Quantifying Avoided Emissions from Renewable Generation

by

Gabriel R. Gomez

[S.B. MIT EECS 1998]

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

June 2009

Author____
Department of Electrical Engineering and Computer Science
May 22, 2009

Certified by____
Stephen Connors
MIT Energy Initiative
Director, Analysis Group for Regional Energy Alternatives
Thesis Supervisor

Accepted by____
Arthur C. Smith
Professor of Electrical Engineering
Chairman, Department Committee on Graduate Theses

Quantifying Avoided Emissions from Renewable Generation
by
Gabriel R. Gomez

Submitted to the
Department of Electrical Engineering and Computer Science

May 22, 2009

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Quantifying the reduced emissions due to renewable power integration and providing increasingly accurate emissions analysis has become more important for policy makers in the age of renewable portfolio standards (RPS) and emerging carbon markets. This study focuses on the method of quantifying the reduced CO_2, NO_x, and SO_2 emissions due to renewable penetration on the electric power grid. A simplified model is used which calculates avoided emissions from the marginal generators that are displaced if generation from a renewable source were to come online. Analyzing the reduced greenhouse gas emissions from these load shape following generators provides a more telling story of the benefits of time varying nature of renewable power sources. This study gives an updated picture and retrospective analysis into 3 major regions of the US electric grid and a look into the emission savings possibilities brought on by renewable generation using a windmill feasibility study as an example.

Thesis Supervisor: Stephen Connors
Title: Director, Analysis Group for Regional Energy Alternatives, MIT Laboratory for Energy and the Environment

Acknowledgements

For Chase, who mostly slept while daddy worked

Table of Contents

List of Figures

List of Tables

1. Introduction

Greenhouse gas reduction is increasingly becoming a topic highlighted by both local and global policymakers with aims at increasing awareness and decreasing emissions either with incentives or imposed limits and caps. Many industries have come about regarding the issue of reducing greenhouse gas emissions and they are increasingly becoming integrated into the daily fabric of society with things such as 'carbon footprints', 'carbon-credits', and demand side management programs. Many actions may lead to essentially 'avoiding' emissions such as turning off unnecessary lights, turning up the thermostat, or buying renewable energy certificates (RECs), especially since all the aforementioned actions involve decreasing electricity use. Ironically in an age where most things are electronic and the reliance on these items increases with time, decreasing electricity use, and therefore the underlying power generation, is the most vital and direct way to decrease the emission of greenhouse gases. Since every load on the electric power grid is balanced by a network of power plants and the storage of electricity is non-economical, decreasing demand has the direct impact of essentially decreasing electric power generation and therefore decreasing the burning of fossil fuels. The burning of fossil fuels for electric power generation is the leading contributor to global green house gas emissions and the compounds CO2, SO2, and NOx. Offsetting emissions on a grander scale, a level up from individual demand and behavior changes, is increasingly becoming a reality with the promotion of cleaner low and zero emissions electricity sources to curb the burning of fossil fuels. Determining the 'avoided emissions' impact of a renewable generator, such as a wind mill or other time varying zero emitting resource, is complicated because of the intricate nature of supply and demand of electricity across the large network composing the electric power grid.

The following is a study into the exercise of determining which generators are the most likely to "ramp down" or "turn off" their generation due to the penetration of a renewable electricity resource. With this knowledge, along with the emissions data from each individual power plant (CO2, SO2, and NOx), avoided emission rates can be determined on an hourly basis and serve as the building block to forming a good quantitative tool to use when calculating avoided emissions for a variety of applications. A massive data mining exercise such as this comes with the added complexity of relational database and programming tasks which are described with the intent of having energy analysts interested in this area of study focus on the energy problems at hand and less on the information technology and programming. Along with the calculation of base avoided emissions rates and their methodologies, this study follows the exercise of avoided emission calculations for three large regions of the US power grid. The assumption is that once an operational process is put in place for determining avoided emissions rates, it can be replicated as many times as needed for multiple regions of study with ease.

In this study, the calculated avoided emission rates then serve as a tool in an exercise of determining the avoided emissions of a proposed windmill in the Town of Scituate, MA. The goal here is once again to show how an energy problem can easily be reduced to an information problem and analyzed through the use of software and database programming.

1.1. What do we mean by "load shape following" (LSF)

The electric power grid is supplied by a wide mix of generators of all fuel types with a widely varying distribution of heat rates (efficiency) and varying levels of resulting SO2, NOx, and CO2, the most common classifications of green house gases. The complex nature of the electric power grid makes its 'load state' (i.e. a snapshot in time of which generators are on, off, or ramping generation) at any given moment difficult to predict and complicates the ability to calculate the effects of reducing emissions due to renewable integration. A combination of both system dynamics as well as the complicated nature of generator dispatch in modern energy markets compose a challenge of determining which units are essentially following the time varying load at any moment in time. There have been a variety of methods that have been applied in the calculation of reduced emissions due to time varying and intermittent resources, with special attention to the issue of wind penetration and its impact on generation [12]. Two specific models of interest are the "fuelsaver" approach and the "forecasted" approach [12] in determining the impact of wind generation on system operation. The fuelsaver approach is the simpler of the two and is premised with a large assumption that the penetration of the renewable resource does not impact generator dispatch. Here as the renewable resource comes on line, the assumption is that the mix of marginal generators at that same moment are the most likely to curtail their output and thus "save" the respective fuel. For studies such as this where the impact of local renewable resources are the focus, the fuelsaver approach is a safe one as the level of generation is not large enough and in the case of a lone windmill have a net capacity of zero and will not participate or bid into the local market.

The "forecasted" approach is a bit more sophisticated and accounts for the dynamics of the way generators bid into the local market to determine which generators are at margin. This reflects more the reality and operation of most deregulated markets and is reflected in the larger independent system operators (ISOs) such as ERCOT, ISO New England, and the New York ISO. This approach is operationally more difficult and prone to another level of error as the method of determining which generators should be dispatched on a least cost basis is limited still by a linear programming minimization algorithm with large assumptions. Though this method is a step in the right direction and deeper in scope and more appropriate for study of larger sources of generation such as those proposed in Texas and fall on the order of tens of thousands of megawatts.

Under the "fuelsaver" methodology, the fuel type of the generators who are dispatched to follow the time varying renewable resource differ from those units running at high capacity and can be considered as non base load and will reflect a broader mix as shown in Figure 1, the fuel stack can be weighted more toward natural gas than coal resulting in a different emission rate average as gas plants burn "cleaner" than coal plants. The system loads in Figure 1 demonstrate how a system peak on a typical summer day varies by control area each with their own supply stack and marginal fuel. The emission rates from these LSF units can be calculated and will vary based on the fuel type (natural, gas, coal, diesel, etc.) and it has been demonstrated that the emissions emitted by these units give a reasonable expectation what will be displaced by zero emission renewable sources. Based on the season and time of day, this mix of fuel type will vary and so will the emission rate thus an hourly resolution is necessary to capture this level of detail. Overlapping this variable emission rate with the variable intermittency of renewable generation such as wind, one can create a clearer picture of the true emissions reduction.

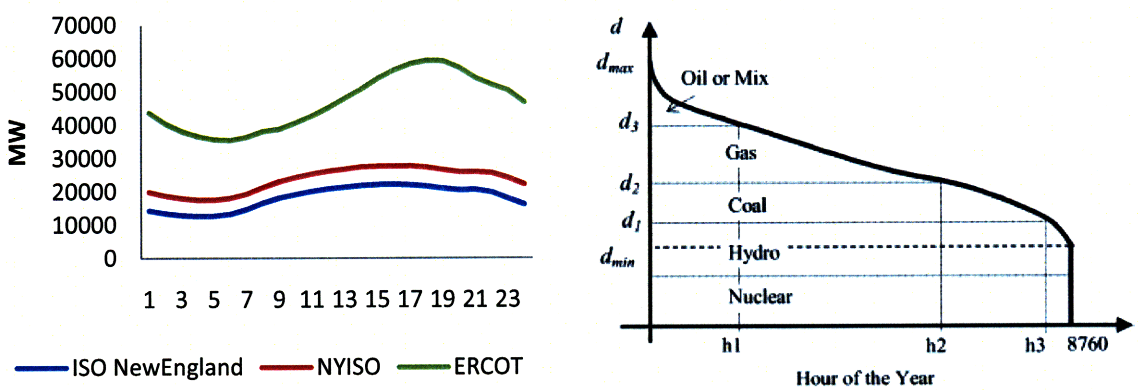

Figure 1 System loads for typical summer day and a typical load duration curve (supply stack) [13]

1.2. General need for avoided emissions analysis

A recent study released by the Electric Reliability Council of Texas (ERCOT) and GE Energy (reference 5, 2008) describes the impact of integrating large capacities of wind generation from the wind rich areas of West Texas. The study analyzes the impact on prices, emissions reduction, and transmission constraints with the addition of 5000MW, 10000MW, and 15000MW of wind generation. The general outcome of this analysis calls for the addition of new high capacity transmission between the west and eastern sections of Texas, an expensive necessity if Texas is to take advantage of its renewable resources in the future. The roles in Regulatory councils of course rely on this information to make decisions on whether to fund large undertakings such as this, and rely on "big ticket" benefits such as consumer fuel cost savings, lower market prices, contributions to the renewal portfolio standards, and emissions reduction. It is interesting that the emission reductions included in this study show that combined cycle or "peaker" gas turbine generation was displaced on the on peak and coal displaced in the off peak. This is consistent with studies that have shown that when wind is at its highest in the evening, it will offset the more "dirty" units such as coal instead of the cleaner burning gas units in the on peak.

The above prospective and feasibility study lacks the transparency of how the emission reductions were calculated (and forecasted) in their simulation, but provides motivation for this study especially with the possibility of very large capacities new wind generation in Texas. The EPA publishes a variety of rich emissions data sources such as the eGrid and Continuous Emissions Monitoring (CEM) data and provides an informational foundation to develop insight into emission reduction analysis. For studies like the above in Texas one can assume that the emissions calculations are based on system wide and geographical averages of generalized EPA data because of the rich library of data that the EPA publishes. A transparent retrospective analysis will give some insight on the emission reduction rates due to the currently installed wind capacity and help with the forecasted outlook large prospective studies as the one above. With the emerging carbon markets as with the early RGGI initiative, in a cap and trade scenario, emissions will have an economic impact as generators or even large consumers (if cap and trade will impact vertical markets) may want more insight into offsetting their generation and emission portfolios. It is possible that the future carbon markets may favor counterparties replacing increasing their renewable generation mix and provide incentives and credits during participation in their carbon markets.

The big picture goal of this study is to provide the processes and data necessary to investigate the avoided emissions that would result from smaller generation sources on the power grid. This includes generation from lone windmill installations, arrays of solar panels, demand response curves, etc, all which act to displace the immediate generation of power from the marginal fossil fuel burning plant.

2. Background and Problem Setup of Data problem

Applying avoided emissions analysis is a useful instrument that can be used to provide insight and guidance into many things including: calculating the environmental impact of lone wind mill installations, emission offsets due to demand response, and fitting the needs of more accurate geographical analyses where system wide avoided emissions averages may be too broad. One may ask of the significance of the impact of smaller sources of generation especially in the vastness of the large electric ISOs which make up large portions of the US electric grid.

With the increasing awareness of environmental conservation the urgency to offset the negative factors of greenhouse gases and overproduction have become more local, where the presence of a windmill can null the electricity usage of a small community, or balance out the portfolio mix if a large user of electricity. Knowledge of the avoided emissions of well placed generation is useful especially during the planning and negotiating stages of project funding and approvals. A more accurate representation not only is more helpful but necessary in the long term as the economics of building renewable generation become tighter. Cap and trade policies and renewable portfolio standards will eventually have an influence on energy prices and the tracking of emissions offsetting may become more important especially if credits are granted in the new initiatives such as the Regional Greenhouse Gas Initiative [20].

A variety of emissions reduction analyses have been done by past members of the MIT AGREA team (Analysis Group for Regional Energy Alternatives) using the LSF method and have been linked together with both photovoltaic and wind resources and avoided emissions calculated. The studies range from the year 2004 [4] to 2006 [1,2]. For example, [1] analyzes the emissions reduction in the ISO-NE grid due to the proposed offshore wind farm off the coast of New England. In this study a cost benefit study of analytical methods was done as well as a comparison between the outcomes of various methods in the scenario of the introduction of wind from this offshore farm.

In a similar fashion this study investigates a similar emission reduction study on the south shore of Massachusetts mainly to highlight and verify the re-engineered process of applying the LSF algorithm and determining the appropriate avoided emission rates. This would entail updating existing data sources in the MIT AGREA database to reflect the latest EPA 2006 emission data, updating ERCOT load curves, and have the latest CEM hourly emissions data up to the end of 2008. Analysis can then be prepared to shed some insight into the conclusions that can be drawn from the magnitude of emissions reduction for policy and decision makers. The intent is to continually publish newer avoided emissions numbers that reflect the direction in which the electricity markets are headed.

2.1. Shifting an energy problem into a data problem

The problem I set out to investigate was whether the data collection, application of a load following algorithm, and calculation of avoided emission rates can be streamlined for ongoing studies in the future. The data mining involved in a study such as this is a can be a very time intensive process and can be expensive requiring substantial requirements both from hardware and software perspective and a general level of technical expertise by the researcher. The data gathering processes, codebase, and database architecture in previous studies have served well but were in need of a fresh look as it is always helpful to keep the database in line with the changes in EPA/EGRID data as it is published. Also as programming code changes multiple hands over a period of years, the original intent of the author is sometimes lost leaving many questions as to why some methods were chosen. I also wanted to "scale up" and tackle calculating the avoided emission rates and running the LSF algorithm on more than one control area. When dealing with hourly generator data across multiple power control areas it is helpful to have a streamlined and simplified process so that the programming and data efforts are minimized if more geographical areas want to be investigated in the future. This also opens the door to studying the impacts of avoided emissions instead of spending an exhaustive amount of time on database programming.

2.2. Avoided Emissions studies from AGREA team and motivations

Team members of the AGREA team have done a variety of studies on the avoided emissions data. The studies include various looks into the avoided emissions due to various renewable sources such as wind and solar installations. The most recent study was mostly focused on the New England region with the application of a proposed wind turbine in the town of Hull, MA [14]. The avoided emissions study was a rigorous look into the site data of the town of Scituate with conclusions that led to a more accurate analysis of the emissions avoided under various scenarios. The study is to aid decision makers in viewing more closely the emissions avoided under a more rigorous method as opposed to using a system wide average.

2.3. The Power Control Areas chosen for this study

The regions of focus in this study are three of the major ISO (Independent System Operators) in the United States. The New York ISO (NYISO) and the Electric Reliability Council of Texas (ERCOT) and the Independent System Operator of New England are responsible for a large portion of the generation and loads and include some of the largest metro areas in the United States. The NYISO and ERCOT ISOs both geographically fall within the boundaries of their own respective states, whereas the Independent System Operator New England (ISO-NE) spans all of the New England states of offering a wider and more complex region of study.

The goal of this section is to show that two of the regions ERCOT and ISO-NE are unique in that they marginal fuel is heavily weighted toward natural gas. As the average emission rates of natural gas plants are lower than coal, I will observe how the resulting avoided emission rates are compare for ERCOT, ISO-NE, and NYISO control areas.

As demonstrated in Table 1, ERCOT, ISO-NE, and NYISO make up one third of the established regional transmission authorities, or ISOs (independent system operators), in North America. Load data for ISOs is readily available and make the exercise of aggregating emissions data easier than other areas of the grid where the electricity markets may still be regulated or composed of smaller authorities such as electric cooperatives. Data transparency is very important especially for regions not under the umbrella of an ISO or RTO, where it can be challenging to gather the total loads for a given utility, which are a vital input into the calculation of marginal avoided emission rates. The EPA eGrid database uses the designation of "power control area" and includes the ISOs as specific strata group in their reporting structures, which make it easier to gather data for and leave the guesswork to a minimum. The ISOs also have their own market dynamics as they were created for reliability purposes and to serve as the foundation of deregulated markets.

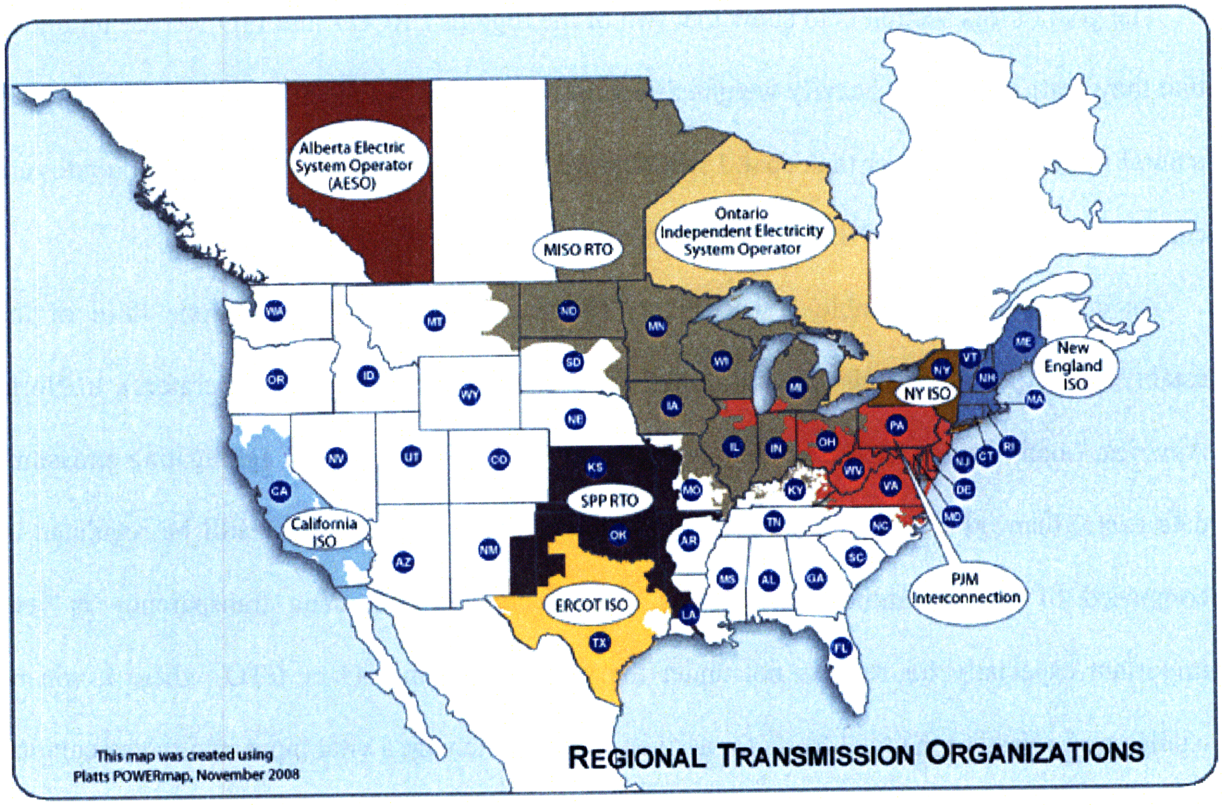

Table 1 Map of US Power Control Areas [17]

The total loads and generation capacity for each of the three ISOs can be seen in Figure 1 and Table 2 respectively. The generation mix in ERCOT is interesting as it has twice the natural gas capacity as New York and almost three times the amount of coal capacity since ERCOT lacks the amount of nuclear generation that is present in the northeast.

