4.0 MODEL PERFORMANCE EVALUATION

4.1 EVALUATION METHODOLOGY

St. Louis $PM_{2.5}$ State Implementation Plan (SIP) attainment demonstration modeling uses the 2002 Base 5b modeling results from the Community Multiscale Air Quality (CMAQ) modeling system Version 4.5.1 (Byun and Ching, 1999) with SOAmods enhancement (Morris et al., 2006). The CAMx Version 4.5 with CB05 chemistry mechanism was also used as part of the modeling selection process. As part of the Visibility Improvements for States and Tribes in the Southeast (VISTAS), the CMAQ treatment of Secondary Organic Aerosols (SOA) was enhanced to treat SOA processes not included in the standard version of CMAQ (SOAmods enhancement). The SOAmods updates include SOA from sesquiterpenes and isoprene and the polymerization of SOA so that it is no longer volatile. EPA intends to address these processes in the fall 2008 release of CMAQ version 4.6, which was too late for this submittal. Thus, St. Louis has adopted CMAQ version 4.5.1 with SOAmods updates as the core model for the modeling analysis. CAMx v4.5 also includes similar but enhanced SOA update which adds two-product absorptive partitioning model of isoprene SOA (Henze and Seinfeld, 2006) and direct SOA formation from oxygenated semi-volatile organic compounds in addition to the SOAmods updates. The St. Louis PM regional model performance evaluation focused on monitoring sites within the St. Louis 12 km modeling grid (Figure 4-1).

In the St. Louis modeling performance evaluation, the CMAQ and CAMx modeling results are compared with observational data from the **I**nteragency **M**onitoring of **PRO**tected **V**isual **E**nvironments (IMPROVE), **S**peciated **T**rends **N**etwork (STN), **C**lean **A**ir **S**tatus **T**rends **Net**work (CASTNet), and **F**ederal **R**eference **M**ethod (FRM) $PM_{2.5}$ mass monitoring networks. In addition, the model performance evaluation also used observational data from the St. Louis Super Site that was operating during the 2002 modeling period, which greatly enhanced the evaluation of the model over just using routinely available data. The St. Louis PM model performance evaluation focuses primarily on the operational model evaluation of the air quality model's performance with respect to individual components of fine particulate matter ($PM_{2.5}$), as good model performance of the PM component species are used in the procedures used to project future-year $PM_{2.5}$ Design Values. The model performance evaluation was conducted for both the regional scale covering the states of Missouri, Illinois and other nearby States, as well as a focused evaluation on monitoring sites within the St. Louis Nonattainment Area (NAA).

4.2 OPERATION MODEL EVALUATION APPROACH

EPA's integrated ozone, $PM_{2.5}$, and regional haze modeling guidance calls for a comprehensive, multi-layered approach to model performance testing, consisting of the four major components: operational, diagnostic, mechanistic (or scientific), and probabilistic (EPA, 2007a). The St. Louis model performance evaluation effort for $PM_{2.5}$ discussed in this chapter focused on the first two components of the EPA's recommended evaluation approach, namely:

- **Operational Evaluation**: tests the ability of the model to estimate $PM_{2.5}$ mass concentrations and the components of $PM_{2.5}$, that is sulfate, nitrate, ammonium, organic carbon matter, elemental carbon, and other inorganic $PM_{2.5}$. This evaluation examines whether the measurements are properly represented by the model predictions but does not necessarily ensure that the model is getting "the right answer for the right reason"; and

- **Diagnostic Evaluation:** tests the ability of the model to predict visibility and extinction, PM chemical composition including PM precursors (e.g., SO_x, NO_x, and NH_3) and associated oxidants (e.g., ozone and nitric acid); PM size distribution; temporal variation; spatial variation; mass fluxes; and components of light extinction (i.e., scattering and absorption).

The diagnostic evaluation also includes the performance of diagnostic tests to better understand model performance and identify potential flaws in the modeling system that can be corrected.

In this model performance evaluation for the St. Louis 2002 Base 5b CMAQ and CAMx 36/12 km base case simulations, the operational evaluation has been given the greatest attention since this is the primarily thrust of EPA's modeling guidance. However, we have also examined certain diagnostic features dealing with the model's ability to simulate gas phase and aerosol concentration distributions.

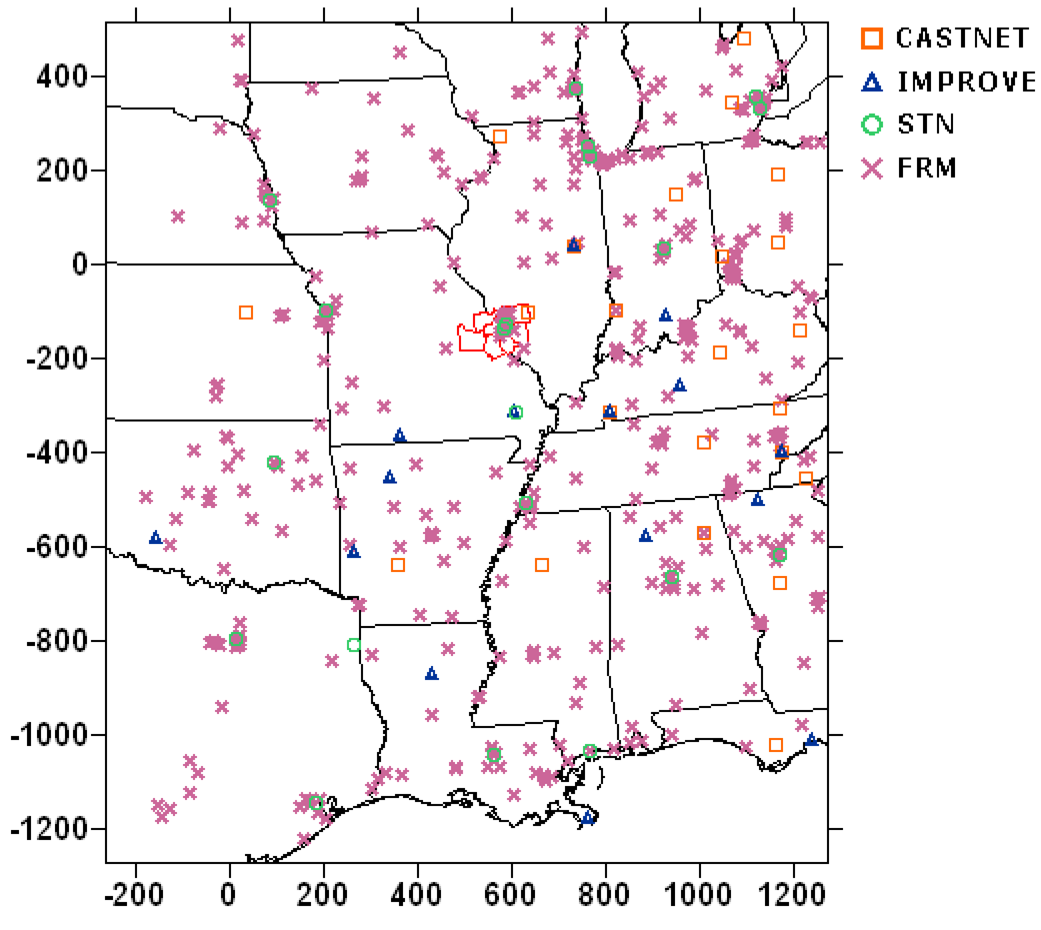


Figure 4-1. CASTNET, IMPROVE, STN, and FRM monitoring sites located in the St. Louis 12 km grid (red boundaries indicate St. Louis $PM_{2.5}$ NAA counties).

4.2.1 Particulate Matter and Component Species

$PM_{2.5}$ attainment is based on $PM_{2.5}$ mass measurements using FRM monitoring devices that consists of the following $PM_{2.5}$ components:

- Fine Sulfate (SO_4)
- Fine Nitrate (NO_3)
- Fine Ammonium (NH_4)
- Organic Carbon Matter (OCM)
- Elemental Carbon (EC) [also called Black Carbon (BC)]
- Other Inorganic $PM_{2.5}$ that is also referred to as SOIL (also known as crustal material, fine soil, major metal oxides, or other $PM_{2.5}$)
- Particle Bound Water (PBW)
- Sea Salt (mostly sodium chloride-NaCl)
- Passive Mass (Blank Correction)

4.2.2 Ambient Air Quality Data for Model Performance Evaluation

A ground-level model evaluation database for 2002 was compiled by the modeling team using several routine and research-grade databases. The St. Louis model performance evaluation focuses on $PM_{2.5}$ mass and its components. The primary monitoring networks available to evaluate this component of the models are: IMPROVE, CASTNET, $PM_{2.5}$ and PM_{10} FRM networks, and STN. These PM monitoring networks may also provide ozone and other gas phase precursors and product species, and visibility measurements at some sites. Table 4-1 and Figure 4-1 summarizes the species collected and locations of the monitoring sites for the IMPROVE, STN, CASTNet, and FRM monitoring networks used in the St. Louis model evaluation.

Table 4-1. Ambient monitoring data available in the St. Louis modeling region during 2002.

Monitoring Network	Chemical Species Measured	Sampling Frequency; Duration
IMPROVE	Speciated $PM_{2.5}$ and PM_{10}	1 in 3 days; 24 hr
CASTNET	Speciated $PM_{2.5}$, SO_2 and HNO_3	Hourly, Weekly; 1 hr, Week
FRM	Only total fine mass ($PM_{2.5}$)	1 in 3 days; 24 hr
STN	Speciated $PM_{2.5}$	Varies; Varies

4.3 MODEL PERFROMANCE GOALS AND CRITERIA

To quantify model performance, several statistical measures were calculated and evaluated for all the IMPROVE, STN, CASTNet and FRM monitors within the St. Louis 12 km modeling domain, individually for each monitoring network. The statistical measures selected were based on the recommendations outlined in section 18.4 of the EPA's Guidance On The Use Of Models And Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze (EPA, 2007a).

For the St. Louis model evaluation, the model performance goals and criteria are used for components of fine particle mass based on previous studies for VISTAS/ASIP model performance for ozone and fine particles (Morris et al., 2004a,b,c). EPA modeling guidance for fine particulate matter noted that PM models might not be able to achieve the same level of performance as ozone models. VISTAS/ASIP reviewed numerous model performance evaluation metrics to evaluate their descriptive capabilities for summarizing the salient features of the model performance evaluation. Although numerous model performance statistics measures are routinely calculated, VISTAS/ASIP have found that the fractional bias and fractional gross error provide the best descriptive power over a wide range of concentrations. The fractional bias and error are expressed as a percentage and are normalized by the average of the predicted and observed values. Consequently, they are bounded statistics, with the fractional bias bounded by -200% to +200%b and the fraction error bounded by 0 to 200%. Table 4-2 summarized the formulas for the fractional bias and error statistics. The model performance goals and criteria used for VISTAS/ASIP are given in Table 4-3. We will use these statistics and performance goals/criteria for the St. Louis PM model performance evaluation.

Table 4-2. Definitions of the fractional bias and fractional error statistical model performance metrics.

Statistical Measure	Shorthand Notation	Mathematical Expression	Notes
Mean Fractional Gross Error (Fractional Error)	MFE	$\frac{2}{N}\sum_{i=1}^{N}\left\|\frac{P_i - O_i}{P_i + O_i}\right\|$	Reported as %
Mean Fractionalized Bias (Fractional Bias)	MFB	$\frac{2}{N}\sum_{i=1}^{N}\left(\frac{P_i - O_i}{P_i + O_i}\right)$	Reported as %

Table 4-3. Model performance goals and criteria for components of fine particle mass**.**

Fractional Bias	Fractional Error	Comment
<u><</u>±15%	<u><</u>35%	Goal for PM model performance based on ozone model performance, considered excellent performance
<u><</u>±30%	<u><</u>50%	Goal for PM model performance, considered good performance
<u><</u>±60%	<u><</u>75%	Criteria for PM model performance, considered average performance. Exceeding this level of performance indicates fundamental concerns with the modeling system and triggers diagnostic evaluation.

4.4 REGIONAL OPERATIONAL MODEL PERFORMANCE EVALUATION

The regional model performance evaluation for the St. Louis 12 km CAMx and CMAQ simulations is summarized below using monthly and seasonal fractional bias and fractional error performance statistics. Scatter plots of observed and predicted concentrations for $PM_{2.5}$ and its constituents as well as additional model evaluation products are provided in Appendix A.

4.4.1 Fine Sulfate (SO_4) Model Performance

Figure 4-2 displays the monthly mean fractional biases and fractional errors for CAMx and CMAQ SO_4 concentration against measurement data at CASTNet, IMPROVE, and STN monitoring networks across the St. Louis 12 km modeling domain (Figure 4-1). SO_4 concentrations are underestimated for all the months by CMAQ and for summer months by CAMx. CAMx overpredicts SO_4 for winter months. CAMx fractional bias stays within ±20% during winter and spring months and achieves the ±30% goal except for June to October. CMAQ consistently shows underprediction bias and achieves the ±30% goal for about half of the months during the year. The biases are within the ±60% of performance criteria for both models for all months. The fractional errors are mostly under 50% except for a few summer months (June to August for CAMx and June for CMAQ) and always under 70%. In general, the SO4 performance is better at urban sites (STN) than at rural sites (IMPROVE), which is encouraging because the focus of the model performance evaluation is on the urban $PM_{2.5}$ NAAs.

Seasonal performance is presented in the form of “Soccer Plots” which plot bias (x-axis) versus error (y-axis) on a single plot so one can quickly see whether the model performance meets goals and criteria (indicated when the bias/error data point falls within the soccer goal). CAMx achieves the performance goal of ±30%/50% except for Quarter2 (Q2) (IMPROVE) and Quarter3 (Q3) (CASTNet and IMPROVE) while CMAQ slightly misses the PM performance goal for Quarter1 (Q1) (CASTNet and IMPROVE), Q2 (CASTNet and IMPROVE), and Q3 (IMPROVE) but shows less scattered statistics than CAMx. Both models meet the criteria (±60%/75%) of PM model performance for all four quarters.

It has been suggested that the summer SO4 underprediction bias is partly due to overstated convective precipitation in the MM5 simulations (e.g., Olerud, 2003c,d). This is somewhat confirmed by a previous study for VISTAS where the SO_4 wet deposition model performance evaluation exhibits near zero bias during the winter when precipitation is dominated by synoptic weather events, but has a positive overprediction bias during the summer when convective precipitation is greatest (Morris et al., 2007).

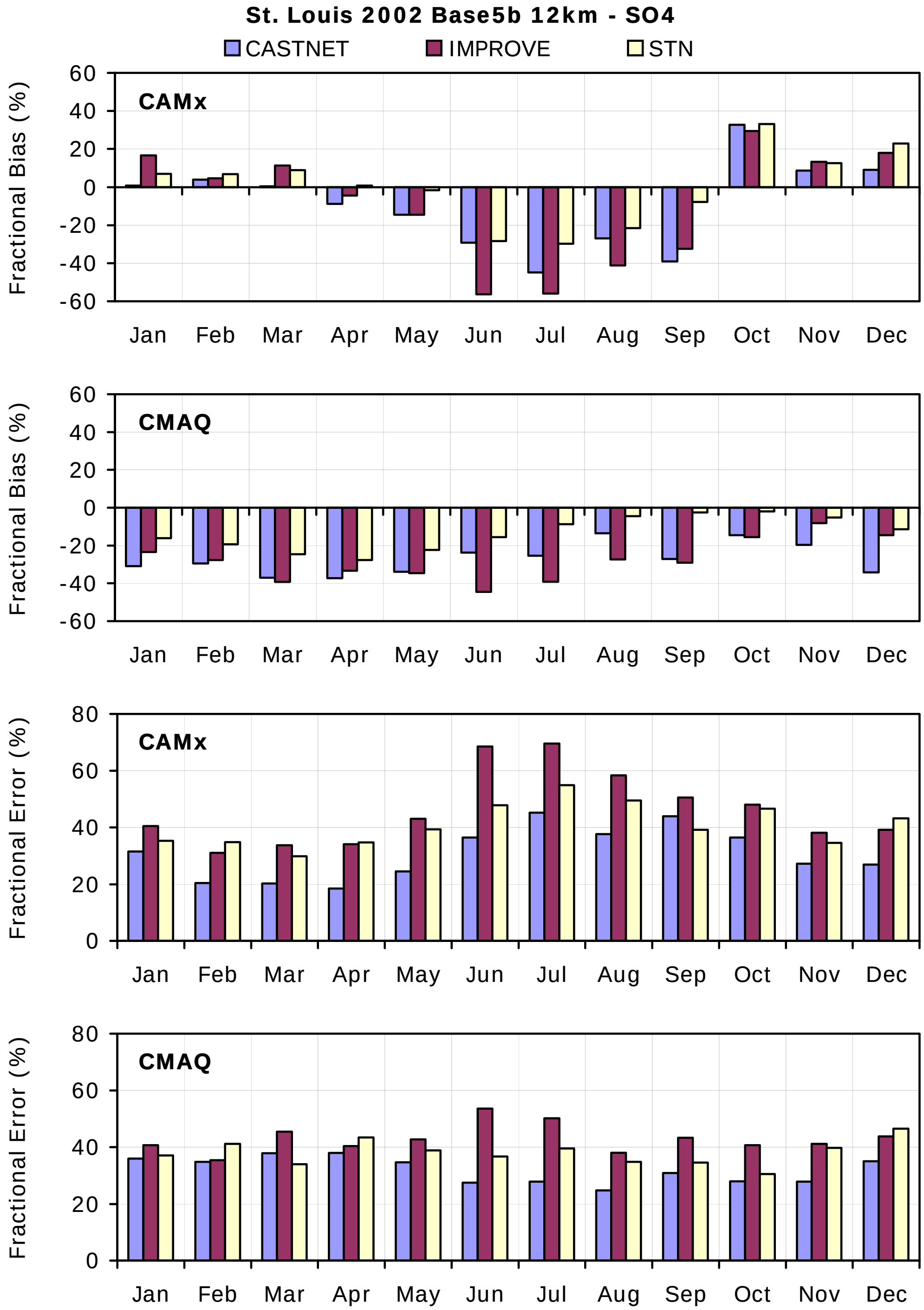


Figure 4-2. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted fine sulfate against measurement data at CASTNet, IMPROVE, and STN monitoring sites in the St. Louis 12 km domain.

