

04-17-18 Preliminary Draft Comments from Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel. These preliminary pre-meeting comments are from individual members of the Panel and do not represent CASAC consensus comments nor EPA policy. Do not cite or quote.

## **Risk and Exposure Assessment for the Review of the Primary National Ambient Air Quality Standard for Sulfur Oxides (External Review Draft - August 2017)**

### **Evaluation of SO<sub>2</sub> Model Performance at Fall River, Indianapolis, and Tulsa Prepared by: James Boylan**

The first step in EPA's Risk and Exposure Assessment was to model SO<sub>2</sub> concentrations in three study areas (Fall River, Indianapolis, and Tulsa) with AERMOD. AERMOD provides 1-hour SO<sub>2</sub> concentrations that vary spatially and temporally. These modeling results are used to estimate 5-minute SO<sub>2</sub> exposures across the study areas and the risk assessment is built upon these results. If the AERMOD model performance is significantly biased high or low, the resulting exposure assessment will be significantly biased resulting in large uncertainties in the risk assessment. Therefore, it is critical that the model has "acceptable" model performance. If the model does not have acceptable model performance, the impacts of the model biases needs to be discussed and accounted for in the REA.

The final Integrated Science Assessment (ISA) states, "For models intended for application to compliance assessments (e.g., related to the 1-h daily max SO<sub>2</sub> standard), the model's ability to capture the high end of the concentration distribution is important. Measures such as robust highest concentration (RHC) (Cox and Tikvart, 1990), and exploratory examinations of quantile-quantile plots (Chambers et al., 1983) are useful. The RHC represents a smoothed estimate of the top values in the distribution of hourly concentrations. **In contrast, for dispersion modeling in support of health studies where the model must capture concentrations at specified locations and time periods, additional measures of bias and scatter are important.**"

All three of the model evaluation methods used in Appendix D are associated with using the model for regulatory compliance assessments. For example, the model's ability to capture the high end of the concentration distribution is evaluated with QQ-plots where the highest data point from the model is compared to the highest data point from the observations even if they occur at a different time-of-day, day-of-week, or season-of-year. In the REA, the model is being used to support health studies where spatial and temporal accuracy is much more important compared with regulatory compliance assessments. Since the Air Pollutants Exposure (APEX) model uses the model results paired in time and space, the model results need to be evaluated against observations paired in time and space. Appendix D does include absolute fractional bias (AFB) paired in space and presents QQ-plots paired in space (Figures 1-3). However, there is no analysis or discussion on the model performance paired in space and time.

During our September 18-19, 2017 meeting, the CASAC panel requested from EPA the 1-hour SO<sub>2</sub> concentrations for 2011, 2012 and 2013 from each of the seven monitors included in Appendix D and the modeled 1-hour SO<sub>2</sub> concentrations from each receptor at the location of the monitor (paired in time). On October 19, 2017 a memo from Erika Sasser (EPA) titled "CASAC Request for data associated with Appendix D of the document titled Risk and Exposure Assessment for the Review of the Primary National Ambient Air Quality Standard for Sulfur Oxides" was sent in response to the request. In this memo, EPA states:

04-17-18 Preliminary Draft Comments from Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel. These preliminary pre-meeting comments are from individual members of the Panel and do not represent CASAC consensus comments nor EPA policy. Do not cite or quote.

The AERMOD output for the three study areas was evaluated in Appendix D, consistent with EPA's model evaluation protocol that focuses on the higher concentrations in the concentration distribution as those represent the concentrations of interest in regulatory applications. These concentrations are also generally most important for the exposure assessment performed in the draft REA. Given the uncertainty in emissions and meteorological inputs (e.g., wind direction) for a given hour, EPA's *Guideline on Air Quality Models* explicitly discourages pairing of model estimates and monitor data on an hour-by-hour basis in evaluating performance of air quality models. Appendix D recognized the uncertainties in available model inputs (i.e., emissions and meteorology) for the three urban study areas, as well as the limited spatial coverage provided by the seven monitors across the three study areas.

Given that the analyses in the draft REA involve a system of models, AERMOD model output should be considered in terms of how it is being used in the Air Pollutants Exposure model (APEX) to generate estimates of exposure as individuals move through time and space. In particular, in evaluating model outputs, it is important to consider whether predicted high values are occurring during times where people may be exercising in microenvironments where ambient SO<sub>2</sub> concentrations contribute to exposure. It is also appropriate to consider broad temporal patterns in hourly concentrations (e.g., seasonal, weekday/weekend, times-of-day), which may be more informative than comparing simultaneous hourly time-series of concentrations (e.g., model-to-monitor comparisons).

The issue in the first paragraph is that EPA continues to rely on a model evaluation protocol that was developed for regulatory applications which have a significantly different focus than health study applications. EPA's *Guideline on Air Quality Models* is appropriate for regulatory applications (PSD permit modeling and SIP attainment demonstration modeling), but should not be applied to health study applications (risk and exposure assessments) since the models are used in different ways. In the second paragraph, EPA acknowledges the importance of timing of the high predicted values and states that it is "appropriate to consider broad temporal patterns in hourly concentrations (e.g., seasonal, weekday/weekend, times-of-day)". I would agree with EPA that broad temporal patterns in hourly concentrations (e.g., time-of-day, day-of-week (weekday/weekend), or season-of-year) are more informative than comparing simultaneous hourly time-series of concentrations because exposures during identical times-of-day, days-of-week (weekday/weekend), and seasons-of-year are treated the same by the APEX model. However, EPA has made no effort to look at whether predicted high values are occurring at the right time-of-day, day-of-week, or season-of-year.

To address this shortcoming, I performed a model performance evaluation comparing 1-hour SO<sub>2</sub> model concentrations against measurements at one monitor in Fall River, MA (250051004); three monitors in Indianapolis, IN (180970057, 180970073, and 180970078); and three monitors in Tulsa, OK (401430175, 401430235, and 401431127). The 1-hour SO<sub>2</sub> model concentrations and measurements were provided by EPA. Since there are very few SO<sub>2</sub> monitors in the study areas, it can be assumed that good performance at the monitoring sites indicates good model performance at the receptors that don't have monitors. Conversely, poor performance at the monitoring sites indicates poor model performance at the receptors that don't have monitors.

To prepare the data for analysis, I removed all data pairs with an invalid measurement value (-999) and/or an invalid modeling value (-999). AERMOD can't calculate a concentration when the wind speed is zero, resulting in a null value (-999). Then, a number of model evaluations were performed to

04-17-18 Preliminary Draft Comments from Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel. These preliminary pre-meeting comments are from individual members of the Panel and do not represent CASAC consensus comments nor EPA policy. Do not cite or quote.

better understand the model performance paired in time and space. After looking at the results, it was determined that the most informative graphical depictions were maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for each monitor site, where the maximum values were obtained for each hour across the three year period that was evaluated (2011-2013). In addition, QQ-plots for each hour-of-day provide additional details for the rest of the distribution. Since time-of-day is an important variable in the APEX model, QQ-plots for each hour-of-day is much more appropriate than QQ-plots that are unpaired in time. Also, the higher concentrations in the distribution are generally more important for the exposure assessment in the REA; therefore, performance of those values should be given more weight than lower concentrations. Since it is not likely that any modeled or measured 1-hour SO<sub>2</sub> concentrations less than 50 ppb will result in a 5-minute SO<sub>2</sub> concentration above 200 ppb, most of the analysis will focus on the modeled and measured 1-hour SO<sub>2</sub> concentrations above 50 ppb. The approach described above is a simple approach and does not account for differences in day-of-week or season-of-year. Therefore, the summary of results presented below can be considered the best case scenario since accounting for day-of-week and season-of-year in addition to time-of-day would show larger biases between measured and monitored SO<sub>2</sub> concentrations. A summary of the MPE results for each monitor location are presented below.

#### Fall River, MA (250051004)

See Figure 4 and Figures 11-16. For Hours 1-6 and 19-24, the 1-hour SO<sub>2</sub> observations are ~4x higher than the modeled values. Smaller biases (both positive and negative) are seen in Hours 7-18, with a few outliers in Hours 9, 10, 11, and 13.

#### Indianapolis, IN (180970057)

See Figure 5 and Figures 17-22. For Hours 1-7 and 18-24, the 1-hour SO<sub>2</sub> modeled values are ~2x higher than the observations. Smaller biases (both positive and negative) are seen in Hours 8-17, with a few outliers in Hours 10, 11, and 12.

#### Indianapolis, IN (180970073)

See Figure 6 and Figures 23-28. For Hours 1-8 and 18-24, all modeled and measured 1-hour SO<sub>2</sub> concentrations are below 52 ppb and the model bias is small. For Hours 9-17, the 1-hour SO<sub>2</sub> observations are ~2-3x higher than the modeled values.

#### Indianapolis, IN (180970078)

See Figure 7 and Figures 29-34. For Hours 1-8 and 16-24, all modeled and measured 1-hour SO<sub>2</sub> concentrations are below 64 ppb and the model bias is small. For Hours 9-15, the 1-hour SO<sub>2</sub> observations are ~2x higher than the modeled values.

#### Tulsa, OK (401430175)

See Figure 8 and Figures 35-40. For Hours 1-9 and 20-24, the model bias is small. For Hours 10-19, the 1-hour SO<sub>2</sub> observations are ~2-3x higher than the modeled values.

#### Tulsa, OK (401430235)

See Figure 9 and Figures 41-46. All modeled and measured 1-hour SO<sub>2</sub> concentrations are below 50 ppb, except for Hours 1 and 24 where the maximum 1-hour SO<sub>2</sub> observations are 64 ppb and 53 ppb. For Hours 1-8 and 17-24, the 1-hour SO<sub>2</sub> modeled values are ~1.5-2x higher than the observations.

04-17-18 Preliminary Draft Comments from Clean Air Scientific Advisory Committee (CASAC) Sulfur Oxides Panel. These preliminary pre-meeting comments are from individual members of the Panel and do not represent CASAC consensus comments nor EPA policy. Do not cite or quote.

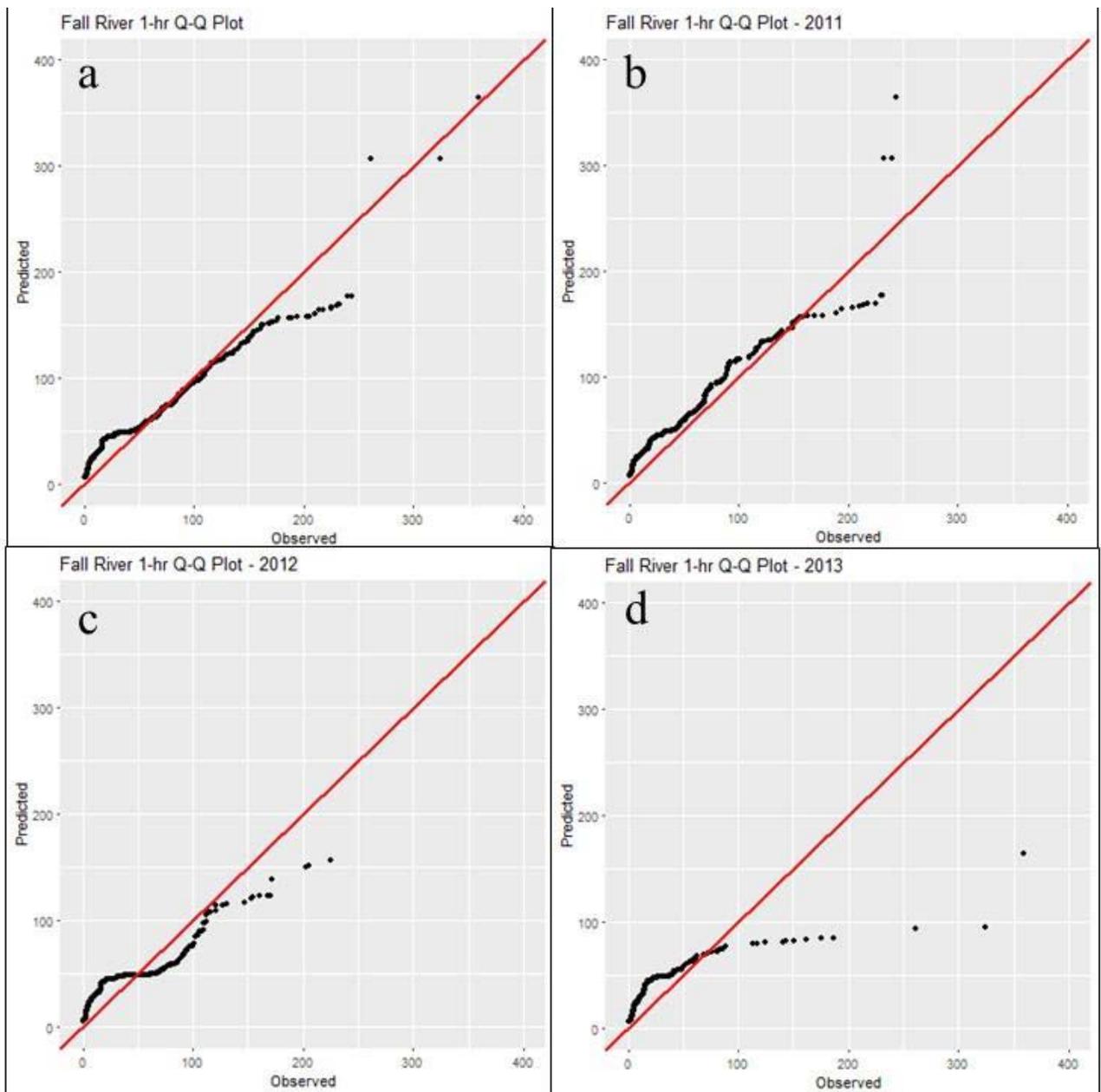
Smaller biases (both positive and negative) are seen in Hours 9-16, with a few outliers in Hours 10 and 11.

#### Tulsa, OK (401431127)

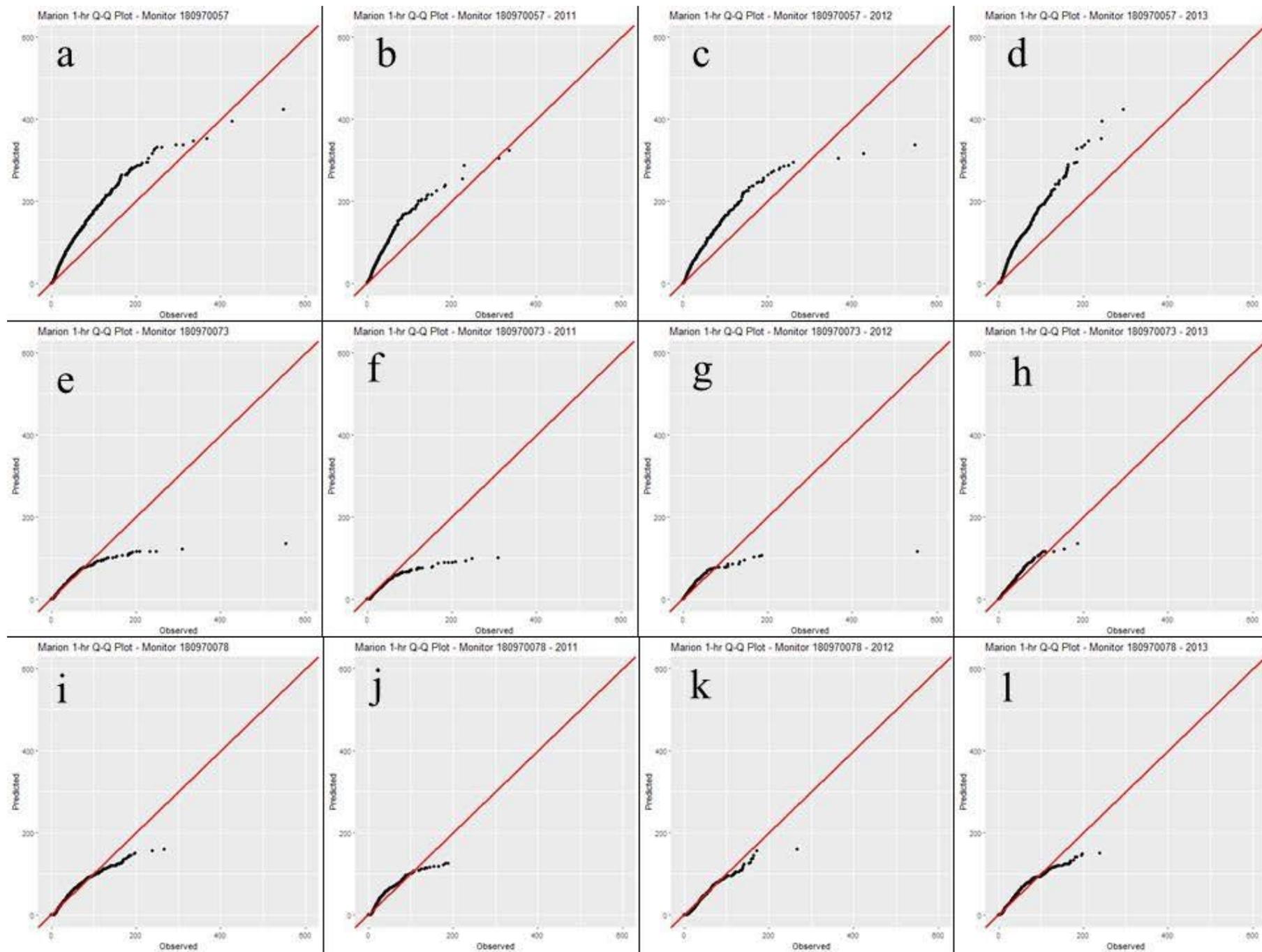
See Figure 10 and Figures 47-52. All modeled and measured 1-hour SO<sub>2</sub> concentrations are below 40 ppb. For Hours 1-8 and 17-24, the 1-hour SO<sub>2</sub> modeled values are ~1.5-2x higher than the observations, except for observations above 20 ppb which match well with the modeled values. Smaller biases (both positive and negative) are seen in Hours 9-16, except for the observations above 20 ppb which the model underestimates by a factor of 1.5-2x.