Control Area	Capacity (MW):	Net Generation (MWh):	Heat Input (MMBtu):
ERCOT	89,456	321,072,020	2,572,493,111
ISO-NE	36,109	133,857,977	875,551,555
NYISO	42,920	148,131,367	775,556,472

Table 2 Control area total Capacity (source eGrid)

Control Area	Coal	Oil	Gas	Other Fossil	Nuclear	Other Fuel	Wind	Solar	Geothermal	Biomass	Hydro
ERCOT	37.1%	0.5%	47.5%	1.2%	11.9%	0.2%	1.2%	0.0%	0.0%	0.1%	0.3%
ISO-NE	15.3%	9.8%	36.9%	1.5%	25.8%	0.0%	0.0%	0.0%	0.0%	4.7%	6.0%
NYISO	13.7%	16.2%	22.5%	0.7%	28.7%	0.0%	0.1%	0.0%	0.0%	1.2%	16.9%
US Total	49.6%	3%	18.7%	0.6%	19.3%	0.1%	0.4%	0.014%	0.4%	1.3%	6.5%

Table 3 Generation Mix of the three power control areas in this study (source eGrid)

ERCOT is unique in that it has a relatively large presence of wind power already installed with non trivial amounts of wind capacity [5] already in the early stages of planning. Although many renewable projects are underway in New York and New England, most all short of the proposed 15,000 MW proposals in ERCOT, but nonetheless avoided emissions still play an important part in the decision making process of approving these projects no matter what the size.

3. Data Mining and the Avoided Emissions Database

The AGREA emissions database used in this study falls in to the medium to large category among similar data stores where data tables can easily get into the hundreds of millions of rows especially if a large number of control areas containing a large number of generators are to be studied. In the current design for the three ISOs (ERCOT, ISO-NE, and NYISO) the total number of rows in the generator tables which house both generator hourly loads and emissions are 14.2M, 6.7M, and 12.9M respectively for a total of 5 years of data per control area. The record count in the hourly generator tables are directly correlated to the number of generators corresponding units are contained in that power control area. Each individual generator can be composed of multiple "units" which further multiplies the data very quickly. The data at an hourly resolution, which originates from the EPA CEM (Continuous Emissions Monitor) data website, contains information for any generator with capacity of more than 25 MW as mandated by the Acid Rain Program [22]. Though the record count is very high which may lead to the assumptions of database performance issues, the emissions database is unlike typical database applications because it is very non-transactional in nature: where database *reads* and database *writes* are so few since the since data updates will only happen when the time horizon needs to be extended and new data is released by the EPA. This imbalance between database reads and writes allows the database architecture to remain fairly thin and straightforward as more normalized data structures are not necessary since the added over head of multiple concurrent users and heavy querying does not exist. Typically once the LSF algorithm is run on a dataset, the user will issue one final query to retrieve the weighted average emissions by hour for the control area of interest and then the database remains inactive until months later when new data is released. Thus the emissions database can remain in a simple relation database format instead of traditional data warehouse formats that are more difficult to maintain, where data would have to be transformed into architecture more apt to viewing for increased performance with tables of large size.

3.1. Database Architecture and Creation of data processes

As demonstrated in Figure 2 and Appendix 3, the AGREA avoided emissions database consists of a set of relational database tables that store the before mentioned data sets in a fashion that is structured for efficient database programming. The large datasets that compose the hourly emission data at the generator level had originally been intended to be stored in one large table for ease in querying and more straightforward SQL joins with other tables in the database. For example, if all hourly emissions for each generator are in one table, then the data can be aggregated into a number of interesting ways when studying various segments of the power grid and not be limited to say control area alone (as would be the case of having a separate table per control area). After much trial and error, it was found that maintaining a table with on the order of 30+ million rows of data becomes cumbersome and more dependent on processor speed and increased hardware requirements. Because of the large size of the data files that result from an architecture such as this, portability also becomes an issue if there is a need to replicate the environment on a local machine for offline use and development. Instead of researching more advanced ways to handle very large tables, a table was created for each individual control area for the storage of the hourly generator level emission data. The data model still remained clean and through the use of stored procedure programming, the disparate nature of the multiple tables becomes a non issue.

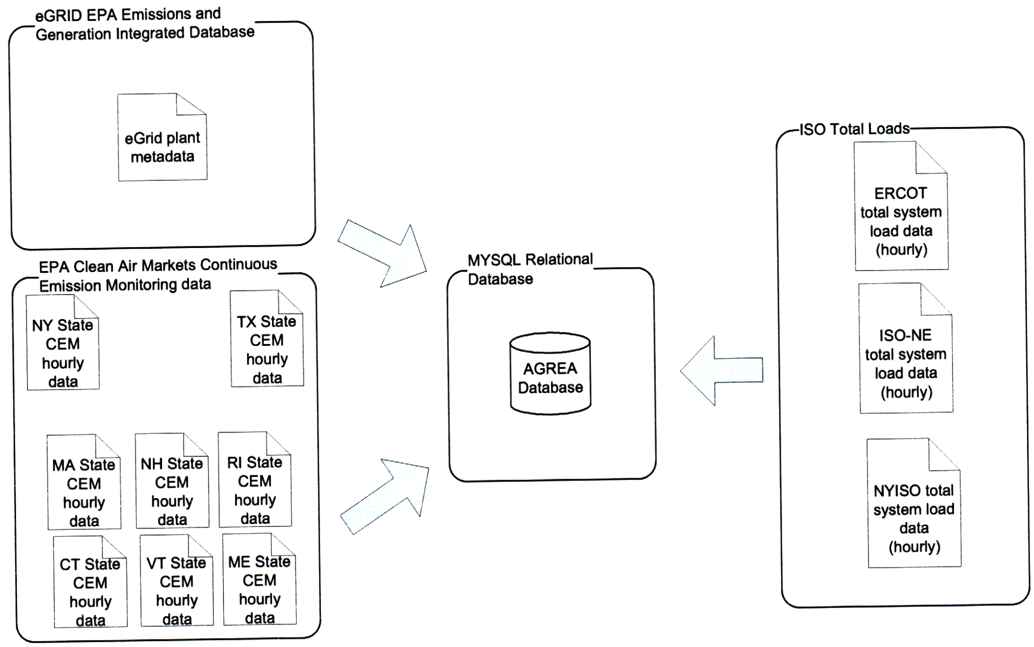

Figure 2 Information needed for calculating marginal avoided emission rates

Several steps were taken in preparing the database for the NYISO, ISO-NE, and ERCOT data to accommodate the data for the years of 2004-2008. These included the loading of the raw data into the associated tables, creating data transformations for data not in the correct format for upload into its associated table, and database tuning of indexes as well as system level parameters that impact performance. The AGREA avoided emissions database utilizes the MySQL open source relational database product and for the purposes of this study, the data was loaded onto a dual processor PC for local processing. The data inputs from Figure 2 can be summarized as follows.

- **Generator "meta" data – from the eGrid (Emissions & Generation Integrated Database) [23,24]**
  - **AGREA Database tables: plant, generator (Appendix 3)**
  - Contains physical details about all generator and units with capacities larger

than 25 MW. The data attributes of interest include capacity, fuel type, geo data, average historical emissions and a variety of other data points used to track total emissions.

- The vital pieces of information here are the generator/unit unique identifier (ORISPL code) and total generation capacity which will be used to match the hourly generation and emission data from a different data source (CEM).
- The data is divided in a hierarchical fashion with two levels, the plant/generator level and the unit level. A plant can have one or many units.
- Data files are provided by year in one large Excel data file. The generator and unit level information were imported directly into the 'plant' and 'generator' table respectively.

- **CEM (EPA Continuous Emissions Monitoring data) [15]**
  - **AGREA Database tables: ny_cem, erct_cem, ne_cem (Appendix 3)**
  - Hourly emission data for CO_2, SO_2, NO_x compounds
  - Uptime of unit during hour, heat rate, generation output
  - Very large dataset as hourly data is published in an 8760 fashion, one row per hour. Thus 5 years of data for one generator quickly adds up to approximately 44,000 and it is typical that a generator can have multiple units, making the data quickly add up to 100,000 rows per generator.
  - Data files are provided in [15] by state, by quarter, in zip files that contain the hourly data. These data files were downloaded, unzipped and loaded into the appropriate "xx_cem" table through the MYSQL command line interpreter.
  -

- **Total Control Area load data**
    - **AGREA Database table: total_load (Appendix 3)**
    - Total hourly load data for the power control area (or ISO) of study serves as the reference in the LSF algorithm to which generators are compared to when determining whether generator is following and responding to load and having a higher chance that it would be the unit "on the margin".
    - Each ISO will publish the total load data both in real time and historically. The data was downloaded from the appropriate ISO website and imported into the total_load table via MYSQL command line interpreter.

3.1.1. How is this data turned into marginal avoided emission rates?

The data inputs as described are the initial building block in calculation marginal avoided emissions. Figure 3 shows a high level overview of how these input data points are used in the overall process of the avoided emissions calculations.

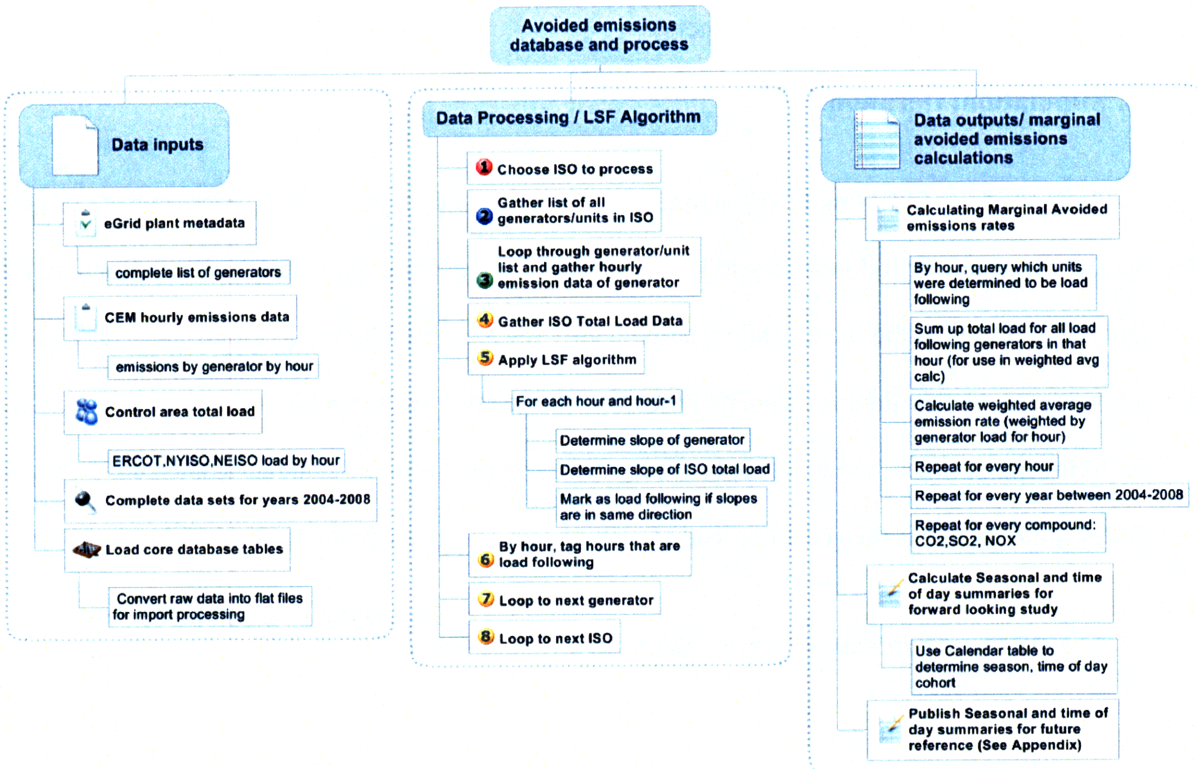

Figure 3 Marginal Avoided Emissions Data Processing – Inputs/LSF Algorithm/outputs

The programming code to implement the LSF algorithm was reprogrammed purely in the Structured Query Language (SQL) inside of a MYSQL database object called a stored procedure. This is a change from the previous work on the AGREA database where the LSF algorithm was programmed using a combination of the PHP programming language with MYSQL merely acting as a data warehouse. The intent here was to encapsulate the LSF programming logic in a simple and easily maintainable form. Using a MYSQL stored procedure for this has the added benefit of programming in the database directly with no external function calls as in the case of PHP. All stored procedure code is compiled locally in the database itself and the code essentially is run in the database management system itself, "closer" to the database tables giving an inherent optimal performance. The stored procedure which performs the LSF algorithm can be seen in Appendix 2. The goal here was to simplify codebase for ease of maintenance and updates by future energy analysts (not programmers).

3.2. Simulation, Iterations

The load shape following (LSF) algorithm is a technique that gives a first order estimate of how closely a unit or generator follows the grid load to see if it should qualify as a marginal unit. Running the LSF simulation results in the identification, at the hourly level, of each hour a generator was deemed as load following. Once this field is calculated the avoided emission rates can be calculated by compound with the following equation which is simply a weighted average of the emission rate against the delta unit load (Δunit load), which is the generator load of current hour minus the generator load of the hour prior.

$$CO2\ marginal\ emission\ rate\ for\ hour_{\ i} = \frac{\sum_{lsf}((\mathrm{CO2_i/load_i}) * \Delta\mathrm{unit\ load})}{\sum_{\mathrm{lsf}} \Delta\mathrm{unit\ load}}$$

$$SO2\ marginal\ emission\ rate\ for\ hour_{\ i} = \frac{\sum_{lsf}((\mathrm{SO2_i/load_i}) * \Delta\mathrm{unit\ load})}{\sum_{\mathrm{lsf}} \Delta\mathrm{unit\ load}}$$

$$NOX\ marginal\ emission\ rate\ for\ hour_{\ i} = \frac{\sum_{lsf}((\mathrm{NOx_i/load_i}) * \Delta\mathrm{unit\ load})}{\sum_{\mathrm{lsf}} \Delta\mathrm{unit\ load}}$$

Figure 4 Marginal avoided emissions rate (by compound): Equations used to calculate weighted average [14]. (lsf => sum across all load shape following generators for particular hour)

To finalize the process, as shown in Figure 4 above, the process of running the LSF algorithm is run for each control area, and the a query is run against the results that calculates the weighted average emission using the three equations above. The implementation of the above equation from a programming perspective can be seen in Appendix 4.

4. Analysis of results – Marginal Avoided Emission Rates

The marginal avoided emission rates for the ERCOT, ISO-NE, and NYISO control areas for years 2004 – 2008 are shown below as calculated with the methodology described in the previous section. The associated marginal avoided emission rates are presented in four formats aimed at highlighting the daily and seasonal variations in the eye of the observer. The descriptions of the formats are described as followed:

- **Hourly avoided emissions rates by year** – 24 hour x 365 day (aka "8760s") color map representation. The daily and seasonal patterns can be observed and aid in deriving qualitative observations. ERCOT, ISO-NE, NYISO charts are presented individually.
- **Average monthly avoided emission** rates – Chart showing average of emission rates across months for years 2004-2008 for ERCOT, ISO-NE, NYISO
- **Average hourly avoided emission rates** – Chart showing average of emission rates across hours for years 2004-2008 for ERCOT, ISO-NE, NYISO
- **Average hourly avoided emission rates w/ seasons** – Chart showing average of emission rates across hours by season for years 2004-2008. ERCOT, ISO-NE, NYISO charts are presented individually.

The charts in the following sections are presented, by compound, in the above sequence with comments and observations following at the conclusion of each respective compound section. This is done in order to present a consistent format for the reader as the number of charts can seem daunting at first but are presented because they indeed tell a story. Any key observations or anomalies in the charts will be commented in the appropriate sections.

