Appendix E
Temporally Allocating Emission with CEM Data for Chemical Transport and SIP Modeling

Temporally Allocating Emissions with CEM Data for Chemical Transport and SIP Modeling

Scott Edick
Michigan Department of Environmental Quality, Air Quality Division
P.O. Box 30260, Lansing MI 48909

Mark Janssen
Lake Michigan Air Directors Consortium (LADCO)
2250 East Devon Ave. Suite 250 Des Plaines IL, 60018

ABSTRACT

This paper describes how Continuous Emissions Monitoring (CEM) data were used to generate temporal profiles based on median CEM heat input values by month, day of week and hour of day for individual National Inventory Format (NIF) emission units, how the profiles were used to generate typical hourly emission records for base and future modeling years, and how emissions temporally allocated by use of these profiles differ quantitatively from the results of more traditional methods.

INTRODUCTION

Continuous Emissions Monitoring (CEM) data has much to recommend it for use in emissions and atmospheric modeling, as it contains directly measured emissions with hourly temporal resolution for a significant fraction of the point source inventory.

CEM data can be challenging to integrate into a modeling inventory. The high-resolution data sets grow large very quickly as the temporal or geographic domain expands. Pollutant coverage is incomplete. Deriving data that is properly representative of a period, e.g. a typical summer day, from the highly resolved CEM data is problematic. Establishing accurate correspondence between CEM reporting units and NIF emission units or release points is a major challenge.

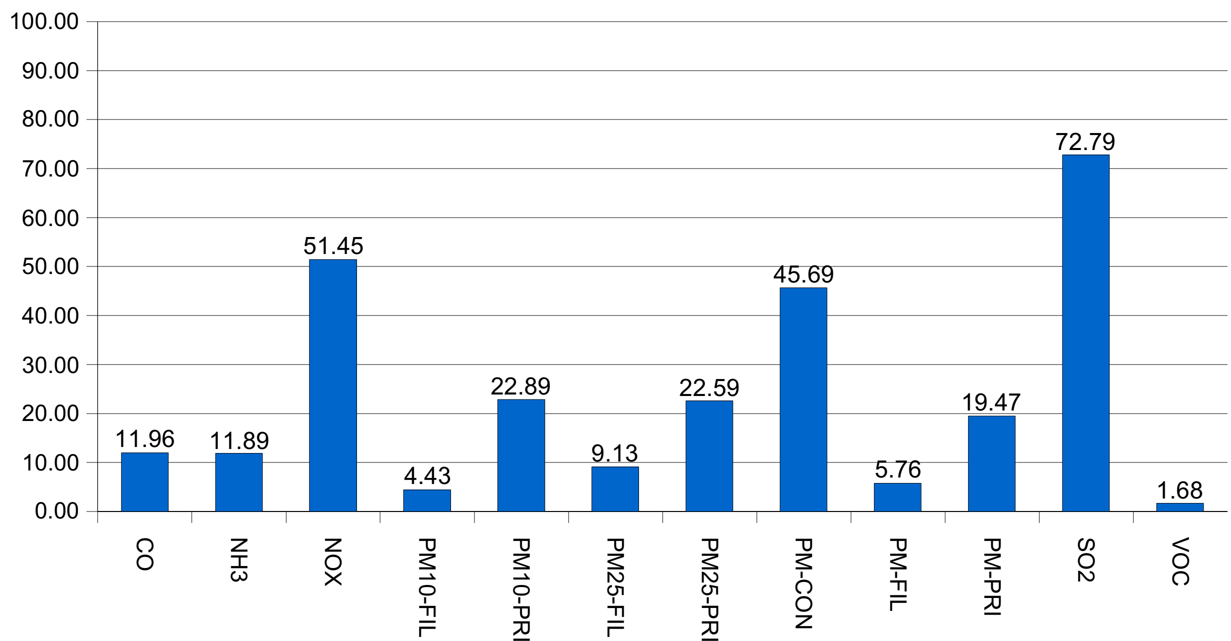
This paper describes how 2001 through 2003 CEM heat input data were used to generate unit-specific NIF emissions records for each hour of a weekday, a Saturday and a Sunday for each month of 2002 for approximately nineteen hundred NIF emission units nationwide, and it describes how the procedure was modified for future year inventories to use the summer and non-summer, rather than annual, emission records derived from output of the IPM model. We review how the above challenges were addressed, how the heat input-based approach compares to unit-specific allocation based on CEM reported SO2 or NOx emission data, and how emissions temporally allocated based on unit-specific CEM heat inputs differ quantitatively from the results of more traditional methods.

APPROACH

A desired improvement, long outstanding[1], to the temporal allocation of emission inventories for atmospheric modeling has been to incorporate improvements based on CEM data. Rather than a direct application of CEM data that would result in an inventory reflective of a specific, historic episode, the goal here was to perform temporal allocations representative of a typical period, which could be applied to base and future year inventories. It was also considered essential that the methods developed be highly automated. This last requirement was dictated by available resources, but also has the virtue of making the methods developed reproducible and readily modified and improved over the inevitable iterations of modeling inventory development.

LADCO recently contracted with E.H. Pechan for development of IPMTOOL, software which can be used to convert IPM model output data to NIF-formatted emission data. One of the results of this project was to make available a cross-reference between the key identifiers for CEM boilers and NIF emission units. This cross-reference provided an opportunity to develop a highly automated application of CEM data to modeling inventories. Figure 1 below shows the fraction of the total 2002 base year inventory that can be matched by the cross-reference, hence the fraction of the total 2002 point source inventory that can be temporally allocated by the approach described below.

Figure 1. Percentages of pollutants in the 2002 point source inventory matched to CEM heat input.



The compiled quarterly CEM data files from Clean Air Markets Division (US EPA) covering the years 2001 through 2003 (bracketing the 2002 base year) were selected for use. This data set contains just over eighty eight million records. The data elements in these tables include facility and unit identifiers, date and hour of the record, CO2 and SO2 emissions, NOx emission rates and heat inputs.

To develop the temporal allocation factors, the approach taken was based on selecting median heat input values corresponding to a given unit/temporal period combination. The rationale here is that the middle of the range of operating values will be typical of the period it represents, and uninfluenced by extremes of operation or upset events. Middle values were selected for each combination of unit, month of year, day type (weekday, Saturday, or Sunday) and hour of day.