#### Conclusion

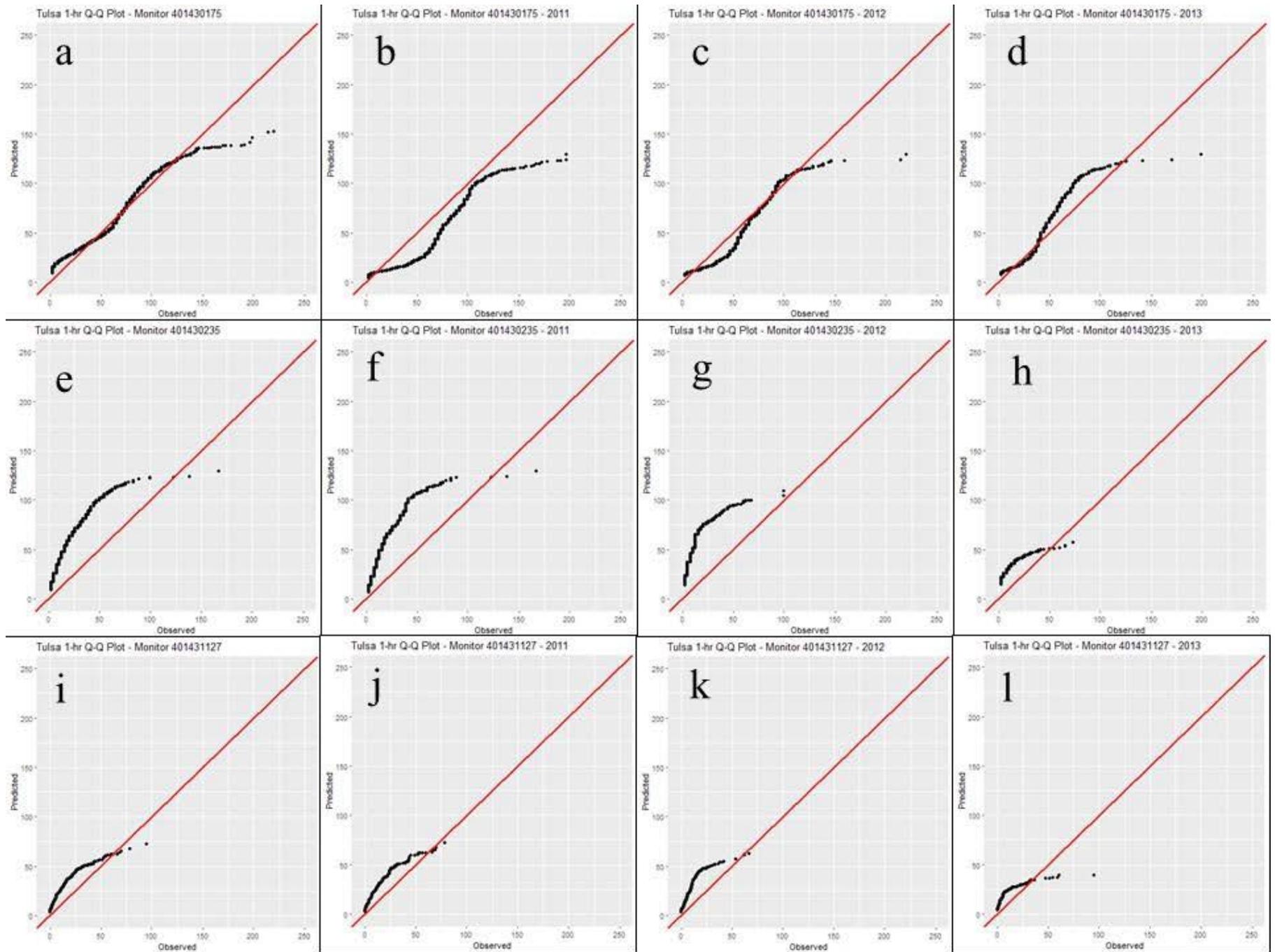
The model performance varies by monitoring site and time-of-day. For Fall River (250051004), the early morning and late evening 1-hour SO<sub>2</sub> observations are ~4x higher than the modeled values. For one Indianapolis monitor (180970057), the early morning and late evening 1-hour SO<sub>2</sub> modeled values are ~2x higher than the observations. For the other two Indianapolis monitors (180970073 and 180970078), the late morning, afternoon, and early evening 1-hour SO<sub>2</sub> observations are ~2-3x higher than the modeled values. For one Tulsa monitor (401430175), the late morning, afternoon, and early evening 1-hour SO<sub>2</sub> observations are ~2-3x higher than the modeled values. For the other two Tulsa monitors (401430235 and 401431127), early morning and late evening 1-hour SO<sub>2</sub> modeled values are ~1.5-2x higher than the observations. These model biases will have a direct impact on the APEX results, possibly calling into question the percent of children and adults experiencing 5-minute exposures at or above 200 ppb. In addition, EPA should examine the day-of-week (weekday/weekend) and season-of-year model performance. Finally, the time-of-day, day-of-week (weekday/weekend), and season-of-year model biases should be discussed and accounted for in the final REA.



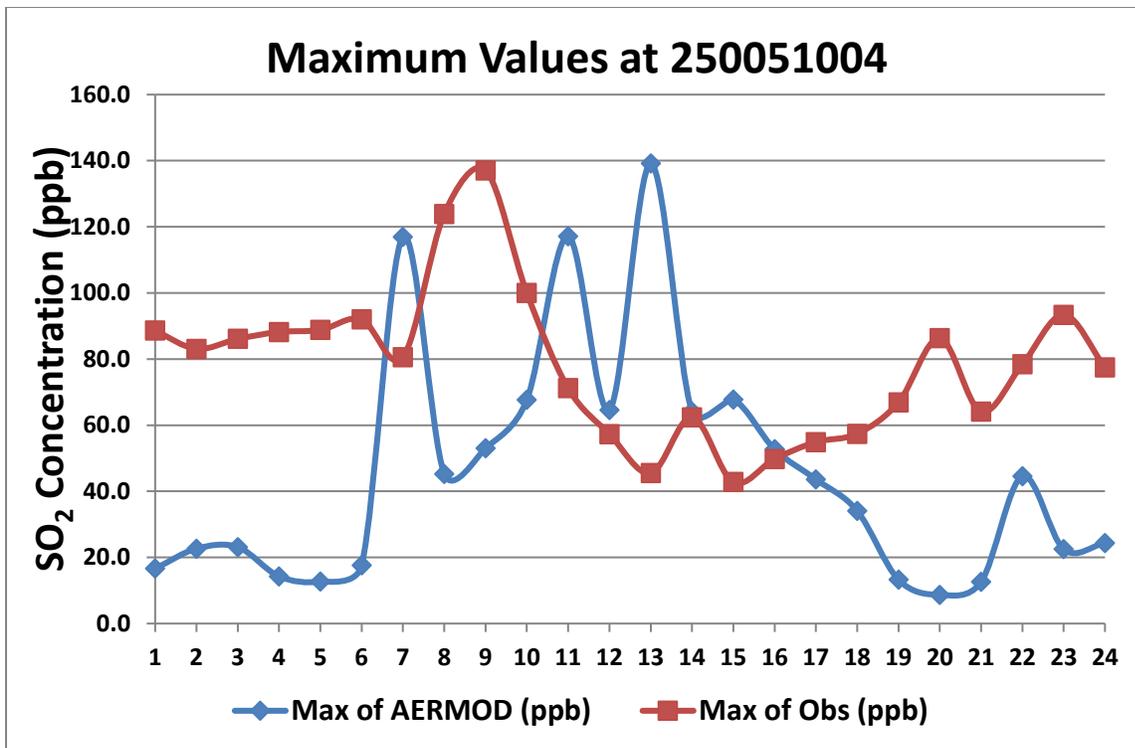
**Figure 1.** 1-hour SO<sub>2</sub> Q-Q plots ( $\mu\text{g}/\text{m}^3$ ) for Fall River, MA (unpaired in time) from EPA's REA (Figure D-1).



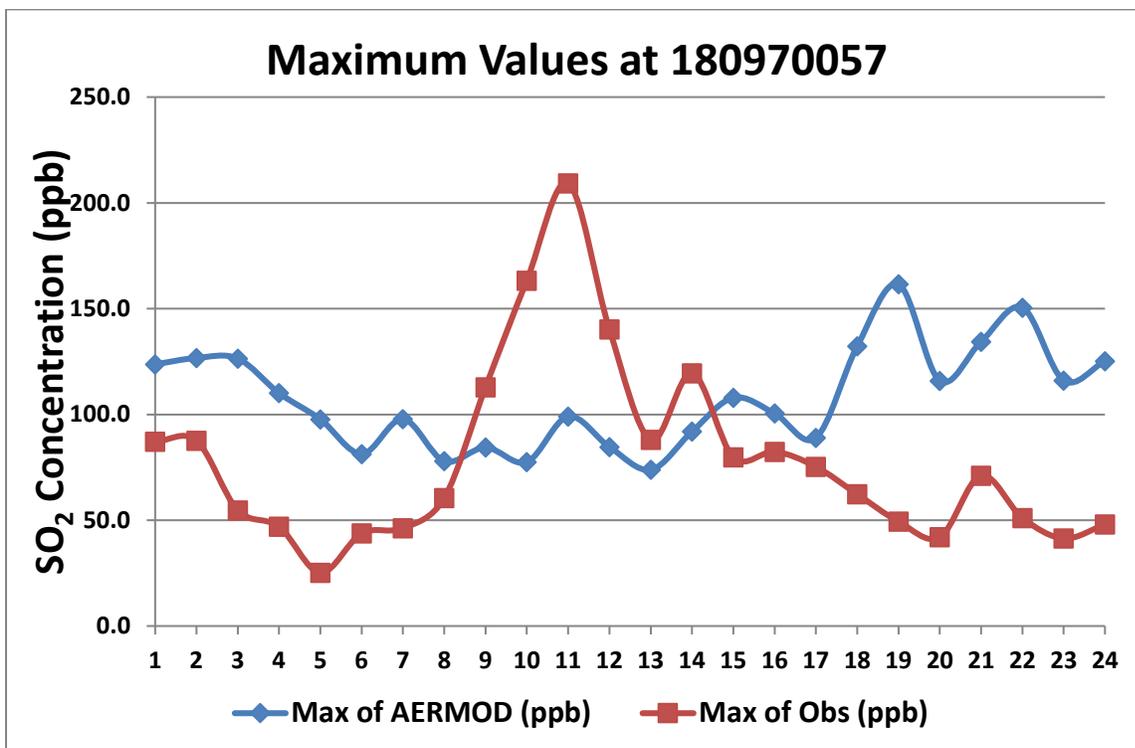
**Figure 2.** 1-hour SO<sub>2</sub> QQ plots ( $\mu\text{g}/\text{m}^3$ ) for individual monitors in Indianapolis, IN (unpaired in time) from EPA's REA (Figure D-7).



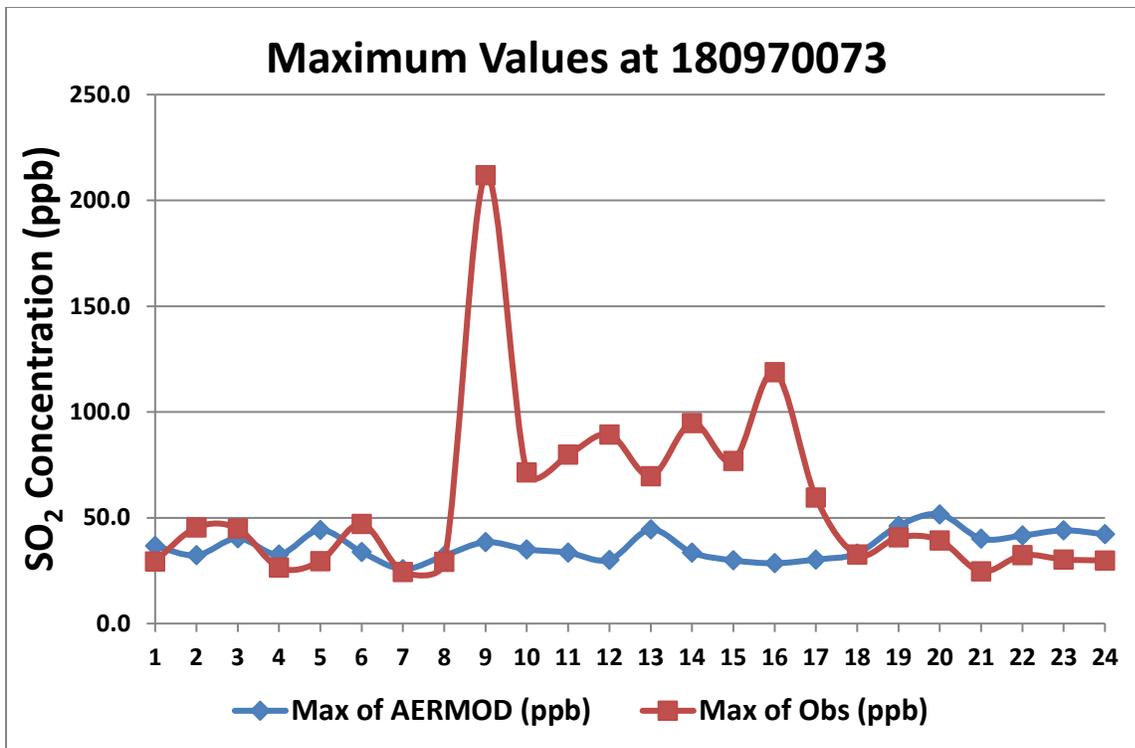
**Figure 3.** 1-hour SO<sub>2</sub> QQ plots ( $\mu\text{g}/\text{m}^3$ ) for individual monitors in Tulsa, OK (unpaired in time) from EPA’s REA (Figure D-13).



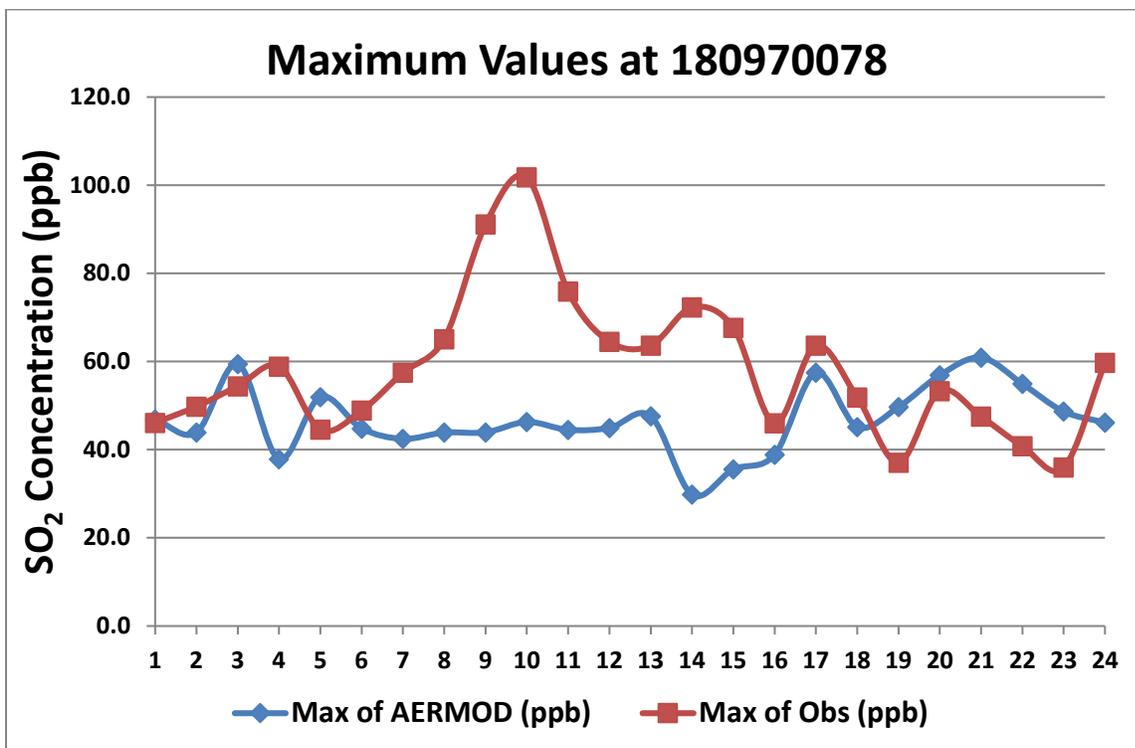
**Figure 4.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).



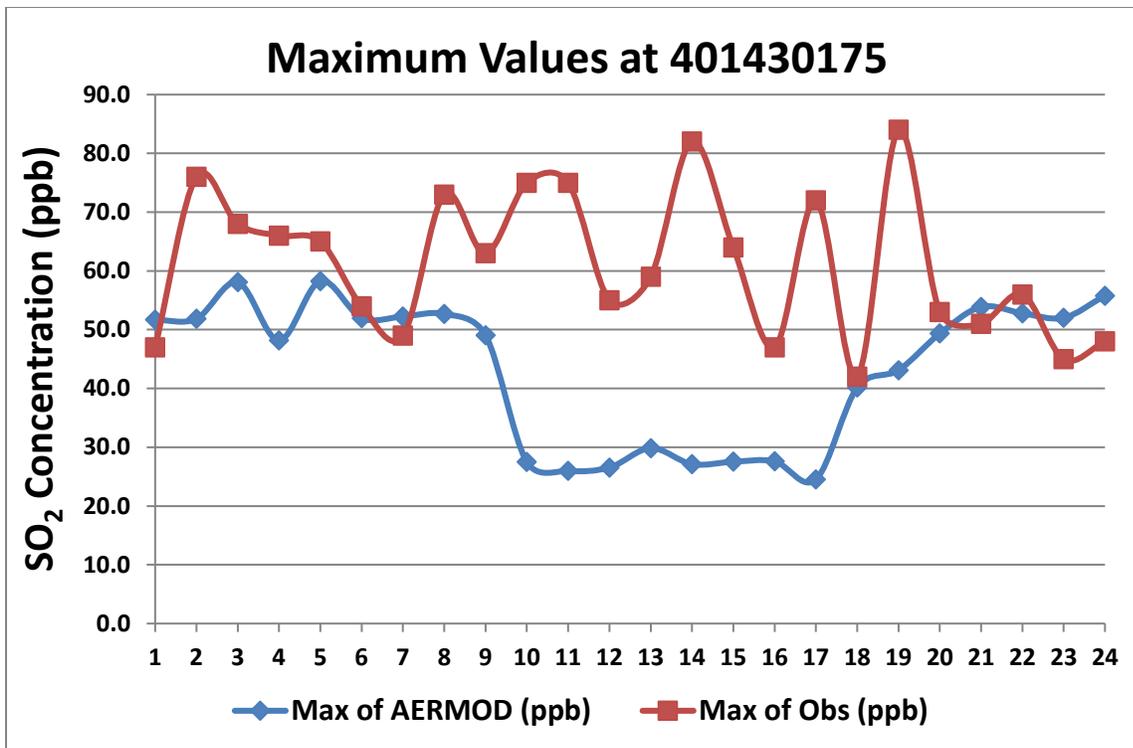
**Figure 5.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).