4.1. CO2 Emission Rates

4.1.1. CO2 Hourly avoided emissions rates by year

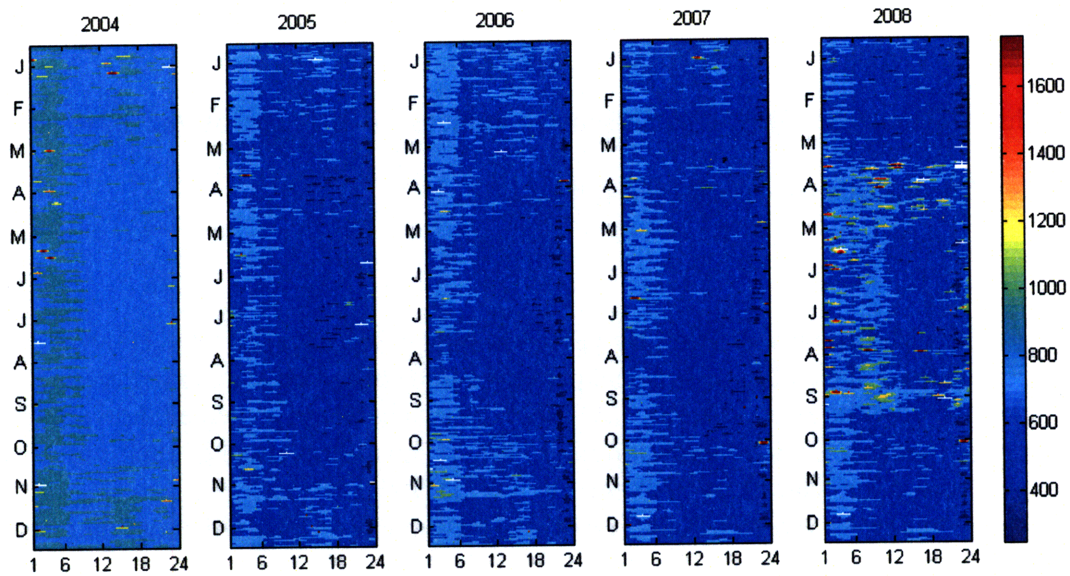

Figure 5 ERCOT Marginal CO2 Emissions Rate [kg CO2/MWh]

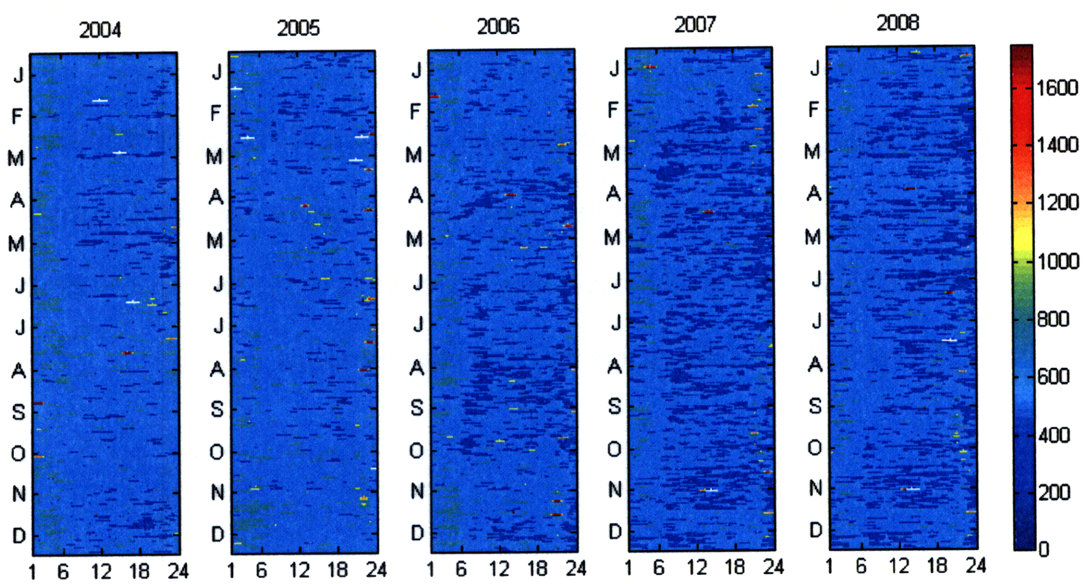

Figure 6 ISO-NE Marginal CO2 Emissions Rate [kg CO2/MWh]

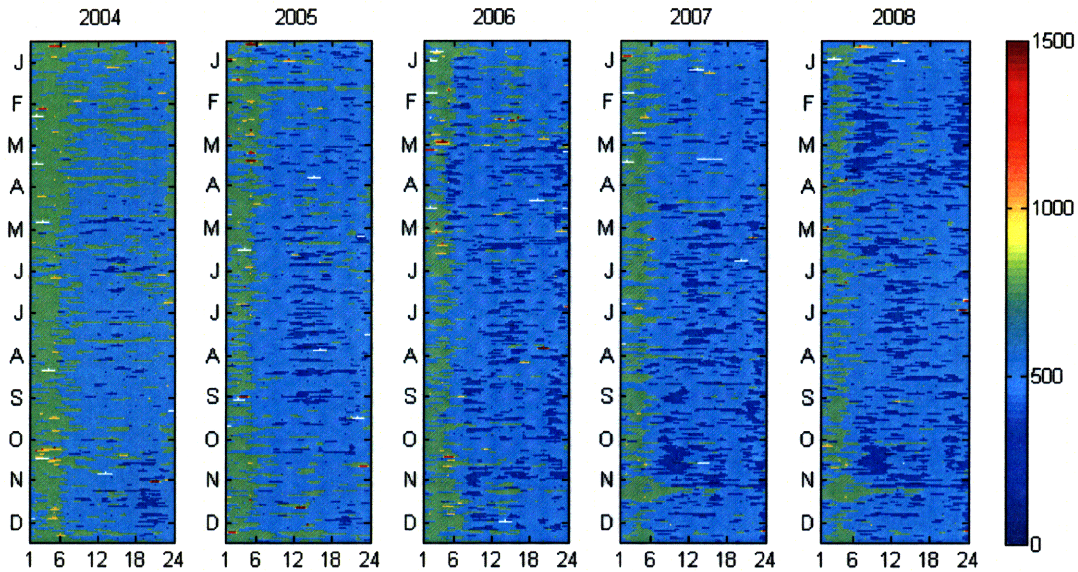

Figure 7 NYISO Marginal CO2 Emissions Rate [kg CO2/MWh]

4.1.2. CO2 Average daily avoided emission rates

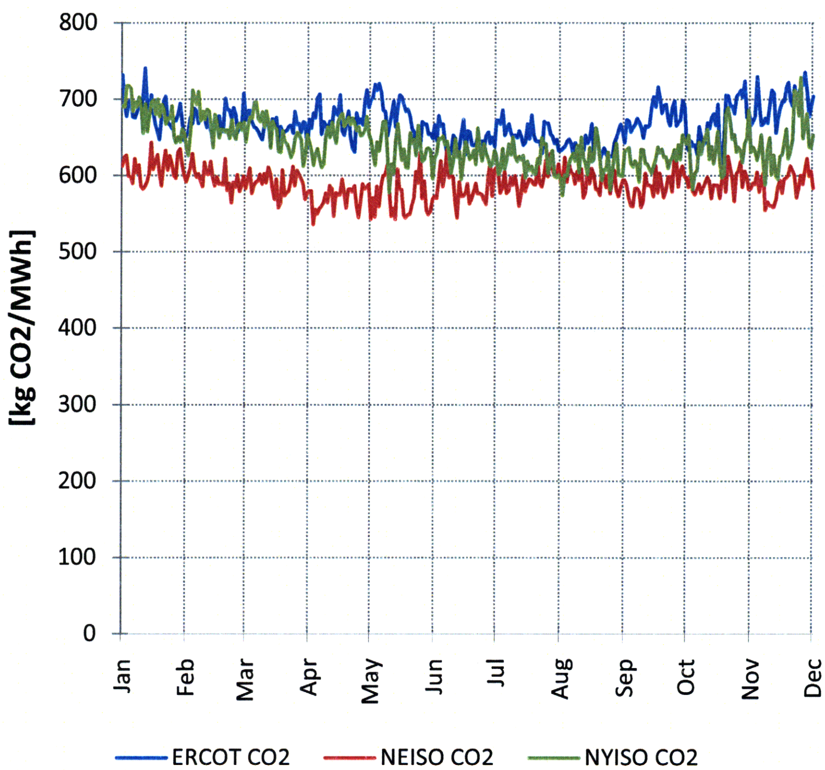

Figure 8 2004-2008 Average daily Marginal CO2 Rates [kg CO2/MWh]

4.1.3. CO2 Average hourly avoided emission rates

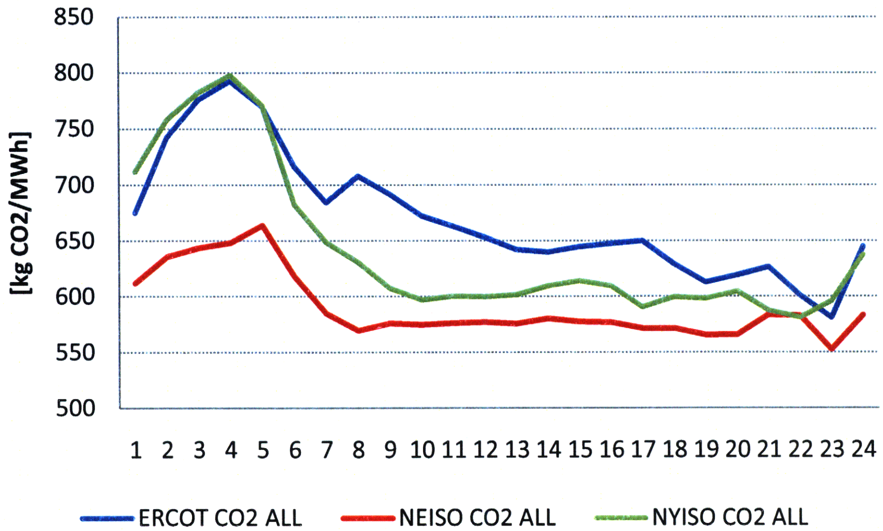

Figure 9 2004-2008 Average Hourly Marginal CO2 Rate

4.1.4. CO2 Average hourly avoided emission rates w/ seasons

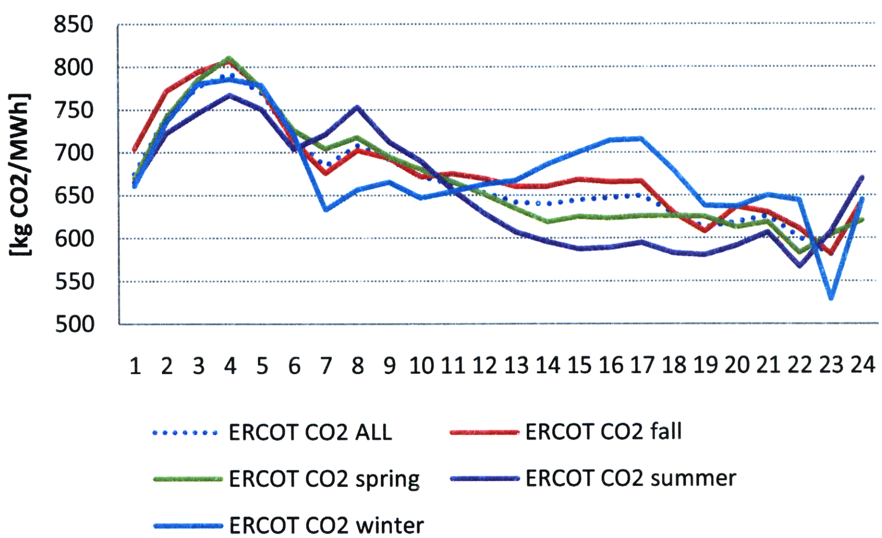

Figure 10 2004-2008 ERCOT Average Hourly Marginal CO2 Rate w/Seasons

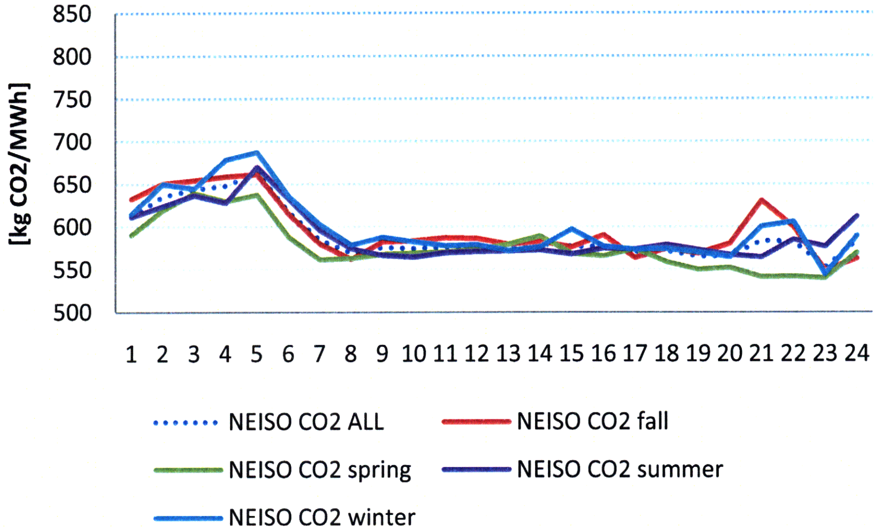

Figure 11 2004-2008 ISO-NE Average Hourly Marginal CO2 Rate w/Seasons

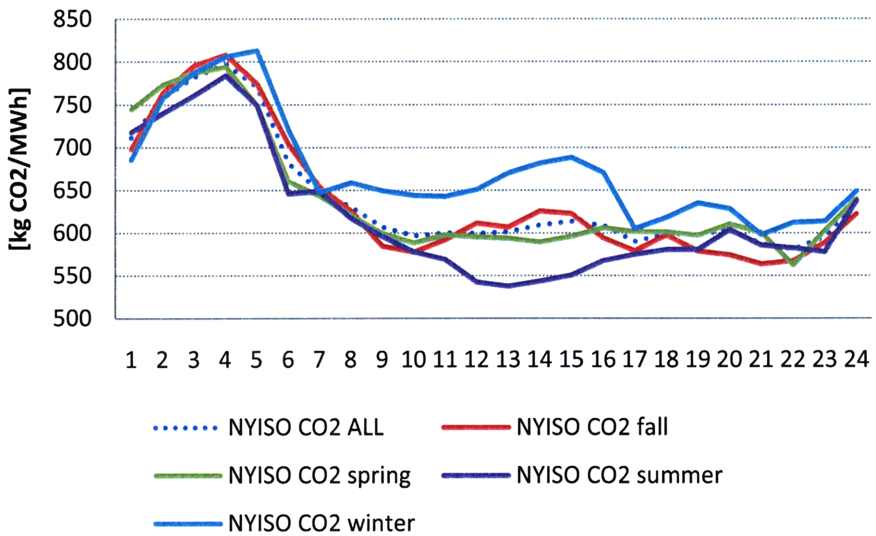

Figure 12 2004-2008 NYISO Average Hourly Marginal CO2 Rate w/Seasons

Overall, the three control areas present limited bimodal diurnal characteristics. The first daily emissions peak in the overnight hours dominates relative to the rest of day. Looking a little more closely there are glimmers of a bimodal diurnal pattern in the winter months with resurgence in the mid-day hours, with ERCOT being the most noticeable. The winter months bimodal peaks for ERCOT and NYISO are most pronounced in Figure 10 and Figure 12 respectively, as the winter average marginal rates show a considerable peak in contrast to the summer curve which exhibits a limited bimodal effect. NEISO is the noisiest of the three with smaller variance in the on peak hours across seasons but still showing a peak in the early hours.

The early morning peak in the marginal emissions rate is clearly visible across all three control areas around hours 4 and 5. In these early hours the phenomena can be thought of taking a "walk" down the supply stack (Figure 1) toward the marginal fuel of coal, with the ramping down of gas units in the early AM, and the ramping up of gas units right before the beginning on peak hours contributing to the emissions rate as well.

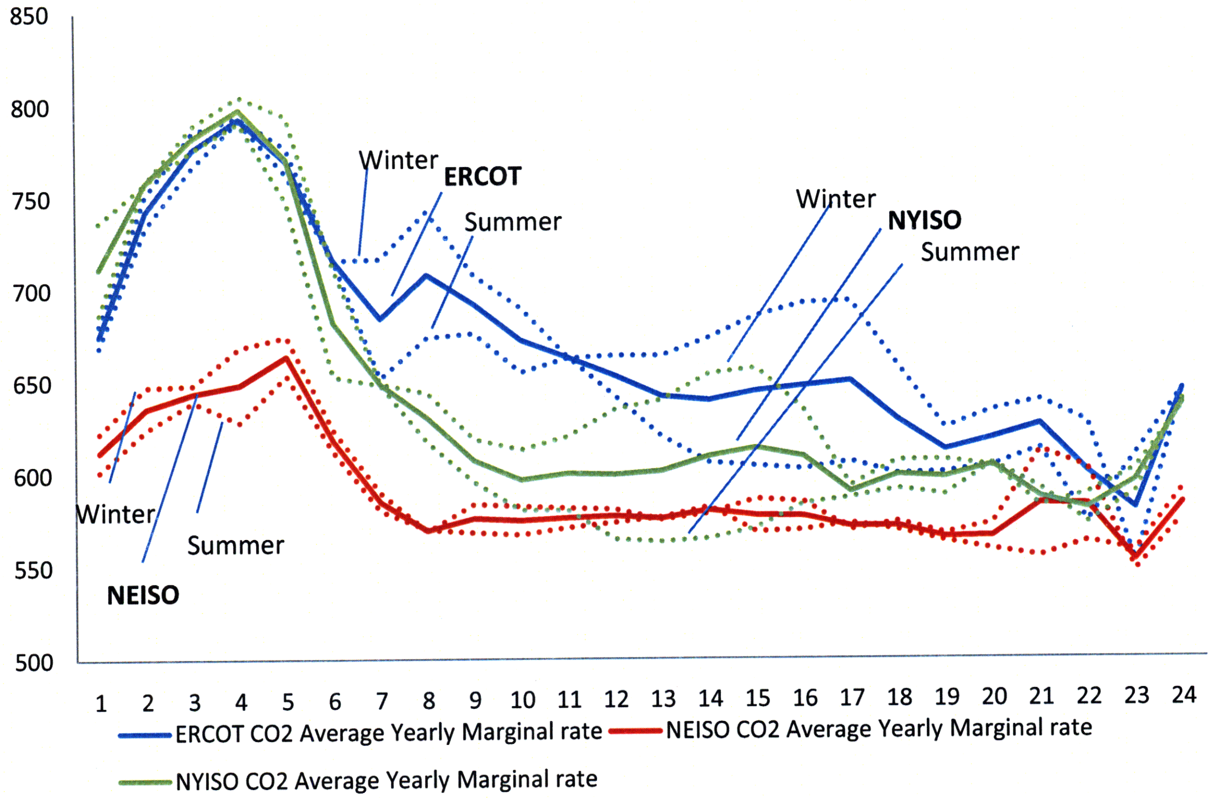

Figure 13 2004 – 2008 Average Hourly Marginal CO2 Rate [kg CO2/MWh]

Breaking up the year into two seasons, summer and winter, allow one to see the pronounced bimodal diurnal rate patterns especially in the winters of both ERCOT and NYISO. ERCOT also exhibits a third peak around hour 7 or 8, the beginning of the on peak period as defined by the ERCOT market. With gas generators exhibiting higher emissions during ramping activity, perhaps this is a contributing factor as gas units are coming online for the load increase in the on peak hours.