The term middle value is used here because we chose to deviate from the definition of median in two ways. First, we chose to assume that null heat input values and missing records indicated no operation. Therefore, it was determined how many records *should* exist for each month and day type over the 2001 through 2003 period, and the middle position of that range was determined based on that number. Second, where an even number of values was expected, rather than taking the mean of the two values nearest the middle, we simply chose the lower of the two.

As an example of how the middle values were selected, note that there were sixty nine January weekdays in 2001 through 2003, hence for any given hour of the day sixty nine opportunities for a unit to report January weekday heat input values. The January weekday middle value for a given hour and unit is taken by sorting the corresponding reported values in descending order, and selecting the value in the thirty fifth slot as the desired middle value. If there are fewer than thirty five values reported, consistent with our assumptions about missing data, the middle value is set to zero.

Once the middle values have been selected, normalizing values are generated. This is accomplished by calculating what the total annual heat input for each unit would be if it operated every hour of the 2002 base year at the middle value appropriate for each hour, day type and month throughout the year.

To this point, all processing has been done on CEM data with its facility and unit identifiers. To integrate this data with a NIF inventory, NIF emission unit identifiers have to be matched to the profiling factors via the cross-reference table described earlier. While this cross reference matches CEM boilers to NIF emission units, the correspondence is not always one-to-one. Addressing this issue is straightforward in the case of one CEM boiler with multiple corresponding NIF emission units; simply duplicate the CEM-based profile data for each corresponding NIF emission unit identifier. In the case of multiple CEM boilers corresponding to a single NIF emission unit, the middle values of each month, day type and hour of day for all corresponding boilers were aggregated, as were the normalizing values. The aggregate values result in a composite profile, weighted by heat input, which is then matched to the NIF emission unit identifier.

Each of these matched NIF emission units now has has one annual normalizing value and a total of eight hundred sixty four middle values (12 months x 3 day types x 24 hours), each of which is specific to a combination of month, day type and hour of day. Generating an hourly emission record is now a relatively straightforward exercise in data matching. Given an annual emission record for a cross-referenced NIF emission unit, the fraction of the annual emissions allocated to a given hour is determined by multiplying the annual emission value by the middle value appropriate to the hour, day type, month and emission unit for which the record is being generated, and dividing by the emission unit's normalizing value.

The hourly NIF emission records generated this way are then appended to the original NIF point source file, and generation of the modeling inventory proceeds as usual, relying on the emission processor's temporal allocation routines to give priority to hourly records over records covering longer periods.

From the outputs of the IPM model, future year emissions can be derived for electrical generation units (EGUs) in both annual and summer/non-summer period emissions. To use the shorter period emission records, rather than the annual records as for the base inventory, the middle values are chosen as above, but two sets of normalizing values are calculated for each unit. One set reflects operation at the heat input middle values appropriate for each hour, day type and month over the summer period, the other set reflects operation over the non-summer period. Emissions are calculated as before, using the period-specific emission records and normalizing values rather than their annual counterparts.

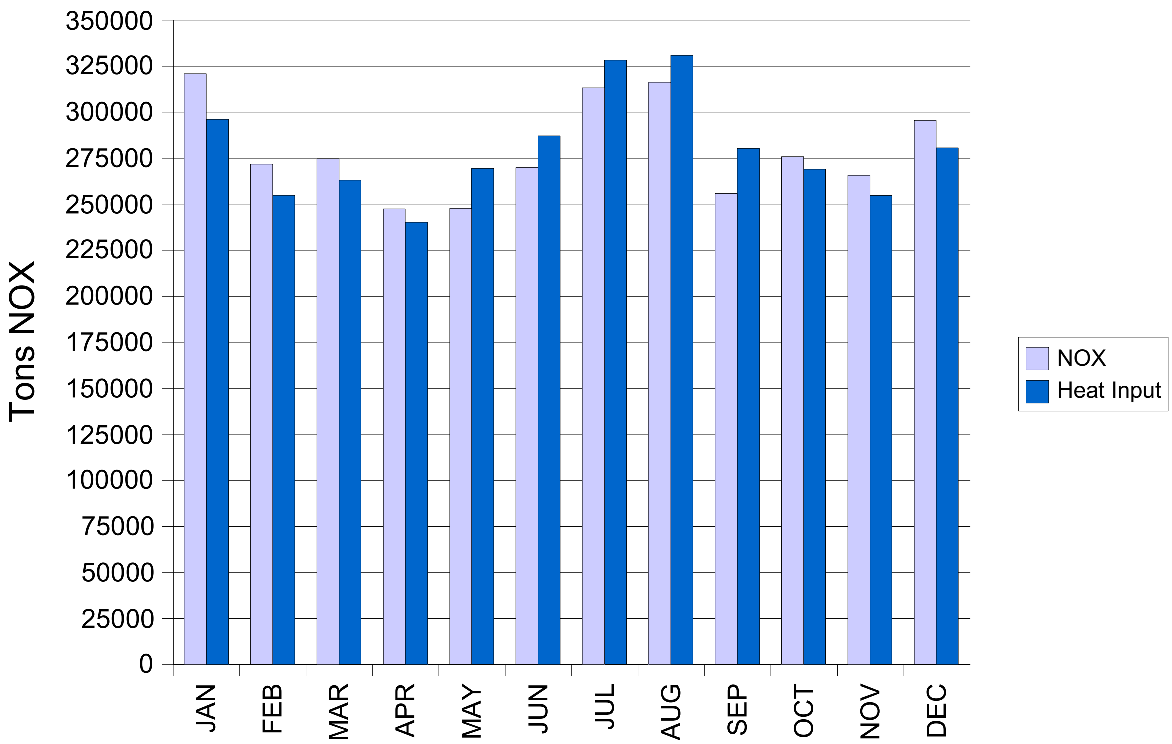
ANALYSIS OF SEASONAL EFFECTS

Other work with CEM-based data[2] has shown that, particularly when aggregated to the state or regional level, monthly CEM heat inputs do not correlate well with CEM SO2 and NOx data. Individual units have also been shown to exhibit this behavior.

To determine the extent of this problem under the median-based, unit-specific approach, the same process of developing unit, month, day and hour specific temporal profiles was performed using the CEM reported SO2 and NOx data in place of heat input. The comparison was performed on units which reported both heat and emissions data, and which could be matched to NIF data via the cross reference table.

2002 CEM reported NOx and SO2 emissions were allocated by unit using both the heat input based profile and the appropriate emissions based profile. Figures 2 and 3 below show the monthly comparisons aggregated to the national level.