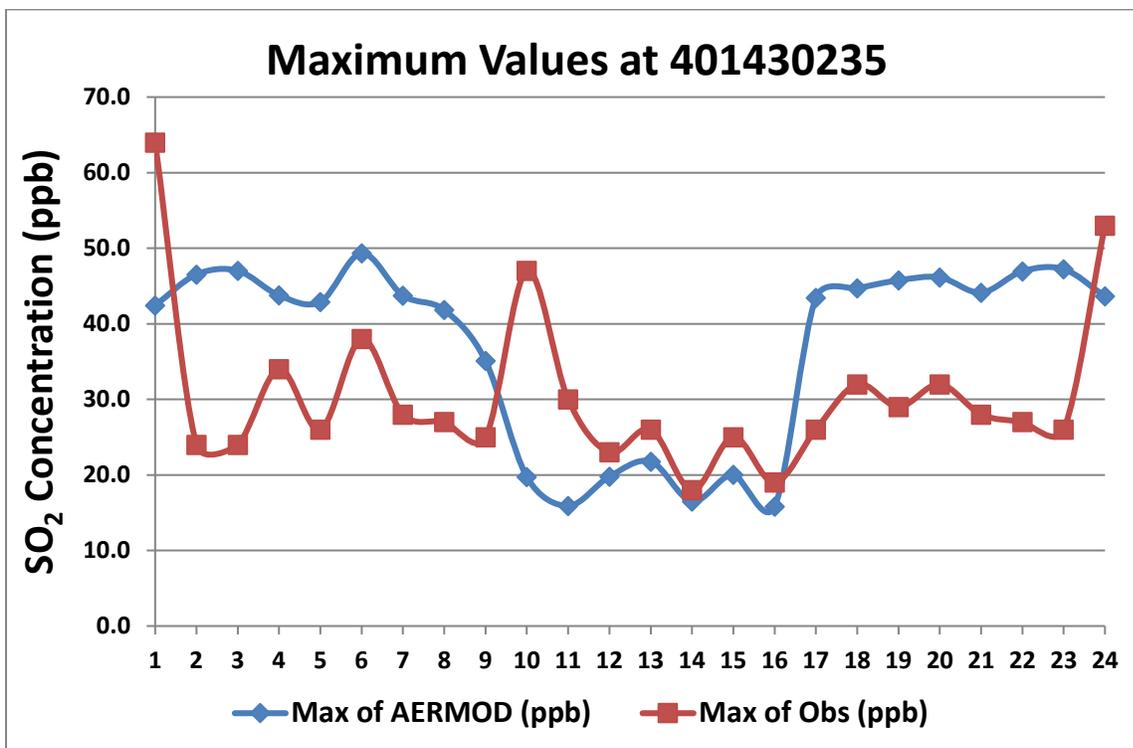
**Figure 6.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).



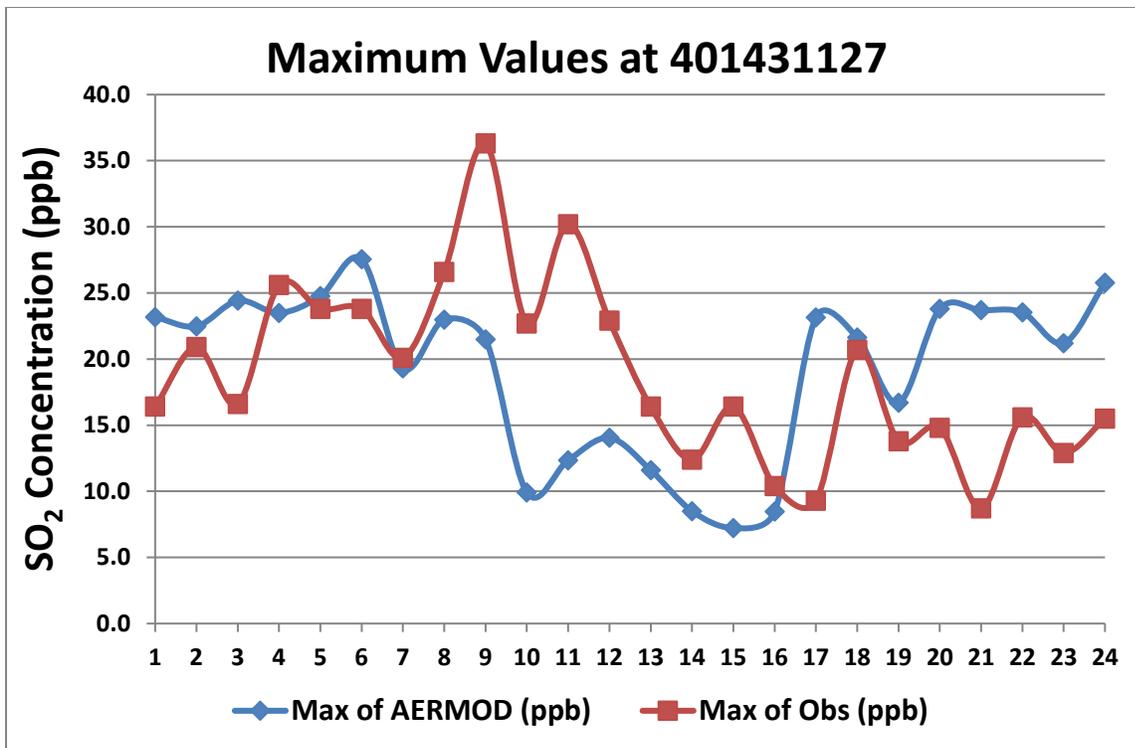
**Figure 7.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 8.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).



**Figure 9.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).



**Figure 10.** Maximum modeled 1-hour SO<sub>2</sub> concentration vs. maximum measured 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).

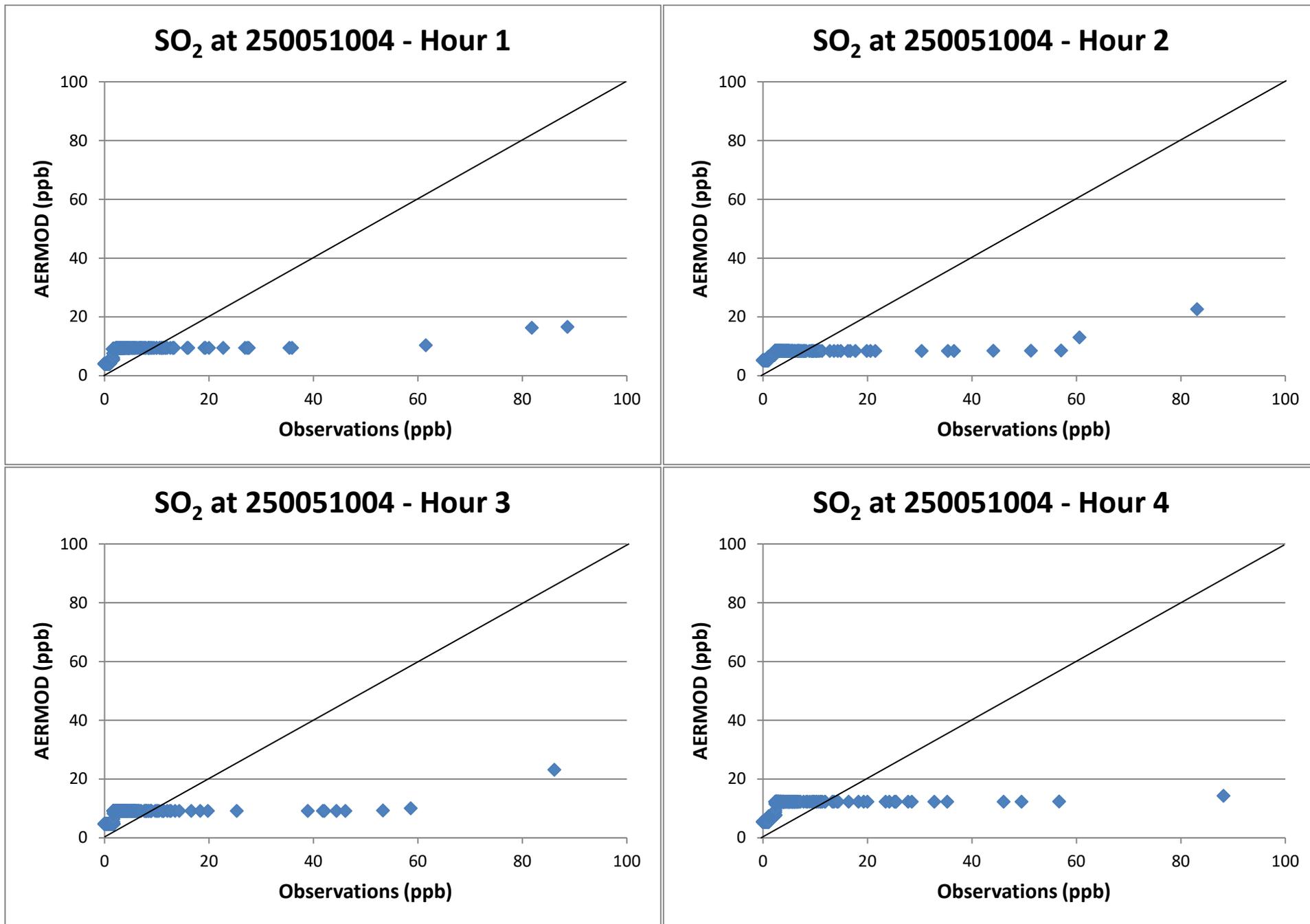
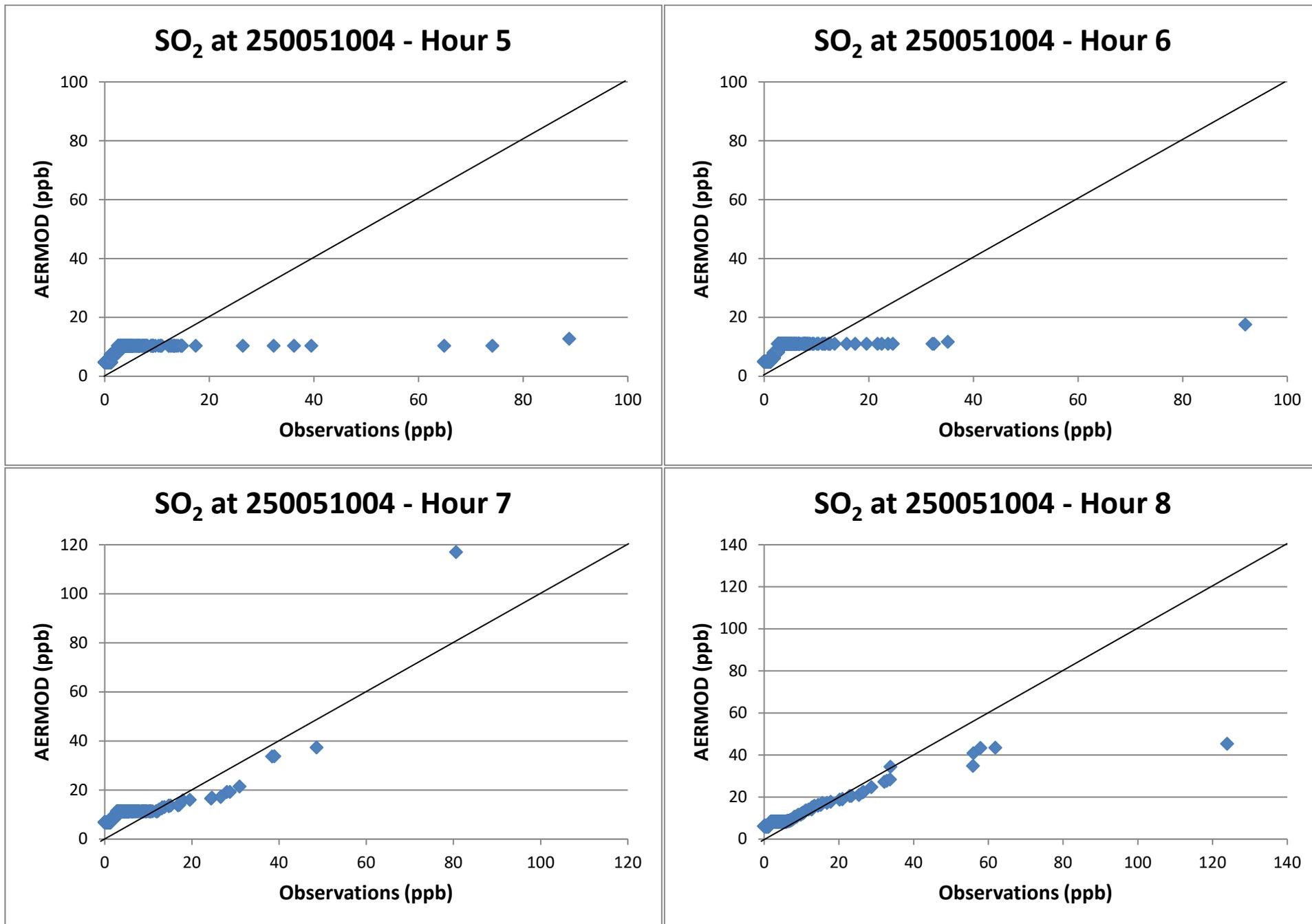
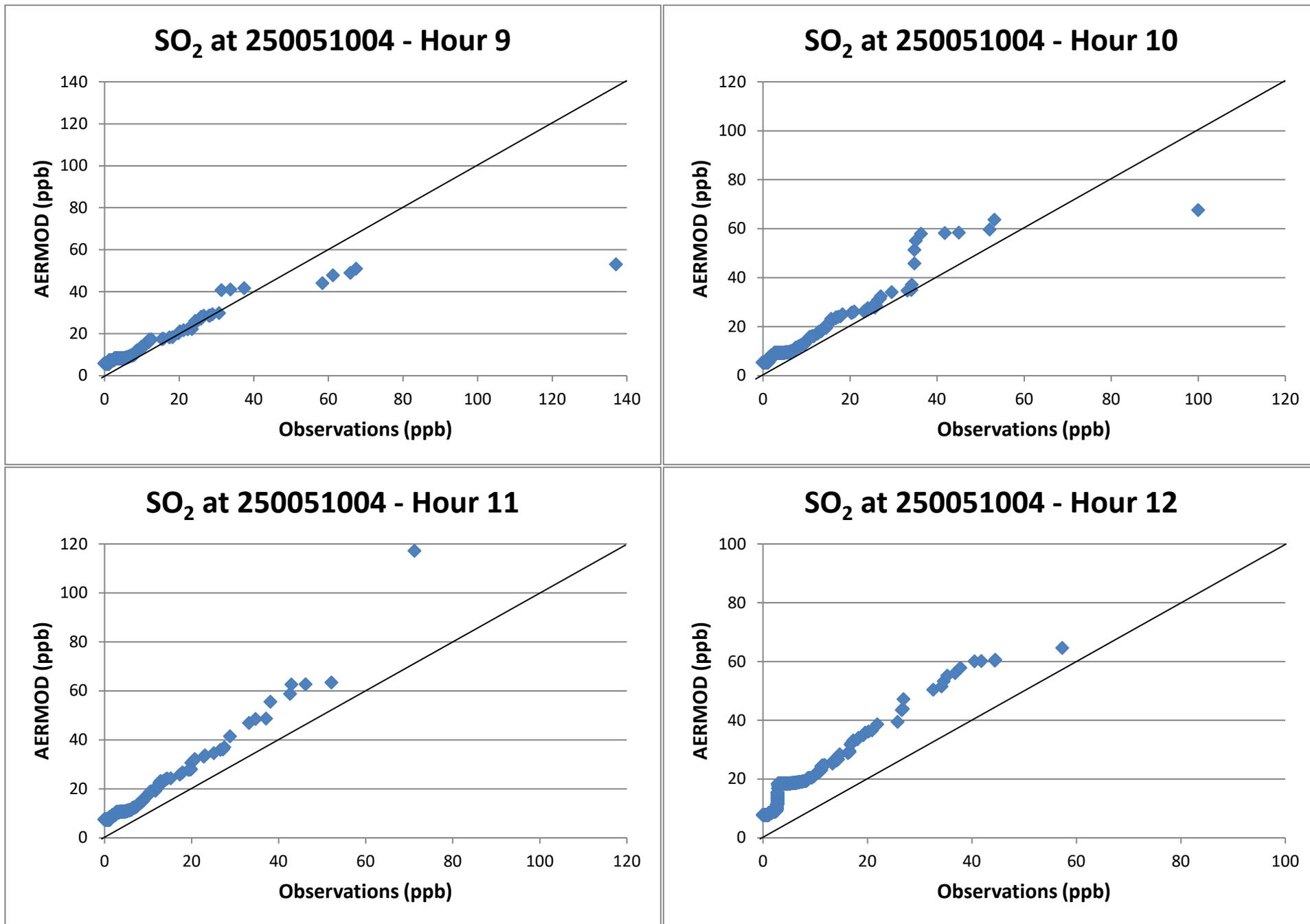


Figure 11. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).



**Figure 12.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).



**Figure 13.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).

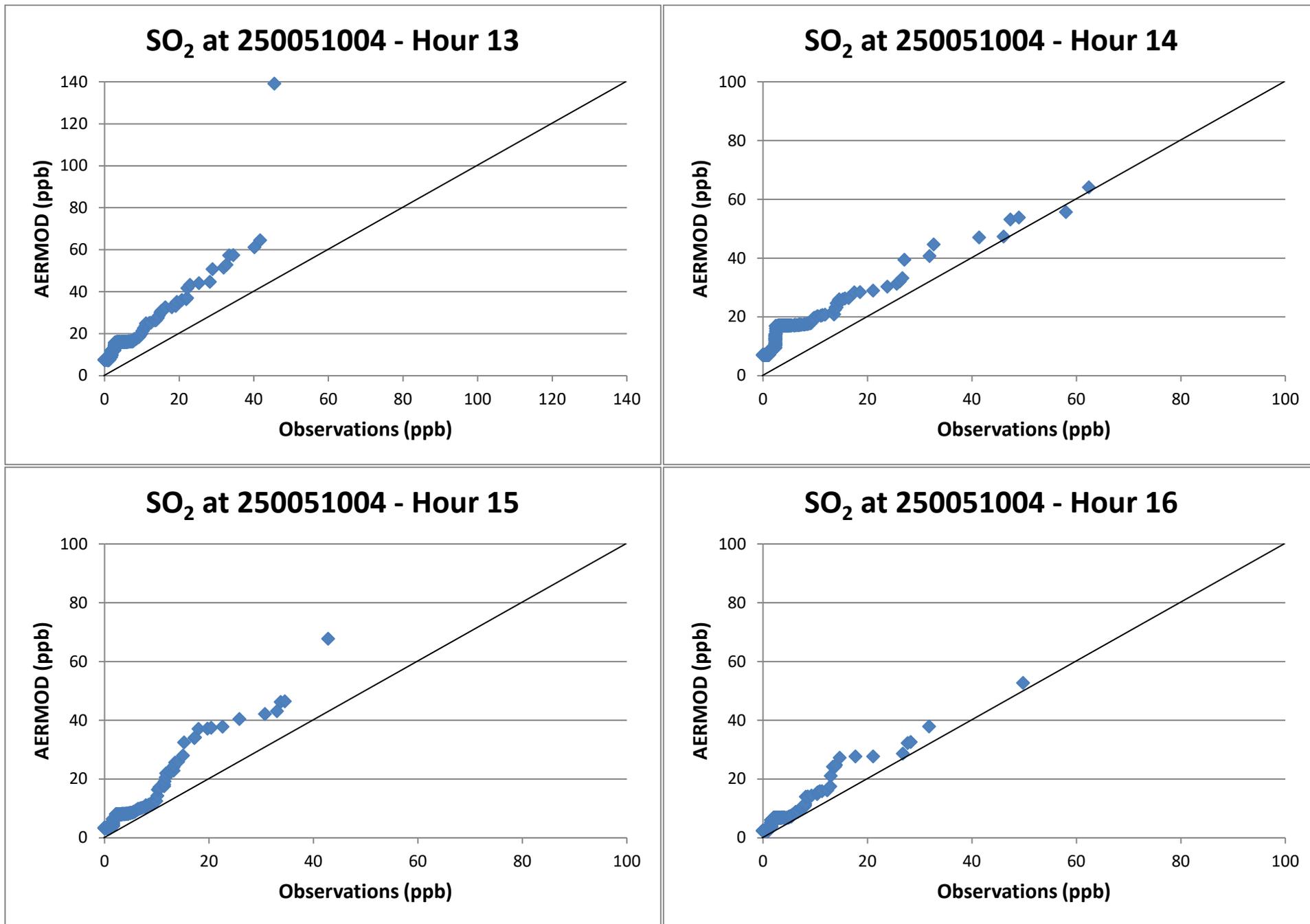


Figure 14. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).