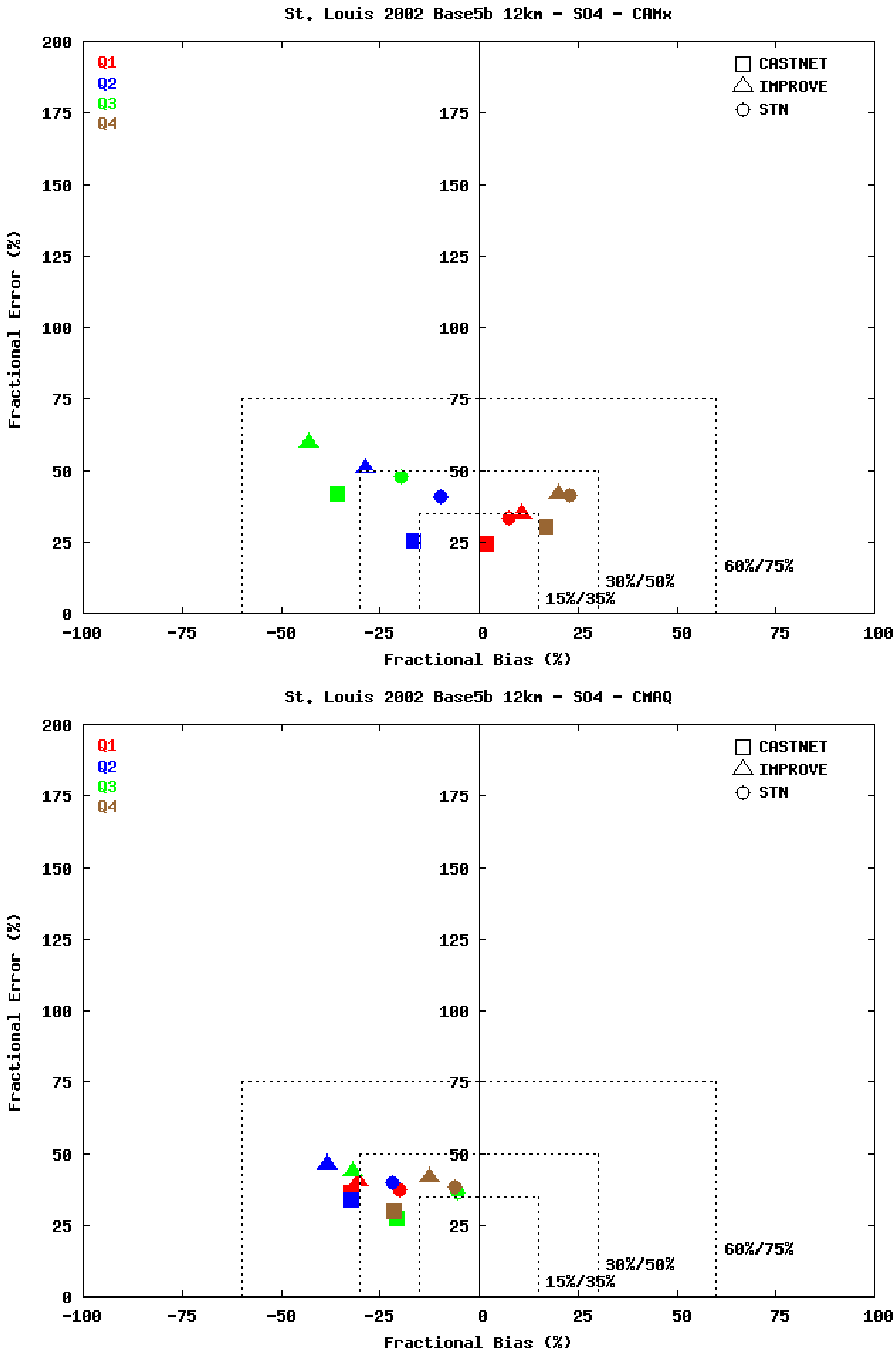


Figure 4-3. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ fine sulfate performance.

4.4.2 Fine Nitrate (NO_3) Model Performance

Among the major $PM_{2.5}$ constituents, NO_3 shows the poorest performance. Both models severely underpredict NO_3 concentrations in the summer with large negative fractional bias of -50% to -150%, with the CAMx fractional bias even goes over -150% at the IMPROVE sites (Figure 4-4). The model performance improves in the winter where CMAQ fractional bias is mostly within ±50%. Except for November and December, CAMx shows negative bias below -50% even in the winter. Fractional errors show similar pattern with larger errors in the summer.

The Soccer Plots clearly show that both models are unable to achieve the PM model performance criteria in most cases (Figure 4-5; note that, in the Soccer Plots, if the symbol plots outside of the range of the plot, then it is plotted on the axis). However, the large summer underprediction bias values occur when NO_3 concentrations are extremely low. The NO_3 will almost completely volatilize off the FRM filter. Thus, these large summer NO_3 underpredictions are not likely a concern to the overall PM performance as NO_3 is not a significant component of $PM_{2.5}$ mass.

4.4.3 Fine Ammonium (NH_4) Model Performance

Ammonium concentrations are closely linked to the availability of sulfate and nitrate. Therefore, when the models show underprediction biases for sulfate and nitrate during summertime, it is expected that NH_4 show the similar pattern (Figure 4-6). However, the NH_4 model performance is better than NO_3 in the summer because sulfate level is much higher than nitrate during summertime, thus most ammonium is mainly associated with sulfate.

The model performance at the STN monitoring sites seems to be better than at the IMPROVE sites. Ammonium is not directly measured by IMPROVE network, rather it is derived by assuming it completely neutralizes the measured sulfate and nitrate. Assuming that nitrate is completely neutralized by ammonium is a valid assumption most of the time, especially for these inland sites. However the same may not be true for sulfate, especially in the summer months. Hence the derived observed ammonium comparison at the IMPROVE sites is likely overstated during the summer months when sulfate is less likely to be fully neutralized and the underprediction bias is more significant than at the STN sites.

The Soccer Plots for ammonium (Figure 4-7) show that the seasonal model performances always meet the PM model performance criteria except for CAMx during Q3 at the IMPROVE sites.

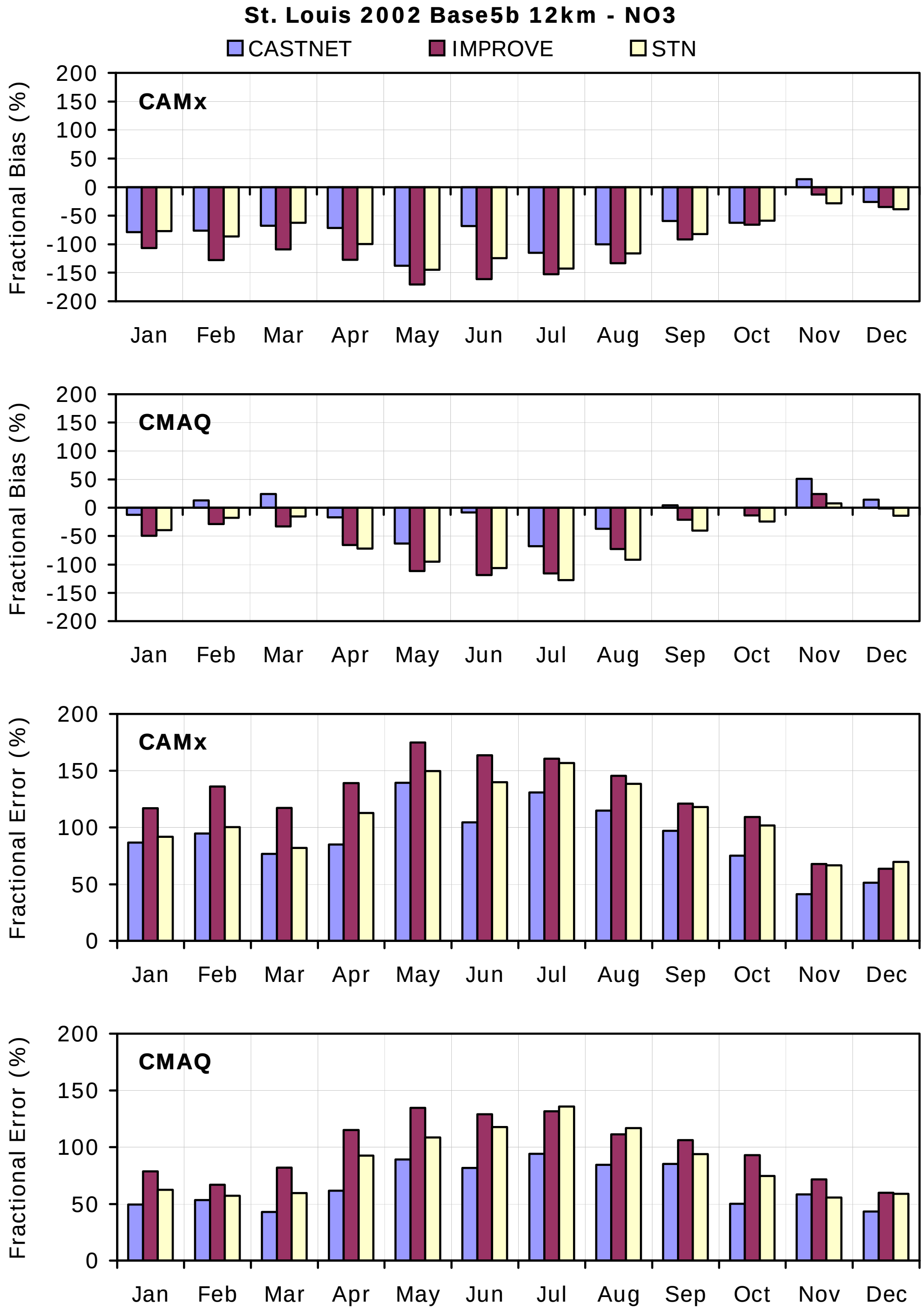


Figure 4-4. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted fine nitrate against measurement data at CASTNet, IMPROVE, and STN monitoring sites in the St. Louis 12 km domain.

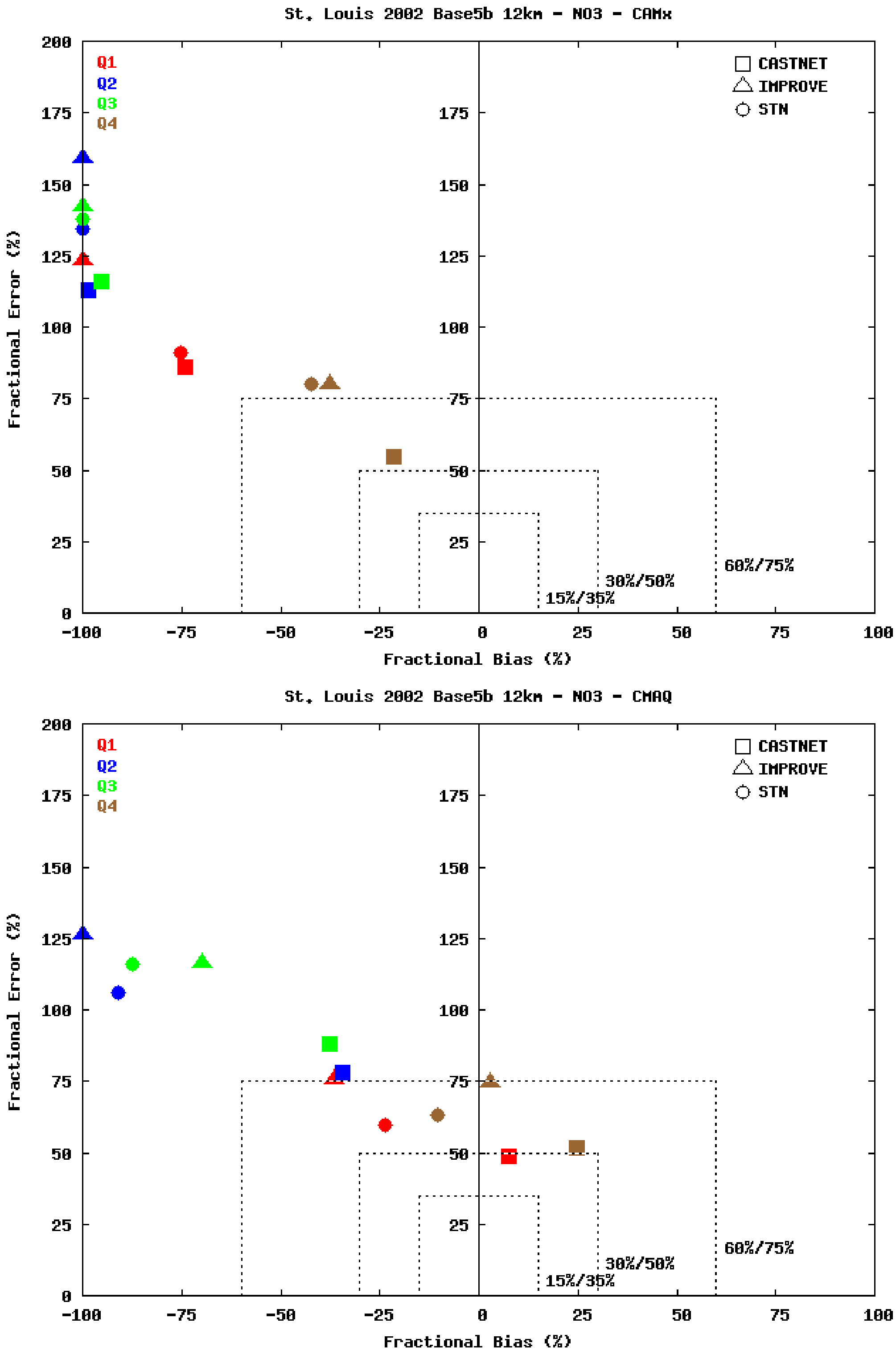


Figure 4-5. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ fine nitrate performance.