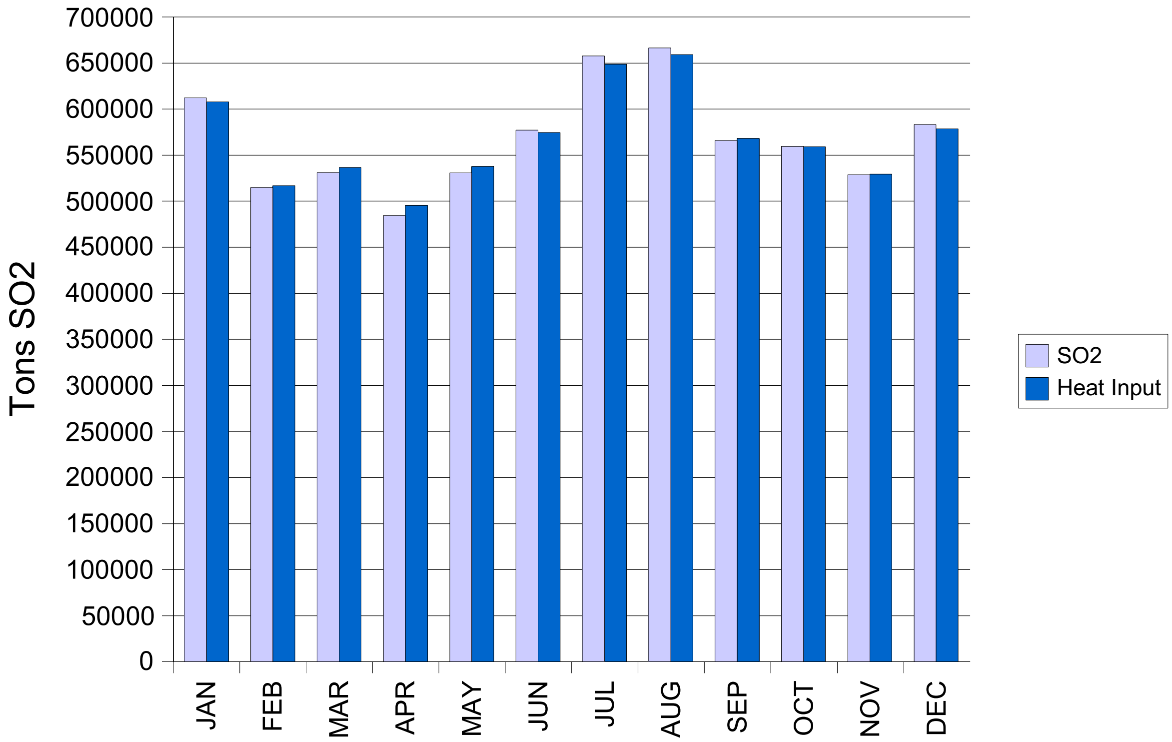
Figure 2. Temporal distribution of 2002 CEM NOx emissions by CEM heat input vs. CEM NOx.



Temporal allocation by heat input, rather than CEM reported NOx, does result in NOx levels that are higher from May through September and lower the rest of the year. This is not surprising, given May through September NOx emission trading budgets and the seasonal application of NOx controls. With the implementation of the CAIR rules and year-around NOx trading budgets, the expectation is for more year-around operation of controls and a resulting attenuation of the seasonal differences in evidence above.

Also, as mentioned earlier, future year emissions for EGUs are generated for summer and non-summer periods, and these summer and non-summer period emission records are what is processed for future years. To the extent that the seasonal emission variation is addressed by these periodic emissions records, normalizing to, and allocating within, these periods will propagate those corrections.

Figure 3. Temporal distribution of 2002 CEM SO2 emissions by CEM heat input vs. CEM SO2.



For SO2, the effect of temporal allocation using heat input is to moderate SO2 emissions slightly. The fact that the summer deviation between heat input and SO2 is generally in the opposite direction of that between heat input and NOx suggests that further improvements would require per-pollutant temporal activity surrogates or improvements in modeling of emission controls.

As another indicator of how well the heat and pollutant based profiles track one another, the fraction of the total 2002 CEM emissions that would be reallocated by unit from one month to another when switching between heat input basis and pollutant basis was calculated according to Equation 1 below.

Equation (1)
$$\sum_{u \in units} \sum_{m=Jan}^{Dec} \frac{|ht_{u,m} - poll_{u,m}| * E_u}{2}$$

where

$ht_{u,m}$ is the heat-based temporal profile for unit *u*, month *m*
$poll_{u,m}$ is the emissions-based temporal profile for unit *u*, month *m*
E_u is the total 2002 CEM-reported emissions for unit *u*

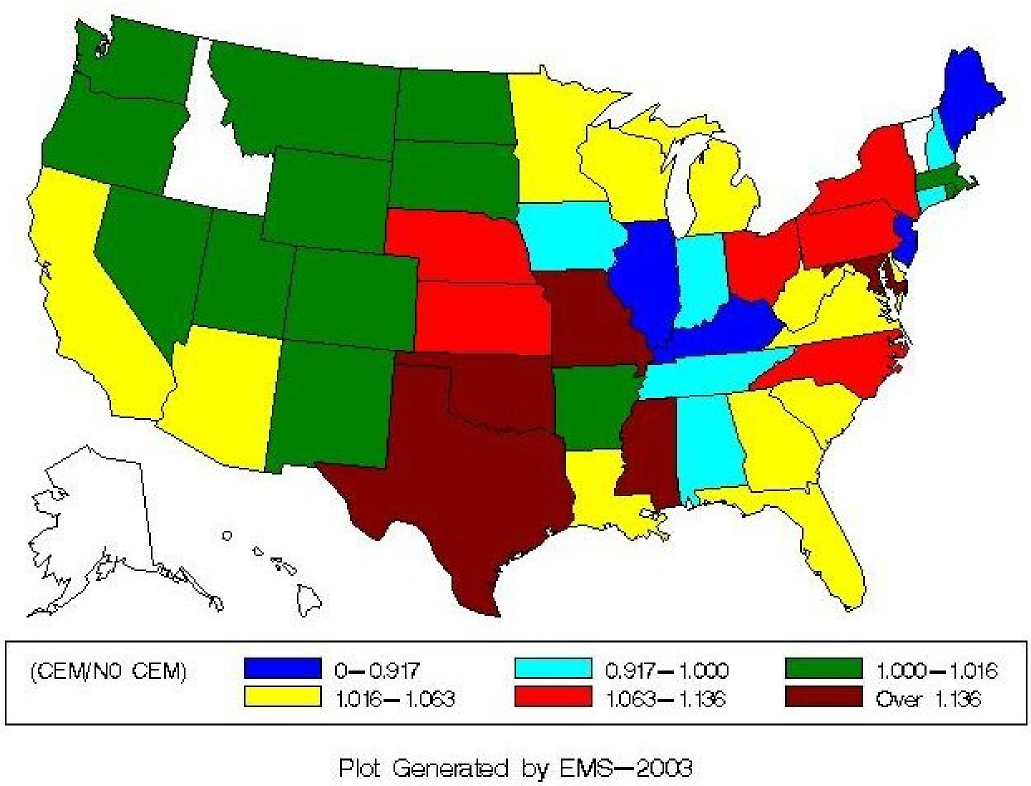
Note that variation is calculated per unit then aggregated to prevent canceling effects of compensating variations across units. For SO2, approximately 2.8% of the total 2002 emissions would be reallocated by switching between heat input and emissions based profiles. For NOx that value is 5.0%.