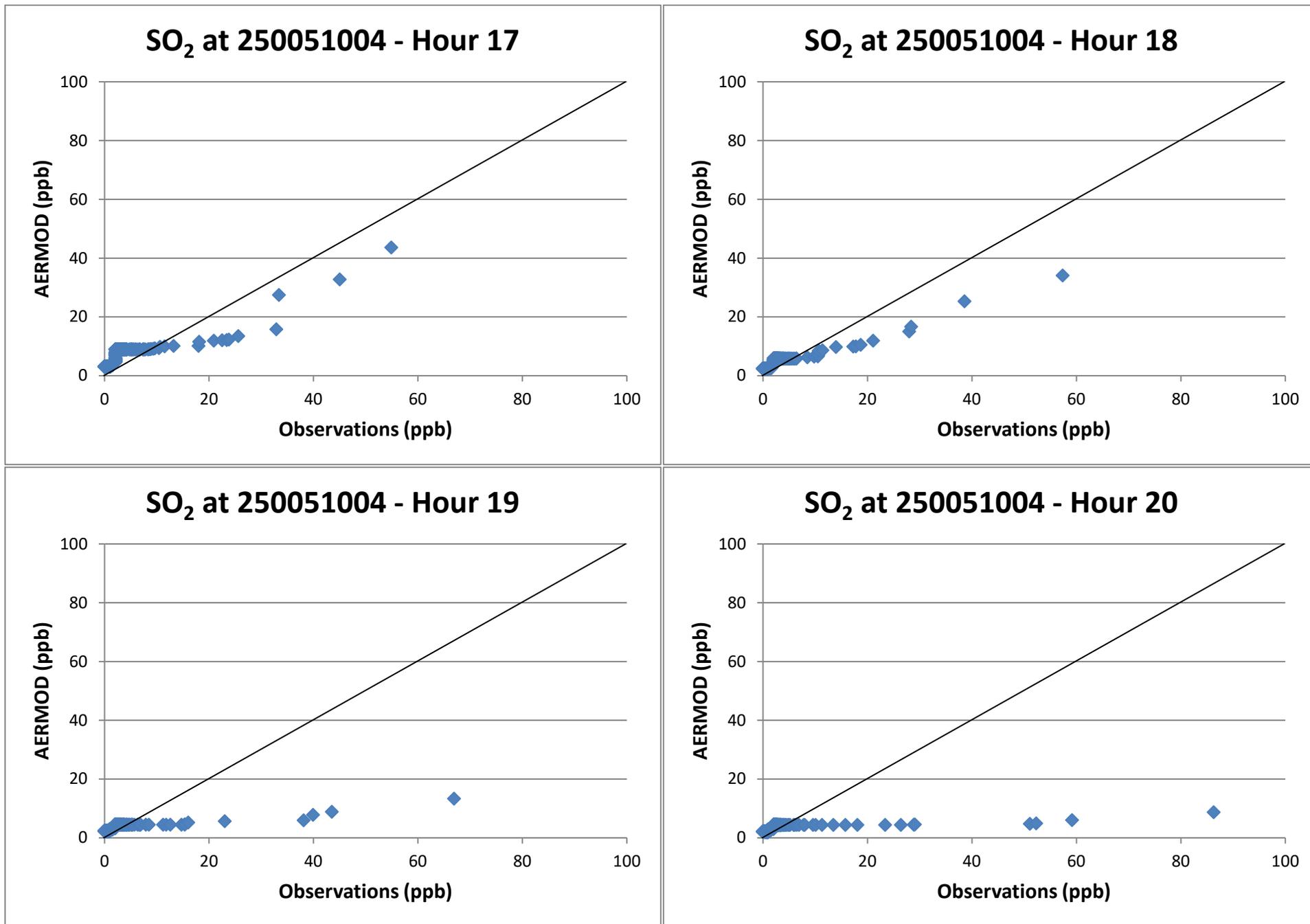


Figure 15. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).

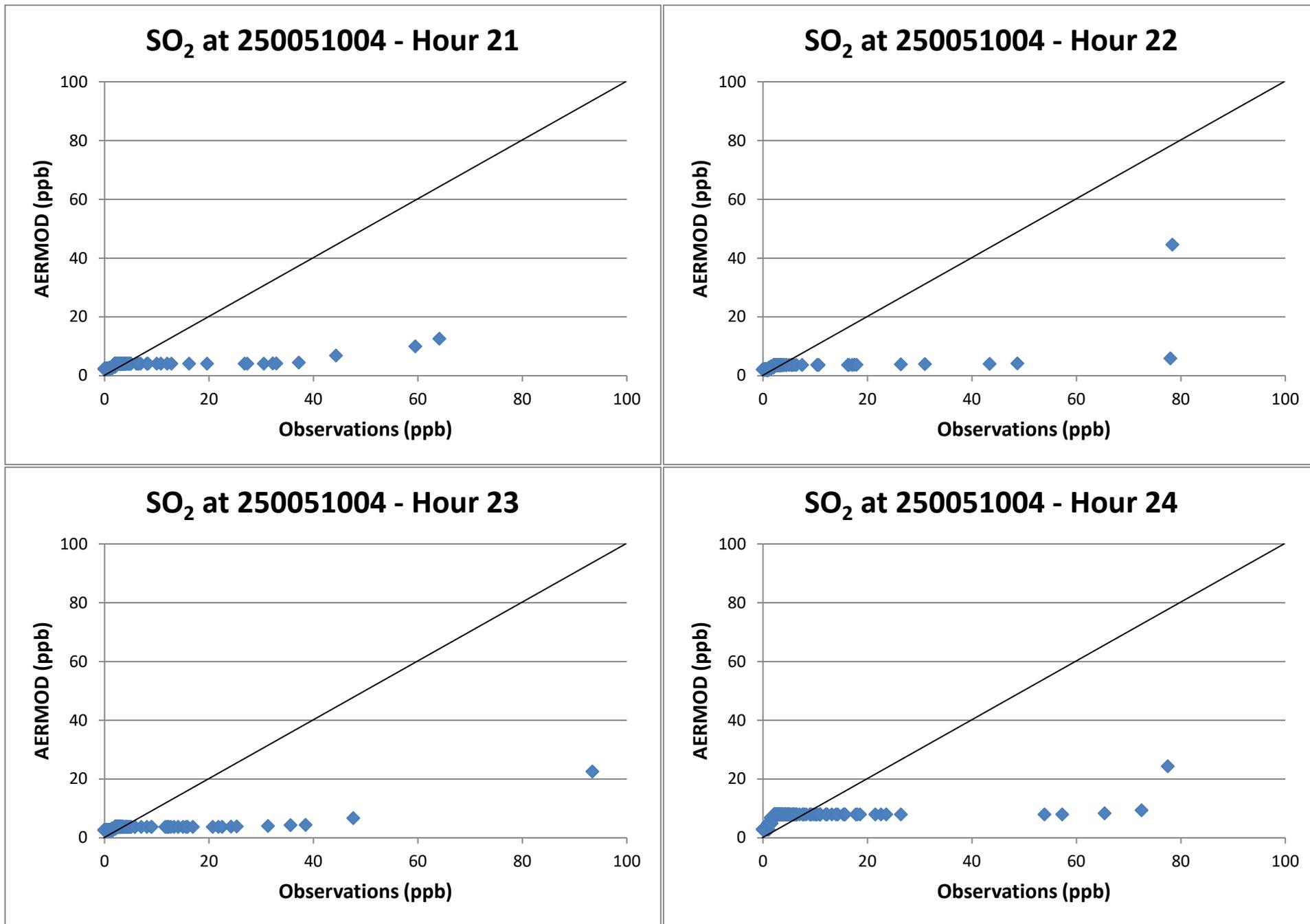


Figure 16. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Fall River, MA (250051004).

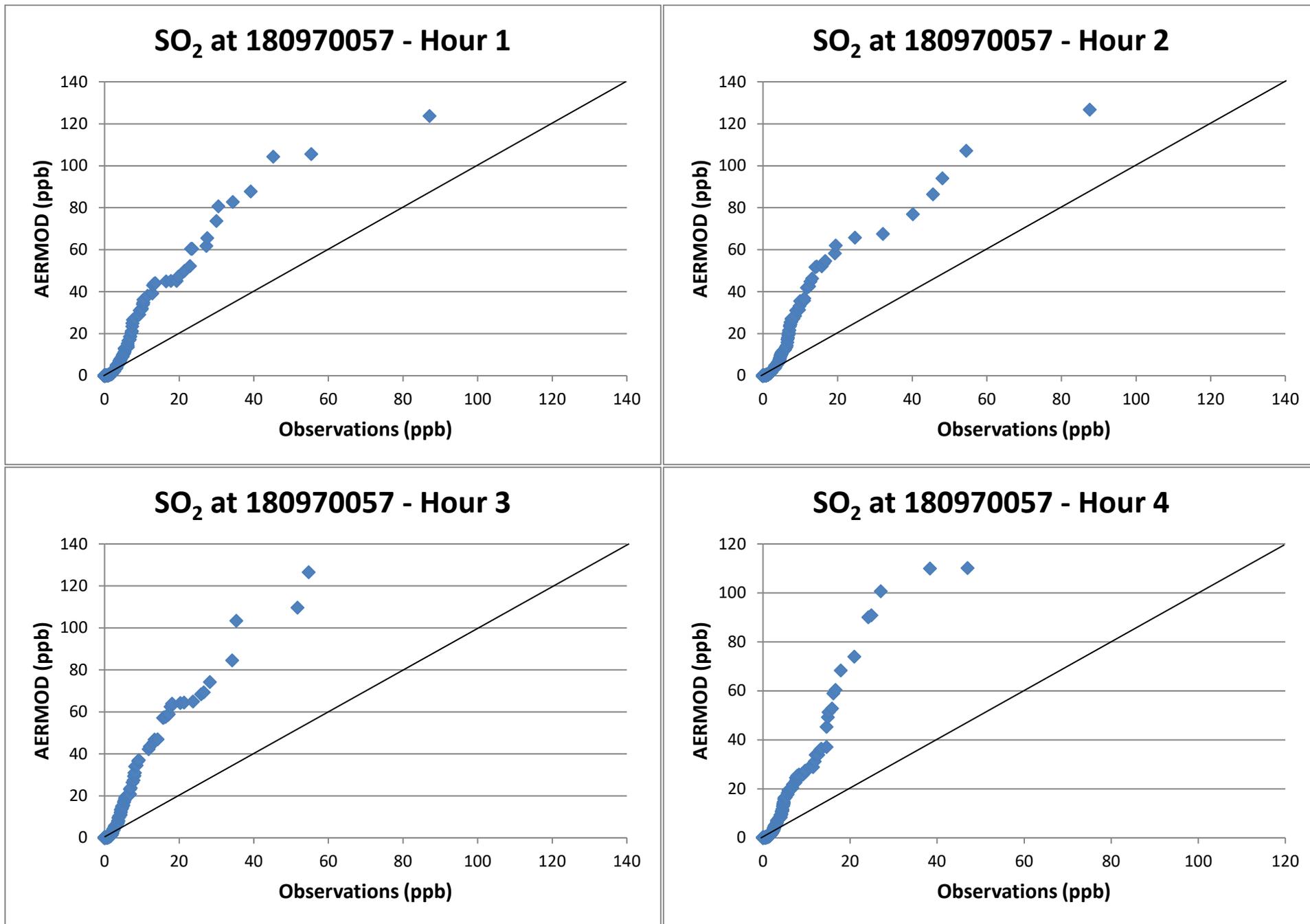
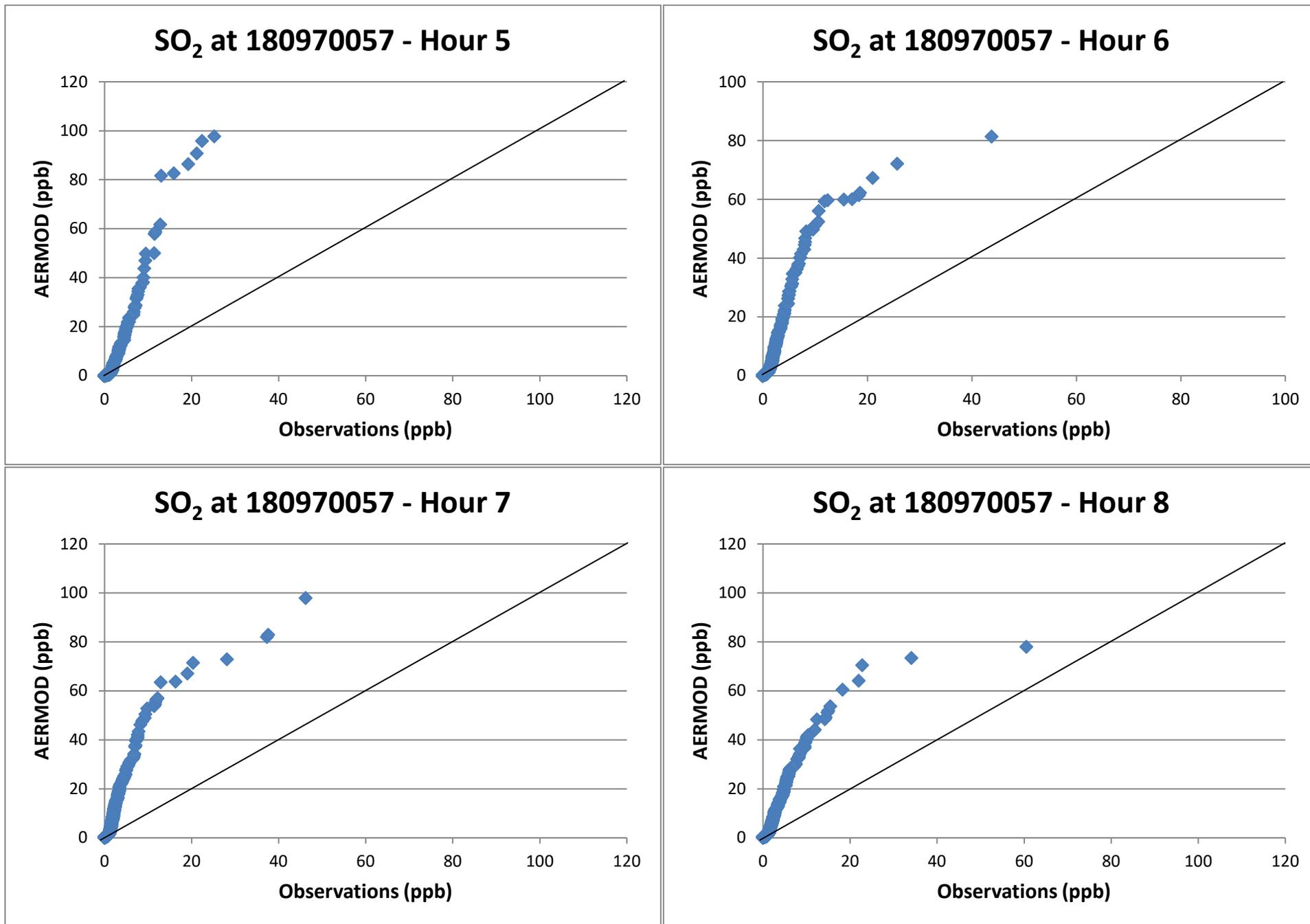
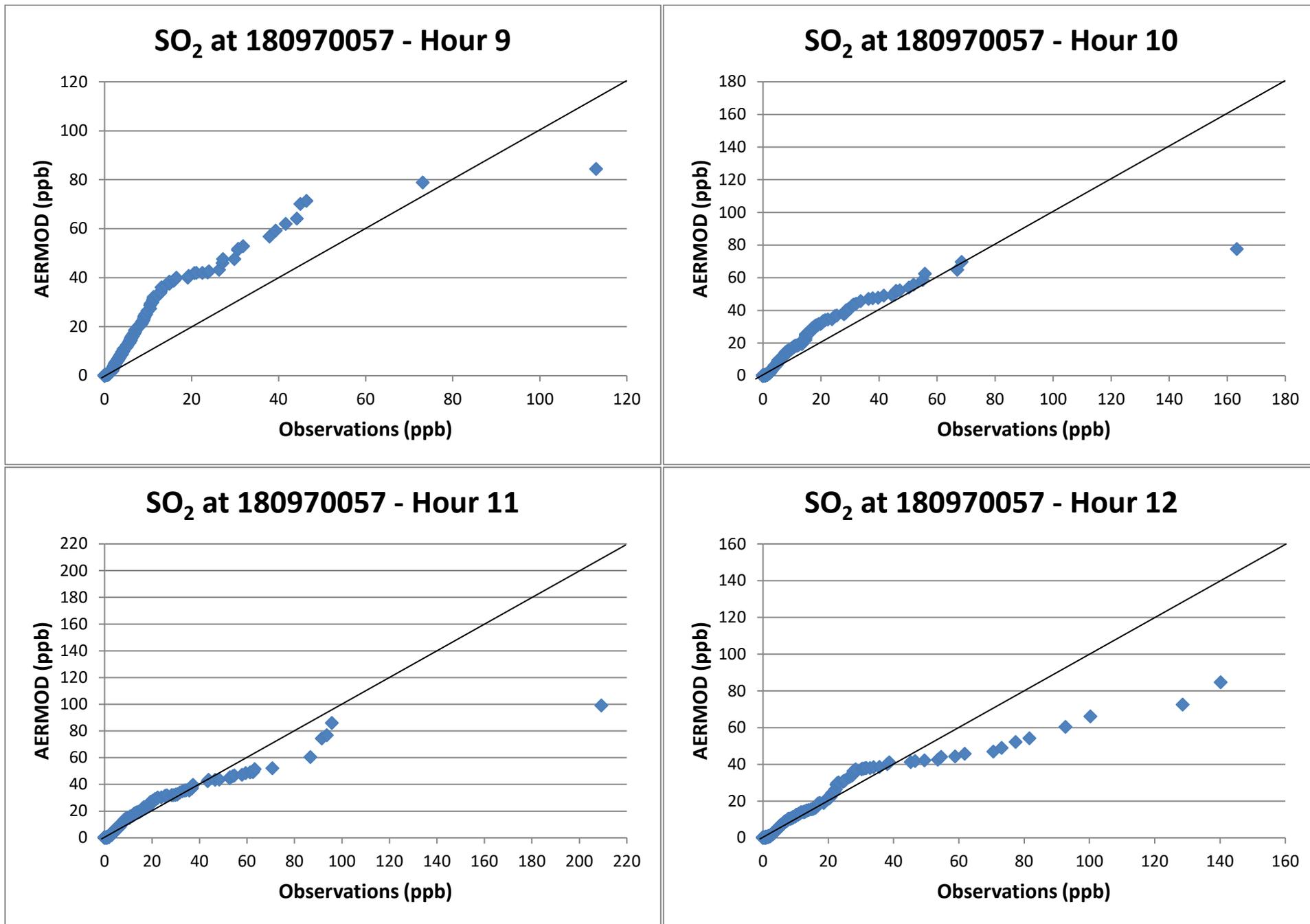