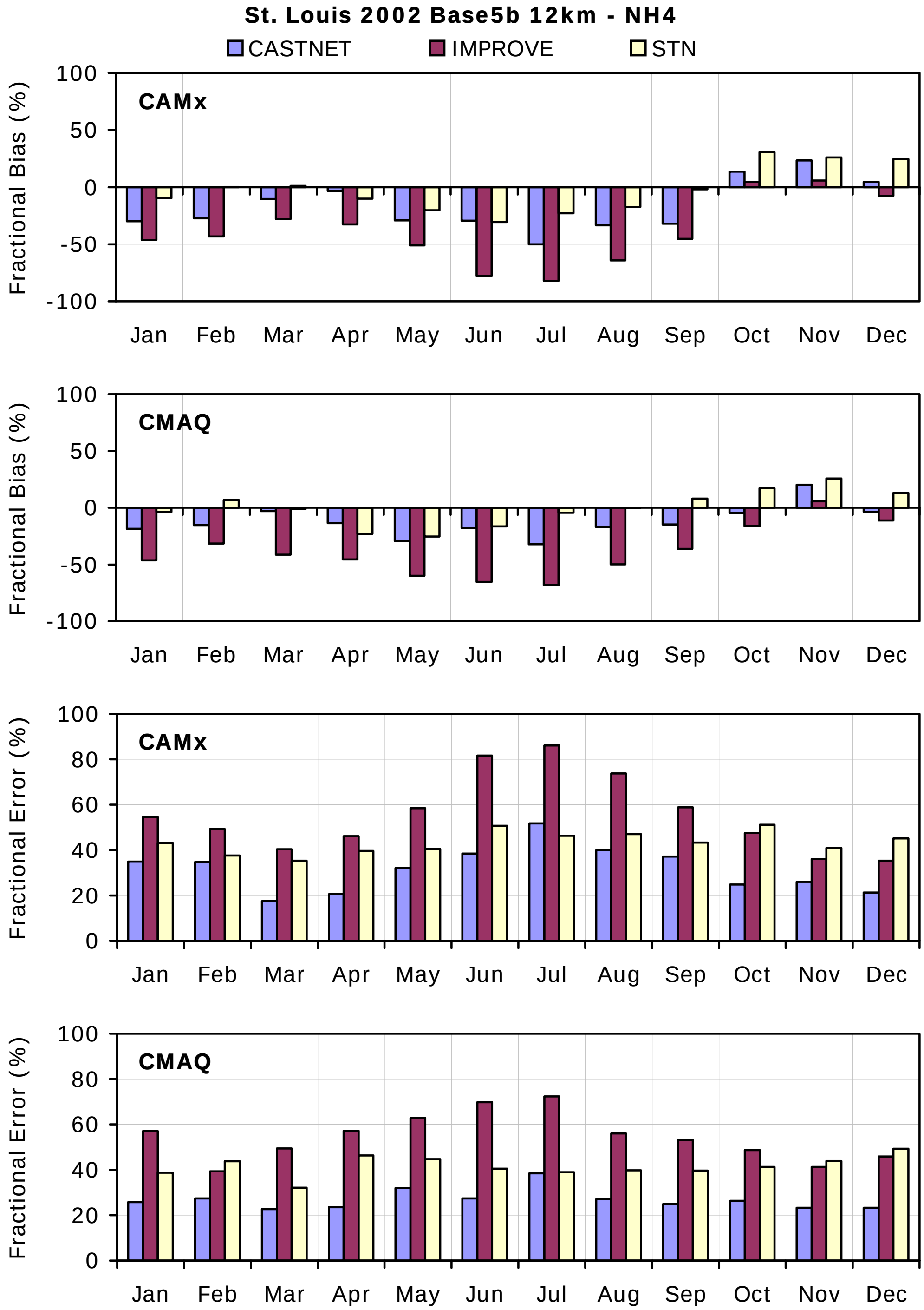


Figure 4-6. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted fine ammonium against measurement data at CASTNet, IMPROVE (derived assuming complete neutralization), and STN monitoring sites in the St. Louis 12 km domain.

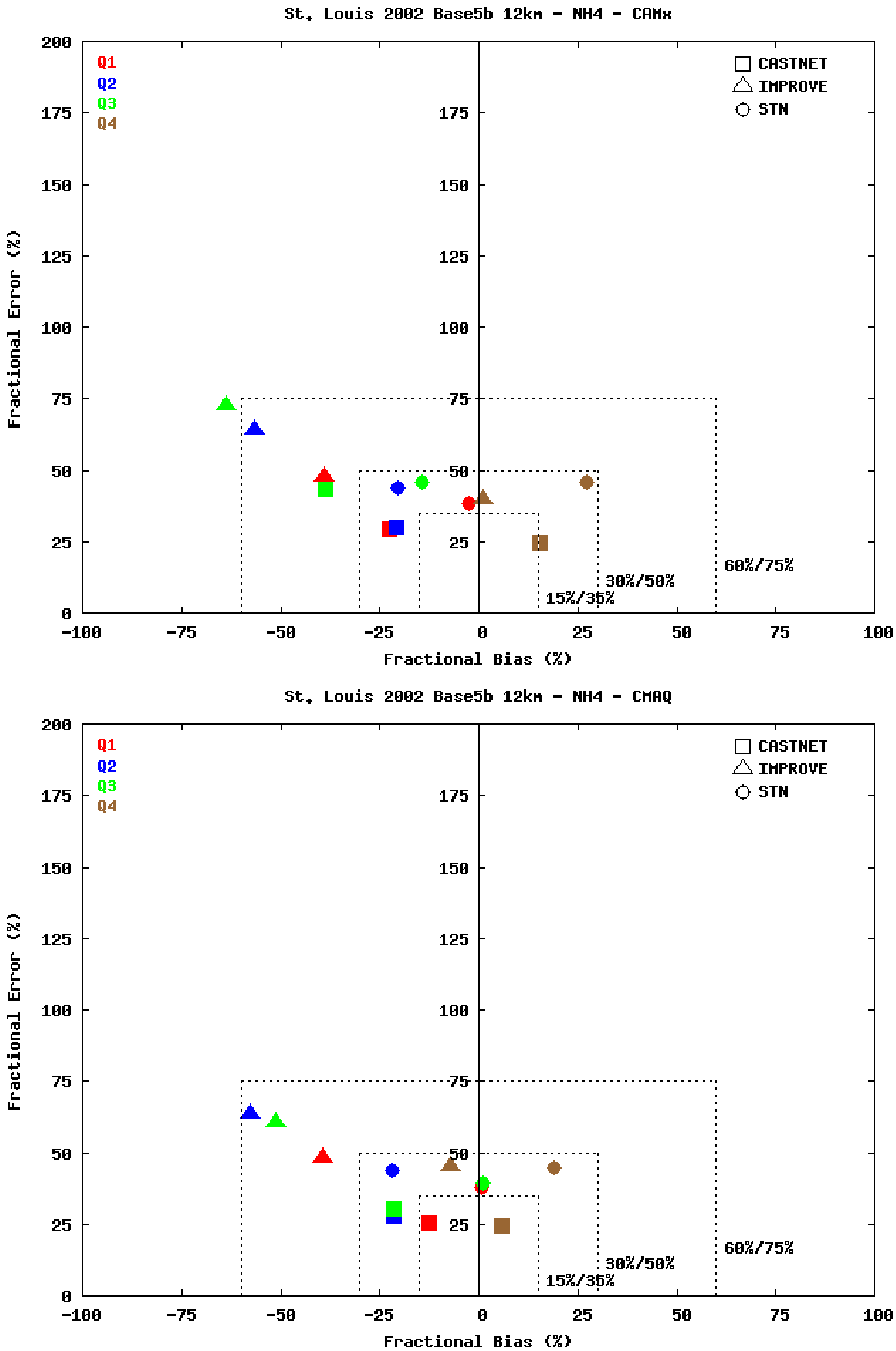


Figure 4-7. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ fine ammonium performance.

4.4.4 Fine Organic Carbon Matter (OCM) Model Performance

Both CAMx and CMAQ models exhibit significant underprediction biases for OCM throughout the year, with slightly larger underprediction for summer months (Figure 4-8). The model performance is better at the IMPROVE (rural) sites where fractional biases range from -24% to -64% and fractional errors from 39% to 68%. At the STN (mostly urban) sites, the models exhibit larger fractional biases (-64% to -96%) and errors (68% to 98%). The Soccer Plots show that the seasonal model performances meet the ±60%/75% bias/error PM criteria for both models at the STN sites while falling outside of the criteria region at the IMPROVE sites (Figure 4-9).

Part of the underprediction bias at the STN sites may be due to measurement uncertainties and artifacts and part may also be due to the model over diluting the urban OCM emissions through the coarse 12 km grid. The underprediction of urban OCM is a common problem in PM modeling and likely also points to uncertainties in the OCM and SOA precursor emission inventories. One source of uncertainty in the OCM measurements is the fact that OCM is actually derived from OC measurements. The amount of additional elements (e.g., oxygen) attached to the OC to form OCM varies with the age and level of photochemical processing of the organic particles with OCM/OC ratios typically ranging from 1.2 to 2.2 with lower ratios for fresh and higher ratios for aged processed OCM. For the OCM model performance evaluation we used a 1.8 OCM/OC ratio that is consistent with the new IMPROVE equation where a higher OCM/OC ratio reflects the fact that OCM that reaches the mostly rural Class I areas will be aged and subject to photochemical processing. For urban sites that are close to emission sources, the ratio of 1.8 may be too high resulting in underprediction bias artifacts. There is no single correct right OCM/OC factor to use in all cases. However, in interpreting the model performance it is important to know which OCM/OC ratio was used and recognize that selection of another ratio could make a ±30% difference in the OCM measurements.

The STN OC measurements are also not blank corrected, which is believed to result in an approximate 0.5 $\mu g/m^3$ positive artifact in the STN OC observations (which would be 0.9 $\mu g/m^3$ positive artifact in OCM with OCM/OC factor of 1.8). The subtraction of 0.9 $\mu g/m^3$ from the measured OCM value would significantly improve the OCM model performance across the STN network for both models. Thus, a large component of the seemingly OCM underprediction bias at the urban STN monitoring sites is measurement artifacts.

4.4.5 Fine Elemental Carbon (EC) Model Performance

The monthly fractional bias and error performance statistics for EC across the 12 km modeling domain are shown in Figure 4-10. Bias and error are lower in the winter months. During the second and third quarters of the year, EC performance across the IMPROVE networks exhibits a large underprediction bias (-33% to -92%) that peaks in June. The EC performance across the STN network is relatively good with the bias and error within the ±30%/50% performance goal for most of the year (Figure 4-11), which suggests that the anthropogenic EC emissions inventory may be adequately characterized in urban areas.

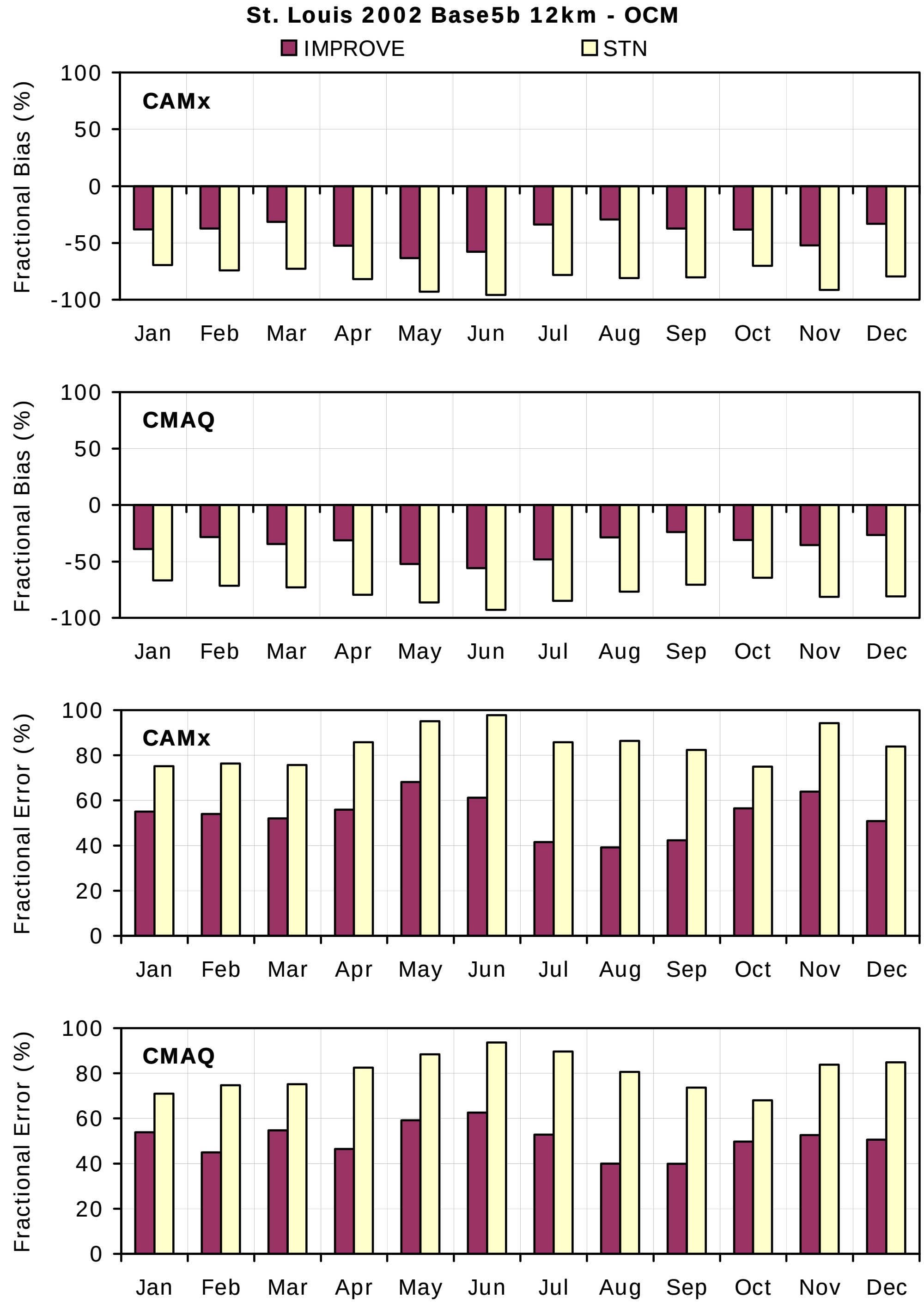
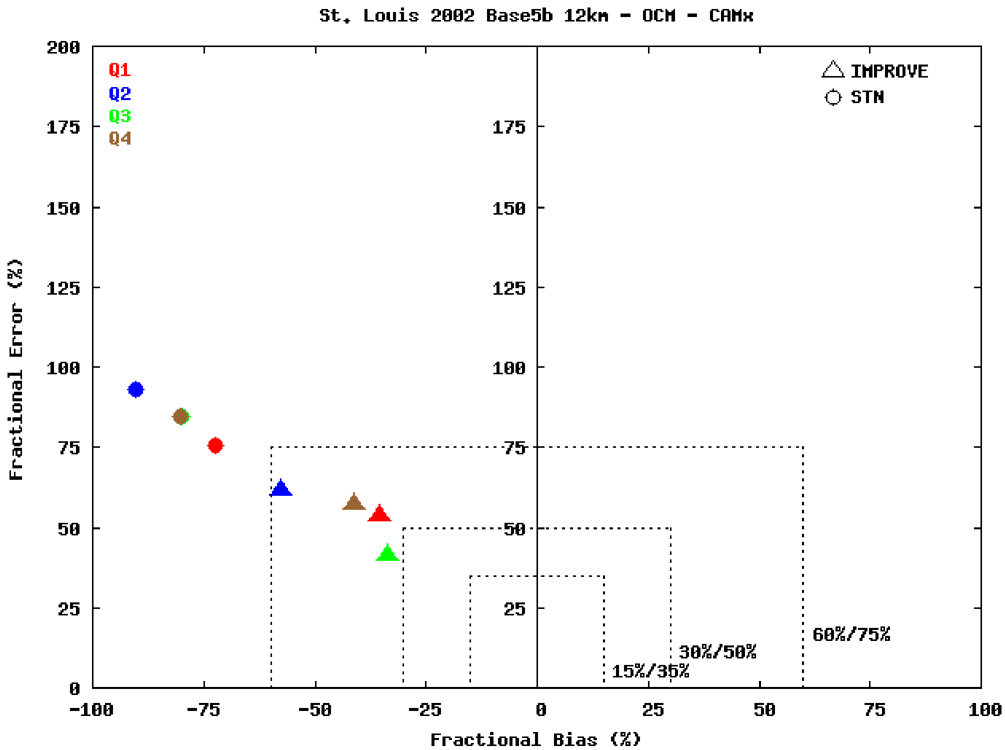


Figure 4-8. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted fine OC matter against measurement data at IMPROVE and STN monitoring sites in the St. Louis 12 km domain.