4.2. SO2 Emission Rates

4.2.1. SO2 Hourly avoided emissions rates by year

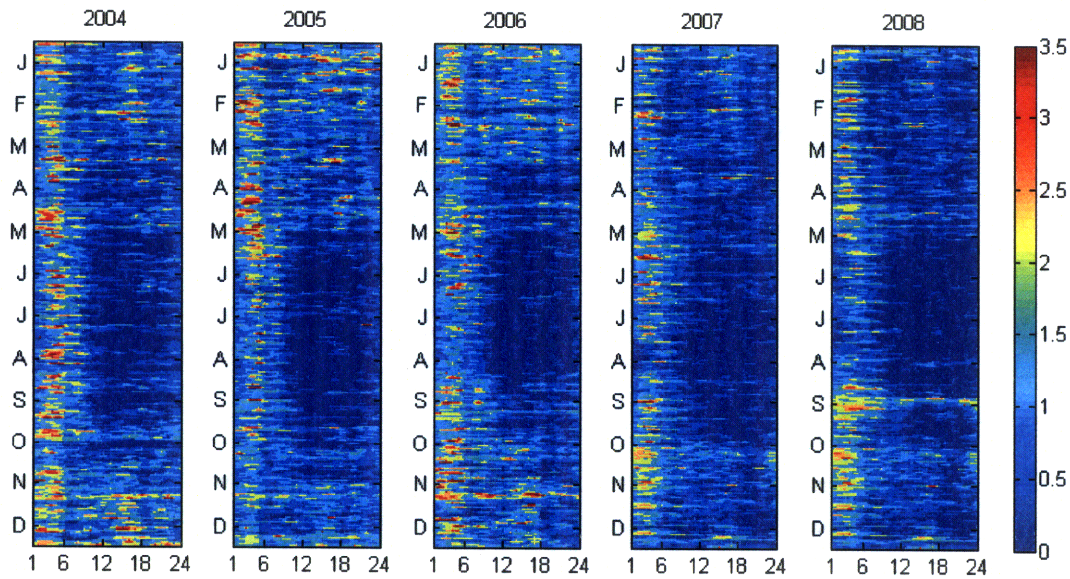

Figure 14 ERCOT Marginal SO2 Emissions Rate [kg SO2/MWh]

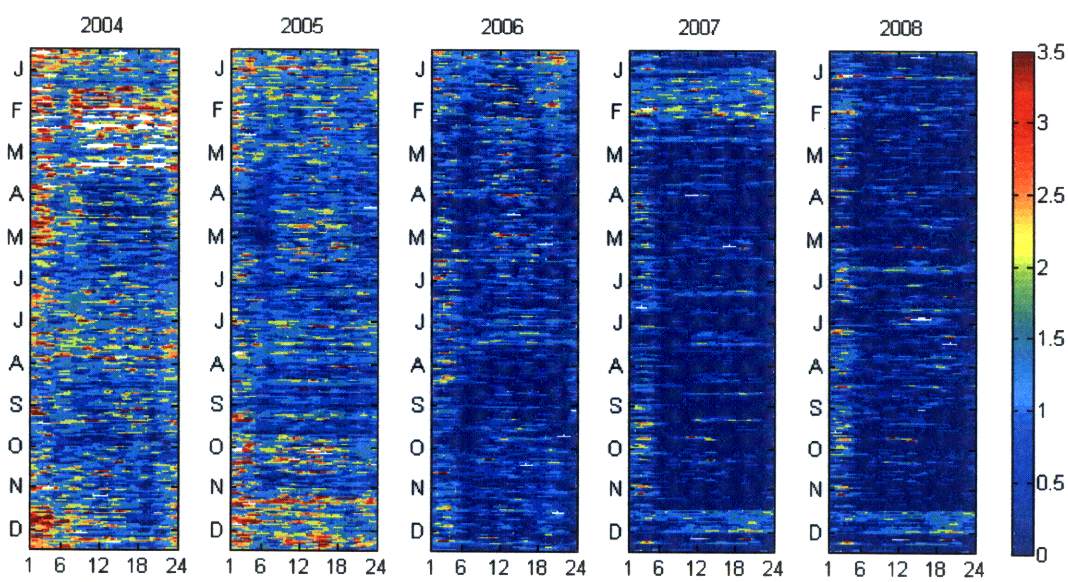

Figure 15 ISO-NE Marginal SO2 Emissions Rate [kg SO2/MWh]

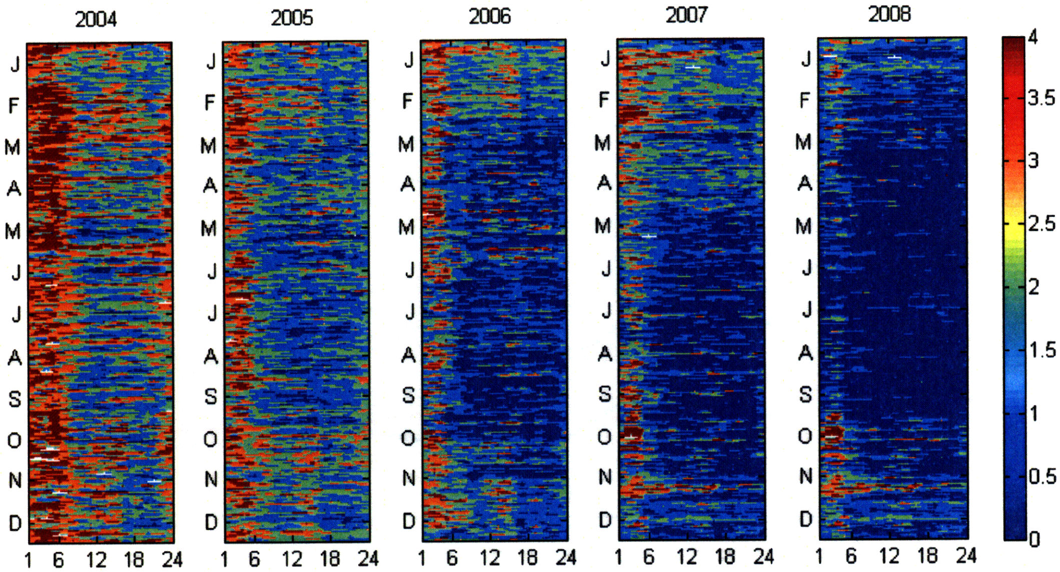

Figure 16 NYISO Marginal SO2 Emissions Rate [kg SO2/MWh]

4.2.2. SO2 Average daily avoided emissions rates

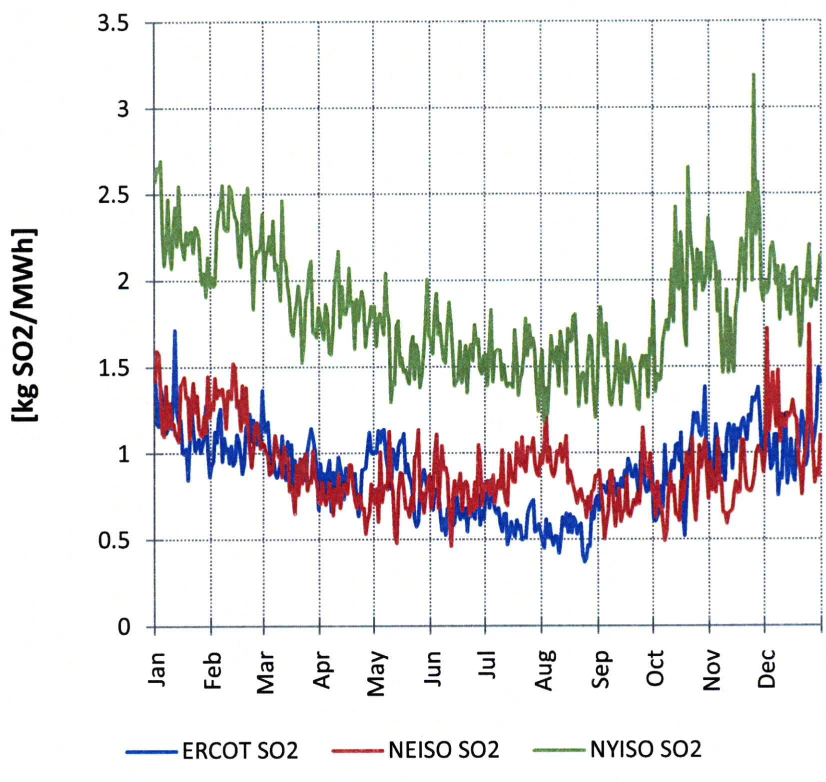

Figure 17 2004-2008Average Daily Marginal SO2 Rate [kg SO2/MWh]

4.2.3. SO2 Average hourly avoided emission rates

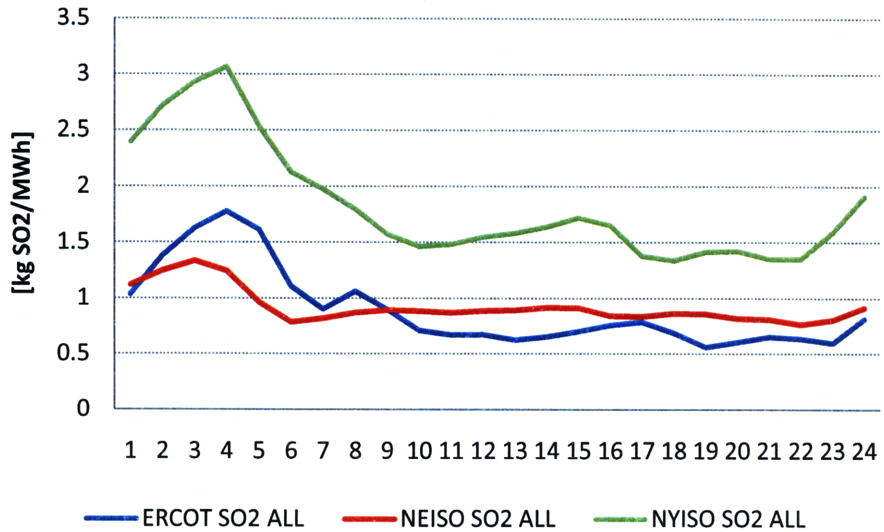

Figure 18 2004-2008 Average Hourly Marginal SO2 Rate

4.2.4. SO2 Average hourly avoided emission rates w/ seasons

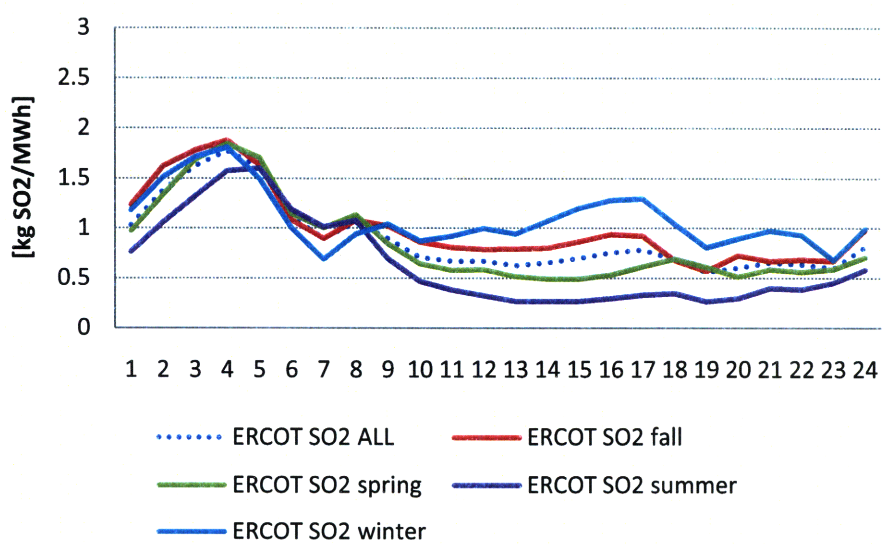

Figure 19 2004-2008 ERCOT Average Hourly Marginal SO2 Rate w/Seasons

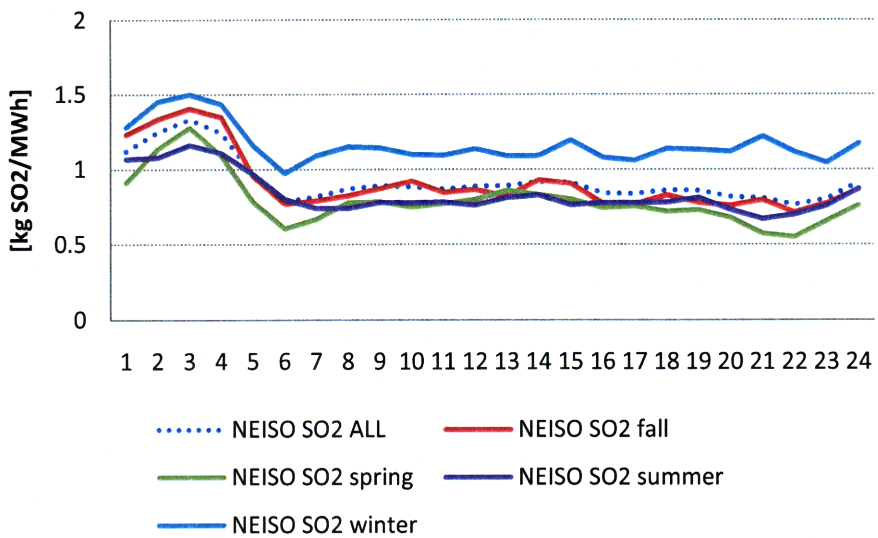

Figure 20 2004-2008 ISO-NE Average Hourly Marginal SO2 Rate w/Seasons

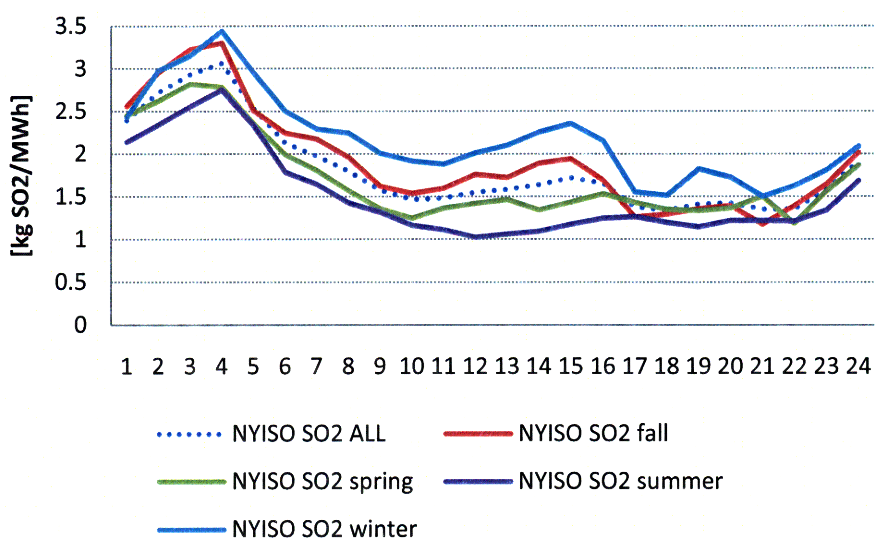

Figure 21 2004-2008 NYISO Average Hourly Marginal SO2 Rate w/Seasons

Out of the three studied compounds, the SO2 rates above provide the most consistent bimodal diurnal pattern across all three control areas. The summer months on peak hours show a consistent drop but then the 2nd daily peak returns in the winter, with the NYISO emission rates reflecting this more strongly. The northeast control areas show consistent yearly reduction of

SO2 and as reported in [14], this may be attributed to the EPA Clean Air Act. The early morning peak exhibits the same hour 4-6 spike as in the CO2 hourly emission rate averages.

4.3. NOx Emission Rates

4.3.1. NOx Hourly avoided emissions rates by year

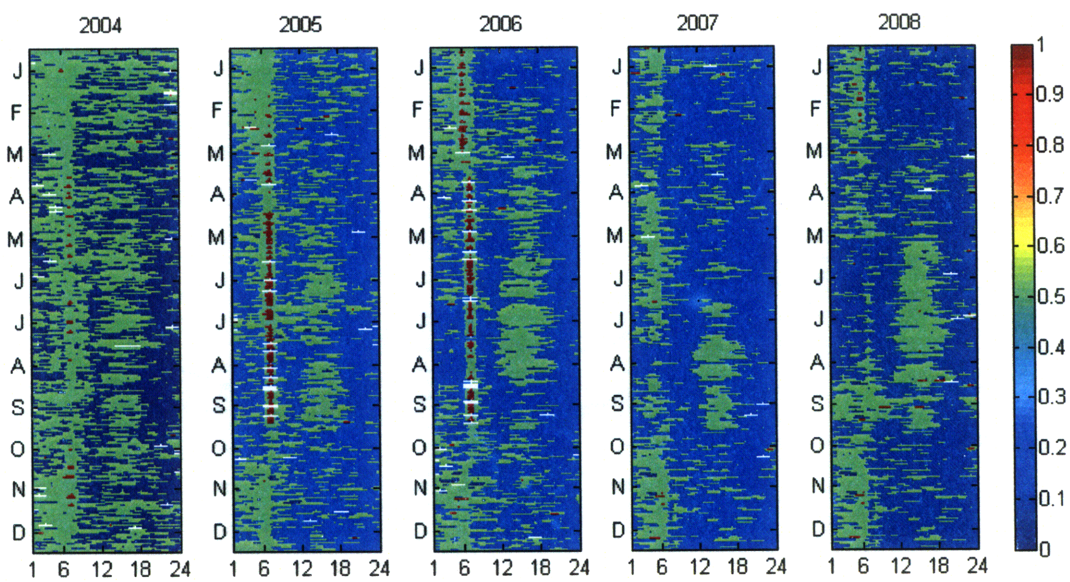

Figure 22 ERCOT Marginal NOx Emissions Rate [kg NOx/MWh]

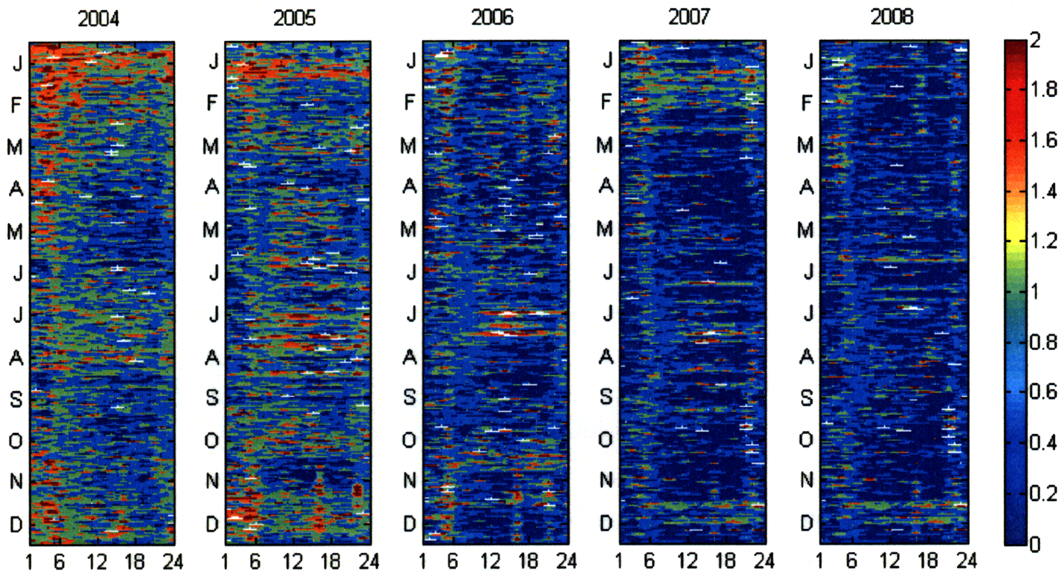

Figure 23 ISO-NE Marginal NOx Emissions Rate [kg NOx/MWh]

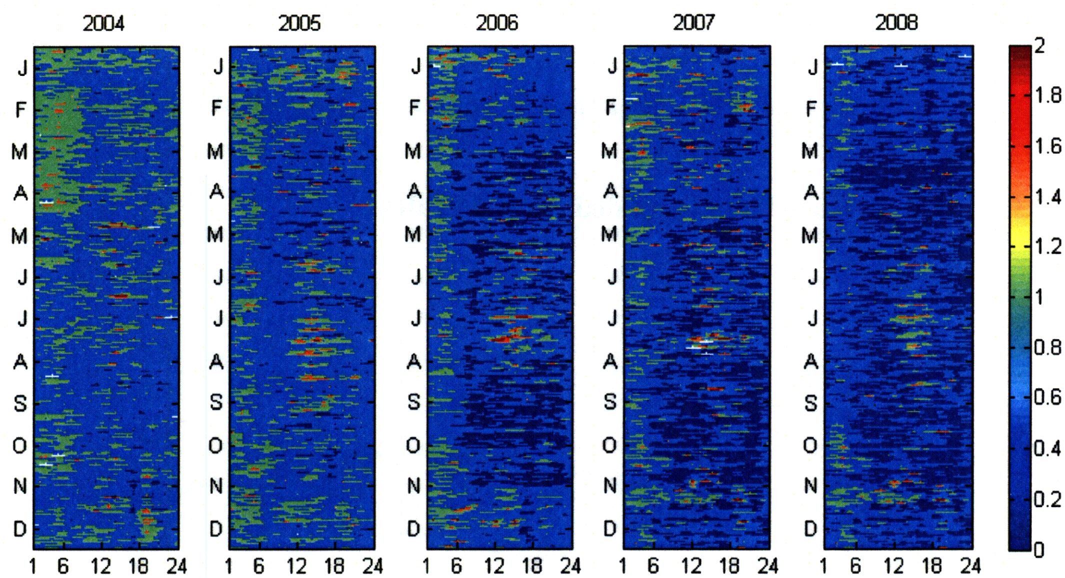

Figure 24 NYISO Marginal NOx Emissions Rate [kg NOx/MWh]