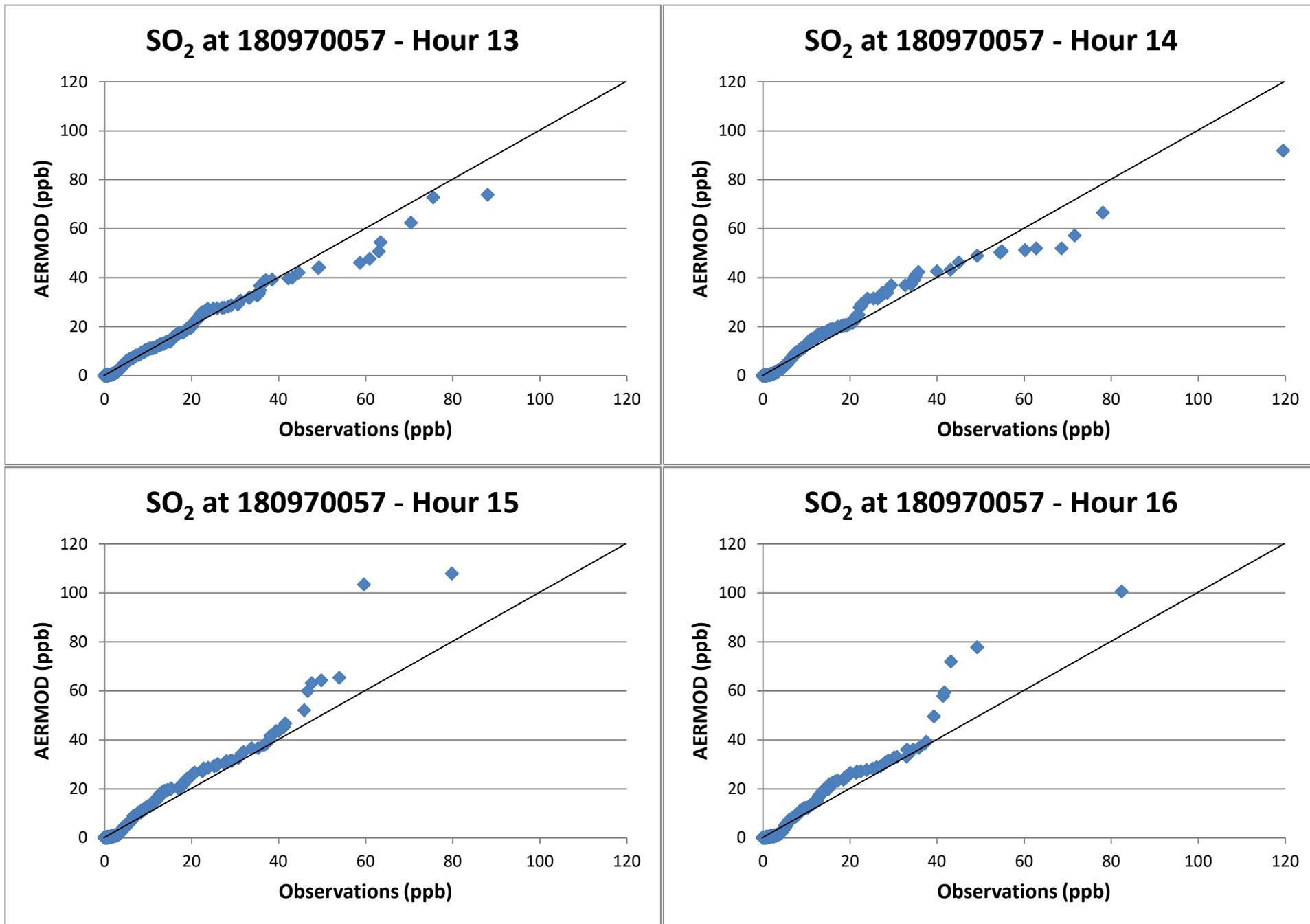
Figure 17. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).



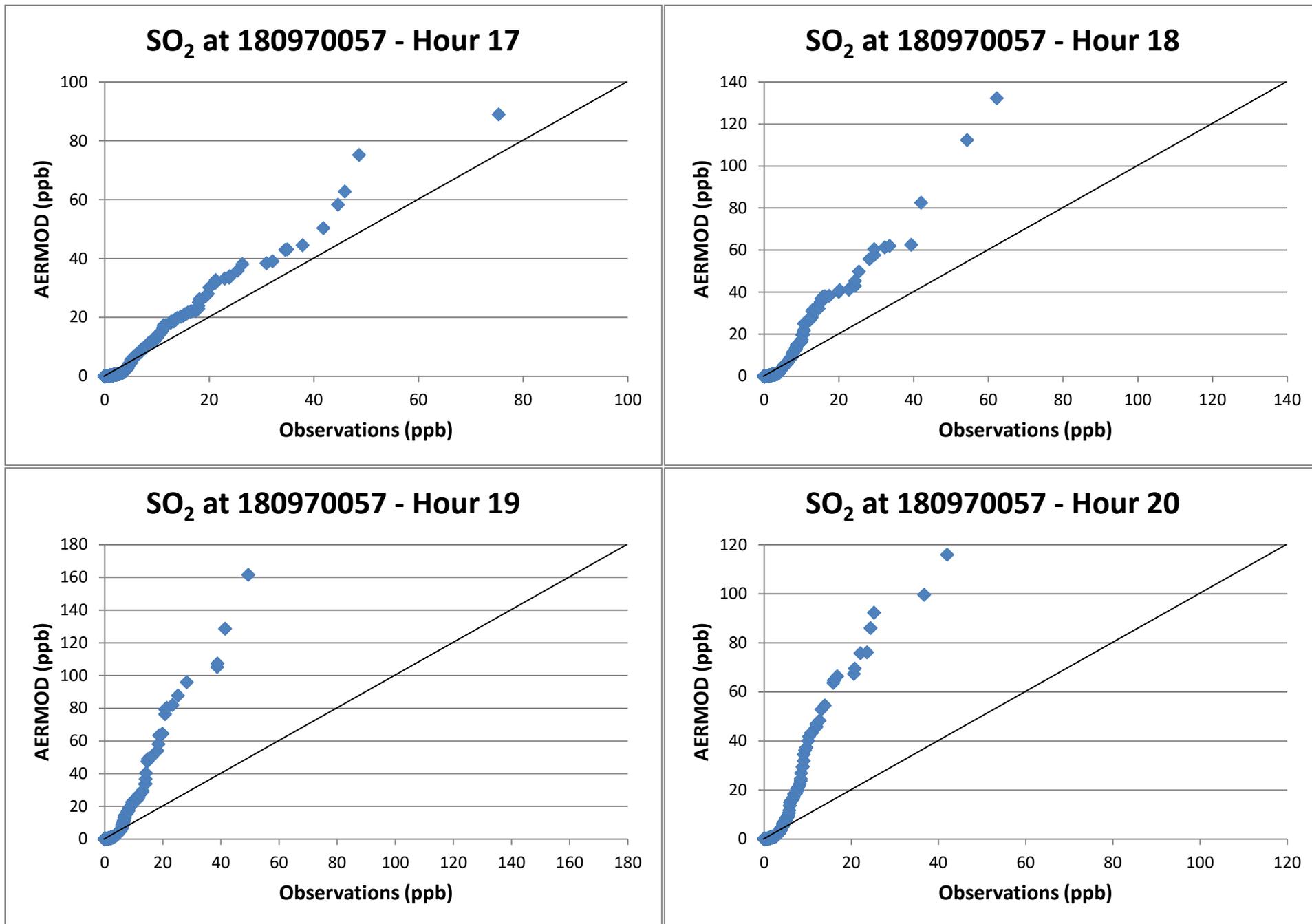
**Figure 18.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).



**Figure 19.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).



**Figure 20.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).



**Figure 21.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).

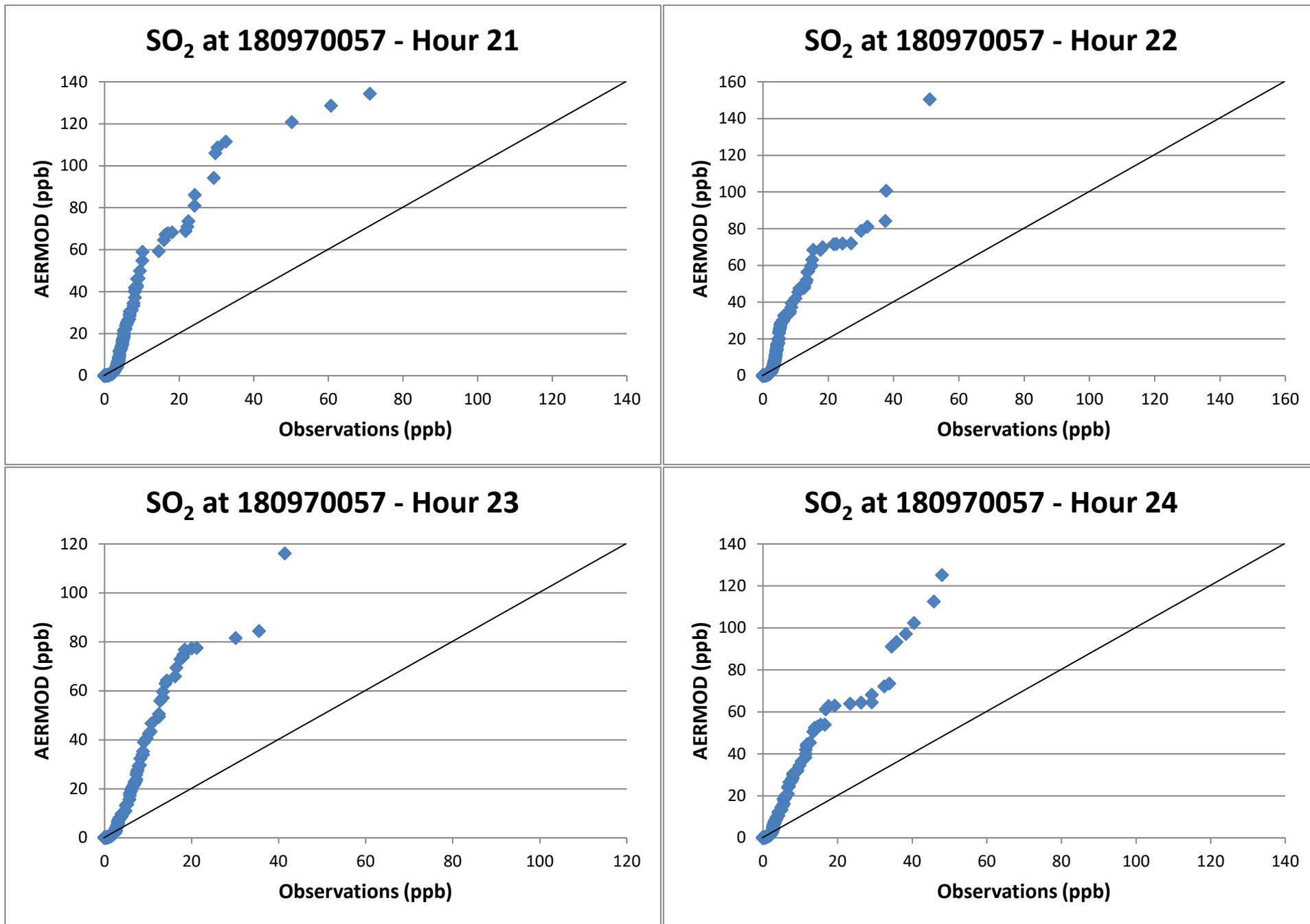


Figure 22. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970057).

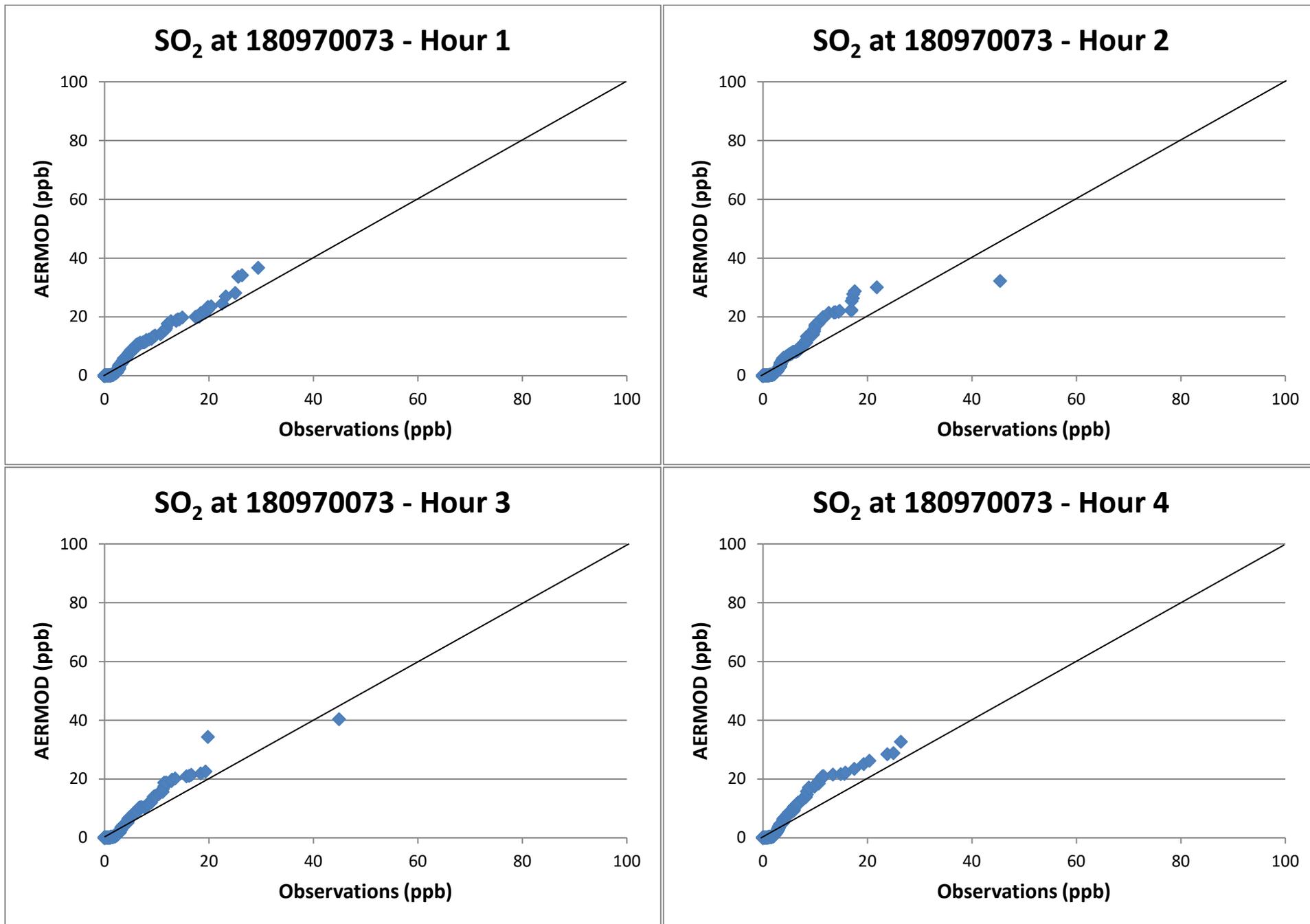


Figure 23. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).

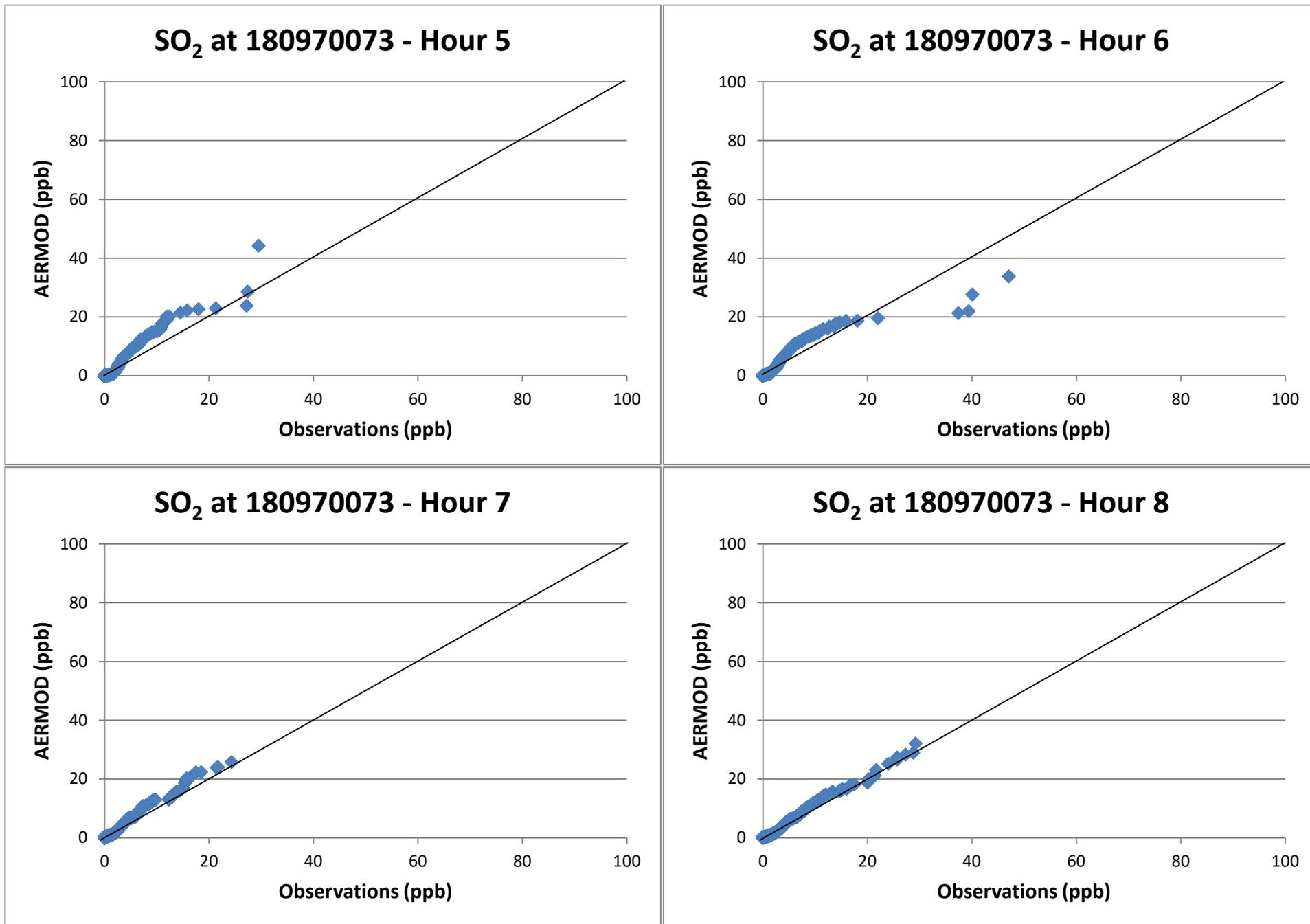


Figure 24. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).