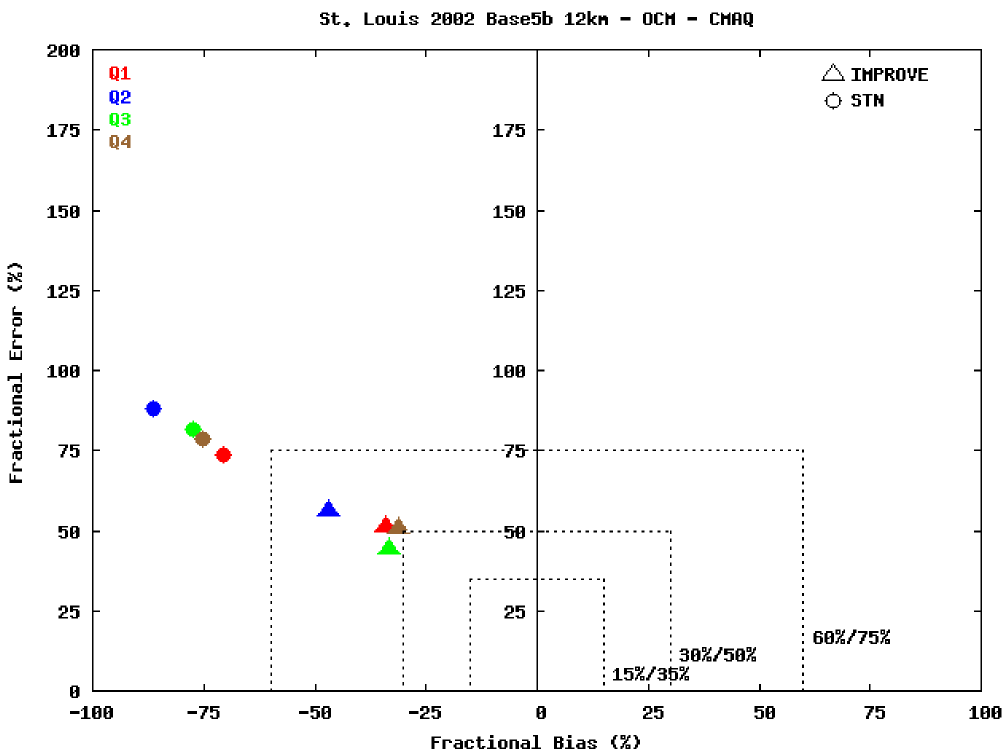


Figure 4-9. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ fine OC matter performance.

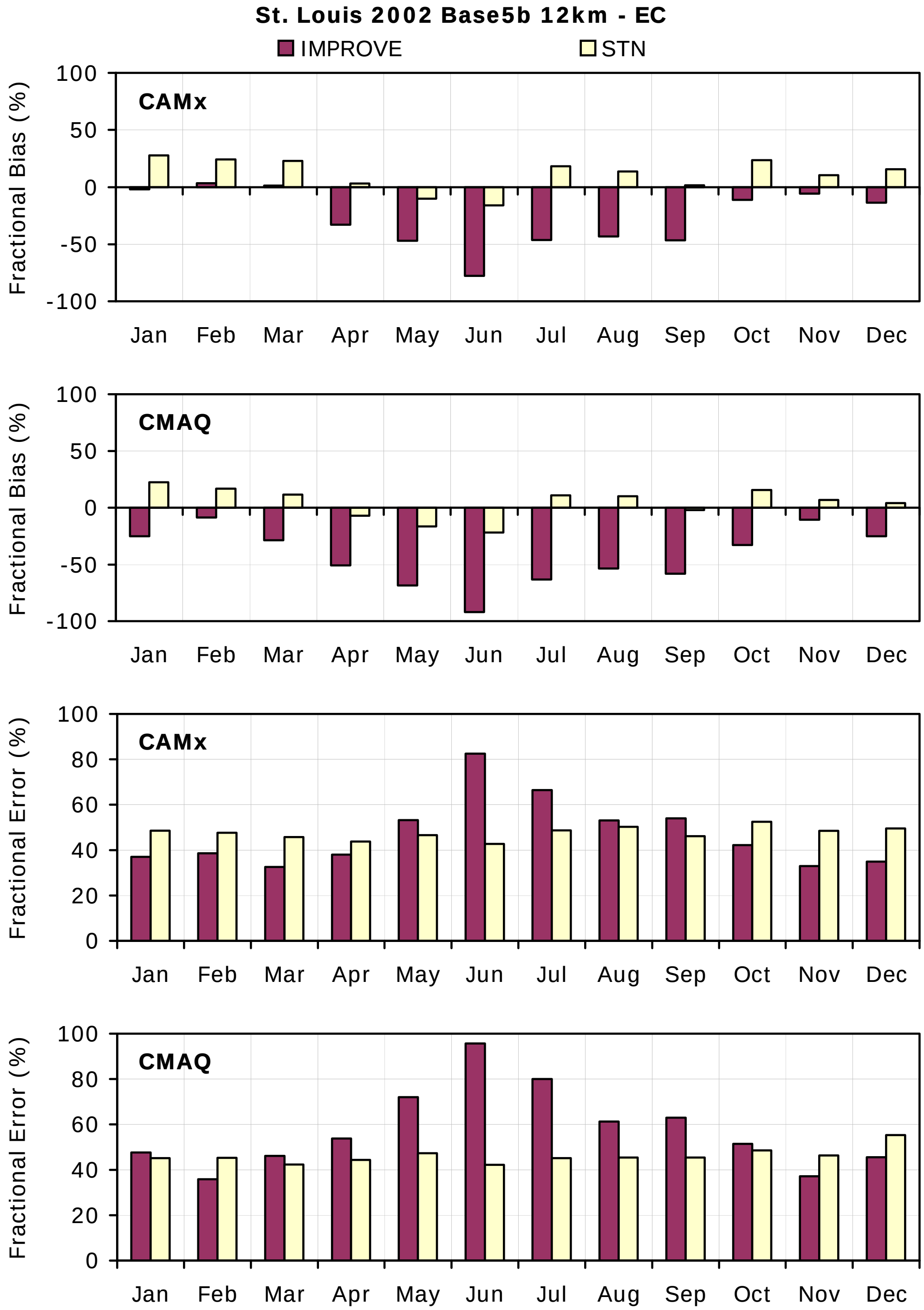
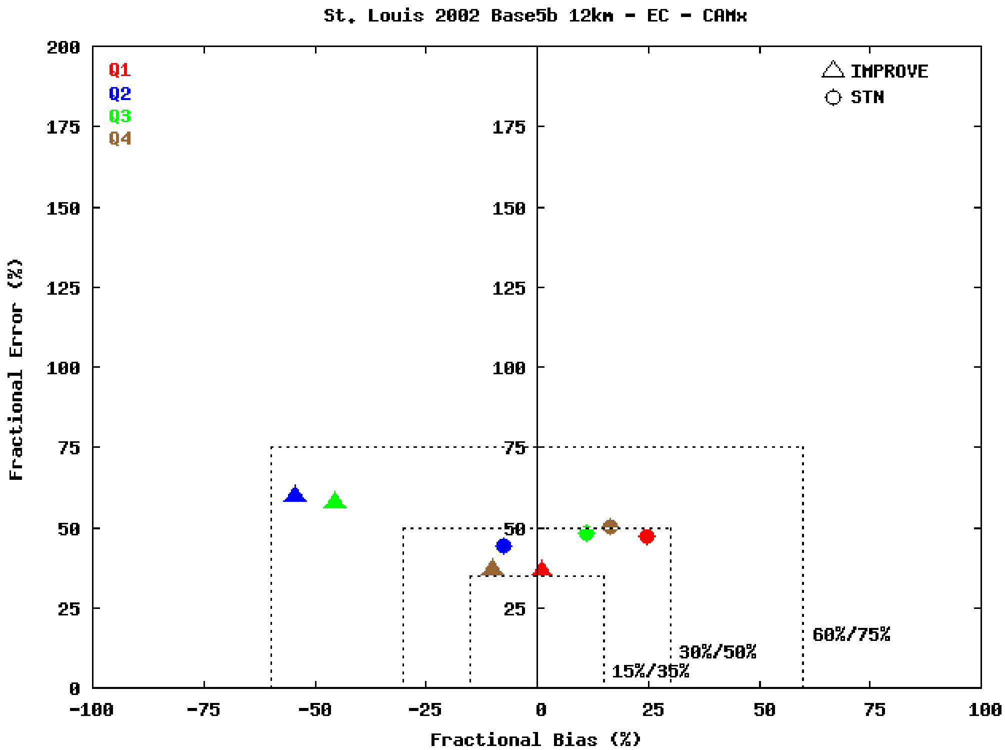


Figure 4-10. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted fine EC against measurement data at IMPROVE and STN monitoring sites in the St. Louis 12 km domain.



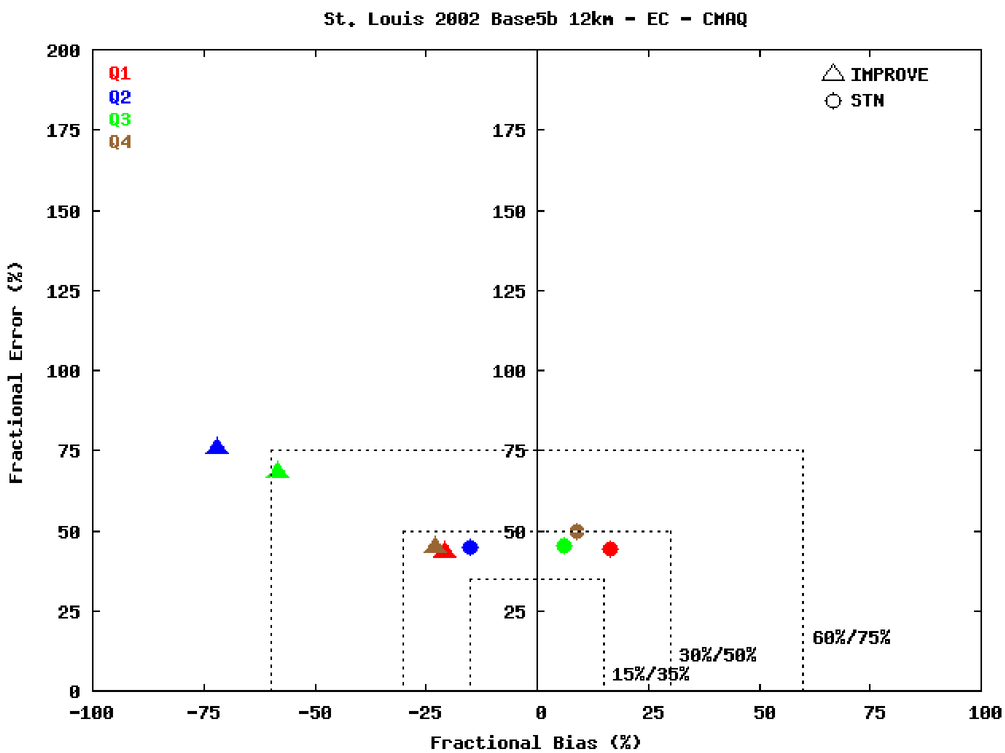


Figure 4-11. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ fine EC performance.

4.4.6 Other Inorganic $PM_{2.5}$ (SOIL) Model Performance

The seasonal dependence of the SOIL model performance across the IMPROVE network in the 12 km modeling domain is clearly evident in the monthly fractional bias plots in Figure 4-12 that show a large (> 100%) overprediction bias in the winter and much smaller bias in the summer. The modeled SOIL values tend to always be between 0 and 5 $\mu g/m^3$ year round, however the observed values are much lower in the winter (0 to 1 $\mu g/m^3$) and comparable to the modeled values in the summer. One notable exception is that several IMPROVE monitoring sites measured 12 to 20 $\mu g/m^3$ of SOIL around July 1, 2002, which resulted in underprediction bias in July.

It has been suggested that the poorer winter SOIL model performance is likely due to incorrect emission temporal adjustment factors. For example, the effect of wetted surfaces that suppresses fugitive dust emissions may not be properly characterized in the seasonal adjustments to the emissions inventory (Morris et al., 2007). Another factor that affects the SOIL model performance is the incompatibilities between the modeled and measured SOIL species. The IMPROVE observed SOIL is built up from measured elements. The "SOIL" in the modeling, on the other hand, is fine particulate matter emissions that have not been explicitly speciated as SO_4, NO_3, OCM or EC in the SMOKE emissions modeling. The emissions PM speciation profiles may have unidentified PM that is lumped in the other PM category that is not the same as the IMPROVE SOIL. Also, impacts of local fugitive dust sources at the monitor that are subgrid-scale to the St. Louis 12 km modeling grid may not be captured by the modeling.

4.4.7 Fine Particulate ($PM_{2.5}$) Model Performance

Figure 4-14 displays the monthly fractional bias and error for total $PM_{2.5}$ mass concentrations for the IMPROVE, STN, and FRM networks in the St. Louis 12 km modeling grid. For the winter months the models perform relatively well while exhibiting considerable underprediction bias for the summer months. The Soccer Plots show both models achieve the PM model performance goal except for the second and third quarters across the IMPROVE network (Figure 4-15). The summer underprediction bias is partly due to the SO_4 and OCM underprediction bias discussed previously as they are the two major components of $PM_{2.5}$ in the summer and the St. Louis 12 km modeling domain.

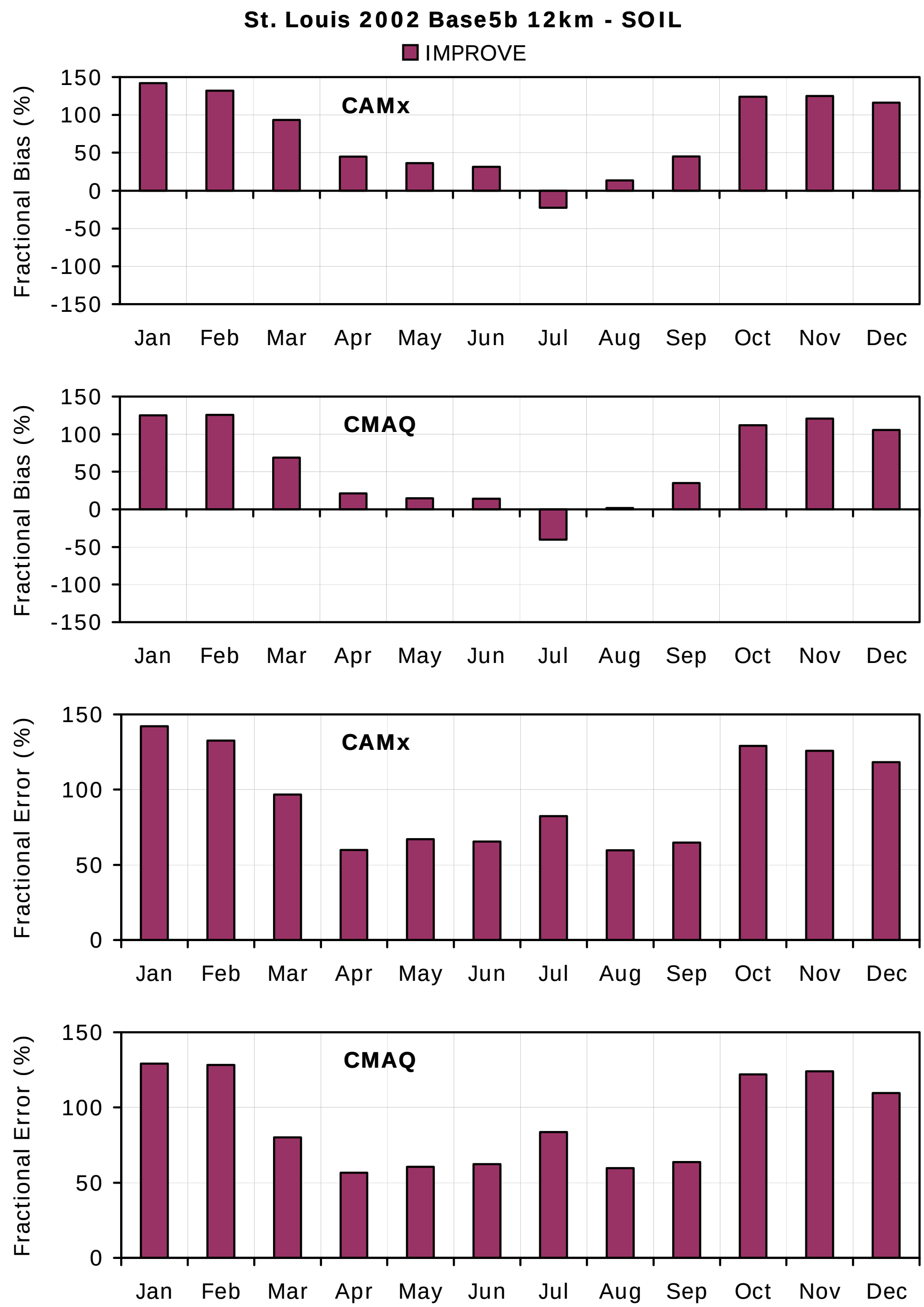


Figure 4-12. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted other inorganic $PM_{2.5}$ against measurement data at IMPROVE monitoring sites in the St. Louis 12 km domain.