For each pollutant, examining the differences between heat and pollutant-based profiles for individual units does reveal a small number of units with significant variation. Most of these have relatively low emissions, however there are some units with both significant emissions and significant variation between heat and pollutant based profiles. Further investigation is planned to try to identify the source of these discrepancies.

RESULTS

The figure below shows how EGU NOx emissions at the state level change for a July weekday when applying the CEM-based unit-specific temporal profiles. Note that the temporal allocation scheme used as a basis for comparison reflects changes made several years ago when LADCO modified the default EGU temporal profiles to better reflect aggregated CEM data, and assigned certain base load EGUs a profile reflecting 24x7 operation.

Figure 4. Ratio of July 12, 2002 EGU NOx with and without CEM-based temporal allocation



Figures 5 through 8 below compare diurnal allocation of emissions in several states for a July weekday according to three different methods. The method designated EPA_DEFA applies the default single set of temporal profiles to all EGUs. The NO_CEM method is the LADCO modifications to the defaults as described above. The WITH_CEM method applies the unit-specific temporal allocation based on the CEM data.

Figure 5. July 12, 2002 hourly NOx for Georgia EGUs

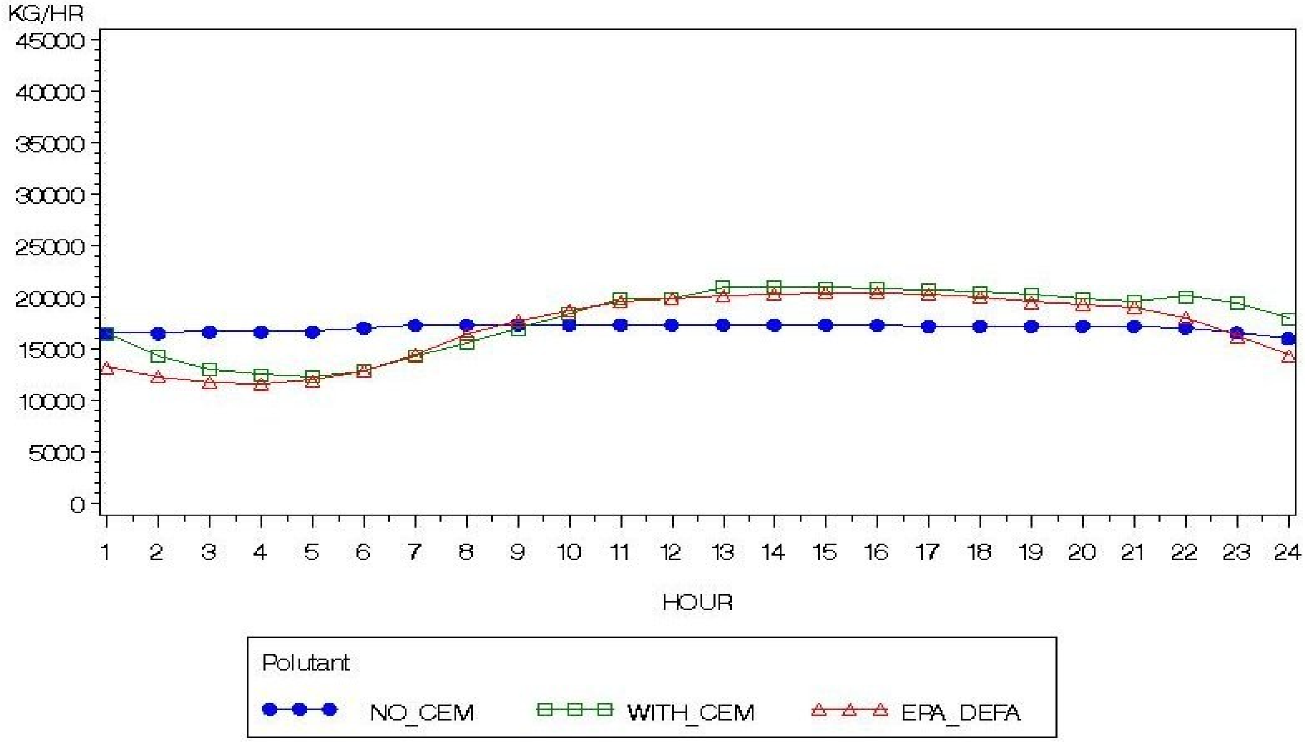


Figure 6. July 12, 2002 hourly NOx for Illinois EGUs

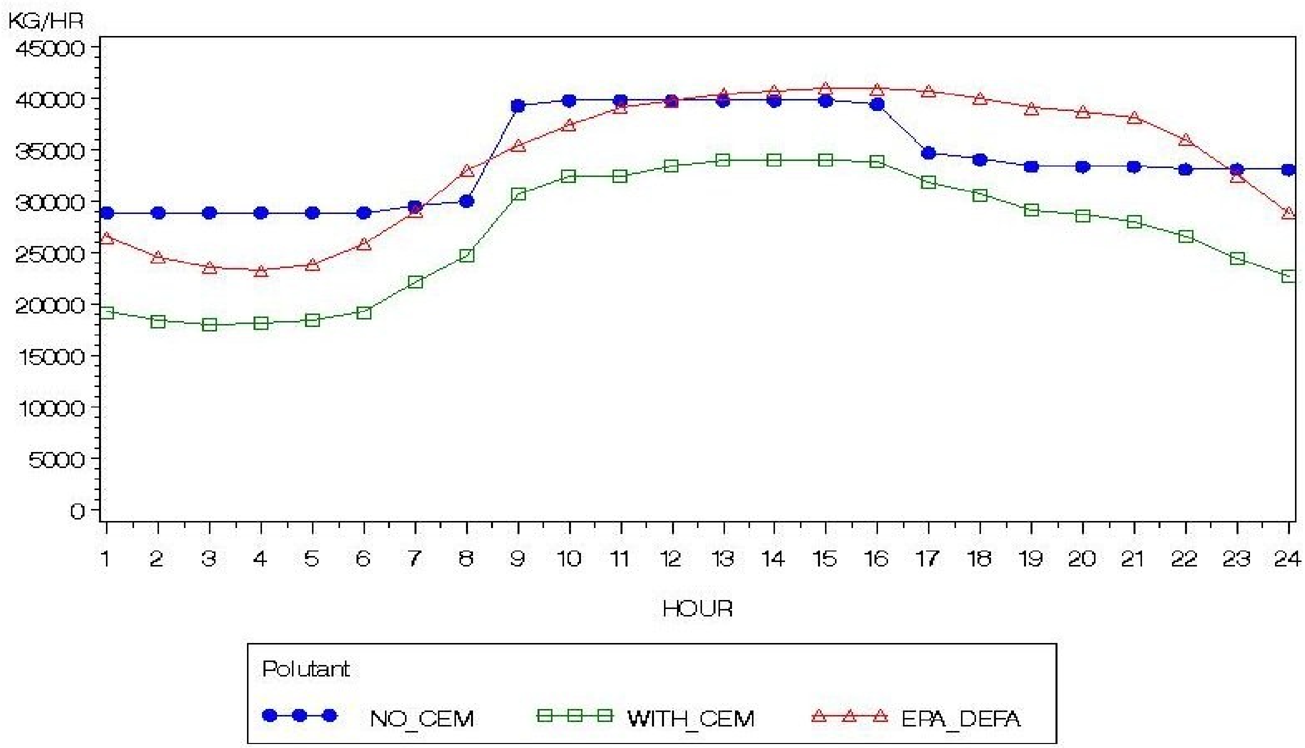


Figure 7. July 12, 2002 hourly NOx for Michigan EGUs

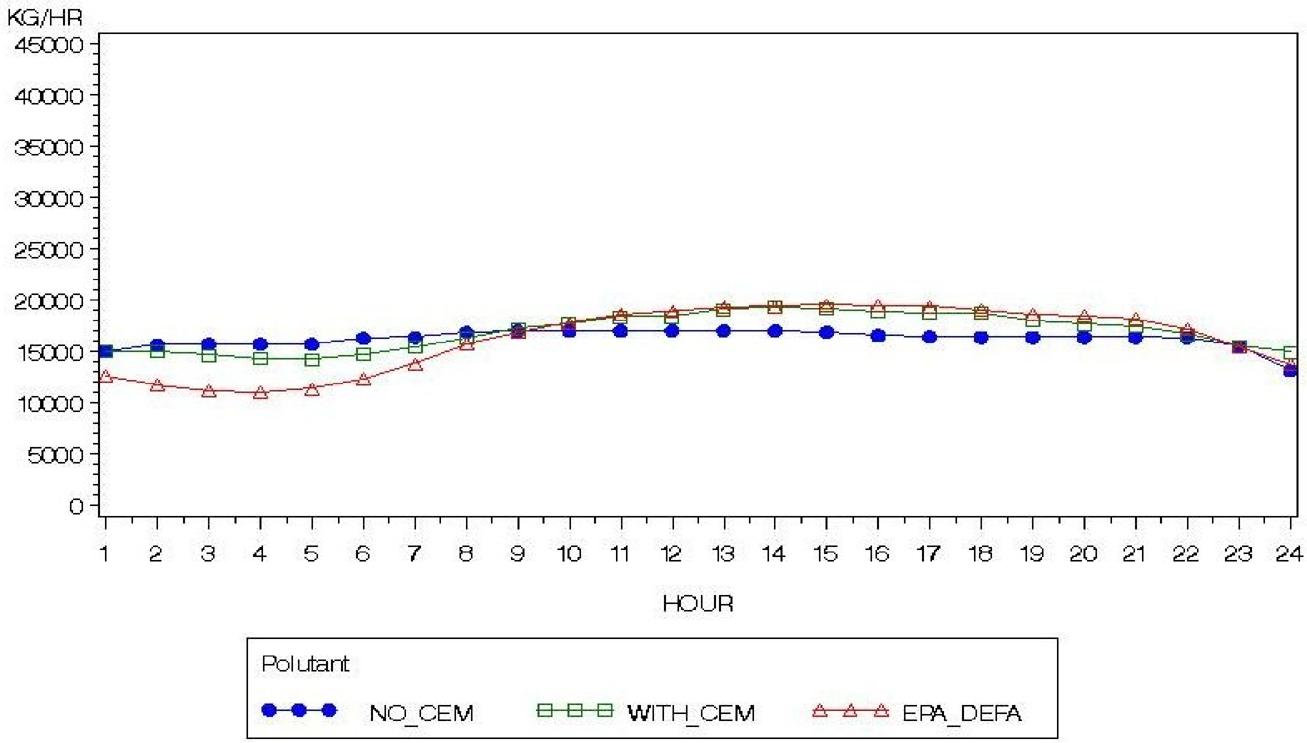
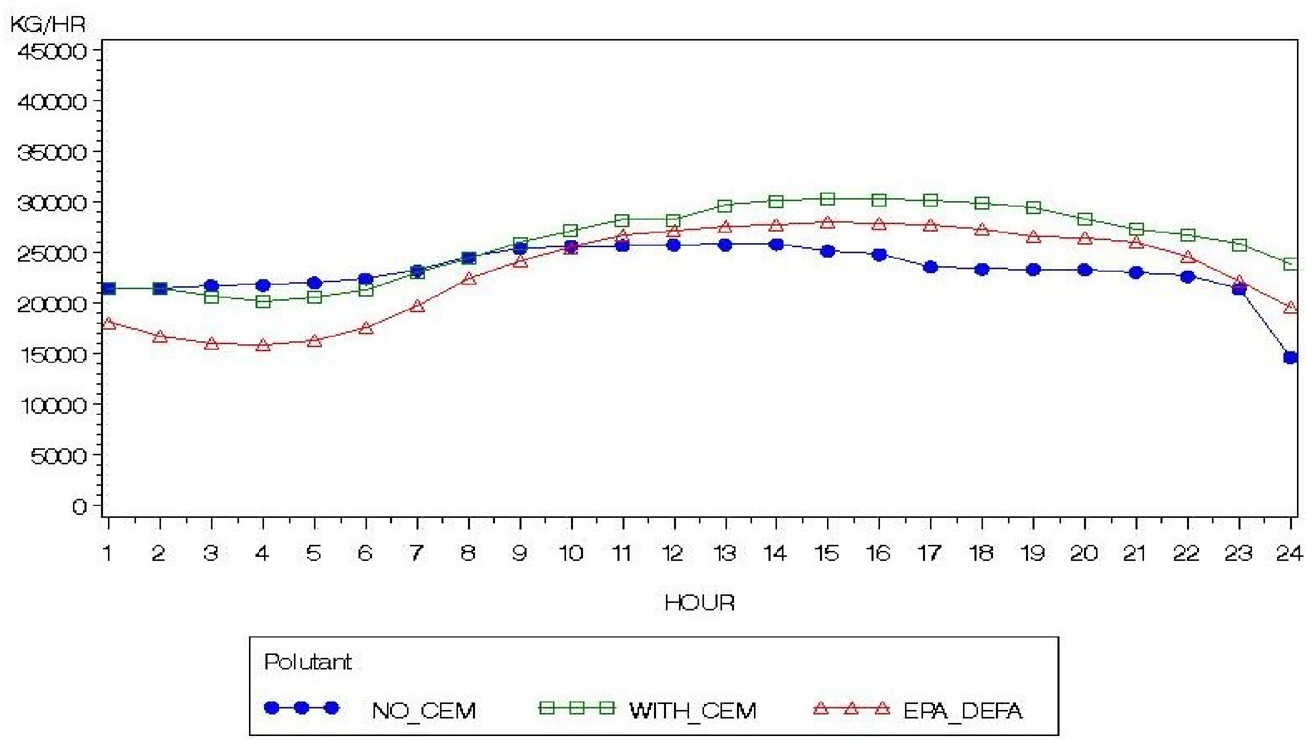