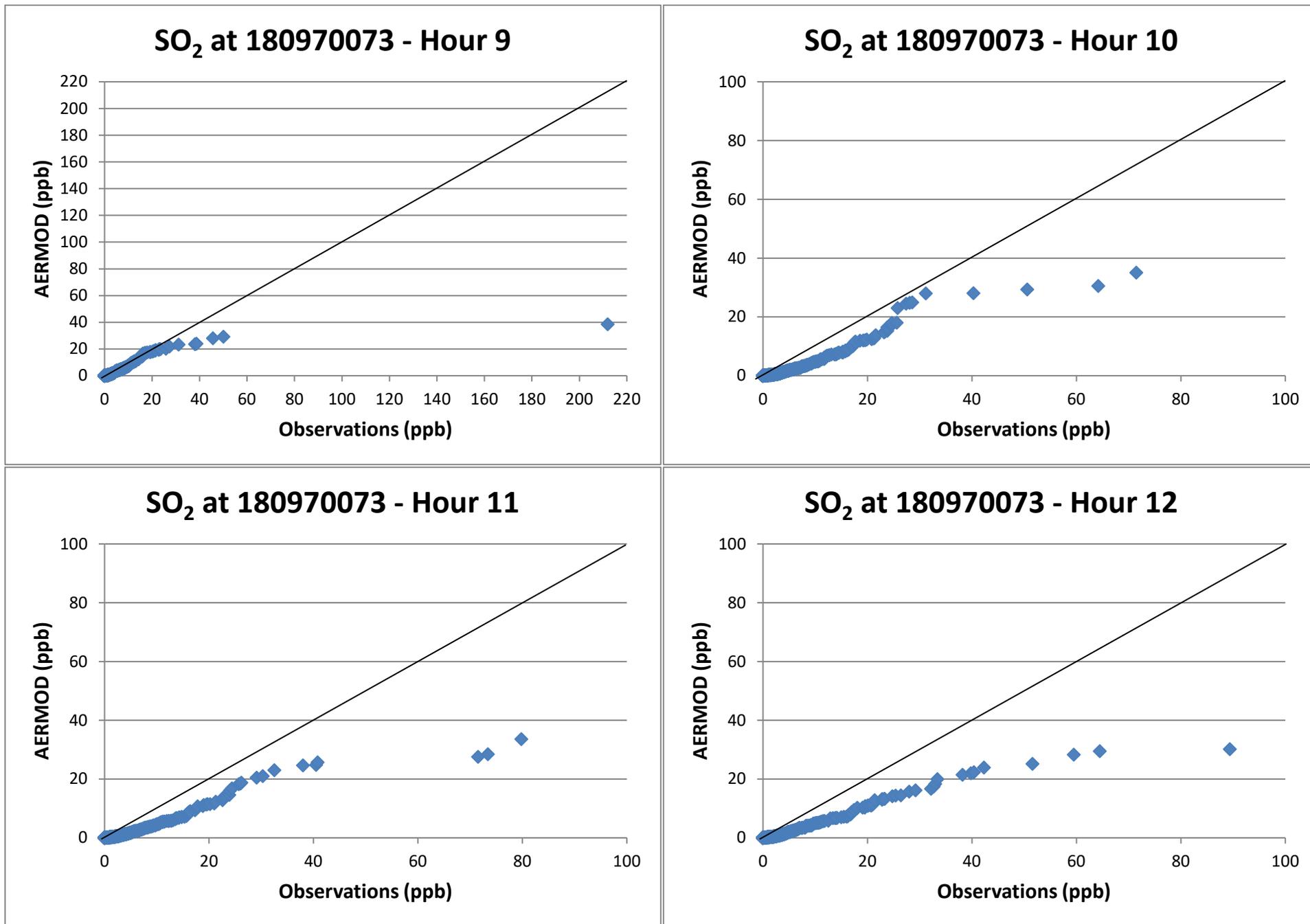
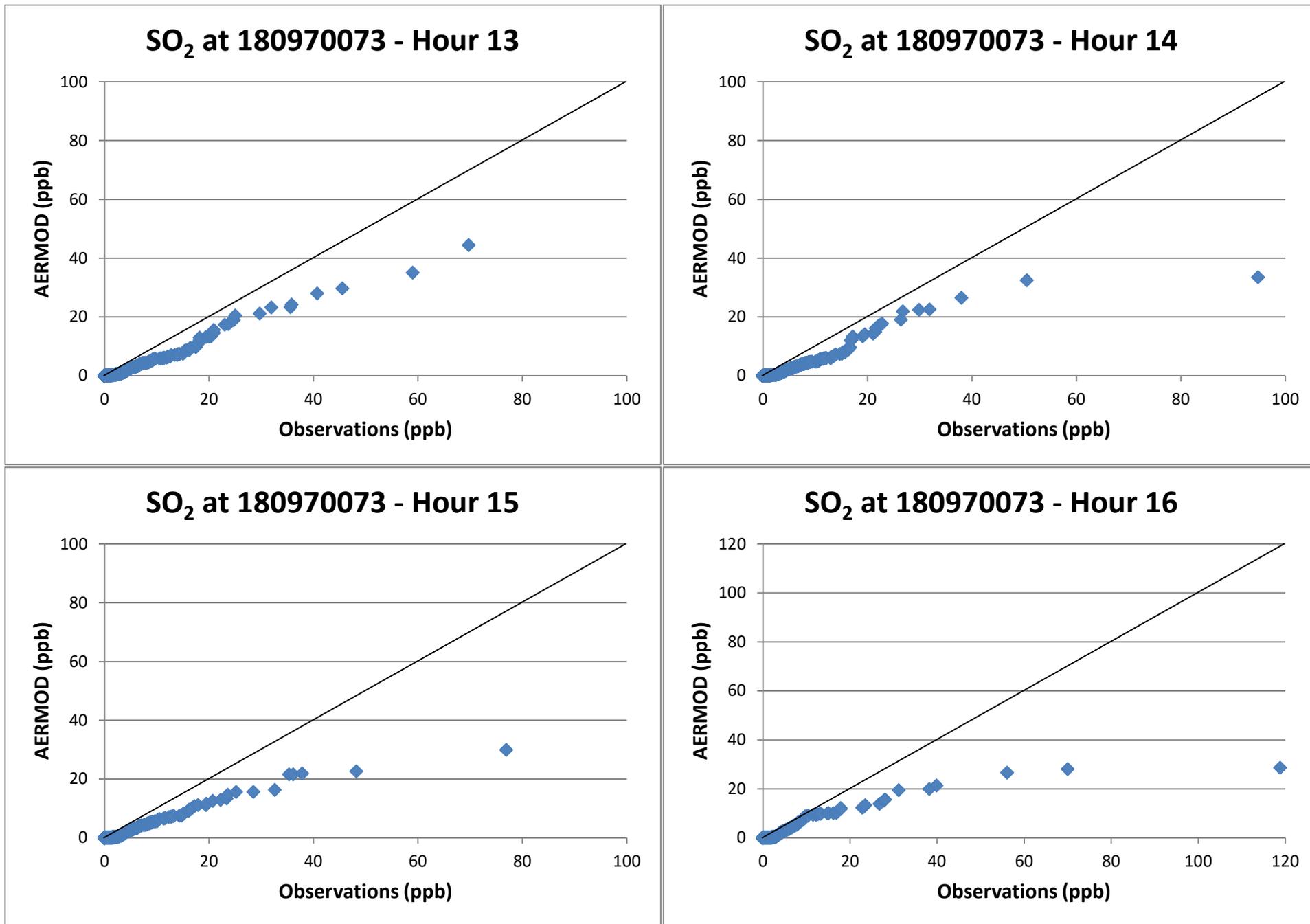
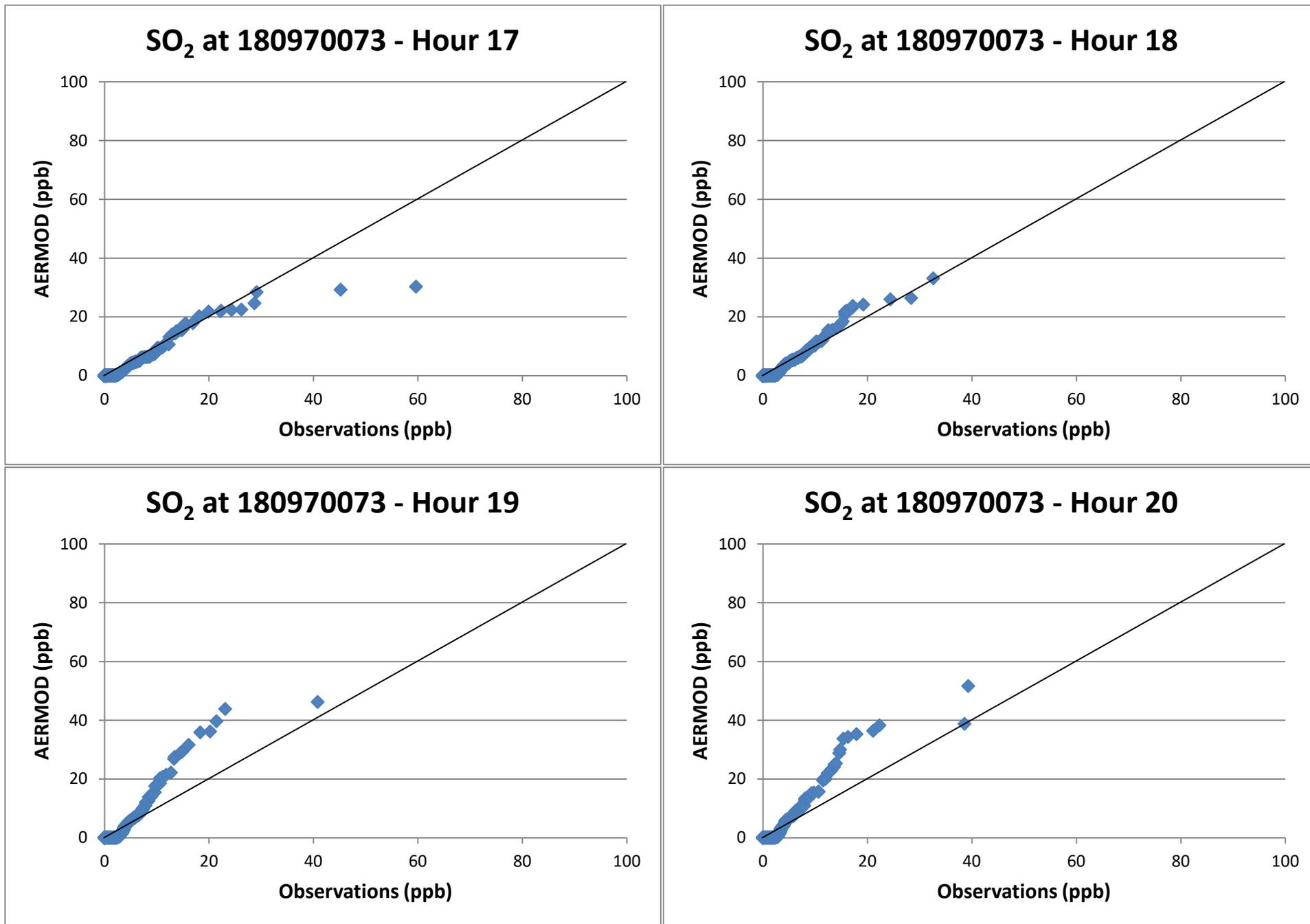


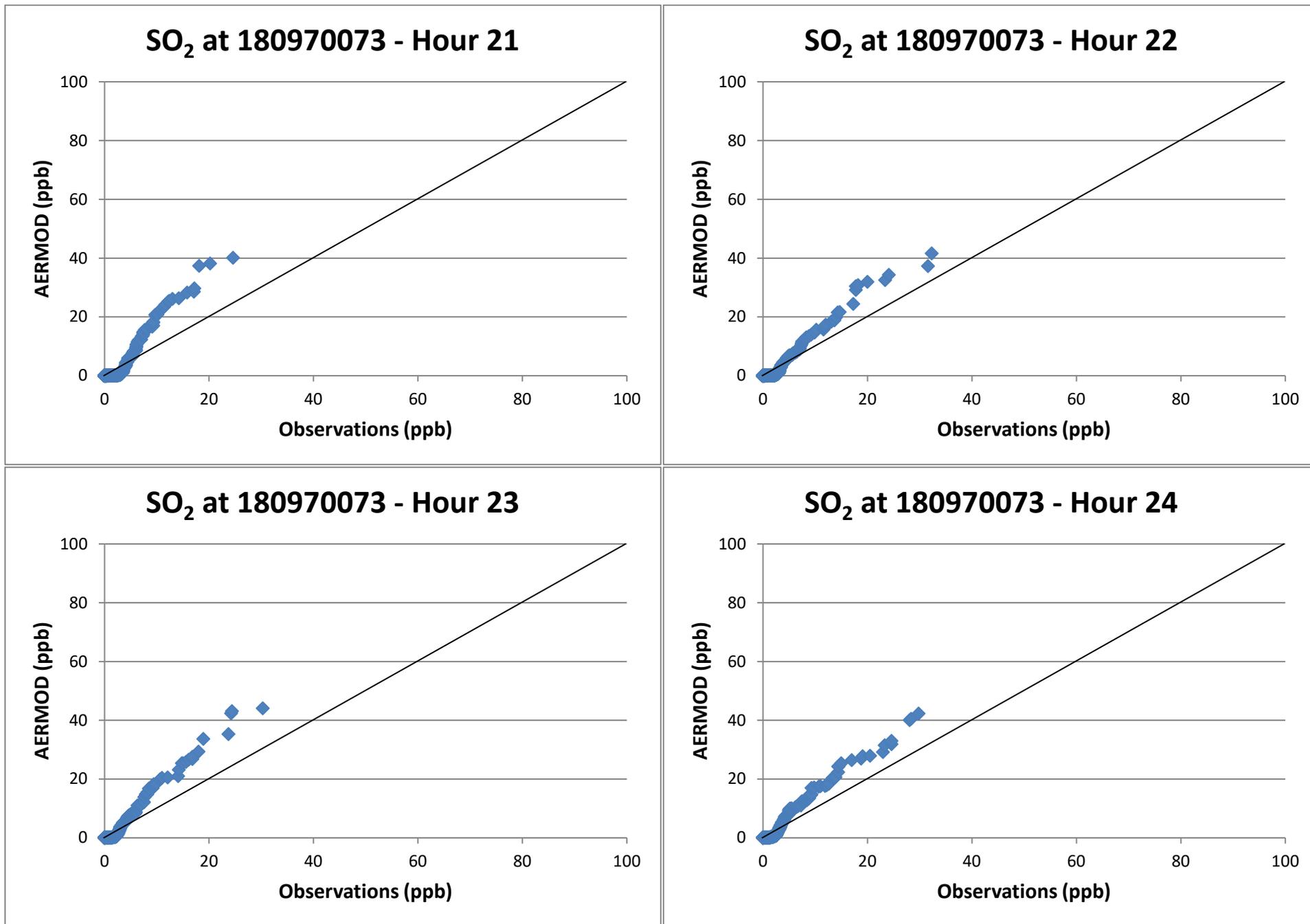
Figure 25. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).



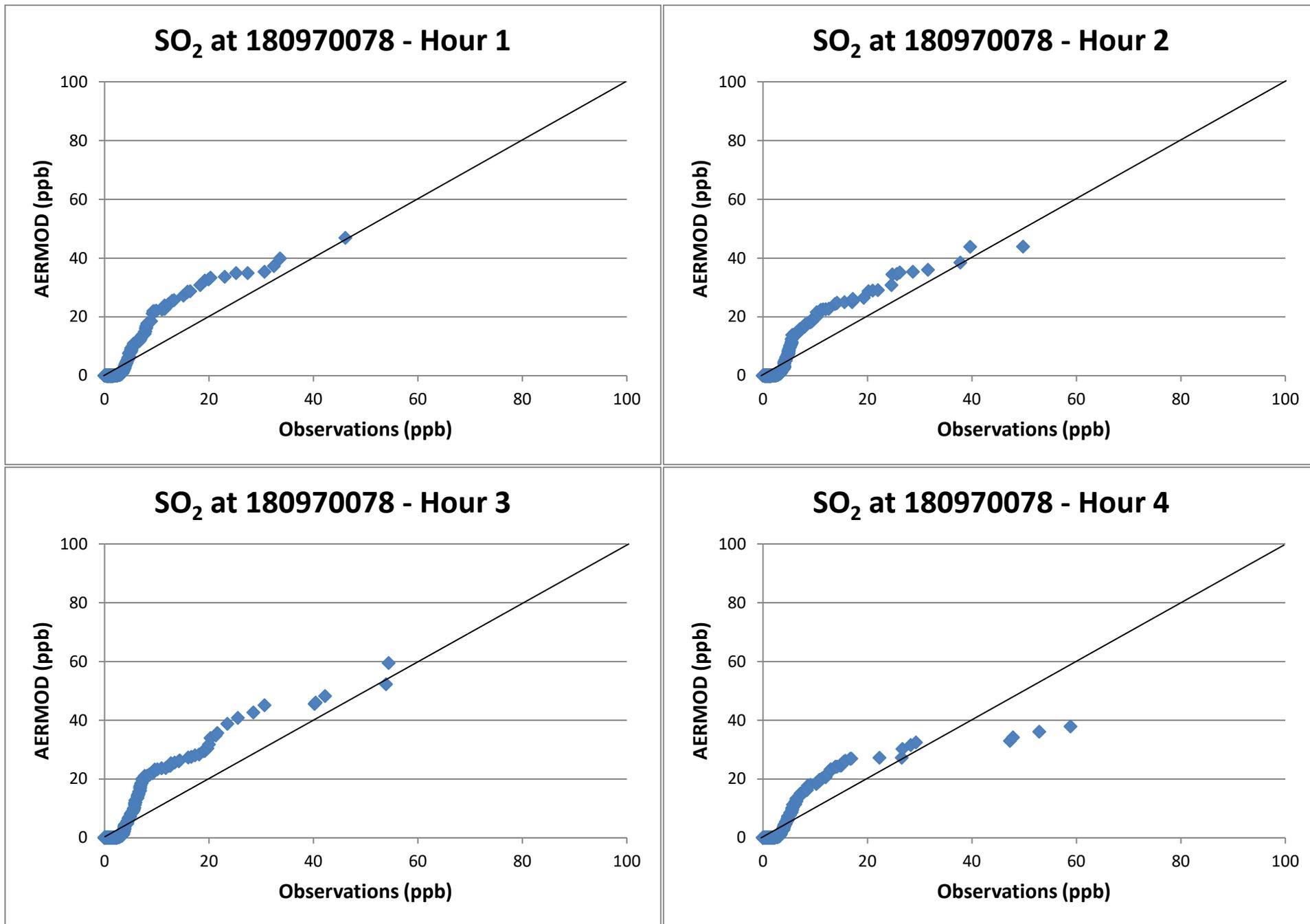
**Figure 26.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).