4.3.2. NOx Average daily avoided emissions rates

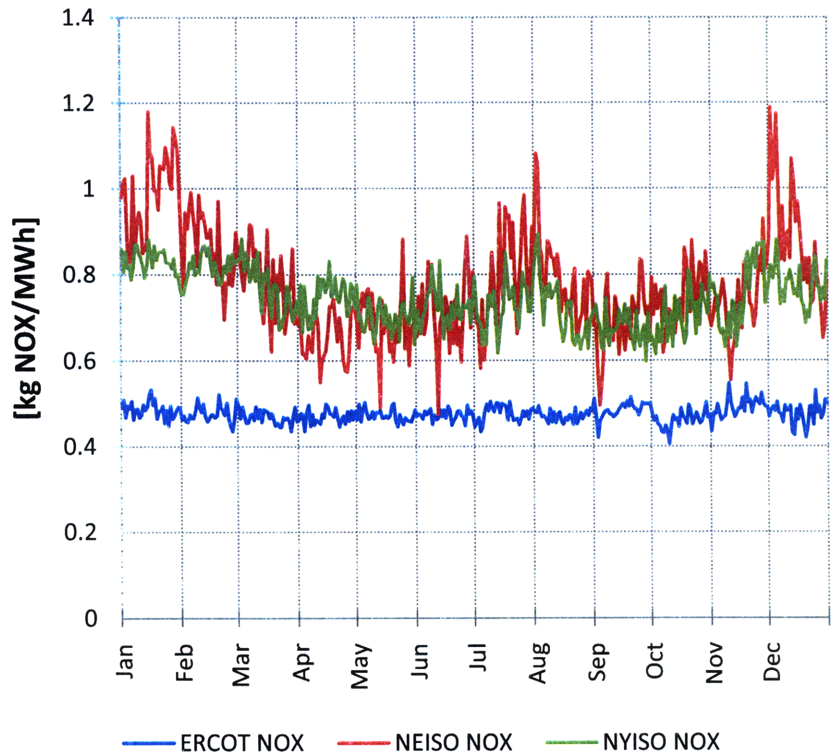

Figure 25 2004-2008 Average daily Marginal NOx Rate [kg NOx/MWh]

4.3.3. NOx Average hourly avoided emission rates

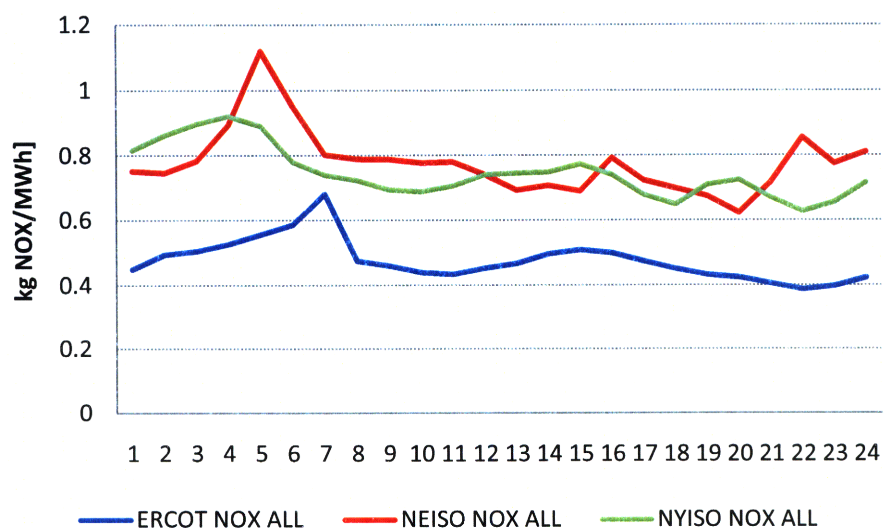

Figure 26 2004-2008 Average Hourly Marginal NOx Rate

4.3.4. NOx Average hourly avoided emission rates w/ seasons

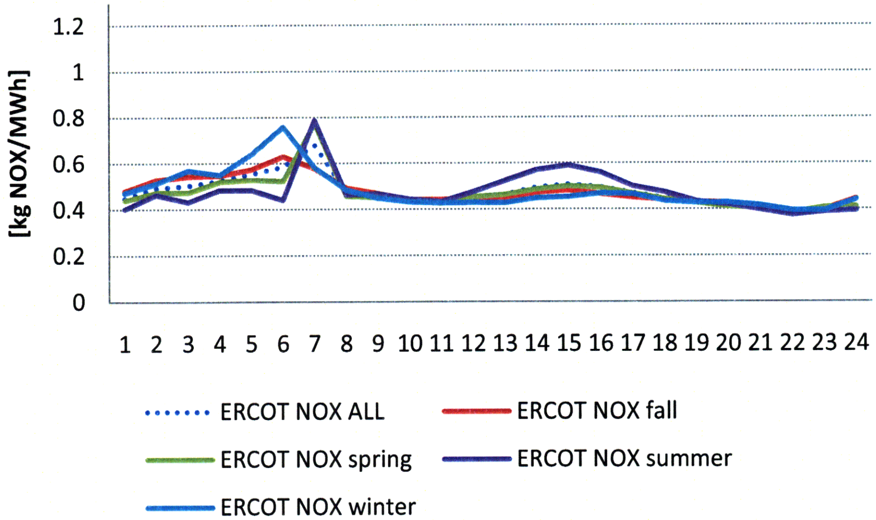

Figure 27 2004-2008 ERCOT Average Hourly Marginal NOx Rate w/Seasons

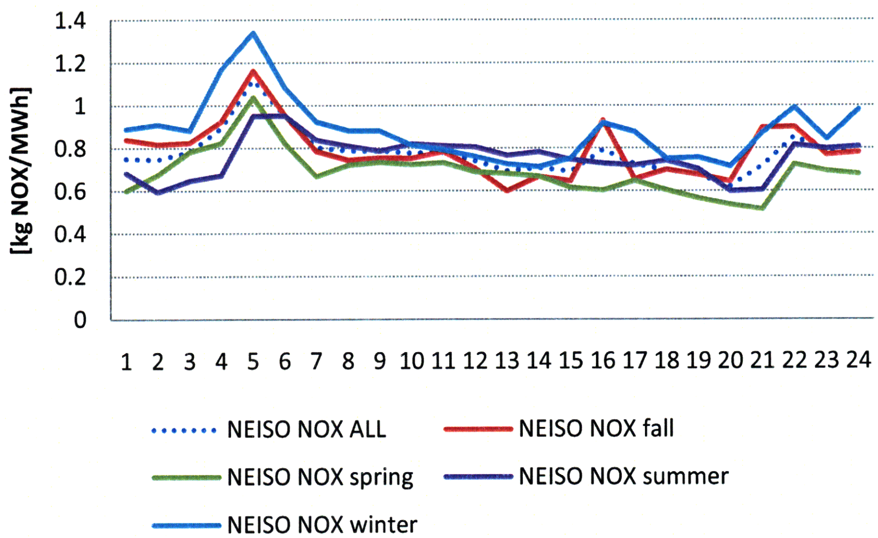

Figure 28 2004-2008 ISO-NE Average Hourly Marginal NOx Rate w/Seasons

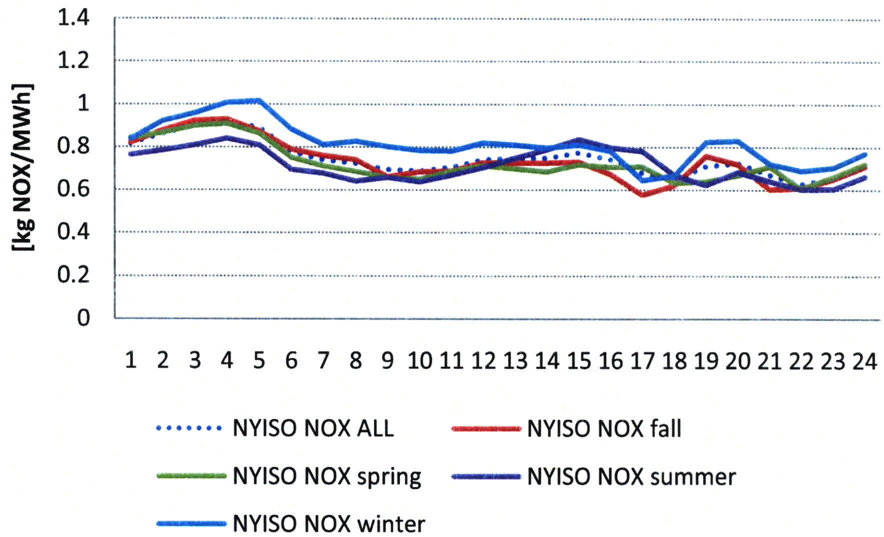

Figure 29 2004-2008 NYISO Average Hourly Marginal NOx Rate w/Seasons

All three control areas exhibit a steady decrease in emission rates between 2004 and 2008 similarly to the SO2 rates as mentioned above. For ERCOT a bimodal diurnal pattern is most pronounced in the summer time and less in the winter, where the on peak spike is relatively negligible. This is contrary to the CO2 and SO2 compounds which dropped in the summer on peak hours and picked up again in the winter on peak hours. Another interesting observation in the ERCOT emission rates is in Figure 25, where the average NOx rates is steady year round at a value of approximately 0.5 [kg-NOx/MWh]. With the high natural gas generation mix leaning toward a higher probability of marginal gas, it is worth noting if there is some significance to this observance. As shown in Figure 30, for natural gas generators with combined gas cycle turbines (CCGT) and open gas cycle turbines (OGCT), the NOx rates are intentionally minimized by introducing lean premixed combustion. This process mixes fuel and air prior to combustion and allows for lower NOx emissions [12]. The process however can render a turbine unstable at operational levels, where the NOx emissions rates jump when the premix system is turned off, but then drops when the unit is ramped up and the premix system is turned on again. As generation capacity approaches 60%-70% , the switching to the premixture flattens the NOx rate at a minimum between 0.4-0.5 [kg NOx/MWh]. The curves in Figure 30 give the impression that if natural gas is the marginal fuel for the load shape following generators then the average NOx rate may fall close to these numbers when averaged. It would be interesting to investigate and strengthen this hypothesis by studying the generation capacities of the load shape following generators to see how many fall in the 60+% range, and to correlate the hours showing higher emission rates with either gas generators at lower capacities or a different fuel type such as coal. The summer on peak hour spike may also be attributed to this phenomenon as gas generators are dispatched to keep up with increased demand on a short term basis, thus having marginal units in the lower capacity ranges as they ramp for service.

The northeast control areas follow similar patterns with an early am peak and on peak spike stronger in the winters. Both average monthly and average monthly patterns are similar both in magnitude and shape but average at least twice as high as the ERCOT emission rates.

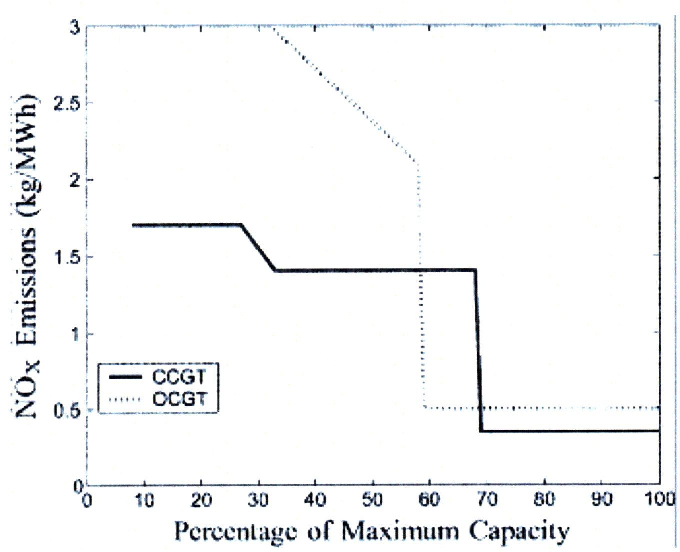

Figure 30 Typical NOx emissions from combined cycle gas turbine (CCGT) and open cycle gas turbine (OCGT) [12]

4.4. Key insights from calculated avoided emissions rates

Model stability and extraneous data

Overall, the calculated avoided for all emission rates results falls within reasonable orders of magnitude. Extraneous data points did occur in some instances however and were left out and did not impact averages. This attributed to a small subset of hours and is not evident in the 8760 color maps as they would present themselves as 'blank' spots. There were, however, milder effects in the results that led to visual patterns attributed to possible model instability. For example, the ERCOT NOx emission rates for the years 2004-2006 (Figure 22) exhibit a "striping" effect across the year occurring in hour-ending 7 (6:00am-7:00am), with values peaking sharply in this particular hour. This may be attributed to a combination of the load shape following algorithm and the turning off of the lean premix as explained below and shown in Figure 30. As described in [7] the LSF approach is subject to numerical instability because "the emissions rates may change from hour-to-hour even when total load is changing only slightly", and for load shape following units the weighted average emission rate calculation, as shown in Figure 4,this may give unreasonably high results.

Time and seasonally varying rates

Examining the time varying behavior of marginal emission rates is worthwhile as the results show true seasonal and hourly variability, even when averaged over a period of the five measured years. When studying the emissions benefits of renewable resources the avoided marginal emission curves allow the time varying and intermittent characteristics of sources like wind generation to be taken into account and weighted more accurately when deriving avoided emissions using this level of detail. As described in the next section data at hourly level grouped into seasonal and less granular time of day buckets are still reasonable at capturing the fluctuations of the marginal emissions rates and based on the type of renewable being studied. Wind for example has certain magnitudes fluctuations between night and day as do the marginal emissions rates themselves and when multiplied together will yield an aggregate number more representative of the avoided marginal emissions than using a system wide calculated average emission rate.

4.5. Using hourly avoided emission rates in forward looking analyses

The above rates can be combined and summarized into a more useful format that compresses the hourly resolution of both the marginal emission rate and renewable generation curves, while preserving enough seasonality and time of day variability. As discussed in [14] a useful way of organizing data is in "marginal emission rate cohorts", which is represented as a table summarizing the emission rate averages over all years grouped by season and time of day (day, evening, night). These are shown for each compound and control area in the following figures in Appendix 8.1. The emission rates above show the behavior of avoided emission rates in the past, but in order to solve future problems in emissions reduction one must prepare the data in such a way as to estimate the future. Summarizing the data into seasonal and time of day buckets by averaging the emissions rates over 2004-2008 is a simplified way of capturing the time varying manner of the avoided emissions rates and easily transportable. In the example of calculating the avoided emissions rates for a windmill project, determining an hourly generation forecast for a future year is a speculative exercise that can be better represented through seasonal and time of use buckets that at least capture high level patterns and shifts in the data.

4.6. Comparing LSF avoided emission rates with eGRID "Non-Baseload" emission rates.

The eGRID database provides two average emission rates in its published data set, the first is the baseload emissions rate as well as a non-baseload emission rate, for the CO2, SO2, and NOx compounds. The derivation of the non-baseload data is based on assuming that all plants with a capacity factor < 0.8 are in a marginal state and non baseload contributing. The output emission rates of the generators that fall below 0.8 are weighted by the generation in a weighted average calculation but based only on plant-level data, not hourly data [23]. Just to see where the yearly average LSF avoided emission rates fall among the eGRID baseload and non-baseload rates, they are shown below in Figures 31-33. It is apparent that the two sets of eGRID rates are fundamentally different by definition and incomparable to the LSF rates, especially since the LSF avoided emissions rates at a bare minimum are most powerful when applied through the seasonal and time of use cohorts to expose the time varying characteristics. Nonetheless, it is interesting to see that for CO2, all three average rates are surprisingly close to each other, especially for the NYISO control area. Another interesting data point is the ERCOT NOx non-baseload rate of approximately 2.5 [kg NOx/MWh], which is contrary to the rate ramp phenomena up for CCGTs below a 70% due to the capacity factor [12]. The average rate is less than half the lowest typical CCGT emission rate, which may mean that averaging the emissions rates in this way do not capture the effect as reflected in the LSF method.