Figure 8. July 12, 2002 hourly NOx for Pennsylvania EGUs



Figures 9 through 12 below show monthly weekday, Saturday and Sunday daily NOX totals for 2002 and 2009 for individual facilities with and without CEM based temporal allocation.

Figure 9. Despite a reasonable overall correspondence, differences between temporal allocation schemes of 10 TPD for weekdays and 20 TPD on weekends are common for this large source in 2002.

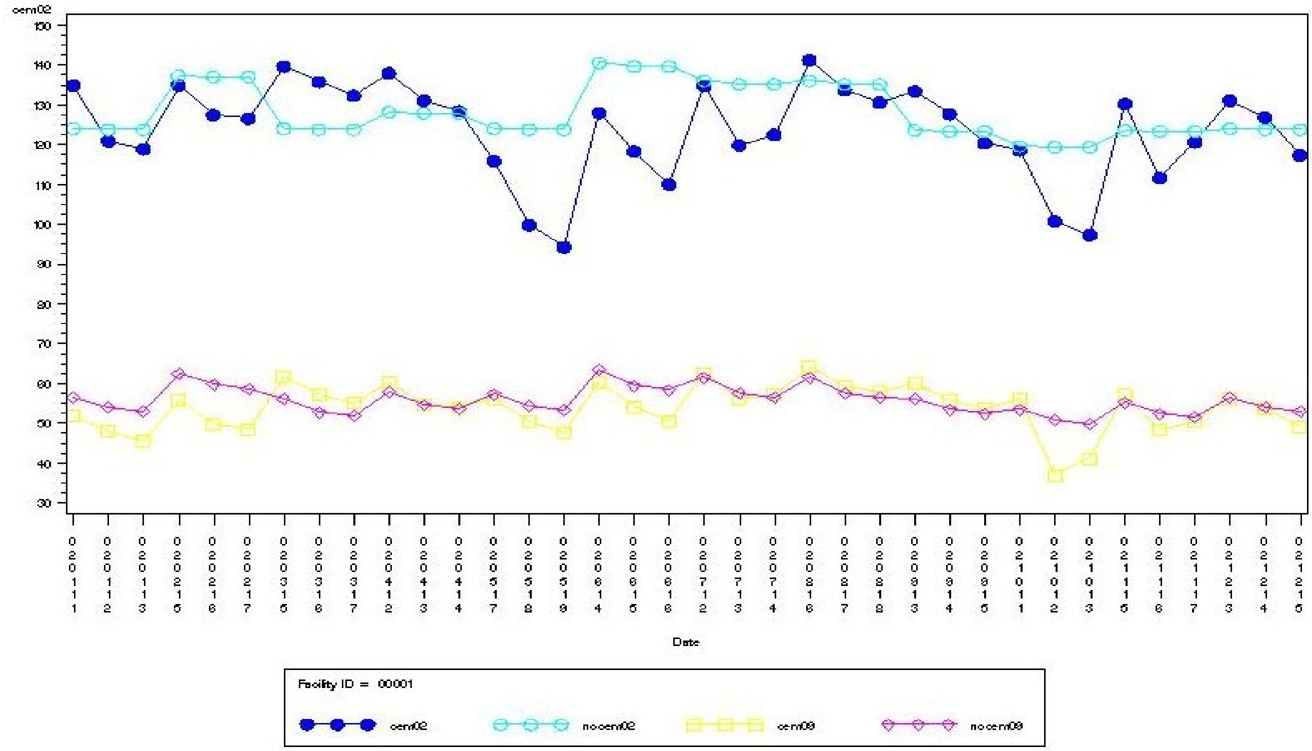


Figure 10. Significant weekday/weekend variability is apparent when allocated with CEM data.