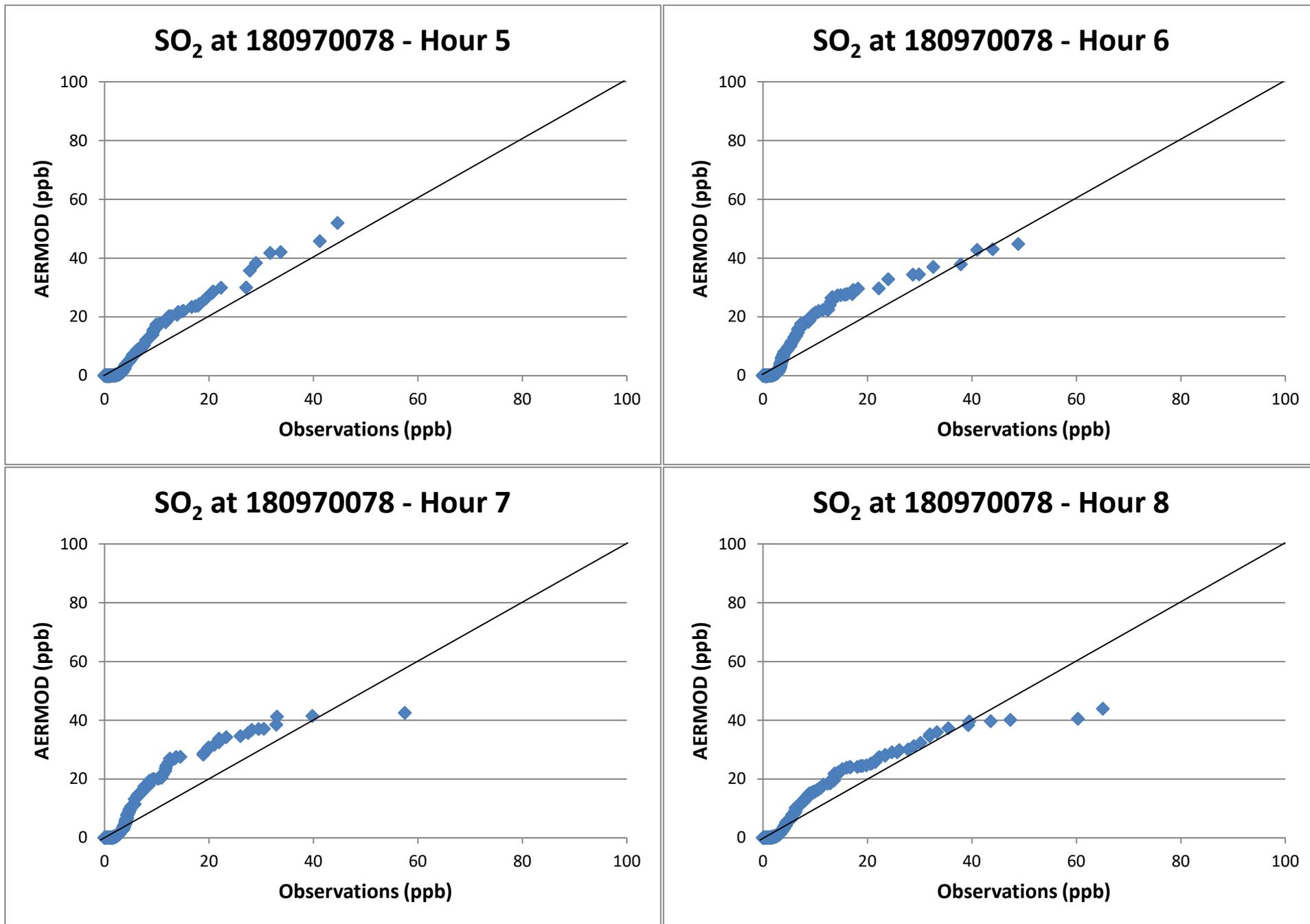
**Figure 27.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).



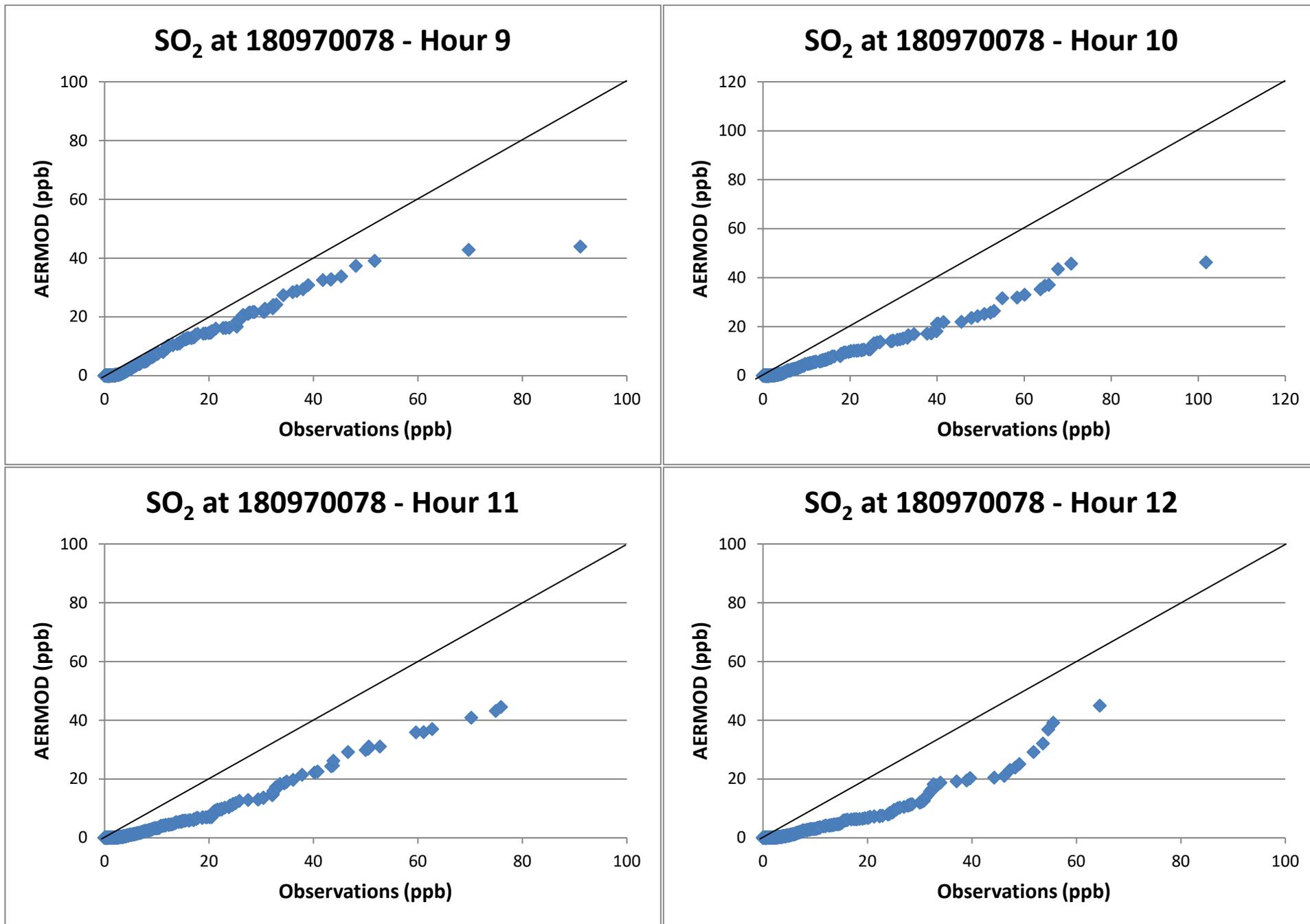
**Figure 28.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970073).



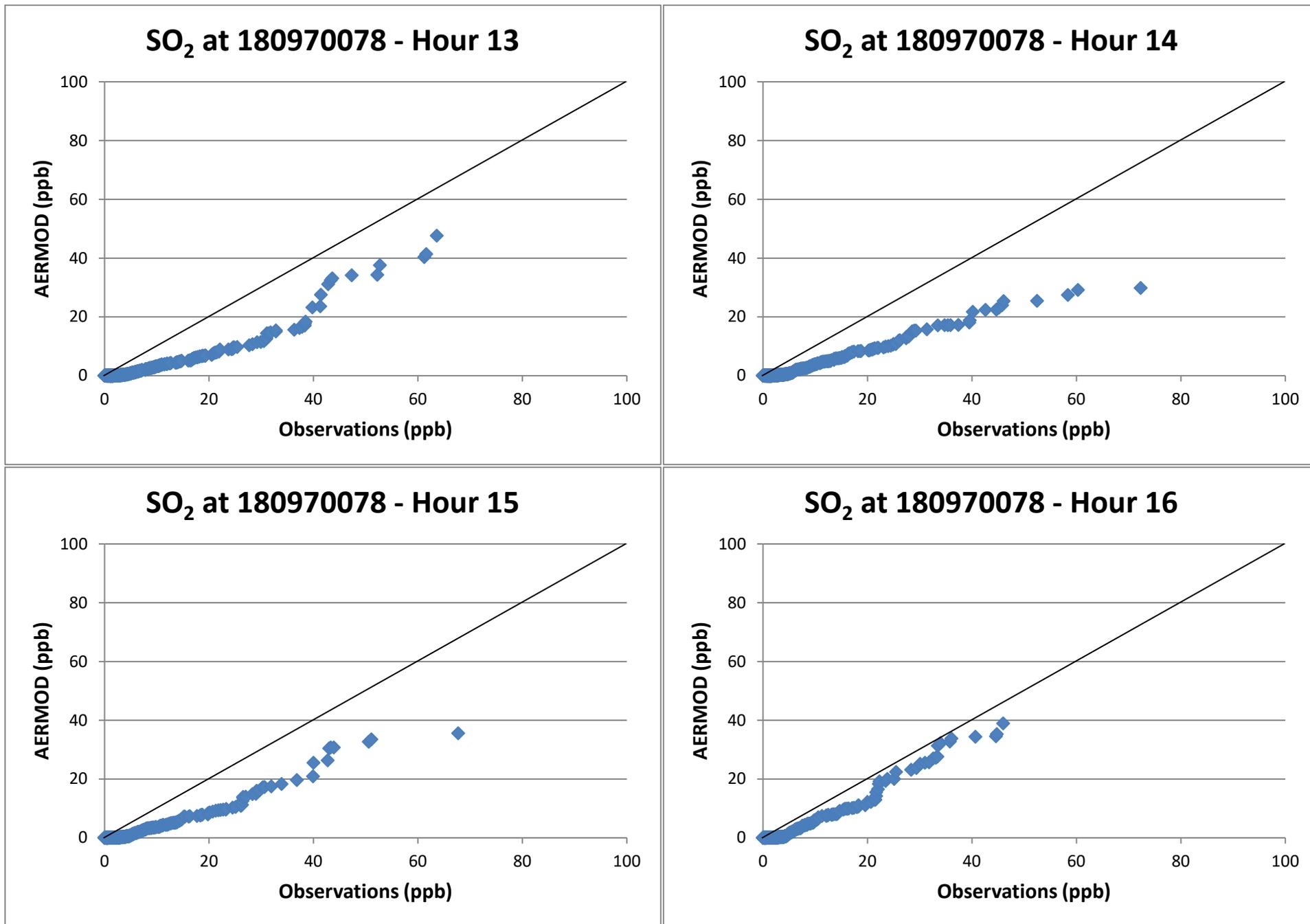
**Figure 29.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 30.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 31.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 32.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).

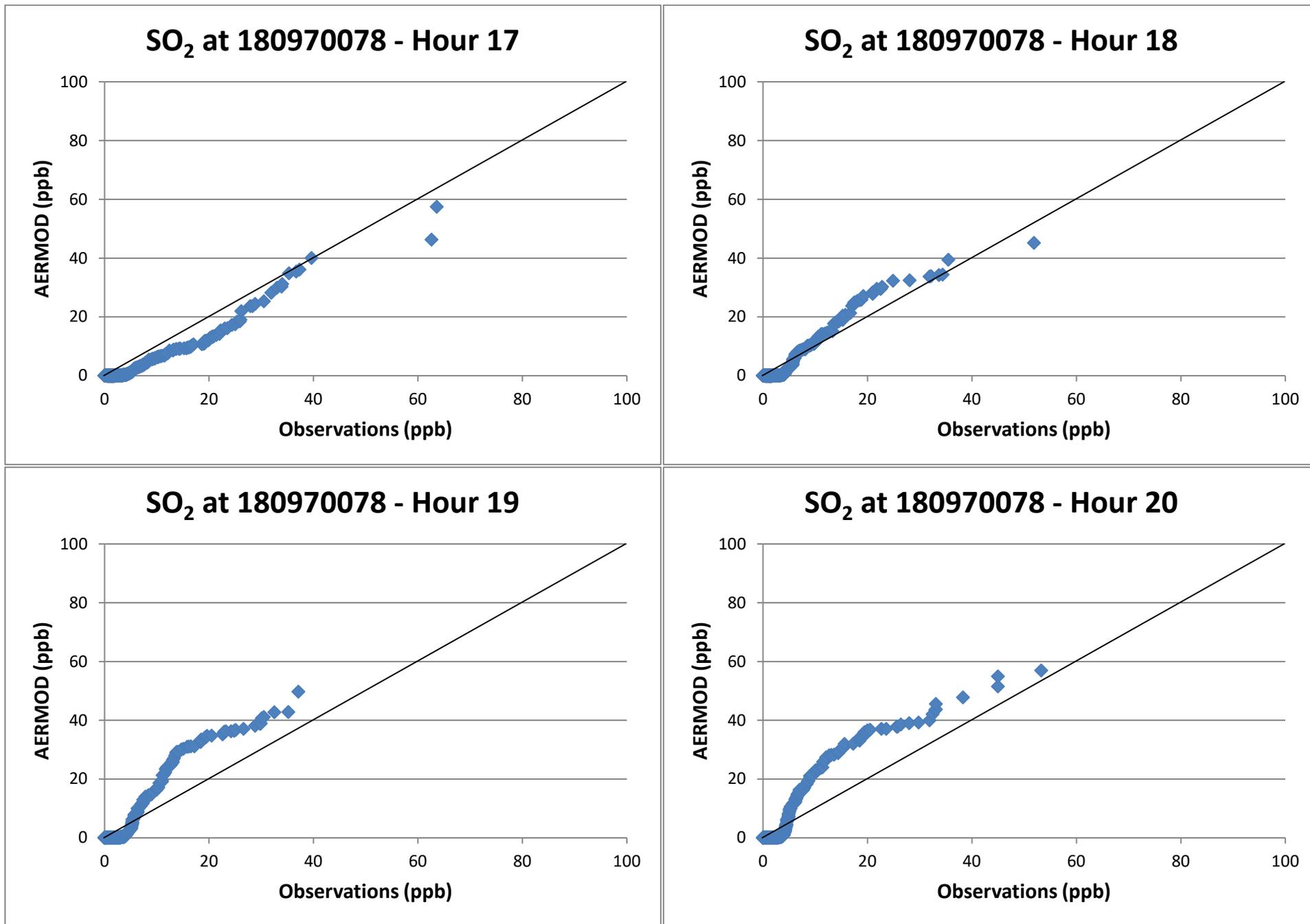
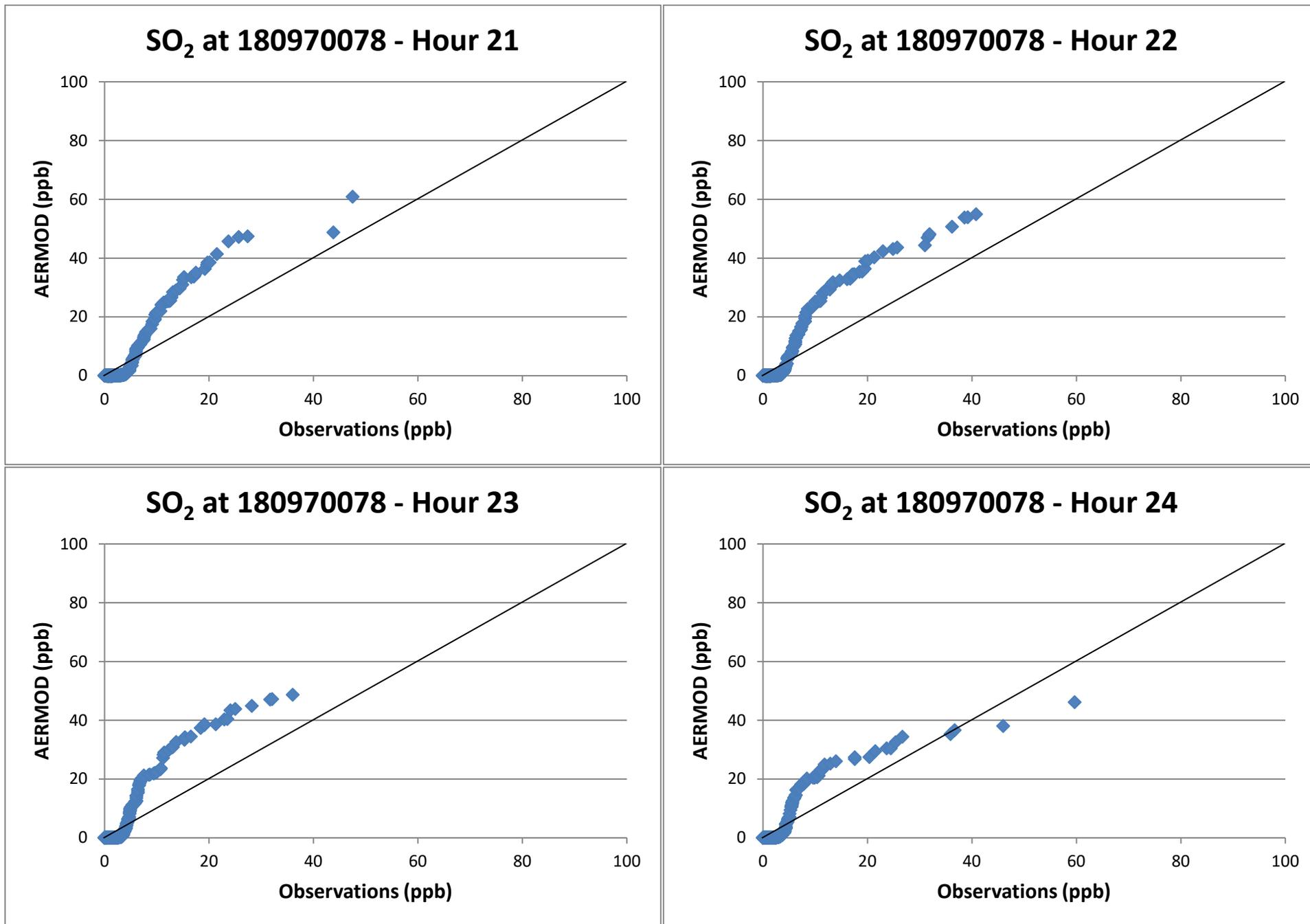
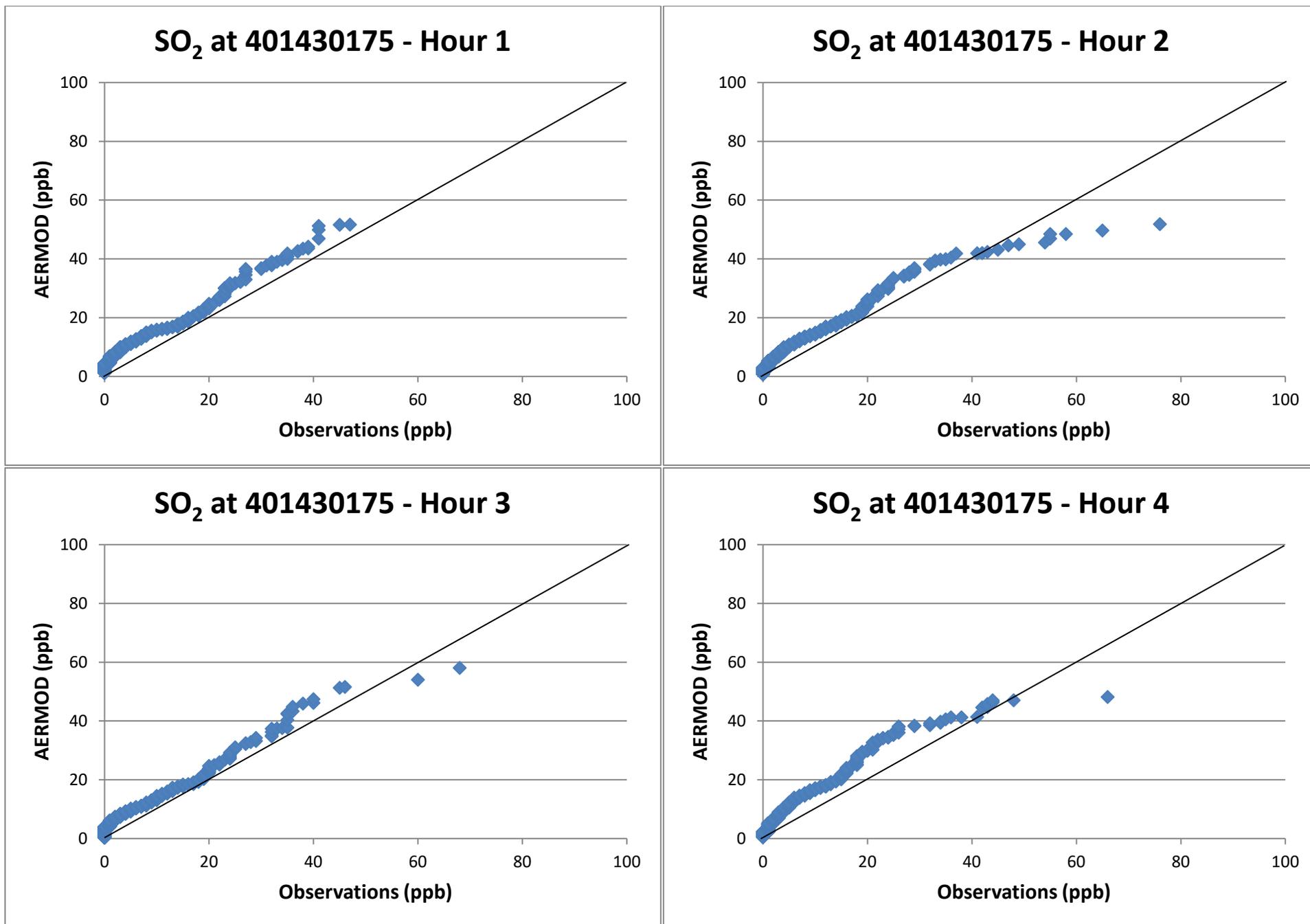