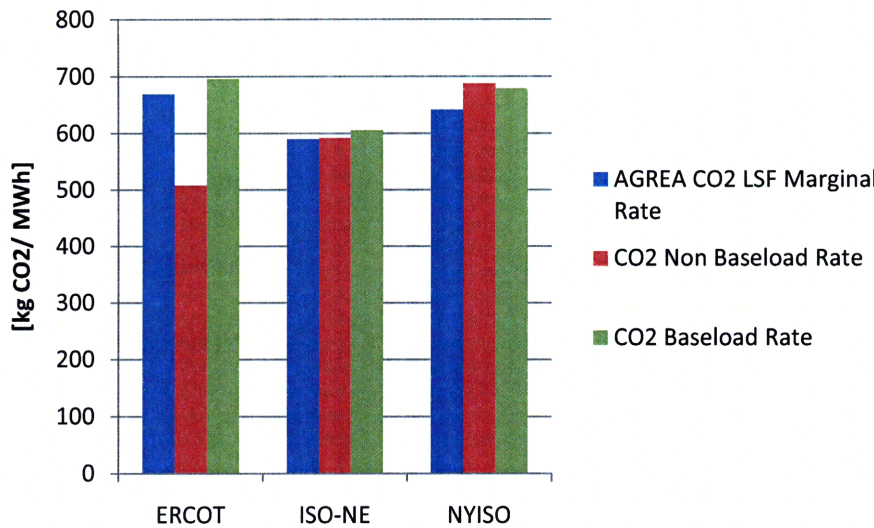

Figure 31 CO2 emission Rate Comparisons LSF marginal vs. eGrid Non-Baseload

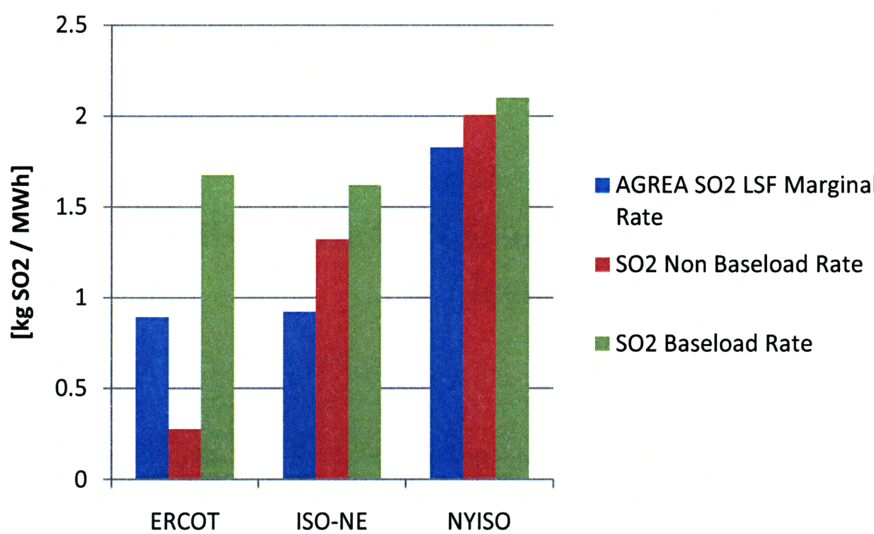

Figure 32 SO2 emission Rate Comparisons: LSF marginal vs. eGrid Non-Baseload

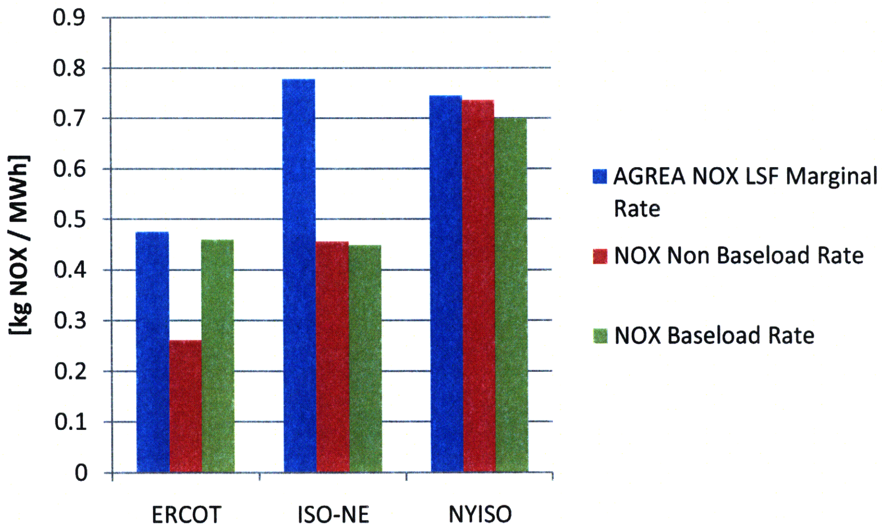

Figure 33 NOx emission Rate Comparisons: LSF Marginal vs. eGrid Non-Baseload

5. Applying marginal emission rates to the Scituate Wind Feasibility Study/Proposal

Background

The Town of Scituate, Massachusetts, which is located along the Massachusetts south shore (Figure 34), has undergone a wind feasibility study for the installation of a proposed lone wind resource. The wind feasibility study provided by Kema Consultants [20] for the Scituate Renewable Energy Committee provides a full scale analysis on the geography, wind resource, and discussion of the viability of the selected location for a lone windmill installation. The results of the feasibility study are lacking any comment on the benefits of avoided emissions by the wind resource so the focus here is to apply the results from the calculated LSF marginal avoided emission rates in a fashion similar to [14] which was a study of a site in Hull, MA. The goal here is to demonstrate the mechanics and data processing involved with applying marginal avoided emission rates to a potential renewable energy resource. In order to calculate avoided emissions for a resource, a generation curve must be derived but can be from any type of resource such as wind, solar, demand response, etc. The benefits of avoided emissions are calculated base on hourly resolutions regardless of how that hourly generation curve is produced. When dealing with a specific application or possible renewable resource, one may not always be provided with an hourly generation curve so one may sometimes need to be derived. The main reason for this is that the future is best told through probabilities, with averages and tolerances sometimes being the best method of describing certain phenomena. This is the case of a potential wind resource which is described via capacity factors and where the *number* of hours of wind at a certain speed are adequate, as opposed to the *actual hour*ly curve of wind speeds by time and date.

As just mentioned, most wind feasibility studies do not provide the hourly resolution necessary to use to calculate total potential avoided emissions, so the following is the method used to extract the hourly resolutions. I will explain the process I created from the base Scituate wind data to calculate the hourly generation curves used in the avoided emissions calculation. I used an industry accepted software created by EMD called Windpro to not only run an MCP (Measure Correlate Predict) algorithm but also to apply wind shear scaling to come up with the hourly output of the proposed Scituate wind site. The resulting hourly resolutions were then summed up into the same seasonal and time of day buckets and were transformed into capacity factors – capacity factors per installed MW. This was done since one of the unknowns in most wind feasibility studies is the actual nameplate capacity (or wind turbine) that will ultimately be chosen. The resulting capacity factor tables can easily be multiplied by the proposed capacity to scale up appropriately.

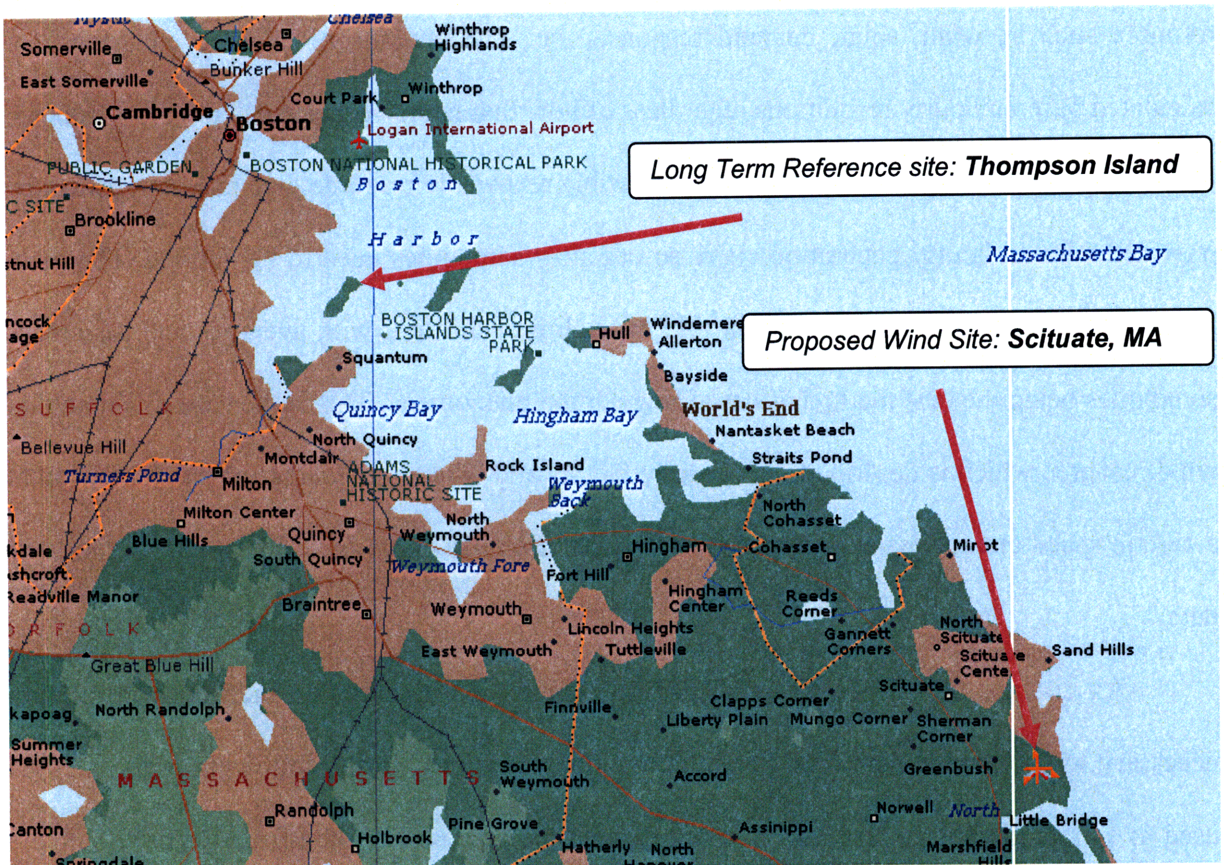

Figure 34 Proposed Site of Scituate Wind Mill in Scituate, MA and long term reference site on Thompson Island (for long term correlation algorithm 'MCP')

5.1. Process of determining capacity factors and avoided emissions for Scituate wind resource

1. Wind Speed Data - Input
   a. The Scituate wind feasibility study is hardcopy/soft copy published by a consulting firm, but all raw wind speed data is published by the UMASS RERL laboratory, as they provided the raw data for the feasibility study. So the data was downloaded from the UMASS RERL website for the following two sites:
      i. Scituate East Coast site wind data (speed/direction), 39 m height – 06/27/2006 – 07/1/2007
         1. Average wind shear factor: 0.37 (for wind shear calculations)

        ii. Thompson Island, Boston wind data (wind speed direction) , 40 m height – 01/01/2002 – 11/27/2007

2. Site – Scituate, MA
    a. As demonstrated in Figure 34, the proposed site near a waste water and treatment facility in Scituate, MA
        i. Geo information
            1. UTM Northing: 4670742
            2. UTM Easting: 357278.8
            3. 42.17581 deg N
            4. 70.72806 deg W
3. Wind turbine data
    a. Three wind turbines are possible candidates in the wind feasibility study
        i. Fuhrlander 600 (600kW)
            1. 50m hub height
        ii. GE 1.5 sle (1.5MW)
            1. 62m hub height
        iii. Vestas V90 2MW
            1. 80m hub height
        iv. 62 meter hub height
            1. Power Curve
            2. Hub height for wind shear analysis
    b. For this study, the GE 1.5 sle was chosen since it fell in the middle range of the 3 turbines as well as the results of the economical study led toward this being the most feasible.
4. Scaling Scituate wind speed data using Measure Correlate Predict (MCP) algorithm and applying wind shear
    a. The Thompson island reference site was used in the wind feasibility study as it was close geographically had a good history of wind data collected.
    b. The MCP algorithm was fed with both the time series of wind speeds form the Scituate site as well as the Thompson Island site, with the only requirement being that the data series overlap for at least a one year period.
    c. Using the Windpro software,
        i. Processing
            1. Data loading
                a. Loaded both Scituate and Thompson island wind speeds and directions.
                b. Data quality was very high due to preprocessing already done by UMASS RERL laboratory. The descriptions of their data cleansing can be found at [21].
            2. MCP exercise
                a. Linear regression was chosen as the method of choice in the MCP exercise. The result was a time series with the same date/time range of the input Scituate date.
            3. Wind Shear

a. Loaded 0.37 as the average wind shear coefficient to scale the results of the MCP exercise time series from the 39m to the 62m hub height.
b. The result of the wind shear scaling was a time series with a higher average wind speed due to the increased hub height at 62m.
c. Average wind speed of adjusted time series: 6.6 m/s

4. Power/generation curve of proposed wind turbine
   a. The power curve for the GE sle 1.5 turbine is shown in Figure 36.

5. Generating avoided emissions
   a. The capacity curve of the wind turbine was divided by the peak capacity to get general number per 1 MW installed.
   b. The scaled year of hourly wind speeds was multiplied against the capacity curve to get a resulting hourly generation curve.
   c. The year of hourly wind generation was summed up into the seasonal and time of day buckets (cohorts) as shown in Table 4. To capture the variability of the wind resource, three "wind year" scenarios were calculated. The medium year is the median wind year from all measured years of wind data and the low and high years were calculated by taking the respective minimum and maximum wind speeds over the same period. The three respective years will provide a respectable variance when describing the possible benefits of a resource in the future.
   d. The results of the above were multiplied against the marginal avoided emissions cohorts for marginal avoided emissions. Then multiply by the number of hours in each bucket to determine the final avoided emissions for that bucket.

5.2. Results

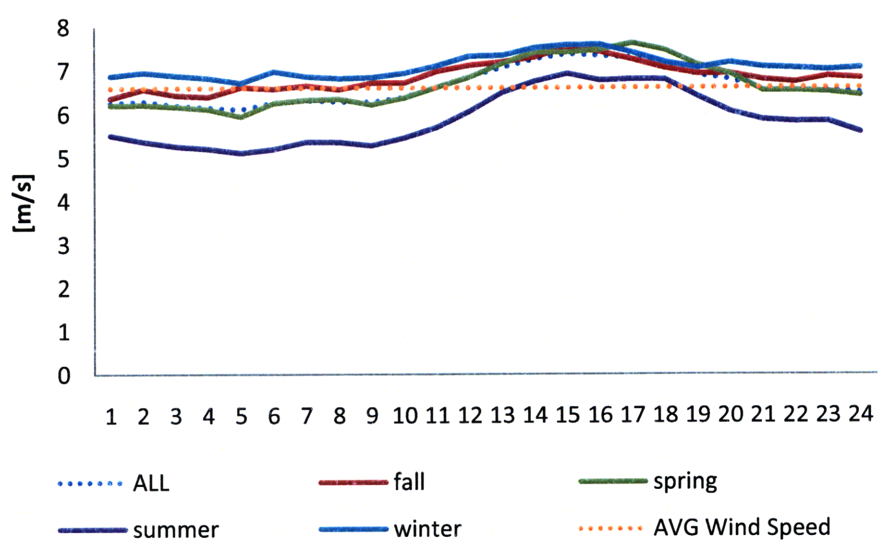

Figure 35 Scituate Average hourly wind speeds for at 65m hub height w/ long term site correlation (2004-2008)

The adjusted Scituate wind speeds are show in Figure 35 with an overall average wind speed of 6.6m/s. For the above analysis, the calculation of the scaled MCP wind speeds as well as the resulting wind shear scaling resulted in the same average wind speed in the Wind Feasibility Study [20], giving validation to the exercise laid out in the above steps. The wind speeds actually peak in the on peak hours and show a small peak in the early morning hours as well. The resulting capacity factors, by seasonal and time of day bucket are shown in Table 4 for the resulting three high, medium, and low wind years. These are then multiplied by equivalent NEISO avoided emission rate buckets in Tables 8,9,10 found in Appendix 8.1 respectively, for each compound. The result is then multiplied against the corresponding bucket hour totals in Table 11 to come up with the avoided emissions.

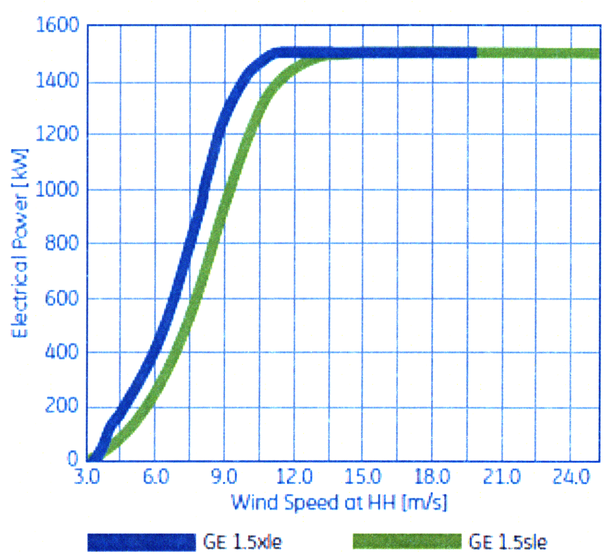

Figure 36 Power Curve of 1.5MW GE sle wind turbine (source EMD Windpro [Ref])

	Season	*All Hours*	*Day*	*Evening*	*Night*
High	**Year-round**	0.34	0.37	0.34	0.29
	Winter	0.40	0.43	0.40	0.37
	Spring	0.35	0.39	0.33	0.27
	Summer	0.25	0.28	0.25	0.19
	Fall	0.36	0.38	0.37	0.32
Medium	**Year-round**	0.31	0.33	0.30	0.26
	Winter	0.36	0.37	0.35	0.33
	Spring	0.31	0.34	0.30	0.26
	Summer	0.23	0.26	0.22	0.17
	Fall	0.33	0.35	0.32	0.29
Low	**Year-round**	0.28	0.30	0.26	0.23
	Winter	0.32	0.32	0.29	0.28
	Spring	0.28	0.31	0.26	0.23
	Summer	0.20	0.24	0.18	0.14
	Fall	0.31	0.32	0.29	0.26

Table 4 Scituate Wind Site Capacity Factors by Season and time of day [MWh/Year of Production]. High, medium, and low typical wind years are shown.

		Total kg	
CO2	High	1,770,147	11.6%
	Median	1,585,732	
	Low	1,403,040	-11.5%
SO2	High	2,795	11.8%
	Median	2,495	
	Low	2,203	-11.9%
NOx	High	2,343	11.7%
	Median	2,098	
	Low	1,85	-11.7%

Table 5 Potential avoided emissions at Scituate site w/ 1.5MW capacity wind mill [kg] for high, medium, and low wind years.

The potential avoided emissions for the Scituate site are summarized in Table 5. For a frame of reference, the resulting median avoided emissions were run through the eGrid Equivalency Calculator for Emissions [24] and a variety of emissions benefit equivalents can be seen in Table 6.

Annual Greenhouse gas emissions from	290 vehicles	
Annual CO2 emissions of	0.0003 power plants	(coal fired)
Carbon sequestered annually by	360 acres	of pine or fir forests
Carbon sequestered annually by	11.1 acres	of forest preserved from deforestation
Carbon sequestered by	40660 tree seedlings	grown for 10 years
CO2 emissions from	179,992 gallons	of gasoline consumed
CO2 emissions from	3,688 barrels	of oil consumed
CO2 emissions from	21.2 tanker trucks	worth of gasoline
CO2 emissions from	66,072 propane cylinders	used for home barbeques
CO2 emissions from burning	8.3 railcars	worth of coal
CO2 emissions from the electricity use of	220 homes	for one year
CO2 emissions from the energy use of	144 homes	for one year
Greenhouse gas emissions avoided by recycling	547 tons	of waste instead of sending it to the landfill

Table 6 Equivalent CO_2 greenhouse gas emissions avoided from 1,585,732 kg due of proposed Scituate windmill [24]

5.3. Summary

Using calculated emissions rates to calculate benefits of renewable resources

The Scituate wind feasibility example shows the simplicity of applying the calculated emissions rates to calculate the benefits of a renewable resource. Once the mechanics of calculating the rates and deriving an hourly profile of the renewable resource are finished, the exercise easily provides a forward looking estimate. The wind example shows how the potential benefits of a quasi-stochastic intermittent resource can be estimated by applying emissions rates that are time varying as well to capture daily and seasonal effects.