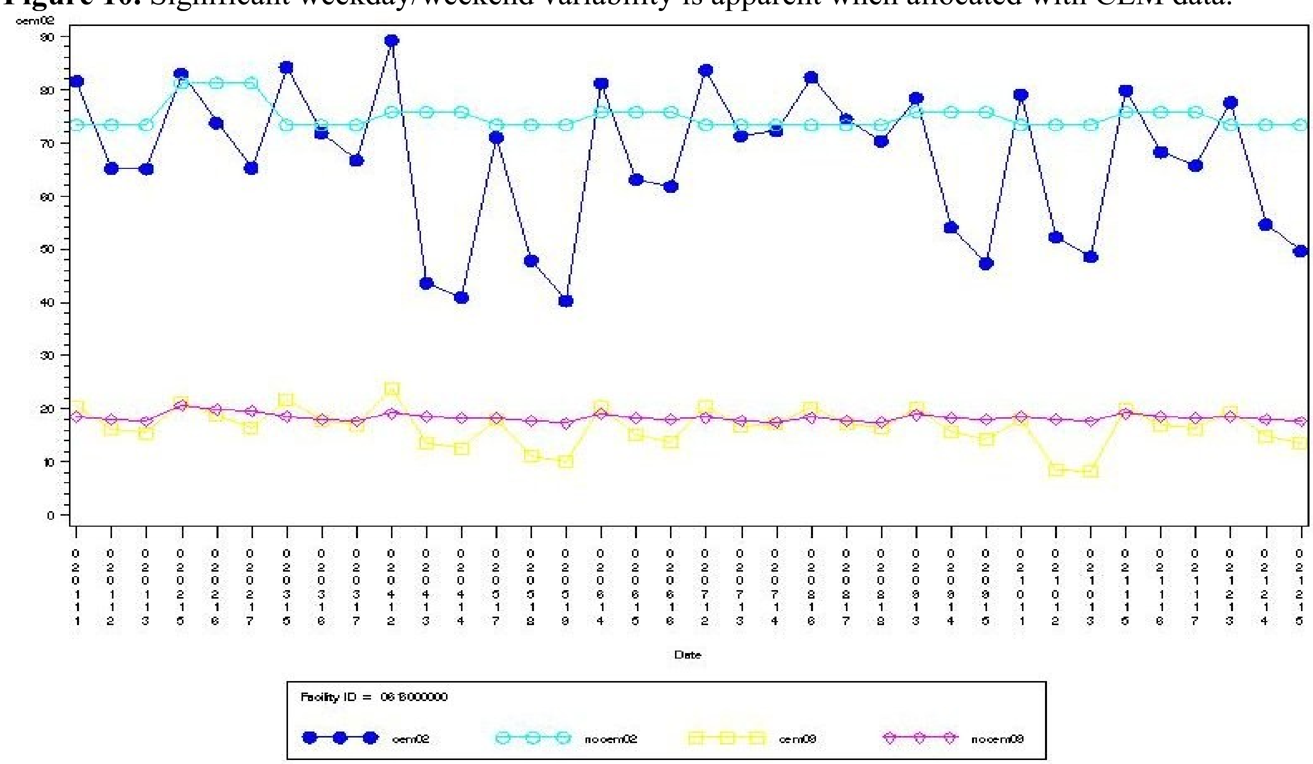


Figure 11. With a median-based approach over three years, the April drop cannot be due to a single outage, but is strongly suggestive of a maintenance pattern.