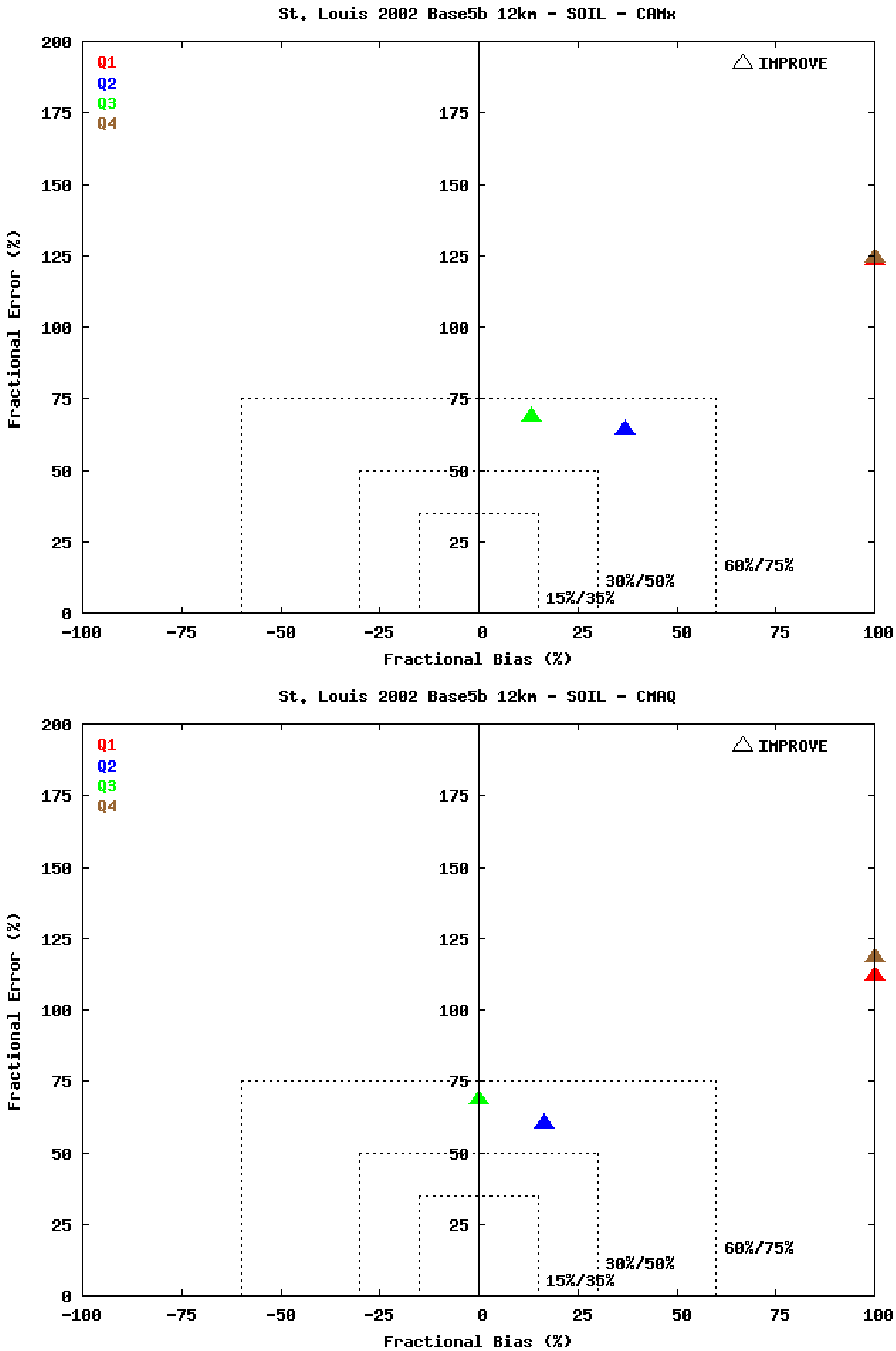


Figure 4-13. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ other inorganic $PM_{2.5}$ performance.

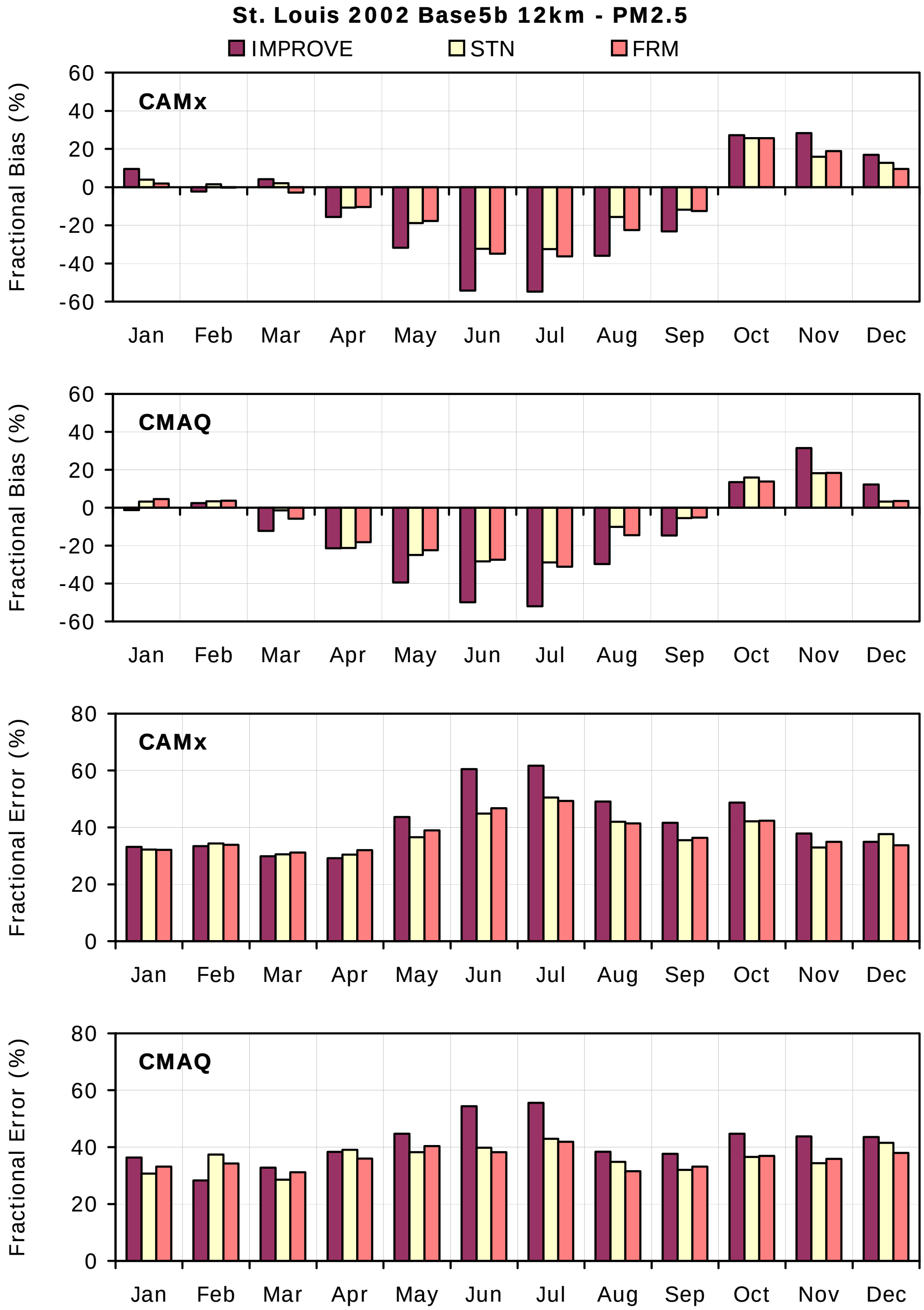


Figure 4-14. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted $PM_{2.5}$ mass against measurement databases at IMPROVE, STN, and FRM monitoring sites in the St. Louis 12 km domain.

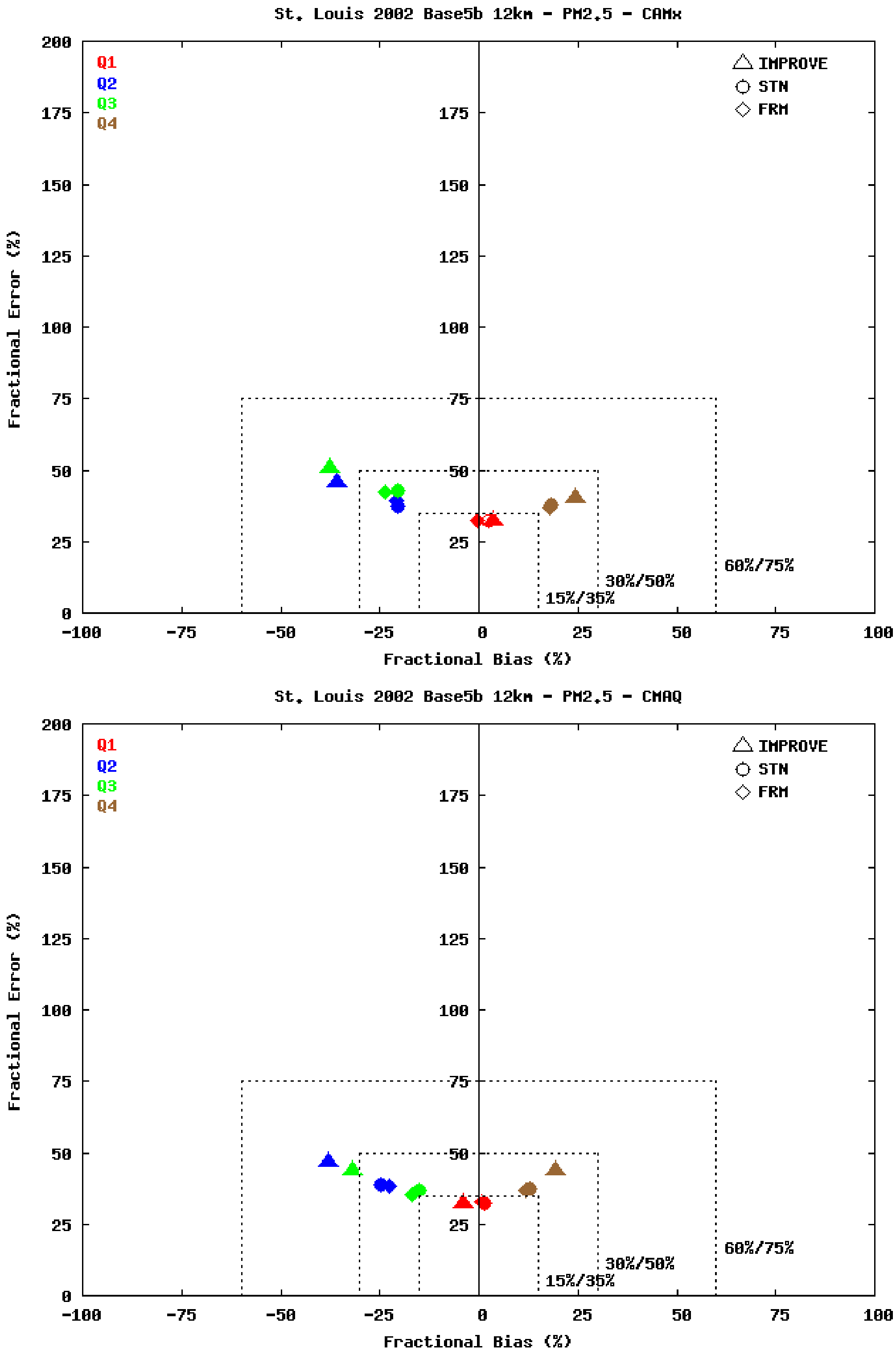


Figure 4-15. Soccer Plot for seasonal fractional bias and error for CAMx and CMAQ $PM_{2.5}$ mass performance.

4.5 URBAN OPERATIONAL MODEL PERFORMANCE EVALUATION

There are four STN monitoring sites in the St. Louis $PM_{2.5}$ nonattainment area (Figure 4-16). The Blair Street site and CAP site are located in the City of St. Louis, MO, while the Alton, IL, and Arnold, MO, are north and south of the city, respectively. The Blair Street, CAP, and Arnold sites operate on a 1-in-3 day schedule. The Alton site operates on a 1-in-6 day schedule but operated on a 1-in-3 day schedule from January to early April 2002. The CAP site operated only for the first half of the year 2002. The 2002 measurement data at these sites was further validated and refined by Dr. Jay Turner's group at Washington University. A blank correction of 0.9 mg/m3 (recommended by Dr. Jay Turner) was applied to the measured OC values and OCM/OC ratio of 1.8 was used.

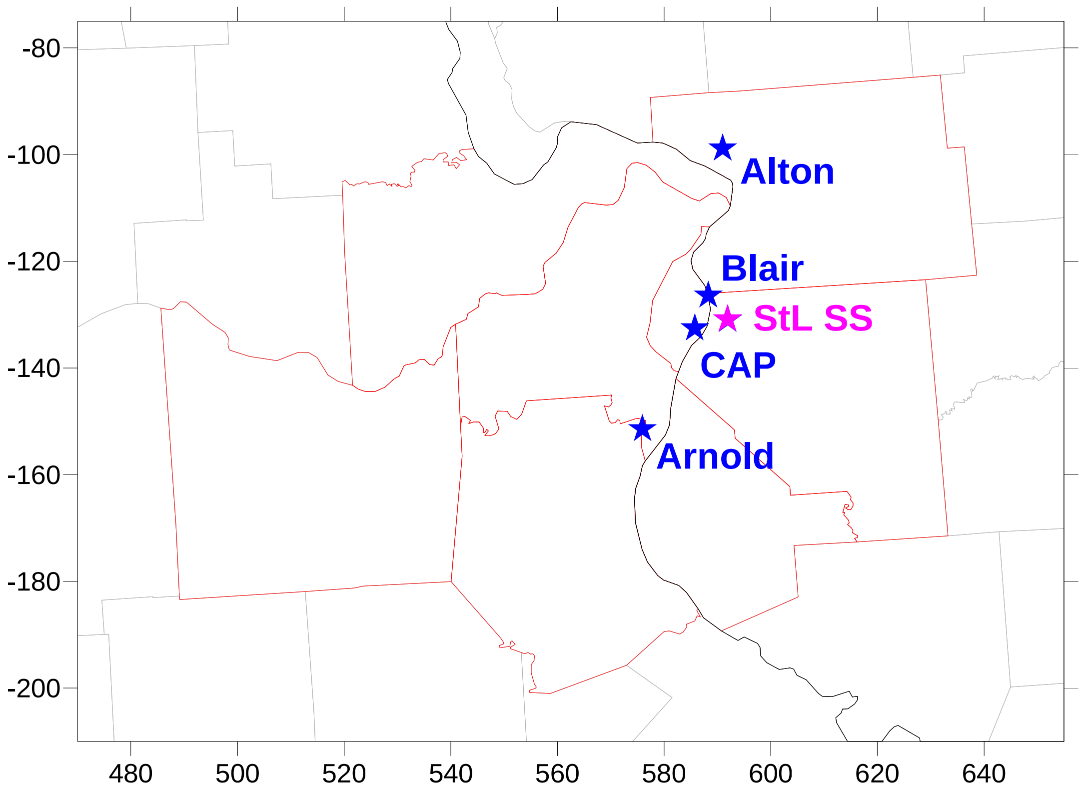


Figure 4-16. St. Louis Super Site and nearby STN monitoring sites in the St. Louis $PM_{2.5}$ NAA.

4.5.1 Fine Sulfate (SO_4)

Figure 4-17 shows time-series plots for observed and predicted SO_4 concentrations at Alton, Arnold, Blair Street and CAP STN sites for the year 2002. CAMx displays very low fractional bias at Alton, Arnold and CAP sites achieving the most stringent <±15% ozone goal while having somewhat larger overprediction bias of 21% at Blair Street site. CMAQ performance is also good achieving the PM model performance goal at all four sites. Location of peak SO_4 concentration is well predicted by both models.