The process of determine the avoided emissions for the Scituate project demonstrate the mechanics of a more general approach that set the framework of determining the emissions benefits of a future renewable resource. Any renewable generation resource can be used as the basis for an avoided emission study using the above calculated marginal emission rates as long as the generation curve is available. If the resource has intermittent characteristics as in the case of wind, it is beneficial to provide a forecasted value and a mechanism of determining the possible variance of the resource when looking into the future. This can be done by calculating low, medium, and high generation curves to calculate a reasonable range of expected avoided emissions.

6. Conclusions and Recommendations

6.1. Methodology and Data

This study provided a design of a system to calculate the marginal avoided emission rates for multiple control areas with the flexibility of adding more in the future and iterating the emission rate calculations in a straightforward manner. A programming codebase for applying the load shape following (LSF) algorithm, as well as calculating the weighted average emission rates was designed and created in a thin set of procedures that run as close to the database as possible for efficiency of processing and in a language understood by a general energy analyst with knowledge of SQL basics. This can open doors in the future if the LSF algorithm were to further be researched or if basic assumptions need to be changed.

The mechanics of calculating the marginal rates for the three regions in this study can be applied to other control areas in a similar fashion with the simple addition of the corresponding regional load data and hourly level generator emission data. Several areas within this framework can be analyzed or changed for future study including the LSF algorithm itself, the weighted average calculation of determining emission rates, and the study of rates over a longer time horizon. Published versions of the resulting avoided emission rates can be collectively updated and maintained through incremental maintenance of the underlying data inputs which will be continually updated by the EPA.

6.2. Insights and comparisons for ERCOT, NY and NE and beyond

The comparison and contrast between the ERCOT, ISO-NE, and NYISO control areas showed some interesting results with ERCOT showing some correlations between the calculated avoided emissions rates with its high dependence of natural gas in the overall ERCOT generation mix. The analysis opened the door to future studies that may look deeper into the subject of marginal natural gas units and their resulting emissions rates in a large pool with generators of other fuels in the load shape following paradigm. ERCOT showed a lower NOx average marginal emission rate likely due to the possibility of many marginal gas generators operating in the generation capacity range of 60% and higher [12]. The NOx results in ISO-NE did not exhibit this behavior even though the mix of natural gas is not substantially lower than that of ERCOT. In future studies, perhaps the LSF method can be transformed into a hybrid model linking supply stacks and bid/demand curves to come up with a rate calculation that mimics the more complex dynamics of the power grid. As stated in [7], the LSF method (referred to in [7] as LFIR- Load Following Incremental Emissions Rate) is discounted in a method comparison study of emission rates calculations, as being numerically unstable and for selecting load following units as "excessively inclusive". Perhaps the latter is true but the results of the current analysis provided calculated emissions rates that were mostly stable with a few outliers, but not enough to warrant the results as unusable. When transforming the hourly emissions rates into the corresponding seasonal and time of day buckets, the hourly resolution is smoothed out in a sense so any outliers that fall under an umbrella maximum will get averaged down, providing a smoother set of data.

6.3. Application to renewable project avoided emissions calculations

The resulting avoided emission rates for compounds CO2, SO2, NOx are presented with reference avoided emission rate tables for future use. As provided in the Scituate Wind Feasibility study, the emissions rates tables can be applied in a systematic process to achieve an end result of a renewable project in a structured manner. In this particular example, an hourly time series of wind speeds was calculated and the resulting avoided emissions calculated after a defined series of steps. In the case of applying the same methods to future renewable projects, all that is needed is either the historical or forecasted hourly generation curve of the renewable source. If it is a seasonally and/or daily varying resource such as wind or solar generation, the hourly resolution provides enough information to calculate the benefits of the avoided emissions using the seasonal and time of day emission rate tables published in this study.

6.4. Recommendations

In closing, the results from calculating avoided emissions rates using the LSF method differ from most avoided emissions rates due to the seasonal and time of day buckets that capture patterns. Even if the results are "bucketed" into a discrete set of compartmentalized averages, the message is clear that when interested in calculating emissions avoided due to a renewable resource, the notion of "when" the resource is running is vital to quantifying the total emissions avoided. In that sense, the seasonal and time of day avoided emissions buckets are unlike the very widely used system-wide average emissions rates published by the EPA in their eGRID databases and market reports. Perhaps through more analysis and even through even more simplification of the LSF methods, institutions such as the EPA, ISOs, or utilities will take note and implement emission rate methods that are more time based.

7. References

[1] Berlinski M., Quantifying Emissions Reductions from New England Offshore Wind Energy Resources, MIT Master of Science Thesis, Engineering Systems Division, June 2006, http://dspace.mit.edu/handle/1721.1/34518

[2] Berlinski M., Connors S., and O. Daly-Jones. Economic and Environmental Performance of Potential Northeast Offshore Wind Energy Resources, In MIT LFEE 2006-02 RP, 1/31/2006.

[3] Connors, S., Adams M.,Martin K, Kern E., Tedd J., Avoided Emissions from Renewables in New England, Where and When Matters., April 2004, http://ww.mtpc.org/renewableenergy/public_policy/dg/resources/economicsDG-renewables.htm

[4] Connors S., and Martin K, Adams M. Kern E., Baafour A., Emissions Reductions from Solar Photovoltaic (PV) Systems, In MIT LFEE 2004-003 RP, August 2004.

[5] General Electric Energy, Analysis of Wind Generation Impact on ERCOT Ancillary Services Requirements, pg. 5-1 thru 5-8. http://www.ercot.com/calendar/2008/02/20080227-WIND.html, 3/21/2008.

[6] Wind Data Report, Scituate Waste Water Treatment Plant http://www.ceere.org/rerl/publications/resource_data/Scituate/Reports/Scituate_2007_FinalReport.pdf , 9/28/2007.

[7] Hausman, E, Fisher, J, Biewald, B. Analysis of Indirect Emissions Benefits of Wind, Landfill Gas, and Municipal Solid Waste Generation, July 23, 2008, Synapse Energy report submitted to Air Pollution Prevention and Control Division, National Risk Management Research Laboratory, (p 30 comment on LSF method).

[8] Petron, G, Tans, P,Frost G, Chao D., Trainer M, High-resolution emissions of CO2 from power generation in the USA, October 2008, Journal of Geophysical Research, VOL. 113, G04008, doi:10.1029/2007JG000602, 2008

[9] Ackerman, KV, Sundquist E T, Comparison of two US power-plant carbon dioxide emissions data sets, August 2008, Environmental Science & Technology, VOL. 42, Issue 15, pp. 5688-5693

[10] Marland, G, Hamal, K, Jonas M, How Uncertain are Estimates of CO2 Emissions, February 2009, Journal of Industrial Ecology, VOL. 13, Issue 1, pp. 4-7.

[11] Denny, E, O'Malley, M, Quantifying the Total Net Benefits of Grid Integrated Wind, May 2007, IEEE Transactions on Power Systems, Vol. 22, No. 2, pp. 605-615.

[12]Denny, E, O'Malley, M, Wind Generation, Power System Operation, and Emissions Reduction, February 2006, IEEE Transactions on Power Systems, Vol. 21, No. 1, pp. 341-347.

[13] Gil, H, Joos, G, Generalized Estimation of Average Displaced Emissions by Wind Generation, August 2007, IEEE Transactions on Power Systems, Vol. 22, No. 3, pp. 1035-1043.

[14] Rached, Tarek, Communicating complexity and informing decision-makers: challenges in the data and computation of environmental benefits of renewable energy, June 2008, Master's Thesis, Master of Science of Technology and Policy, MIT.

[15] CEM data (Continuous Emissions Monitoring) database Provided by the EPA Clean Air Markets division, http://camddataandmaps.epa.gov/gdm/index.cfm?fuseaction=emissions.wizard

[16] EMD WindPro Software, for wind study and analysis, Measure-Correlate-Predict algorithm: http://www.emd.dk/WindPRO/Frontpage

[17] FERC Website – ISO control area maps: http://www.ferc.gov/industries/electric/indus-act/rto.asp

[18] AWEA response to RGGI Model Rule, RGGI website, http://www.rggi.org/docs/awea.pdf

[19] RGGI Regional Green House Gas Inititative website http://www.rggi.org , auction results, Model rule, supporting documents.

[20] Town of Scituate Wind Feasiblity Study, Provided by Kema Consultants and Renewable Energy Committee of Town of Scituate: http://www.town.scituate.ma.us/renewable/index.html

[21] Renewable Energy Research Laboratory Wind speed data, Scituate East Coast, http://www.ceere.org/rerl/publications/resource_data/Scituate/index.html

[22] EIA- Energy Information Administration – Official Energy Statistics from the US Government, http://www.eia.doe.gov

[23] eGrid, Emissions and Generation Resource Integrated Database http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html

[24] eGrid Equivalency Calculator for Emissions - http://www.epa.gov/cleanenergy/energy-resources/calculator.html

[25] ERCOT, Electric Reliability Council of Texas website- http://www.ercot.com

[26] ISO-NE, Independent System Operator of New England website - http://www.iso-ne.com

[27] NYISO, New York Independent System Operator website - http://www.nyiso.com

8. Appendices

8.1. Appendix 1 Marginal Emission Rate Time of Day/Season Cohort Tables

The seasonal and time of day average avoided emission rates from years 2004-2008 are shown below by compound.", avoided emission rate buckets are calculated by averaging all rates in each "bucket" or "cohort" first and then multiplying the resulting average by the sum of total hours in each "bucket" shown in Table 11. This is the output of the load shape following avoided emission rate exercise and is what sets this method apart from similar avoided emission rate categories which only provide only ONE average number *for all seasons and hours*. If anything is to be taken away from this study it is that avoided emissions rates must take into account seasonality and time of day, as the marginal generation is time varying due to a dynamic power system (and market).

8.1.1. Seasonal and time of day hour counts and definitions used in table calculations

The following table is useful in calculating the marginal avoided emissions using the above seasonal/time of day tables. It contains the number of hours in each seasonal and time of day bucket. Very simply, the numbers of hours of each day portion were summed across all days within a season to arrive at the total. The definitions of the time of day values are as follows (Hour ending 1 = 00:00-01:00 am)

- Day : Hour ending 7 – Hour ending 20
- Evening: Hour ending 21 – Hour ending 24
- Night: Hour ending 1 – Hour ending 6

season	all_hours_average	day_average	evening_average	night_average
winter	2136	1246	356	534
spring	2232	1302	372	558
summer	2232	1302	372	558
fall	2160	1260	360	540

Table 7 Seasonal and time of day hour count for Avoided emissions calculations

ControlArea	*season*	*All_Hours_Average*	*Day_Average*	*Evening_Average*	*Night_Average*
ERCOT	ALL	669.890997	653.8631138	613.0076256	745.2116389
ERCOT	winter	678.3144778	668.0273716	616.5927331	743.4655555
ERCOT	spring	668.3055433	650.4021183	606.5885878	751.2248388
ERCOT	summer	653.740301	634.7678781	612.3123398	725.6279285
ERCOT	fall	679.8681188	663.1007517	616.8266661	761.0196107
ISO-NE	ALL	590.1598005	574.3873575	575.4327489	636.7802019
ISO-NE	winter	598.4592267	579.4395479	584.9368133	651.8534194
ISO-NE	spring	576.4955603	567.2236863	547.6987879	617.3277813
ISO-NE	summer	590.1034653	573.0458671	584.4833464	633.6512736
ISO-NE	fall	596.0972967	578.1631981	585.2757499	645.1578913
NYISO	ALL	642.0203454	607.5959743	600.1236896	750.2749819
NYISO	winter	672.3290132	649.5601193	618.3822468	761.4209432
NYISO	spring	639.6006937	602.888956	601.1112334	750.9210551
NYISO	summer	619.495892	577.7618513	595.8893455	732.6130179
NYISO	fall	637.7062729	601.622429	585.3293795	756.8198375

Table 8 CO2 Average marginal avoided emission rates by season and time of day [kg CO2/MWh]

ControlArea	season	All_Hours_Average	Day_Average	Evening_Average	Night_Average
ERCOT	ALL	0.89389	0.731802	0.67264	1.419596
ERCOT	winter	1.092771	0.997181	0.89098	1.45034
ERCOT	spring	0.849095	0.660518	0.610396	1.448243
ERCOT	summer	0.647621	0.446323	0.450856	1.248495
ERCOT	fall	0.997102	0.83685	0.749252	1.536258
ISO-NE	ALL	0.921678	0.867639	0.822924	1.113607
ISO-NE	winter	1.163519	1.113181	1.136514	1.298978
ISO-NE	spring	0.790809	0.760929	0.634325	0.96485
ISO-NE	summer	0.833203	0.773889	0.747889	1.028479
ISO-NE	fall	0.909641	0.832371	0.785944	1.172404
NYISO	ALL	1.828603	1.567259	1.548532	2.625117
NYISO	winter	2.178513	1.986287	1.75593	2.908761
NYISO	spring	1.712096	1.427014	1.526674	2.5009
NYISO	summer	1.517752	1.220325	1.364189	2.314123
NYISO	fall	1.922847	1.654632	1.555445	2.793616

Table 9 SO2 Average marginal avoided emission rates by season and time of day [kg SO2/MWh]

ControlArea	season	All_Hours_Average	Day_Average	Evening_Average	Night_Average
ERCOT	ALL	0.474758	0.476759	0.402387	0.518337
ERCOT	winter	0.479261	0.454107	0.411472	0.583147
ERCOT	spring	0.468832	0.477896	0.402041	0.492209
ERCOT	summer	0.475123	0.510309	0.388242	0.450941
ERCOT	fall	0.476302	0.463885	0.408314	0.550599
ISO-NE	ALL	0.77787	0.73357	0.789345	0.873585
ISO-NE	winter	0.883975	0.804067	0.921298	1.045546
ISO-NE	spring	0.687775	0.654753	0.651616	0.788931
ISO-NE	summer	0.757136	0.760686	0.75613	0.749524
ISO-NE	fall	0.787252	0.716648	0.83557	0.919783
NYISO	ALL	0.744422	0.717507	0.665867	0.859595
NYISO	winter	0.811497	0.783836	0.720694	0.936576
NYISO	spring	0.723809	0.682065	0.674134	0.854329
NYISO	summer	0.713248	0.707613	0.62779	0.78337
NYISO	fall	0.731338	0.698477	0.642189	0.867447

Table 10 NOx Average marginal avoided emission rates by season and time of day [kg NOx/MWh]

8.2. Appendix 2 Description of Scituate Wind Data

The preparation of Scituate Feasibility Wind was based on wind speed data provided for the Scituate, MA site by the Renewable Energy Research Laboratory (RERL) at UMASS. The following is a description of the Scituate data collection site provided by the RERL [21]. The hourly wind data is provided on their website and will not be provided here as it can be downloaded on the RERL website. The Thompson island reference data was downloaded from the RERL website as well.

Site Name: SCITUATE
Location: Scituate, MA
Latitude [N]: 42.17581
Longitude [W]: 70.72806
Time Zone [hrs from GMT]: -5
Elevation [m]: 7
Time Step of Data [seconds]: 600
Logger Sample Interval [seconds]: 2
Report Time Period: 2006-06-27 16:00:00 to 2006-12-31 23:50:00
Data Collection By: RERL @ Univ. of Massachusetts
Report Generated: 2007-04-18 13:02:17
Comments:

Sensors: Sensor Name, Type, Designation, Height Above Ground [m],Measurement Units
Sensor #1: Etemp2aDEGC,Etemp2aDEGC,Primary,0,Units
Sensor #2: Etemp2SDaDEGC,Etemp2SDaDEGC Standard Deviation,Primary,0,Units
Sensor #3: Anem39aMS,Anemometer,Primary,39,m/s
Sensor #4: AnemSD39aMS,Anemometer Standard Deviation,Primary,39,m/s
Sensor #5: Anem39yMS,Anemometer,Consolidated data (larger value selected),39,m/s
Sensor #6: AnemSD39yMS,Anemometer Standard Deviation, Consolidated data (larger value selected),39,m/s
Sensor #7: Vane39aDEG,Wind Direction Vane,Primary,39,degrees
Sensor #8: VaneSD39aDEG,Wind Direction Vane Standard Deviation,Primary,39,degrees
Sensor #9: Anem30aMS,Anemometer,Primary,30,m/s
Sensor #10: AnemSD30aMS,Anemometer Standard Deviation,Primary,30,m/s
Sensor #11: Anem30yMS,Anemometer,Consolidated data (larger value selected),30,m/s
Sensor #12: AnemSD30yMS,Anemometer Standard Deviation, Consolidated data (larger value selected),30,m/s
Sensor #13: Vane30aDEG,Wind Direction Vane,Primary,30,degrees
Sensor #14: VaneSD30aDEG,Wind Direction Vane Standard Deviation,Primary,30,degrees
Sensor #15: Anem10aMS,Anemometer,Primary,10,m/s
Sensor #16: AnemSD10aMS,Anemometer Standard Deviation,Primary,10,m/s
Sensor #17: Vane10aDEG,Wind Direction Vane,Primary,10,degrees
Sensor #18: VaneSD10aDEG,Wind Direction Vane Standard Deviation,Primary,10,degrees

Applied Validation Filter Code, Filter Name
-988, Dissimilar sensors
-989, Out of Range
-991, Icing or wet snow event
-999, Missing Data

