we generated the hourly emission of each source in the models.

Step 4: This step involved following the protocol laid out in the CMAQ documentation to produce PTHOUR file.

Appendix II:

Journal and Proceedings Publications

Y. Chen and B.F. Hobbs, "*An Oligopolistic Power Market Model with Tradable NO_x Permits*," *IEEE Transactions on Power Systems*, 20(1): 119-129, 2005.

Y. Chen, B.F. Hobbs, S. Leyffer and T. Munson, "*Solution of Large-Scale Leader-Follower Market Equilibria Problems: Electric Power and NO_x Permit Markets*," Computational Management Science (accepted)

Y. Chen and B.F. Hobbs, " *An Oligopolistic Electricity Market Model with Tradable NO_x Permits*," Proceedings of IAEE North America Conference, Mexico City, Mexico, 2003

Presentations

Y. Chen, B.F. Hobbs, "*Implications of Climate Change for Regional Air Pollution, Health Effects and Energy Consumption Behavior: Selected Emission Results*," IAEE Annual Conference, Taipei, Taiwan, June 3-6, 2005

Y. Chen and B.F. Hobbs, "*Oligopolistic Power Markets with Transmission, Forward Contracts, Reserve Markets and NOx Permits*", INFORMS Annual Conference, Denver, October 24-27, 2004

Y. Chen and B.F. Hobbs, "*Interaction of Oligopolistic Transmission Portfolio Standards, Green Pricing Programs, & Emission Allowances*," 24th Annual North American Conference of the US Association for Energy Economics and International Association for Energy Economics, July 8 - 10, 2004, Washington, DC

Y. Chen and B.F. Hobbs, "*An Oligopolistic Electricity Market Model with Tradable NOx Permits*," 23rd International Association for Energy Economics, North American Conference / VI Congreso Anual de la AMEE, October 19-21 2003, Mexico City

Y. Chen, B.F. Hobbs, and F.A.M. Rijkers, "*US NO_x Permits and BE-NL Market Integration: Applications of Complementarity Oligopoly Power Models*," Annual Meeting, Institute for Operations Research and Management Science, Atlanta, GA, Oct. 19-23, 2003.

Appendix III: Summary of two published articles associated with this project

Y. Chen and B.F. Hobbs, "*An Oligopolistic Power Market Model with Tradable NO_x Permits*," **IEEE Trans. on Power Systems**, 20(1): 119-29, 2005

Summary of analysis and findings:

Around the world, the electric sector is evolving from a system of regulated vertically-integrated monopolies to a complex system of competing generation companies, unregulated traders, and regulated transmission and distribution. One emerging challenge faced by environmental policymakers and electricity industry is the interaction between electricity markets and environmental policies. We use computational models formulated as linear complementarity problems (LCP) to study such interaction. Models based on LCP formulations have been applied previously to assess the potential for exercise of market power in transmission constrained electricity markets. In this particular application, we use the approach to assess the potential market power resulting from the interaction of pollution permit markets, i.e., the USEPA OTC NO_x Budget Program, with electricity markets in the presence of market power. Because the permits program is regional rather than national in scope, some power producers are relatively large consumers of permits (either in long or short position), and there could be interactions between market power in the permits and energy markets. In our model, the producers with substantial capacity share are designated to exercise Cournot (quantity) strategy in electricity markets, and to utilize conjectured price response, where they conjecture their deviations from market equilibrium will affect prices in NO_x permit markets. The application is to the Pennsylvania – New Jersey – Maryland Interconnection (PJM) market during 2000, which is represented by a fourteen-node, eighteen-arc linearized DC load flow model. A total of 730 generators are included in our analysis and the simulated period is approximated by 5-block step functions. The results show that PJM market is relatively competitive during this period, as its prices are closer simulated competitive levels than to Cournot (oligopoly) prices. The NOx market and Cournot energy markets influence each other in several ways. One is that Cournot competition lowers the price of NOx permits, which in turn results in some high emissions producers actually expanding their output, contrary to simple Cournot energy-only markets. The Cournot producers could be worse off if they naively behave as price takers in the NOx permits markets. Meanwhile, higher energy prices and lower NO_x permit prices provide two reasons for smaller price-taking producers to expand energy generation. Total NO_x emission declines as a consequence of restraining output by Cournot producers. In general, because pollution permits are an important cost and their price is volatile, high concentration in the market for such permits can exacerbate the effects of market power in energy markets.

Y. Chen, B.F. Hobbs, S. Leyffer and T. Munson, "Solution of Large-Scale Leader-Follower Market Equilibria Problems: Electric Power and NO_x Permit Markets," Accepted for **Computational Management Science**.