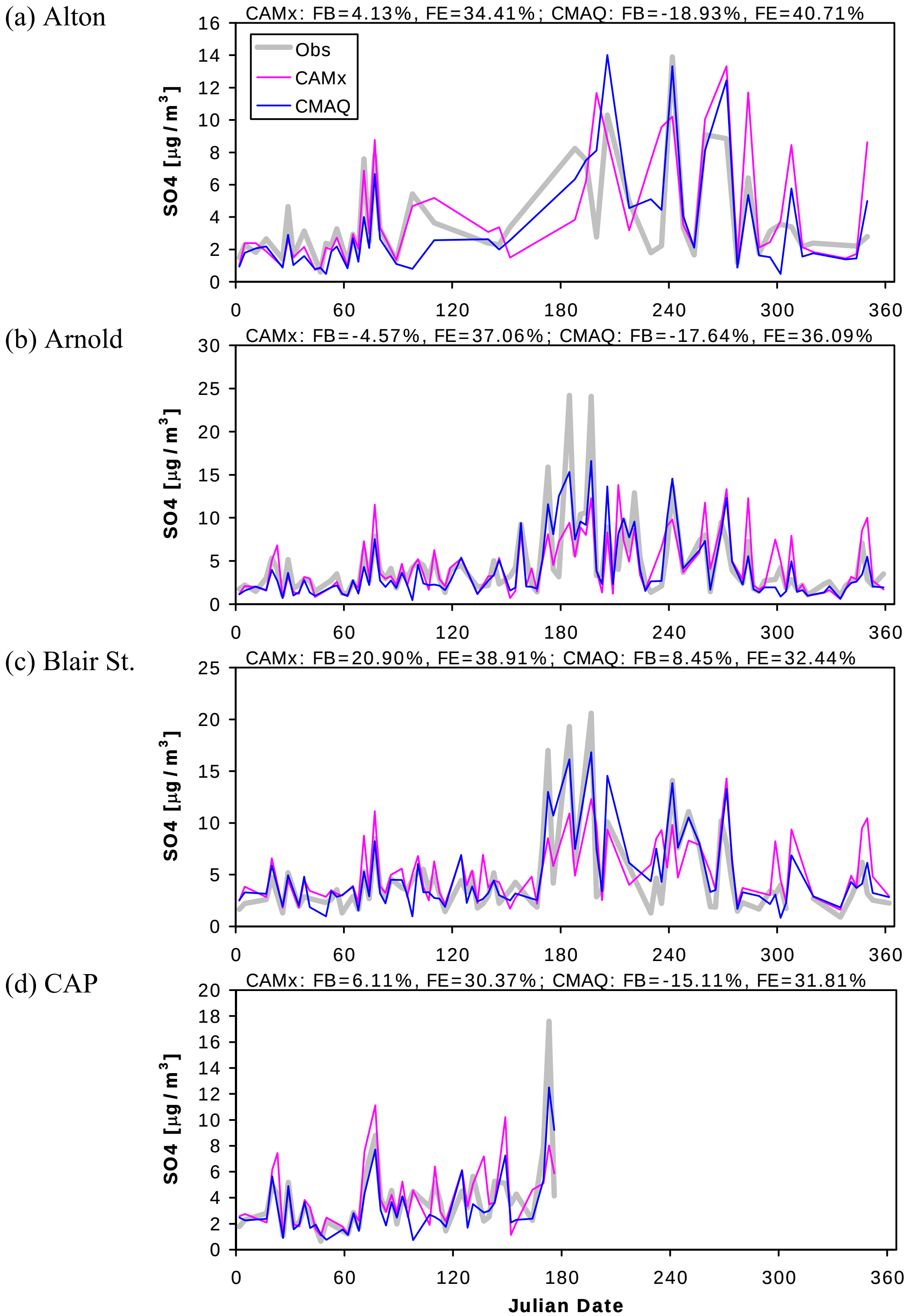


Figure 4-17. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of SO_4.

4.5.2 Fine Nitrate (NO_3)

As observed in the regional performance evaluation, NO_3 is significantly underpredicted at all four STN sites, with the CMAQ model showing less underprediction bias than CAMx (Figure 4-18). It appears that the models almost completely evaporate ammonium nitrate during summer (except for the Arnold site) leading to large underprediction bias as observed NO_3 values do not go that low. During winter time, however, the models follow the observed NO_3 concentrations relatively well.

4.5.3 Fine Ammonium (NH_4)

The NH_4 model performance is much better than NO_3 indicating that ammonium is mostly tied to sulfate rather than nitrate in this region. The time-series plots display reasonable agreement between the model predictions and observations for NH_4 at these STN sites (Figure 4-19). Fractional bias and error values are within ±15% and less than 40%, respectively, for both CAMx and CMAQ.

4.5.4 Organic Carbon Matter (OCM) and Elemental Carbon (EC)

Figure 4-20 shows that both models systematically underpredict OCM concentrations with CMAQ producing slightly larger underprediction bias than CAMx. The OCM annual mean fractional bias ranges from -48% to -81%. Note that blank correction of 0.9 $\mu g/m^3$ was made to the OCM measurement data, thus giving better performance than the regional evaluation where no blank correction was applied.

EC does not contribute much to the total $PM_{2.5}$ mass at these sites and observation values remain below 1.5 $\mu g/m^3$ for most of the year (Figure 4-21). Both models slightly overpredict the EC concentrations with fractional bias ranging 4% to 31% and fractional error less than 42%, which is within the PM model performance goal.

4.5.5 Fine Particulate ($PM_{2.5}$)

The total $PM_{2.5}$ mass model performance at the St. Louis STN sites is quite good with fractional bias of -1% to -22% and fractional error less than 39% (Figure 4-22). Both models meet the PM model performance goals. Again, with SO_4 and OCM being major components in $PM_{2.5}$ mass, the $PM_{2.5}$ time-series plots exhibit similar summer underprediction bias.

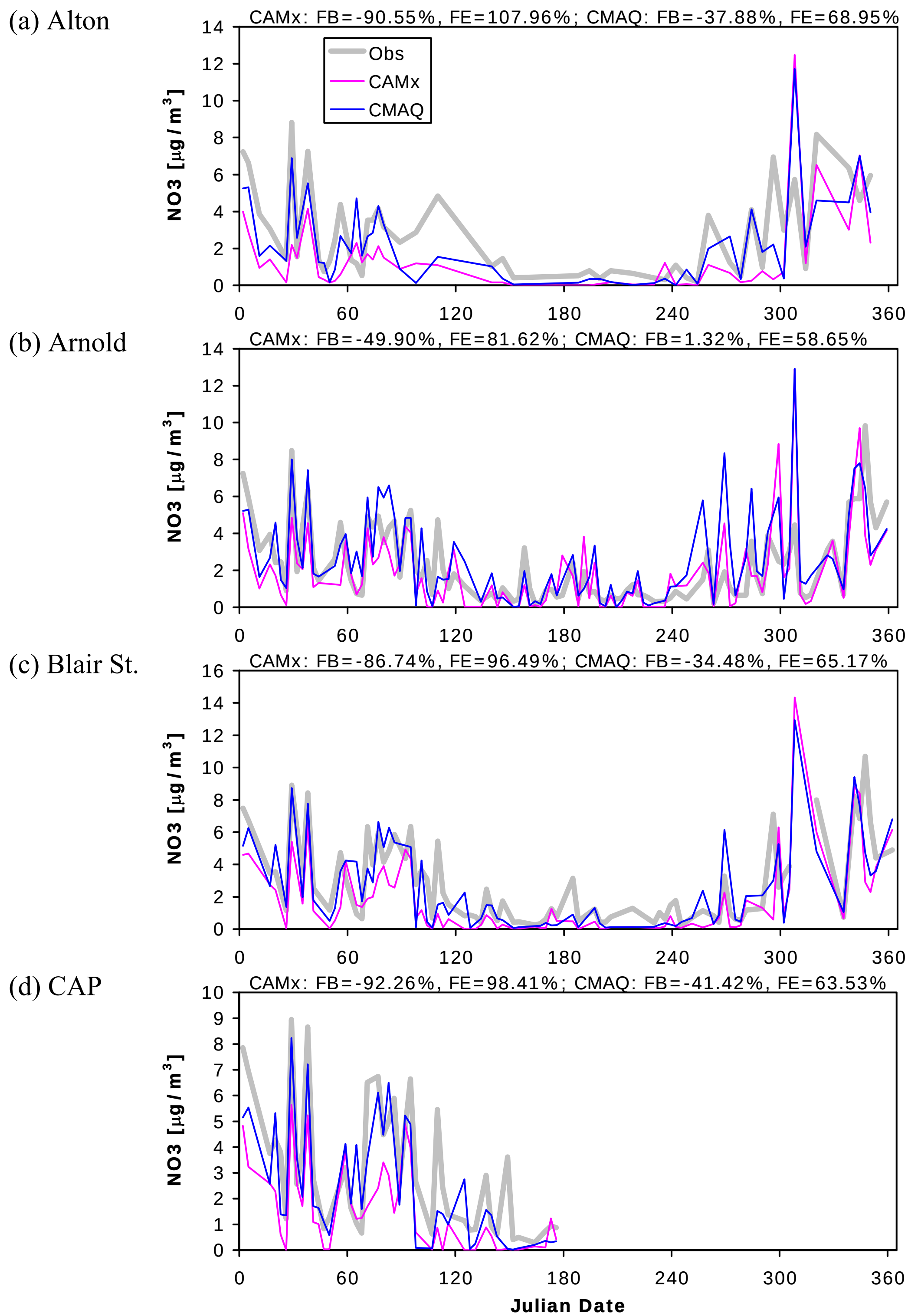


Figure 4-18. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of NO_3.

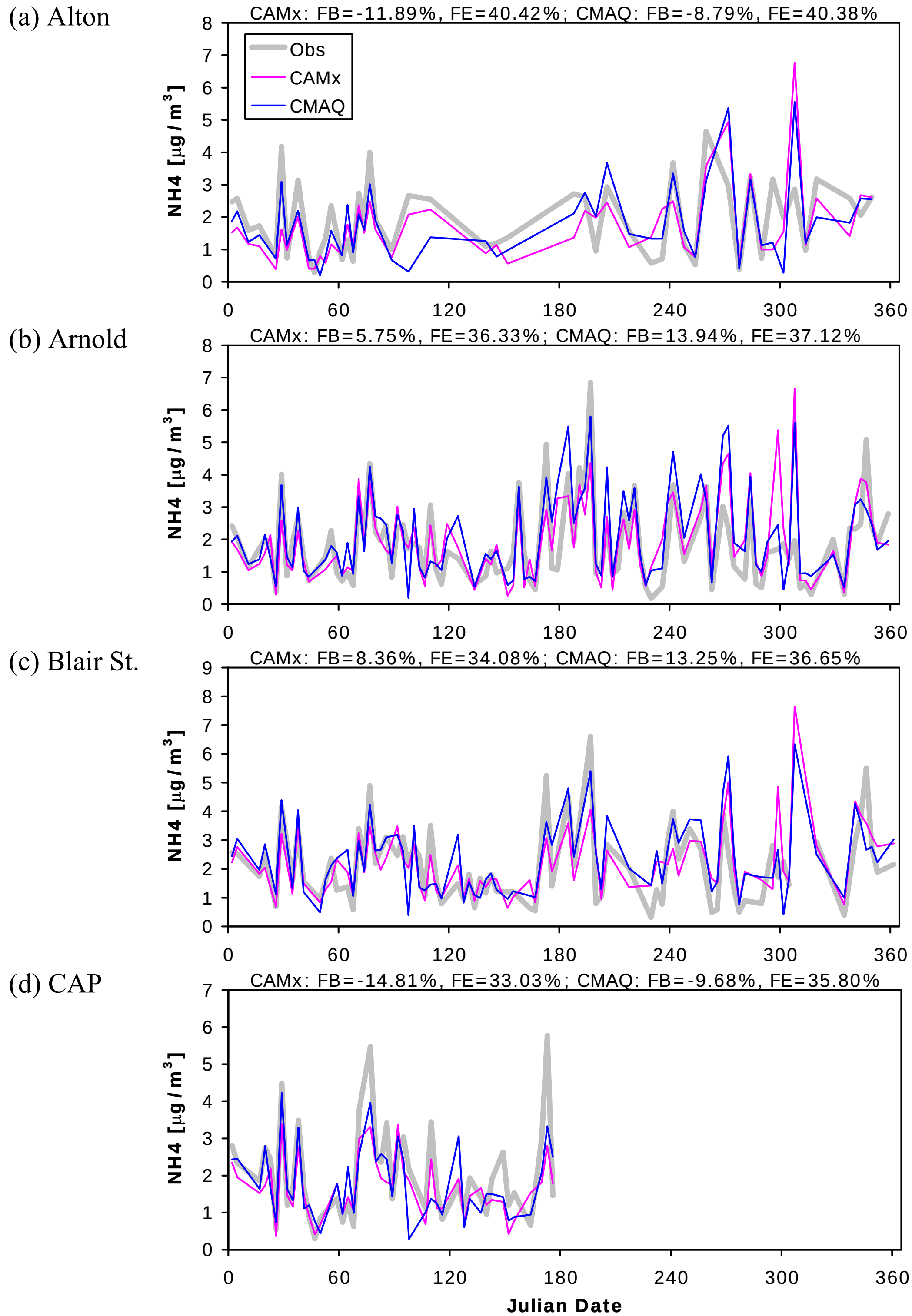


Figure 4-19. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of NH_4.

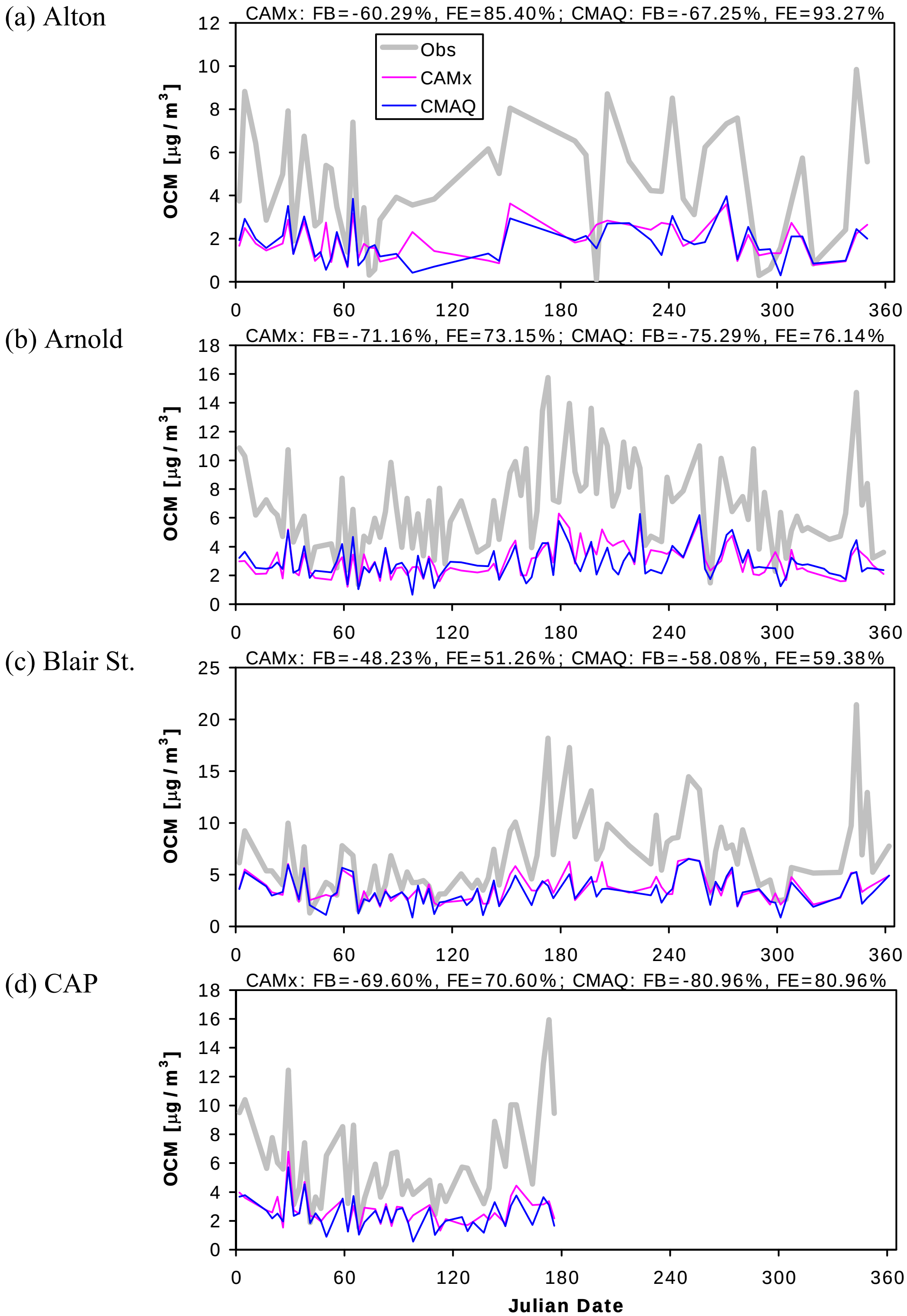


Figure 4-20. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of OCM.