8.3. Appendix 3 – AGREA Database & Programming

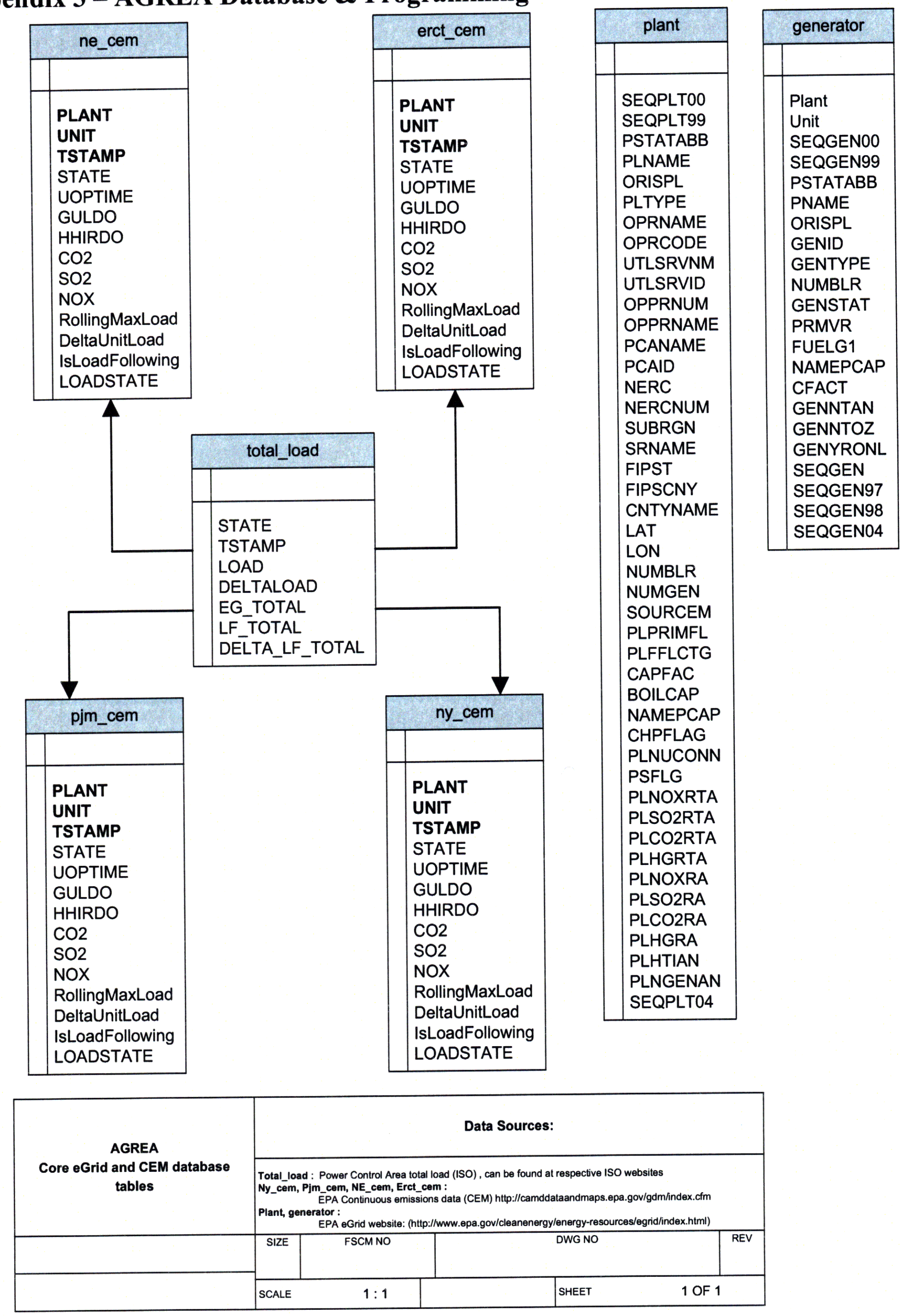

Figure 37 AGREA Database Tables and Structure

8.4. Appendix 4 SQL – Structured Query Language

The following is the core set of SQL (Structured Query Language) code use to process the control area specific generator data and to calculate the load shape following identifier(s).

```
DELIMITER $$

DROP PROCEDURE IF EXISTS `agrea-egrid`.`ControlAreaLoop` $$
CREATE DEFINER=`root`@`localhost` PROCEDURE `ControlAreaLoop`()
BEGIN

DECLARE l_plantid int;
DECLARE l_unitid varchar(6);
DECLARE l_loop_count int;
DECLARE no_more_plant_unit int;
DECLARE l_sql varchar(4000);
DECLARE l_sql2 varchar(255);
DECLARE l_namepcap float;
DECLARE l_tblPlantUnit varchar(64);
DECLARE l_sessionid varchar(45);
DECLARE l_sessiondate datetime;


/*
--Author: Gabriel R. Gomez
          ggomez@mit.edu
          With much credit also to AGREA team members from previous years who originally
programmed LSF algorithm and data processing in PHP/Perl.
--Methods and Pseudocode:

          Goal here is simplification with a focus on backend processing (SQL) with no middle
tier such as PHP or Perl. This is necessary for the millions of rows that are processed for each
individual control area. "Trimming the fat" allows the energy analyst to focus on the problem at
hand and less on the intricacies of data programming.
          SQL is an ANSI standard and can be understood by energy analysts with ease and
practice.SQL can serve as an API (application programming interface) or foundation for a future
web front end (in PHP) for LSF and avoided emission rates.

# Begin Pseudocode:

 Create SQL cursor to create list of generators to loop through during load following code
algorithm
  A "cursor" is another name for a 'loop' in the SQL language.
  Join the regional table (erct_cem,ny_cem,ne_cem, etc) with generator table (eGrid)
  Join on the following fields:
  plant = unique identifier assigned by DOE/EIA and serves (ORISPL)
           as unique identifer for many EPA electric power databases
  unit = unique identifier given to unit (one to many relationship with plant)

  The query will pull all valid eGrid generators for a given control area. The SQL code (via
cursor) will loop through each generator individually and run the LSF algorithm against the total
regional load.
  SQL by nature is a 'set' based language, but set processing in this case is not favorable for
scalability. Creating
  SQL joins to very large tables like these is processor/memory intensive and the solution here
is to loop creating smaller 'temporary' tables for each generator, crunching the numbers, then
saving the result.
  For future research, the LSF algorithm can be tweaked and investigated here, and as more
control areas are added this can serve as the base for the larger datasets.


*/

DECLARE plant_unit_csr CURSOR FOR
SELECT distinct e.plant,e.unit,g.namepcap FROM ERCT_CEM e, generator g
where e.plant = g.plant and e.unit = g.unit #and e.plant = '298' #and e.unit = 'CTG2'
#and e.plant < '3439'
```

```
order by e.plant,e.unit;

DECLARE CONTINUE HANDLER FOR NOT FOUND SET no_more_plant_unit=1;


# agrealog is a log table to store actions and times from code processing. This aids in debugging
# and is a good indication of how long it takes to process a region
#
#ERCOT : 191 minutes
#NYISO : 76 minutes
#ISO-NE: 210 minutes

SET l_sessiondate = now();
SET l_sessionid =
concat(year(l_sessiondate),month(l_sessiondate),day(l_sessiondate),hour(l_sessiondate),minute(l_s
essiondate),second(l_sessiondate));

            insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'Begin pre steps';


## Prepare Temporary Table : tmp_plant_unit

## Using a "temporary table" here for ease of use and performance.
DROP TABLE IF EXISTS tmp_plant_unit;

CREATE TABLE  `agrea-egrid`.`tmp_plant_unit` (
  `Region` varchar(16) NULL,
  `PLANT` mediumint(10) NOT NULL default '0',
  `UNIT` varchar(6) NOT NULL default '0',
  `TSTAMP` datetime NOT NULL default '0000-00-00 00:00:00',
  `YEAR` int NULL,
  `GULDO_N` float default NULL,
  `GULDO_Nminus1` float default NULL,
  `Total_Load_N` float default NULL,
  `Total_Load_Nminus1` float default NULL,
  `LOADSTATE` smallint(6) default '-1' NULL,
  `RollingMaxLoad` float default NULL,
  `IsLoadFollowing` tinyint(4) default NULL,
  `prev_IsLoadFollowing` tinyint(4) default NULL,
  `PeakYearly` float default NULL,
  `namepcap` float default NULL,
  PRIMARY KEY  (`PLANT`,`UNIT`,`TSTAMP`),
  KEY `ix_cem1_tstamp` (`TSTAMP`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

DROP TABLE IF EXISTS tmp_plant_unit_pre;

CREATE TABLE  `agrea-egrid`.`tmp_plant_unit_pre` (
  `Region` varchar(16) NULL,
  `PLANT` mediumint(10) NOT NULL default '0',
  `UNIT` varchar(6) NOT NULL default '0',
  `TSTAMP` datetime NOT NULL default '0000-00-00 00:00:00',
  `YEAR` int NULL,
  `GULDO_N` float default NULL,
  `GULDO_Nminus1` float default NULL,
  `Total_Load_N` float default NULL,
  `Total_Load_Nminus1` float default NULL,
  `LOADSTATE` smallint(6) default '-1' NULL,
  `RollingMaxLoad` float default NULL,
  `IsLoadFollowing` tinyint(4) default NULL,
  `prev_IsLoadFollowing` tinyint(4) default NULL,
  `PeakYearly` float default NULL,
  `namepcap` float default NULL,
  PRIMARY KEY  (`PLANT`,`UNIT`,`TSTAMP`),
  KEY `ix_cem1_tstamp` (`TSTAMP`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;

DROP  TABLE IF EXISTS `agrea-egrid`.`regional_total_load`;
CREATE  TABLE  `agrea-egrid`.`regional_total_load` (
  `REGION` varchar(16) NULL,
```

```
  `TSTAMP` datetime NOT NULL default '0000-00-00 00:00:00',
  `Total_Load_N` float default NULL,
  `Total_Load_Nminus1` float default NULL,
  PRIMARY KEY  (`TSTAMP`)
) ENGINE=MyISAM DEFAULT CHARSET=latin1;


DROP  TABLE IF EXISTS `agrea-egrid`.`tmp_plant_unit_peaks`;
CREATE  TABLE  `agrea-egrid`.`tmp_plant_unit_peaks` (
  `year` int NULL,
  `peakyearly` float NULL,
  `namepcap` float default NULL
) ENGINE=MyISAM DEFAULT CHARSET=latin1;


            insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'End Pre Steps';

INSERT INTO regional_total_load(region,tstamp,total_load_n,total_load_nminus1)
select e2.state,e2.tstamp,e2.LOAD as LOAD_N,e1.LOAD as LOAD_Nminus1
from erct_total_load e1
inner join erct_total_load e2
on e1.tstamp = DATE_ADD(e2.tstamp, INTERVAL -1 HOUR)
order by e1.tstamp;

      insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'End Total Load';

SET no_more_plant_unit = 0;
OPEN plant_unit_csr;

plant_unit_loop:REPEAT

  FETCH plant_unit_csr INTO l_plantid, l_unitid,l_namepcap;
      # CAll Cursor and begin looping

      If no_more_plant_unit THEN
        LEAVE plant_unit_loop;
      End If;

      insert into agrealog (plant,unit,tstamp,session)
      select l_plantid,l_unitid, now(),l_sessionid;

      # Turn data processing into a two part query
      #  1. Pull data from large CEM table and insert hourly data for generator
      #  2. Process data on smaller dataset and update


      #Step 0
        # Empty temp table before loop begins to empty previously processed generator data
      Truncate table tmp_plant_unit_pre;
      Truncate table tmp_plant_unit;

      insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'Step1_PRE';

      #Step 1
        # Insert raw generator data into "pre" temp table.
        # "pre" table is created for easy self-joining for hour and hour-1 in later calculations.
      INSERT INTO tmp_plant_unit_pre(plant,unit,tstamp,year,guldo_n,region)
      select e1.plant,e1.unit,e1.tstamp,YEAR(e1.tstamp),e1.guldo,'ERCT'
      from erct_cem e1
      where e1.plant=l_plantid and e1.unit=l_unitid;

      insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'Step1_Post';

      #Step 2
        # Insert generator data from "pre" table,

      INSERT INTO tmp_plant_unit(region,plant,unit,tstamp,year,guldo_n,guldo_nminus1)
```

```
        select e1.region,e1.plant,e1.unit,e2.tstamp,YEAR(e2.tstamp),e2.guldo_n,e1.guldo_n
        from tmp_plant_unit_pre e1
        inner join tmp_plant_unit_pre e2
        on e1.plant = e2.plant
        and e1.unit = e2.unit
        and e1.tstamp = DATE_ADD(e2.tstamp, INTERVAL -1 HOUR)
        order by e1.plant,e1.unit,e1.tstamp;

              insert into agrealog (plant,unit,tstamp,session,step)
        select l_plantid,l_unitid, now(),l_sessionid,'Step2_Post';


        # Use join for total load as opposed to using temp table
        # Update temp table with total load
        UPDATE tmp_plant_unit,regional_total_load
        SET tmp_plant_unit.Total_Load_N = regional_total_load.Total_Load_N,
tmp_plant_unit.Total_Load_Nminus1 = regional_total_load.Total_Load_Nminus1
        WHERE tmp_plant_unit.tstamp = regional_total_load.tstamp;


        #Get Total Load -- for now just use ERCT_Total_Load
        #Join table realtime -- join by TSTAMP
        # Select Load as totalGULDO from ERCOT_Total_Load;

        #Get nameplate capacity from Generator table
        #Place nameplate capacity in l_namecap variable

        #Determine yearly maximums
        TRUNCATE TABLE tmp_plant_unit_peaks;
        INSERT INTO tmp_plant_unit_peaks(year,peakyearly,namepcap)
        SELECT year(tstamp),MAX(guldo_n),max(l_namepcap) FROM tmp_plant_unit GROUP BY YEAR(tstamp);


        UPDATE tmp_plant_unit,tmp_plant_unit_peaks
        SET tmp_plant_unit.PeakYearly = tmp_plant_unit_peaks.peakyearly, tmp_plant_unit.namepcap =
tmp_plant_unit_peaks.namepcap
        WHERE year(tmp_plant_unit.tstamp) = tmp_plant_unit_peaks.year;

        #Determine loadstate
        #Use tables:
        # tmp_plant_unit, tmp_plant_unit_peaks


        ## The following values correspond to
        # Turning Off / On = 4 ( 0 thru 0.05 )
        # Standby = 3 ( 0.05 thru 0.55 )
        # Spinning Reserve = 2 ( 0.55 thru 0.9 )
        # Full Load = 1 ( 0.9 thru 1.0 )

        # For future reference, in PHP these were previously modelled as arrays:
        #  $loadstatearray = array(4, 3, 2, 1);
        #  $loadstatebounds = array(0.05, .55, .9);


        # percent of max
        update tmp_plant_unit
        set loadstate =
          CASE
            WHEN peakyearly = 0 THEN NULL
            WHEN guldo_n/peakyearly < 0.05 THEN 4
            WHEN guldo_n/peakyearly < 0.55 THEN 3
            WHEN guldo_n/peakyearly < 0.90 THEN 2
            ELSE 1
          END;


/*

        # LSF Algorithm --

          # If total load = 0 then set to 1
```

```
        # IsLoadFollowing = 0  is the default value (binary : either 0 or 1 )
        # IsLoadFollowing = 1 means that the generator is non baseload generation for that
particular hour


        deltaunitload = guldo_n - guldo_nminus1
        percentDeltaLoad = (guldo_n - guldo_nminus1)/guldo_n
        percentDeltaTotalLoad = (total_load_n-total_load_nminus1)/total_load_n

*/
# Load following code in two steps
  #step 1
    # Mark hour as loadfollowing if meets the first two conditions


# BEGIN LSF Algorithm

update tmp_plant_unit
  SET IsLoadFollowing =
  CASE
    WHEN (guldo_n = 0 or total_load_n =0)  THEN NULL
    WHEN
      (
        (guldo_n - guldo_nminus1)/guldo_n > 0 AND (total_load_n-total_load_nminus1)/total_load_n
> 0

      )
      OR
      (
        (guldo_n - guldo_nminus1)/guldo_n < 0 AND (total_load_n-total_load_nminus1)/total_load_n
< 0
      )
      OR
      Loadstate = 2
      THEN 1
    ELSE 0
  END;



  #step 2
    # Look at l_prevloadfollowing variable and mark as loadfollowing where appropriate
    #requires prevloadfollowing field in temp table for recording

      UPDATE tmp_plant_unit e1
              JOIN tmp_plant_unit e2
              ON e1.plant = e2.plant
              AND e1.unit = e2.unit
              AND e1.tstamp = DATE_ADD(e2.tstamp, INTERVAL -1 HOUR)

      SET e2.isloadfollowing = 1
      WHERE e2.isloadfollowing = 0
      AND e1.isloadfollowing
      AND ABS((e2.guldo_n - e2.guldo_nminus1)/e2.guldo_n) < 0.025
      AND ABS((e2.total_load_n-e2.total_load_nminus1)/e2.total_load_n) < 0.025;

# End LSF Algorithm

    # Update ERCT or CEM table of interest

          insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'Update_ERCT_PRE';
      UPDATE erct_cem e1
              JOIN tmp_plant_unit e2
              ON e1.plant = e2.plant
              AND e1.unit = e2.unit
              AND e1.tstamp = e2.tstamp
      SET e1.isloadfollowing =
e2.isloadfollowing,e1.loadstate=e2.loadstate,e1.deltaunitload=(e2.guldo_n - e2.guldo_nminus1);

    # End Update ERCT or CEM table of interest
```

```
            insert into agrealog (plant,unit,tstamp,session,step)
      select l_plantid,l_unitid, now(),l_sessionid,'Update_ERCT_Post';



      SET l_loop_count = l_loop_count+1;
UNTIL no_more_plant_unit
END REPEAT plant_unit_loop;
CLOSE plant_unit_csr;
SET no_more_plant_unit = 0;


END $$

DELIMITER ;
```

8.4.1. SQL Query: Example to calculate weighted average marginal avoided emissions rate from base AGREA tables

```
# The following example shows how to calculate the weighted average marginal
avoided emission rates for ERCOT.

select
date_format(e1.tstamp,'%m/%d/%Y') as mydate,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 0 THEN
  e1.NOx_rate_total
ELSE 0 END) hour1,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 1 THEN
  e1.NOx_rate_total
ELSE 0 END) hour2,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 2 THEN
  e1.NOx_rate_total
ELSE 0 END) hour3,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 3 THEN
  e1.NOx_rate_total
ELSE 0 END) hour4,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 4 THEN
  e1.NOx_rate_total
ELSE 0 END) hour5,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 5 THEN
  e1.NOx_rate_total
ELSE 0 END) hour6,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 6 THEN
  e1.NOx_rate_total
ELSE 0 END) hour7,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 7 THEN
  e1.NOx_rate_total
ELSE 0 END) hour8,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 8 THEN
  e1.NOx_rate_total
ELSE 0 END) hour9,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 9 THEN
  e1.NOx_rate_total
ELSE 0 END) hour10,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 10 THEN
  e1.NOx_rate_total
ELSE 0 END) hour11,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 11 THEN
  e1.NOx_rate_total
ELSE 0 END) hour12,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 12 THEN
  e1.NOx_rate_total
ELSE 0 END) hour13,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 13 THEN
  e1.NOx_rate_total
ELSE 0 END) hour14,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 14 THEN
  e1.NOx_rate_total
ELSE 0 END) hour15,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 15 THEN
  e1.NOx_rate_total
```

```
ELSE 0 END) hour16,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 16 THEN
  e1.NOx_rate_total
ELSE 0 END) hour17,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 17 THEN
  e1.NOx_rate_total
ELSE 0 END) hour18,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 18 THEN
  e1.NOx_rate_total
ELSE 0 END) hour19,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 19 THEN
  e1.NOx_rate_total
ELSE 0 END) hour20,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 20 THEN
  e1.NOx_rate_total
ELSE 0 END) hour21,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 21 THEN
  e1.NOx_rate_total
ELSE 0 END) hour22,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 22 THEN
  e1.NOx_rate_total
ELSE 0 END) hour23,
SUM(CASE CAST(date_format(e1.tstamp,'%H') as unsigned) WHEN 23 THEN
  e1.NOx_rate_total
ELSE 0 END) hour24
from
(
select
e1.tstamp,
sum(
ABS(((e1.NOx*e1.HHIRDO)/e1.GULDO))*
ABS((e1.DeltaUnitLoad/e2.DeltaUnitLoadTotal))) as NOx_rate_total
from erct_cem e1 join erct_DeltaUnitLoadTotal e2
on e1.tstamp = e2.tstamp
where
e1.isloadfollowing=1
group by e1.tstamp
) e1
group by date_format(e1.tstamp,'%m/%d/%Y')
order by mydate;
```