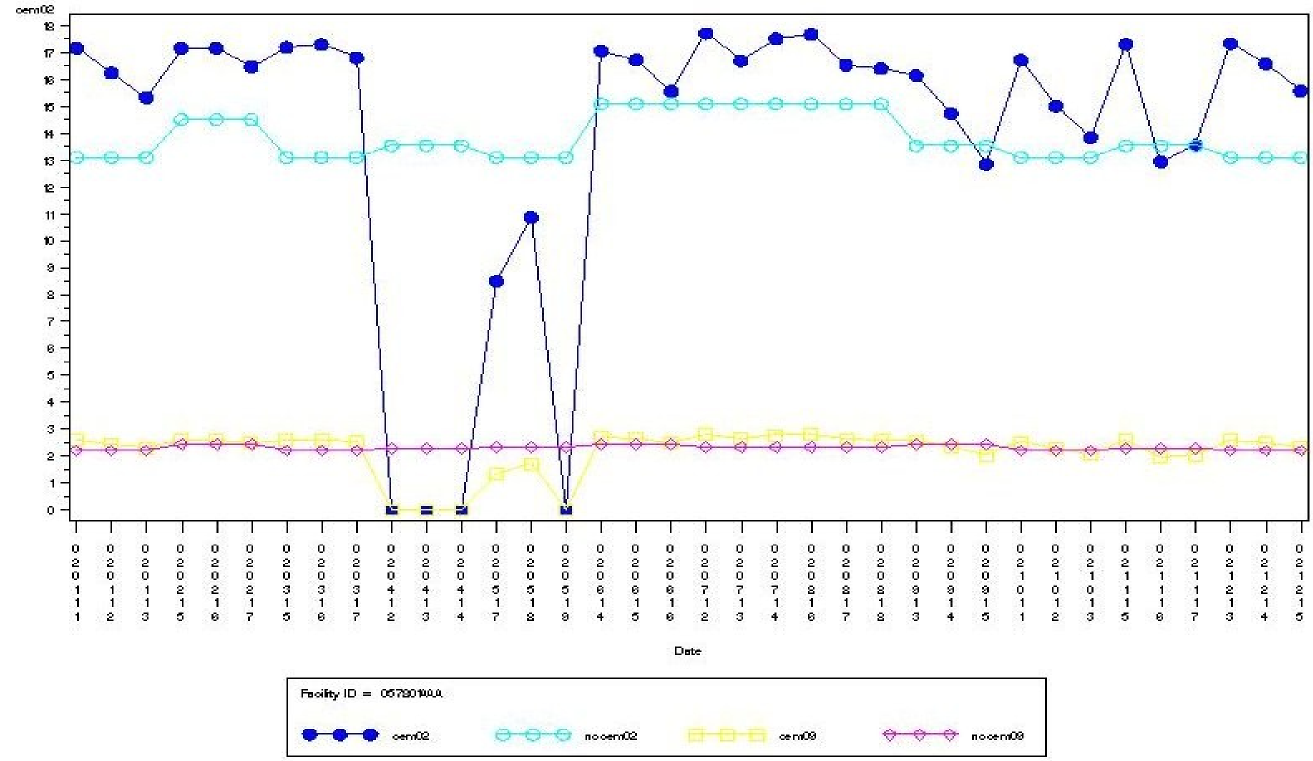
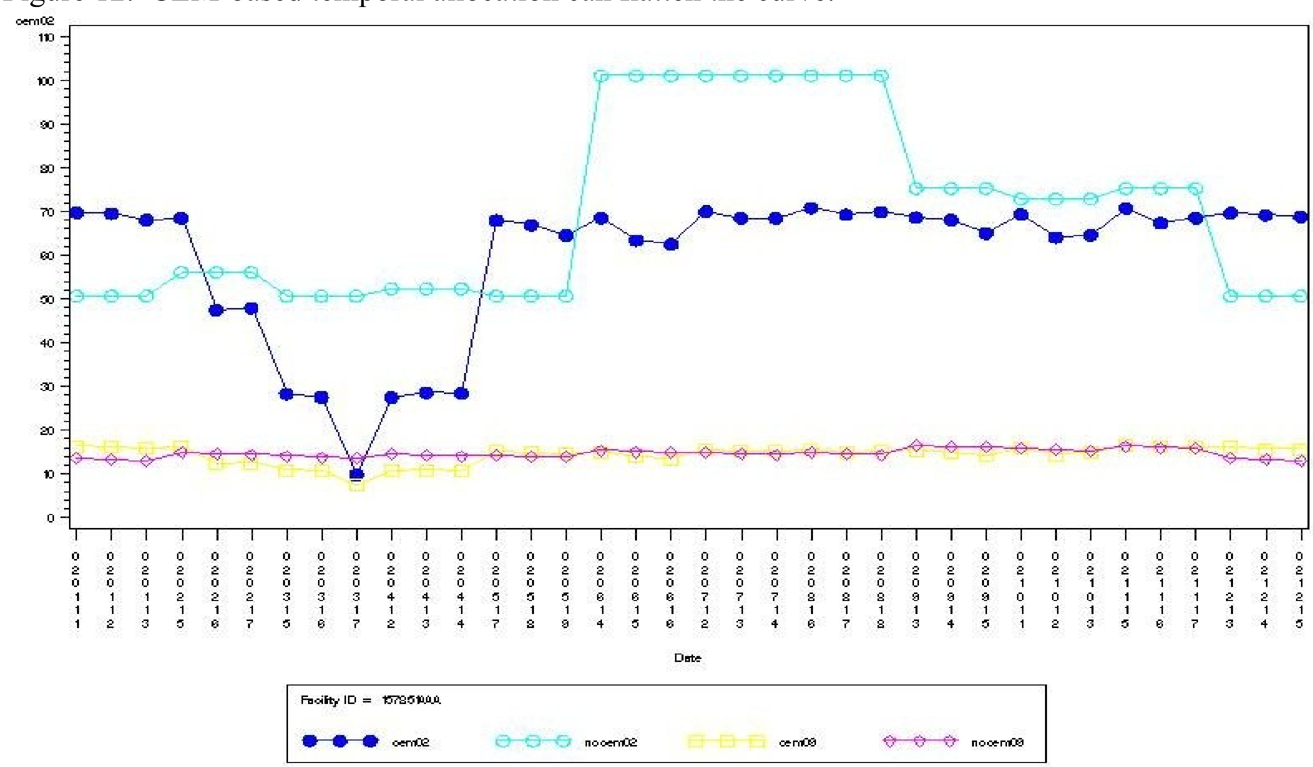


Figure 12. CEM-based temporal allocation can flatten the curve.



CONCLUSIONS AND RECOMMENDATIONS

As a proof of concept for automated application of CEM data to national emission inventories, several significant success can be noted. Starting with the load of eighty eight million CEM records, through generation of unit-specific temporal profiles and matching to a nationwide NIF point source inventory, and finishing with the output of 13.7 million NIF-format hourly emission records, the run takes well under a week on a sub-$1000 computer. Most of the processing time is spent developing the unit-specific temporal profiles, processing which need not be repeated unless the underlying CEM data, the normalizing period(s), or the CEM-to-NIF cross reference changes. Applying the profiles to an updated national inventory and generating updated hourly emissions records takes less than a day. Future year emissions data generated from alternative nationwide IPM runs has been processed in a couple hours.

Given the relatively modest resource requirements, the fraction of the modeling inventory for ozone precursors and particulates that can be matched to CEM data, and the observed quantitative differences between CEM-based and traditional temporal allocation methods, a strong case exists for more widespread adoption of a CEM based approach to temporal allocation.

Heat input is the best available single surrogate for multiple pollutants, though improvement might be realized via a CEM reported emissions-based approach for SO2 and NOx, or improvements to modeling of emission controls.

The median-based approach minimizes the effects of upset events or extremes of operation, and generally results in temporal profiles that can be considered representative of typical operating patterns. The exceptions are units used so infrequently as to rarely or never give non-zero middle values. We have reverted to traditional methods for these units. While the total annual emissions from these infrequently used units is not large, the correlation between days with high electric demand and high ozone potential suggests this as an area for further improvement.

The most obvious opportunity for improvement lies in improving the cross-reference between CEM and NIF data. Improved temporal allocation is only one of many aspects of inventory development and modeling that would benefit from improved integration of the high-quality, high-resolution CEM data with emission inventories. Though it is the best currently available, the cross reference used in this project must be seen as an independent, one-time effort. Provision should now be made for a CEM to NIF cross-reference to be developed and maintained as a cooperative effort among the interested parties.

One possible approach would be to add fields for CEM unit identifiers to the NIF emission unit and release point records (the cross reference used in this project matched data at the emission unit level, but much of the CEM data corresponds better to a release point). Incorporating the cross-reference into the NIF data would make data matching more direct and transparent for users of inventory data, and it would provide opportunity and motivation for emitting facilities and inventory developers to review and improve the correlation between CEM and NIF data.

It is hoped this demonstration of the improvements currently available from integrating high-resolution CEM data with emission inventories, and the low resource requirements for doing so, will provide further impetus for process changes that will allow better utilization of CEM data by inventory developers and modelers.

REFERENCES

1. Janssen M. "CEM Data Analysis and Use" Presented at the RPO National Workgroup Meeting St. Louis, MO November 2003.

2. Stella G. "Development of Hourly Inventories Utilizing CEM-Based Data" Presented at the 14th Emission Inventory Conference of the U.S. Environmental Protection Agency, Las Vegas, NV April 2005

KEYWORDS

Continuous Emissions Monitoring
EGU
Emissions
Modeling
Temporal