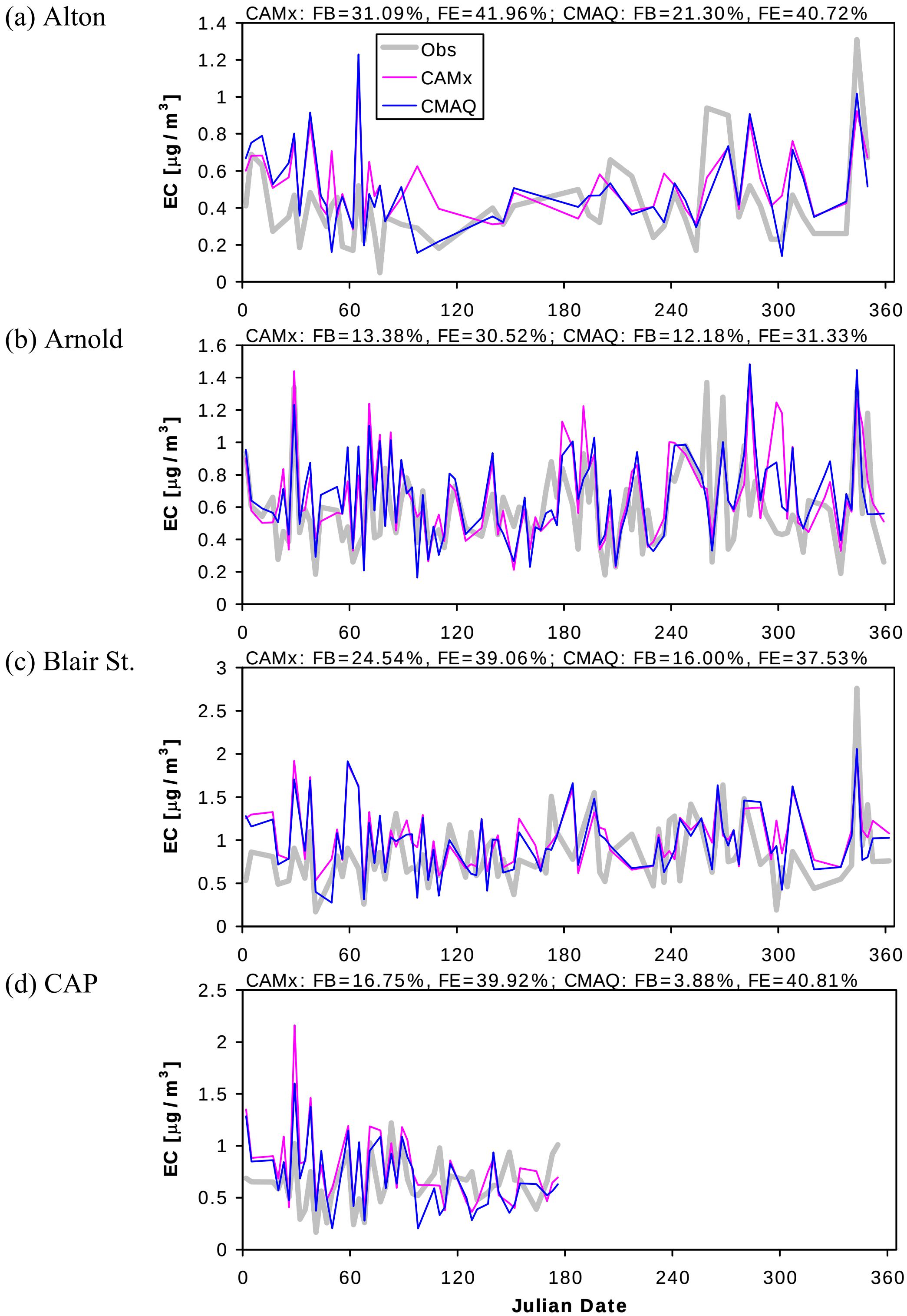


Figure 4-21. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of EC.

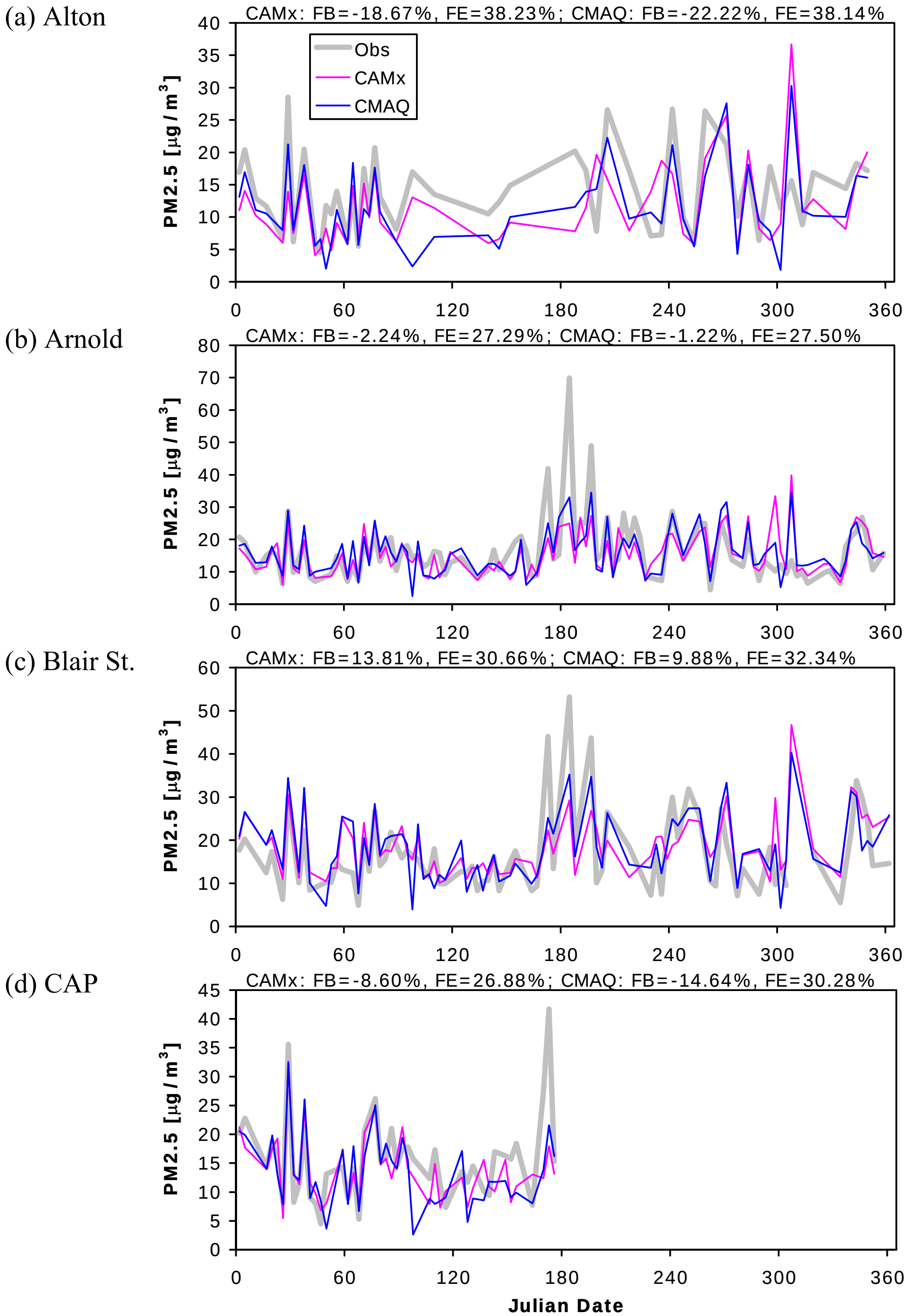


Figure 4-22. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations of $PM_{2.5}$.

4.6 ST. LOUIS SUPER SITE MODEL EVALUATION

The St. Louis – Midwest Super Site performed an extensive set of measurements for fine PM and its components from April 2001 through June 2003. Its core site is located in East St. Louis, IL, which is about 3 km east of the City of St. Louis (Figure 4-16). Unlike the STN's 1-in-3 days sampling schedule, the Super Site collected 24-hour samples everyday and has a richer model evaluation database. Dr. Jay Turner provided St. Louis Super Site "best estimate" observation database. Figure 4-23 shows annual time-series plots of the Super Site observations and model predictions by CAMx and CMAQ for PM species concentrations.

CAMx SO_4 performance is quite good with the annual mean fractional bias of 1.5% and fractional error of 37%. CMAQ slightly underpredicts SO_4 concentrations. The observation data has a spike of high SO4 concentration on Julian day 185 (July 4) due to Independence Day fireworks, which was not included in emissions inventory, thus not captured by the models.

Both models show poor NO_3 performance with underprediction bias during summer days when NO_3 concentration is very low. CMAQ performed better than CAMx for nitrate.

As discussed in the previous section, the NH_4 model performance is mostly related to the SO_4 and NO_3 performances. At the Super Site, the NH_4 performance is as good as that of SO_4, and both models show slight underprediction bias during summertime.

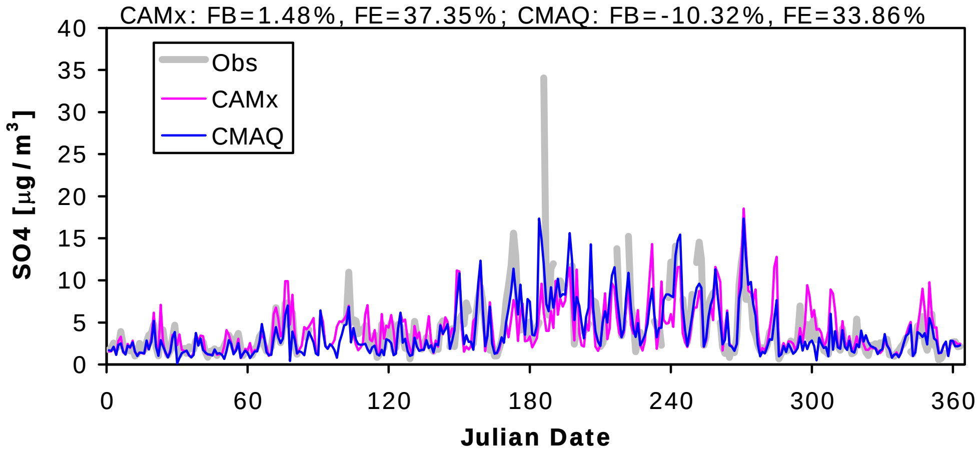
OCM is largely underestimated by both models showing around -90% of annual mean fractional bias. Inclusion of the new biogenic SOA pathway doesn't seem to help much the OCM model performance as the Super Site is located in an urban area. The OCM underprediction bias seems to persist throughout the year suggesting that it may be related to primary OCM emissions underrepresented in the inventory rather than missing SOA processes that would exhibit a summer peak.

Modeled EC is slightly overpredicted for the first quarter but overall has a good agreement with observation.

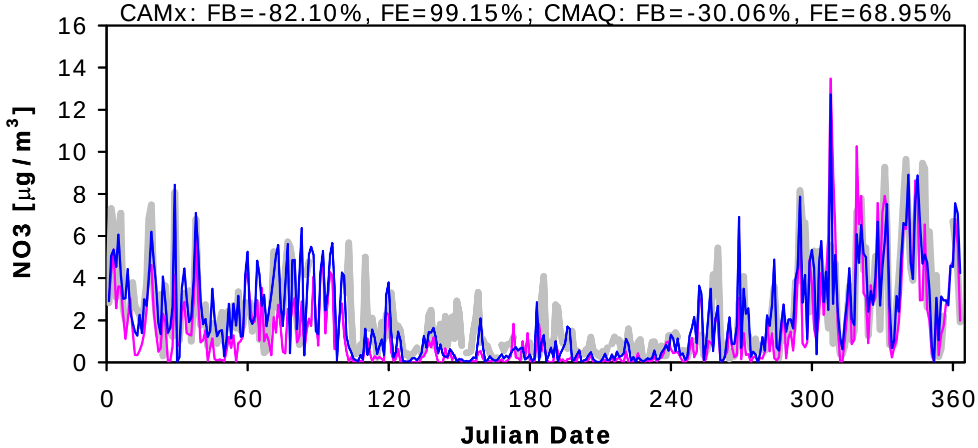
Total $PM_{2.5}$ mass concentration is similarly underpredicted by both CAMx and CMAQ with annual mean fractional bias around -30%. Both OCM and SO_4 are major components to $PM_{2.5}$ mass at the Super Site, which drives the underprediction tendency. Again, impact of the July 4^{th} firework emissions is not captured by the models, which also contributes to the summer underprediction bias.

In general, both models do adequate job of simulating $PM_{2.5}$ components except for NO_3 and OCM. The summer NO_3 underprediction is not likely affect much to the overall model performance because it occurs when NO_3 concentration is low while the OCM underestimation is more of a concern because it contributes significant mass to total $PM_{2.5}$.

(a) Fine Sulfate (SO_4)



(b) Fine Nitrate (NO_3)



(c) Fine Ammonium (NH_4)

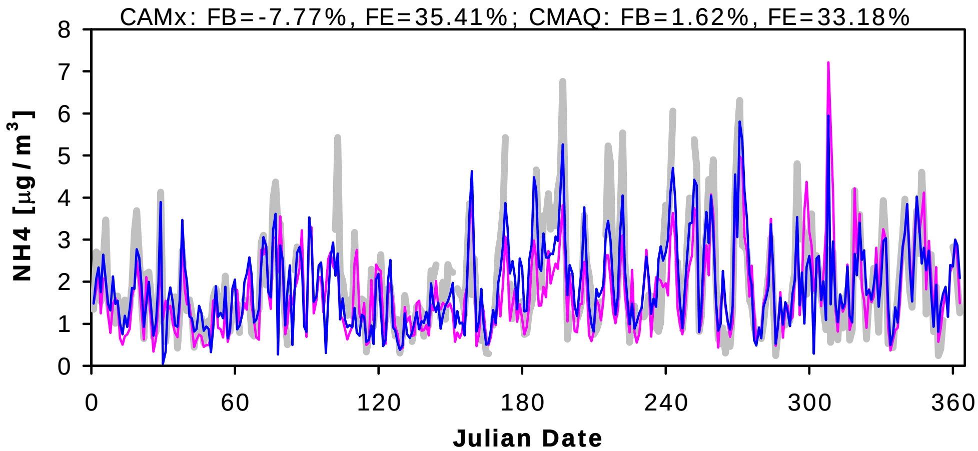
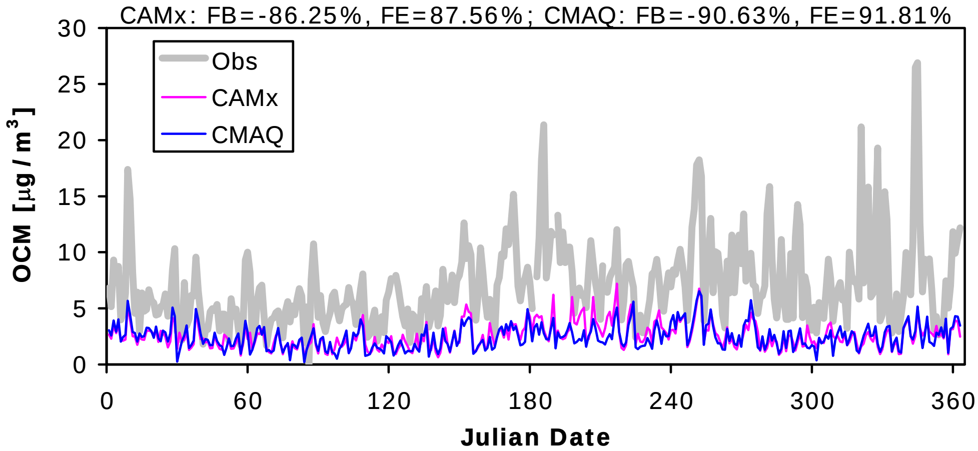
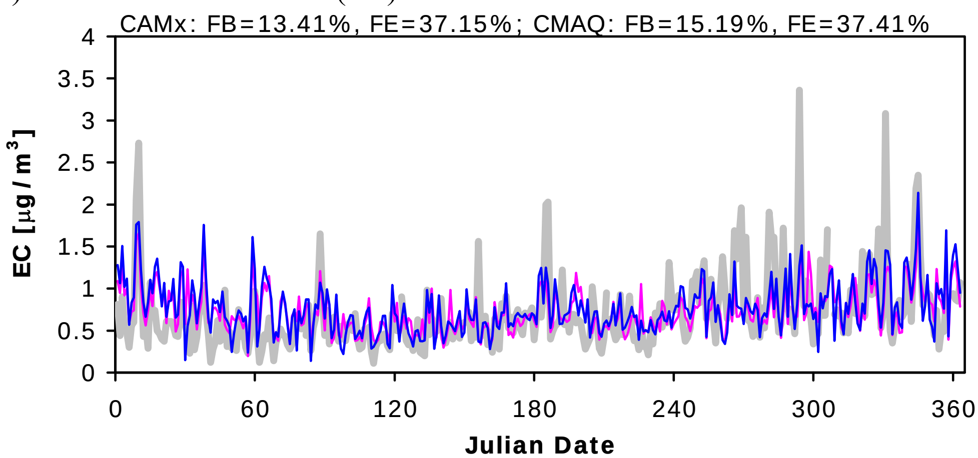


Figure 4-23. Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations at St. Louis Super Site.