Summary of analysis and findings:

The models formulated as Stackelberg game is especially appropriate to simulate sequential-moved or leader-and-follower games. However, the real-world large-scale applications are generally limited. One reason is that the resulting problems violate several regularity conditions, and it makes computational extremely difficult. In the previous analysis, an ad-hoc conjectural price response, which is based on a parameterized linear function to represent an agent's belief about its rivals' reply to its own action, is used to assess the interaction of pollutions emissions and electric market. Although the resulting linear complementarity problems are relatively easy to solve using commercial solver PATH under the GAMS modeling environment, the solutions from models are not entirely satisfactory on the economic ground. This paper investigates the ability of the largest producer in an electricity market to manipulate both the electricity and emission allowances markets to its advantage. A Stackelberg game to analyze this situation is constructed in which the largest firm plays the role of the leader, while the medium-sized firms are treated as Cournot followers with price-taking fringes that behave competitively in both markets. Since there is no explicit representation of the best-reply function for each follower, this Stackelberg game is formulated as a large-scale mathematical program with equilibrium constraints. The best-reply functions are implicitly represented by a set of nonlinear complementarity conditions. Analysis of the computed solution for the Pennsylvania - New Jersey - Maryland (PJM) electricity market shows that the leader can gain substantial profits by withholding allowances and driving up NOx allowance costs for rival producers. The PJM regional market is represented by a fourteen-node, eighteen-arc linearized DC load flow model. A total of 730 generators are included in our analysis and the simulated period is approximated by 5-block step functions. The allowances price is higher than the corresponding price in the Nash-Cournot case, although the electricity prices are essentially the same. In particular, the largest generator can withhold 7% of its total allowances in 2000 and maintains the allowances price at the competitive level. Such an action increases its rivals' production cost considerably. Consistent with oligopoly theory, the competitive generators generally benefit from higher power prices, while other non-Stackelberg strategic generators experienced shrinkage of the market share and profit losses.

References

Amundsen, E. S. and L. Bergman (2002). "Will Cross-ownership Re-establish market power in the Nordic Power Market?" The Energy Journal **23**: 73-95.

Brooke, A., D. Kendrick, et al. (1998). GAMS: A User's Guide. Washington, DC, USA, GAMS Development Corporation.

Brooks, F. J. (2000). GE Gas Turbine Performance Chacacteristics, GE Power System GER3567h**:** 8.

Chen, Y. and B. F. Hobbs (2005). "An Oligopolistic Power Market Model with Tradable NOx Permits." IEEE Transactions on Power Systems **20**(1): 119-129.

Chen, Y., B. F. Hobbs, et al. (2005). "Solution of Large-Scale Leader-Follower Market Equilibria Problems: Electric and NOx Permit Markets." Computational Management Science **Accepted**.

EIA (2004). Assumptions to the Annual Energy Outlook. Washington, DC, USA, Energy Information Agency**:** 67-68.

EIA (2005). Form EIA-860 Database Annual Electric Generator Report.

Gabriel, S. A., A. S. Kydes, et al. (2001). "The National Energy Modeling System: A Large-Scale Energy Economic Equilibrium Model." Operations Research **49**(1): 14-25.

Hobbs, B. F. (1995). "Optimization Methods for Electric Utility Resource Planning." European Journal of Operational Research **83**(1): 1-20.

Ji, Y. D. and B. F. Hobbs (1998). "Including a DC Network Approximation in a Multiarea Probabilistic Production Costing Model." IEEE Transactions on Power Systems **13**(3): 1121-1127.

NERC (2005). Generating Availability Reports and Historical Availability Statistics, North American Reliability Council.

PowerWorld (2005). PowerWorld Corporation: The Visual Approach to Analyzing Power Systems.

Schweppe, F. C., M. C. Carmanis, et al. (1988). Spot Pricing of Electricity, Kluwer.

Turvey, R. and D. Anderson (1997). Electricity Economics. Baltimore, MD, The Johns Hopkins University Press.

USEPA. (2005a, 2005). "Clean Air Interstate Rule." Retrieved Dec. 10th, 2005, http://www.epa.gov/interstateairquality/.

USEPA. (2005b, 2002). "Clean Air Markets - Progress and Results: Compliance Reports." Retrieved Dec. 10th, 2005, http://www.epa.gov/airmarkets/cmprpt/.

USEPA. (2005c, 2002). "Emissions and Generation Resource Integrated Database (eGRID)." Retrieved Dec. 10th, 2005, http://www.epa.gov/cleanenergy/egrid/index.htm.

USEPA. (2005d, 2002). "EPA Modeling Applications Using The Integrated Planning Model." Retrieved Dec. 10th, 2005, http://www.epa.gov/airmarkets/epa-ipm/.

USEPA. (2005e, 2002). "Technology Transfer Network Emission Measurement Center." Retrieved Dec. 10th, 2005, http://www.epa.gov/ttn/emc/cem.html.