(d) Fine Organic Carbon Matter (OCM)



(e) Fine Elemental Carbon (EC)



(f) $PM_{2.5}$

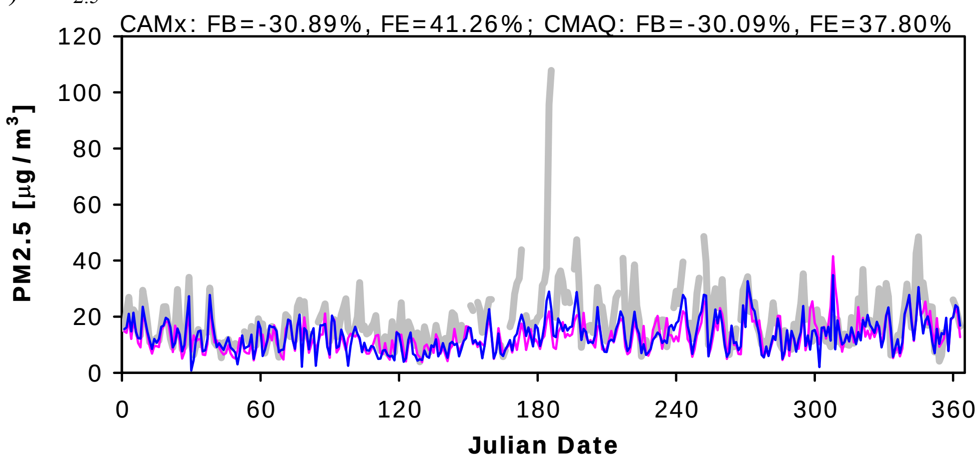


Figure 4-23. (continued). Annual time-series plots for observed (STN) and predicted (CAMx and CMAQ) concentrations at St. Louis Super Site.

4.7 DIAGNOSTIC MODEL PERFORMANCE EVALUATION

As part of diagnostic evaluation, model performance statistics for two gas-phase species, sulfur dioxide (SO_2) and nitric acid (HNO_3) observed by the CASTNet monitoring network, are examined in this section. Figure 4-24 shows the monthly fractional biases and errors for SO_2 at the CASTNet sites. Both models overestimate SO_2 throughout the year. CMAQ shows fractional bias less than 50% all months while the CAMx fractional bias ranges from 42% to 89%.

The monthly fractional biases and errors for HNO_3 and total (gas+particulate) NO_3 are shown in Figure 4-25. The performance statistics for total NO_3 are somewhat better than those for NO_3, which suggests part of the particulate NO_3 performance problem is related to incorrect partitioning between nitric acid and nitrate. The partitioning would depend on the availability of ammonia, thus this may be point to possible problem in the ammonia emissions inventory. The partitioning is also dependent on temperature and relative humidity, which suggests meteorological representation by the MM5 model may also play a part in the poor NO_3 performance. However, during the course of the St. Louis $PM_{2.5}$ Study, the HNO_3/NO_3 partitioning was recalculated using observed hourly meteorology which did not significantly change the NO_3 performance.

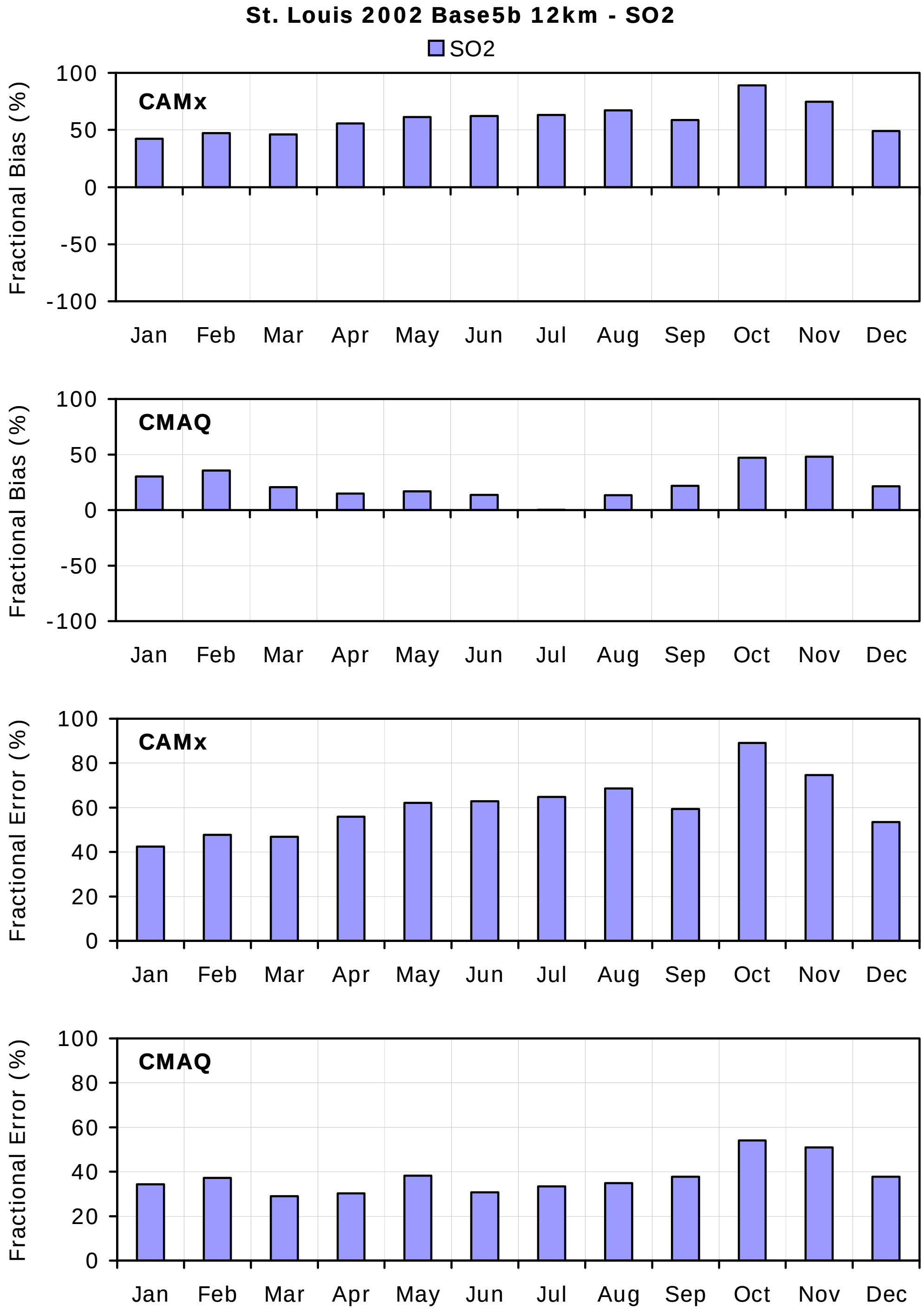


Figure 4-24. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted SO_2 against measurement data at CASTNet monitoring sites in the St. Louis 12 km domain.

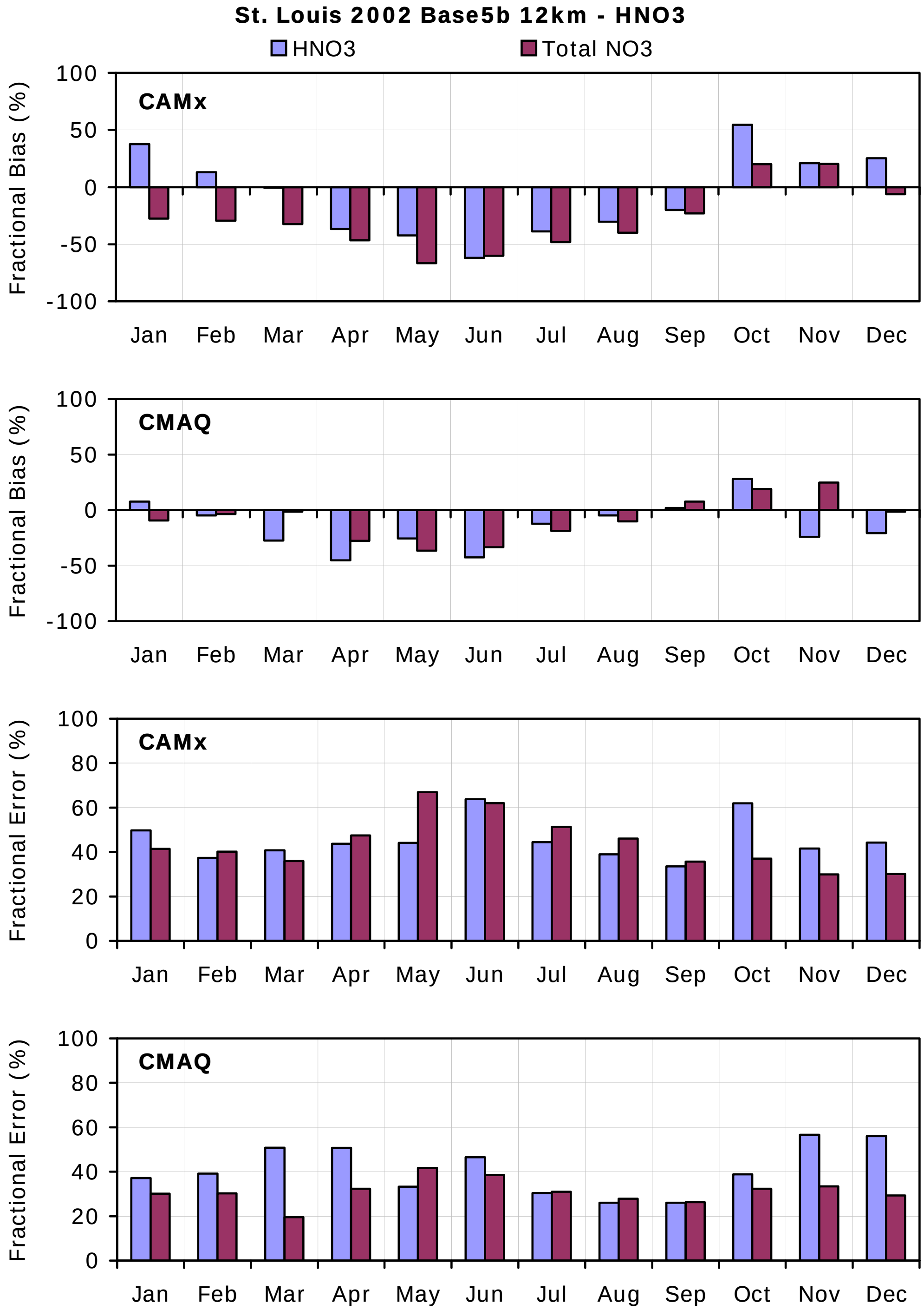


Figure 4-25. Monthly mean fractional biases and fractional errors for CAMx- and CMAQ-predicted HNO_3 and total (gas+particulate) NO_3 against measurement data at CASTNet monitoring sites in the St. Louis 12 km domain.

4.8 MODEL PERRFORMANCE EVALUATION CONCLUSIONS

The quarterly average modeled $PM_{2.5}$ component species results near the FRM monitors are used in the Speciated Model Attainment Test (SMAT) to project the current year $PM_{2.5}$ Design Values to the future year for comparison with the $PM_{2.5}$ NAAQS. SMAT uses the relative changes between the base and future year modeling results to scale each of the $PM_{2.5}$ components of the current year $PM_{2.5}$ Design Values. These model derived scaling factors are referred to as Relative Response Factors (RRF). In Chapter 5 of this TSD, we present the results of the SMAT $PM_{2.5}$ Design Value projections for FRM monitoring sites in and near the St. Louis nonattainment area. These results indicate that most of the reductions in the $PM_{2.5}$ Design Values between the base and future year are due to reductions in sulfate concentrations. Thus, performance of the model for sulfate is of most importance in the model performance evaluation.

Below we discuss the implications of the major findings in the St. Louis PM model performance evaluation in the context of the modeling results are used to project future year $PM_{2.5}$ attainment through SMAT.

- Nitrate Underprediction Bias: NO_3 is routinely underpredicted during the summer and adjacent months throughout the St. Louis 12 km modeling region. This underprediction is due to modeled NO_3 concentrations near zero, when observed values are low, but above zero (typically < 1 $\mu g/m^3$). However, NO_3 is generally a very minor to insignificant contributor to total $PM_{2.5}$ mass at FRM monitors in the St. Louis NAA, especially in the summer when the underprediction bias is largest. Thus, the NO_3 performance issues are not a big concern in the $PM_{2.5}$ projections.

- OCM Underpredictions Bias: The OCM underprediction bias is a cause for concern since it is a major component of the $PM_{2.5}$ mass at St. Louis FRM monitoring sites with maximum contributions to the 2012 $PM_{2.5}$ Design Values of ~8 $\mu g/m^3$, minimum values of ~3 $\mu g/m^3$ and a mean value of ~4 $\mu g/m^3$. The reasons for the underestimation of OCM are unclear, but the fact that the underpredictions are higher in the urban than rural areas and persists year-round suggest that there may be missing anthropogenic emission sources, or possibly the urban OCM emissions are over diluted across the 12 km grid resolution used in the St. Louis modeling. This conceptual model (Chapter 2) notes that several monitoring sites are influenced by local sources that would be poorly represented using a 12 km grid resolution. The changes in projected OCM concentrations between the current and projected $PM_{2.5}$ Design Values are mostly less than 2% (i.e., $0.98 < RRF_{OCM} < 1.02$). Thus, the changes in OCM between the current and future year are having a minor influence on the projected $PM_{2.5}$ Design Values in the regional modeling.

- EC Performance Issues: For the most part, both models performed well for EC at the urban sites and the slight overprediction would not affect the relative changes in the model response to anthropogenic EC emissions changes. Therefore, any EC performance issues are not a cause for concern, although the model performance for EC was generally good.

- Sulfate Underprediction Bias: Although SO_4 is performing well, it does have an underprediction bias that is largest in the summer months. But this underprediction is not severe and the model appears to be capturing the temporal variations in the observed sulfate well and is responding to the SO_2 emission reductions between 2002 and

2009/2012 in a manner as expected. Thus the model performance indicates that the modeled relative changes in SO_4 concentrations are likely a valid response.

- SOIL Performance Issues: The model performance for the SOIL species is quite poor. This SOIL component of the 2012 projected $PM_{2.5}$ Design Value ranges from 0.7 to 1.1 $\mu g/m^3$. The RRFs for SOIL indicate that it is mostly increasing, with summer (Q3) SOIL RRFs typically ranging from 1.1 to 1.2, which is relatively insignificant compared to SO_4 contributions. Therefore, the SOIL performance issues will not significantly affect the projected 2009/2012 $PM_{2.5}$ Design Values.

SO_4 reductions dominate the changes in $PM_{2.5}$ Design Values between 2002 and 2009/2012. SO_4 performance is good in the 2002 Base 5b simulation by both models almost always achieving the PM performance goal at urban sites and achieving the PM performance criteria for rural sites. These factors provide confidence in the Design Value projections using the Base5b modeling results.