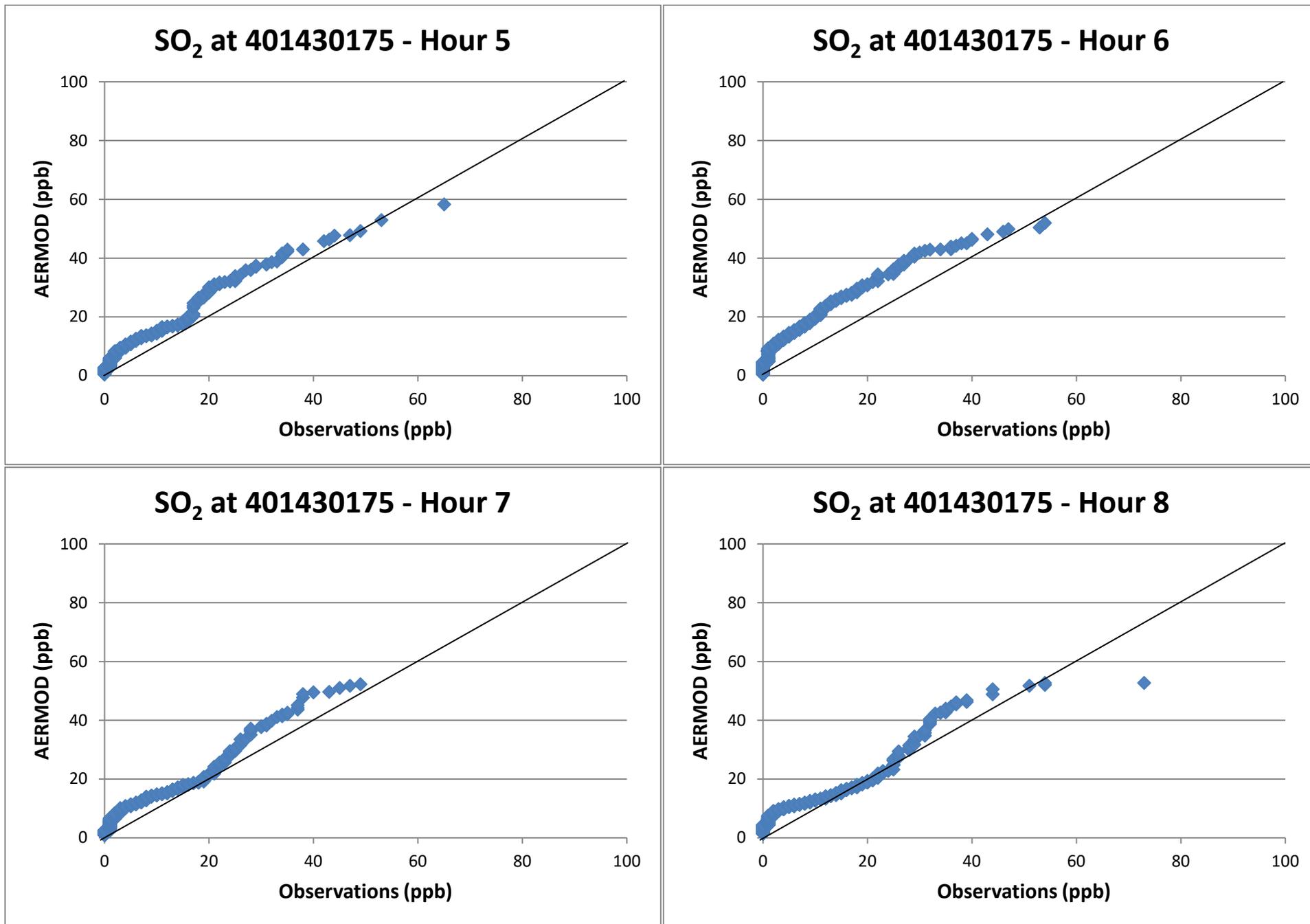
Figure 33. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 34.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Indianapolis, IN (180970078).



**Figure 35.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).



**Figure 36.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).

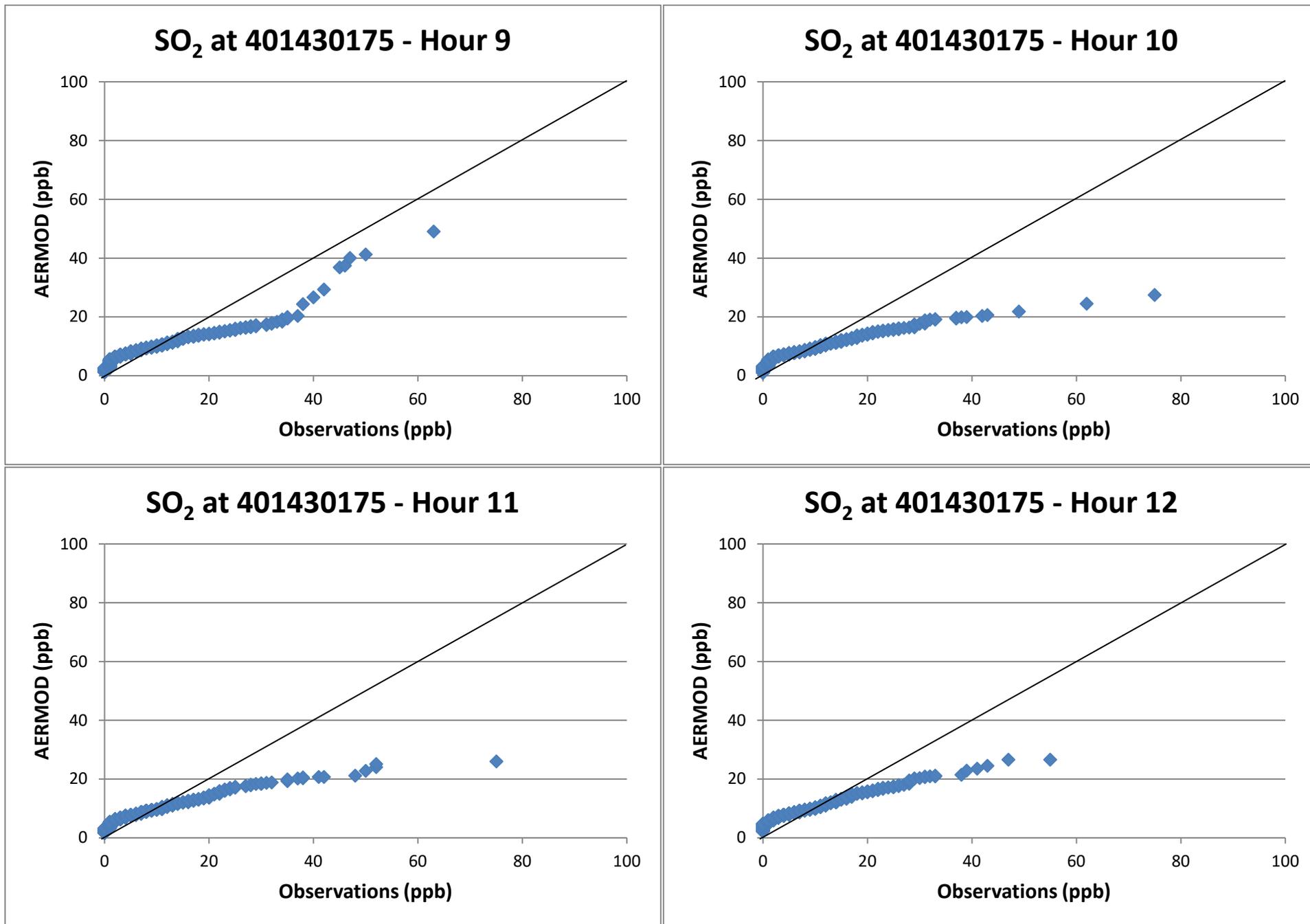
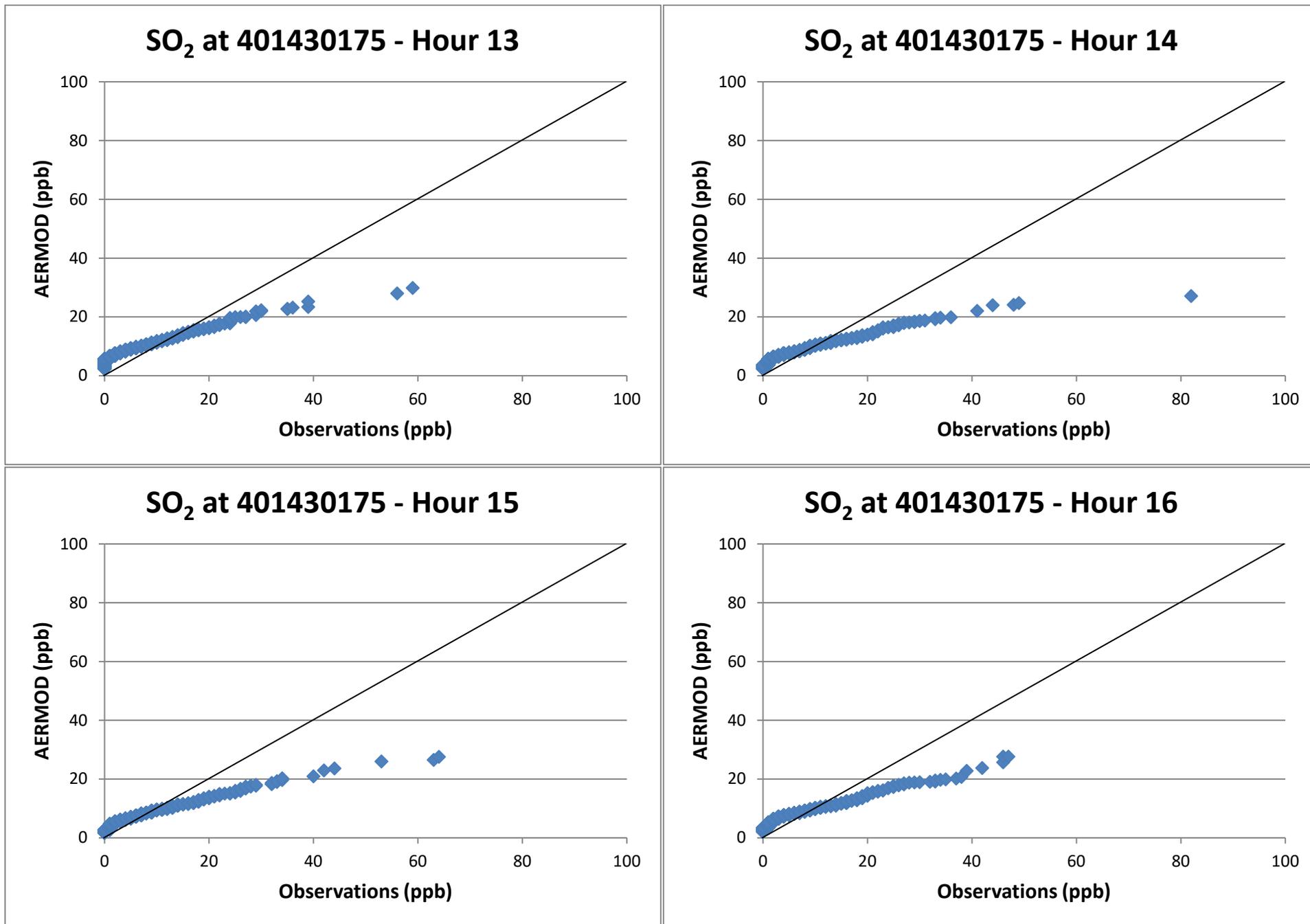
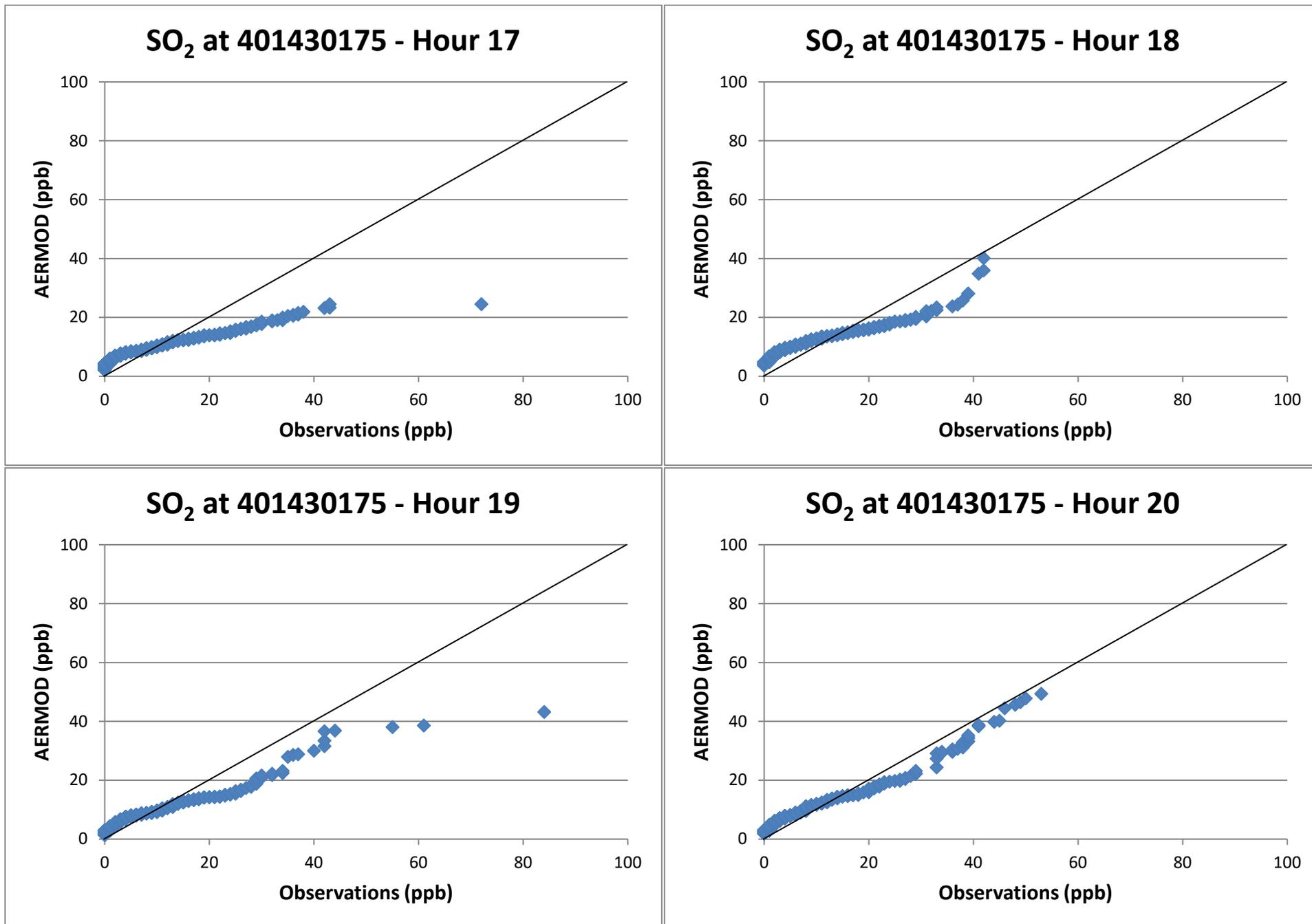


Figure 37. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).



**Figure 38.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).



**Figure 39.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).

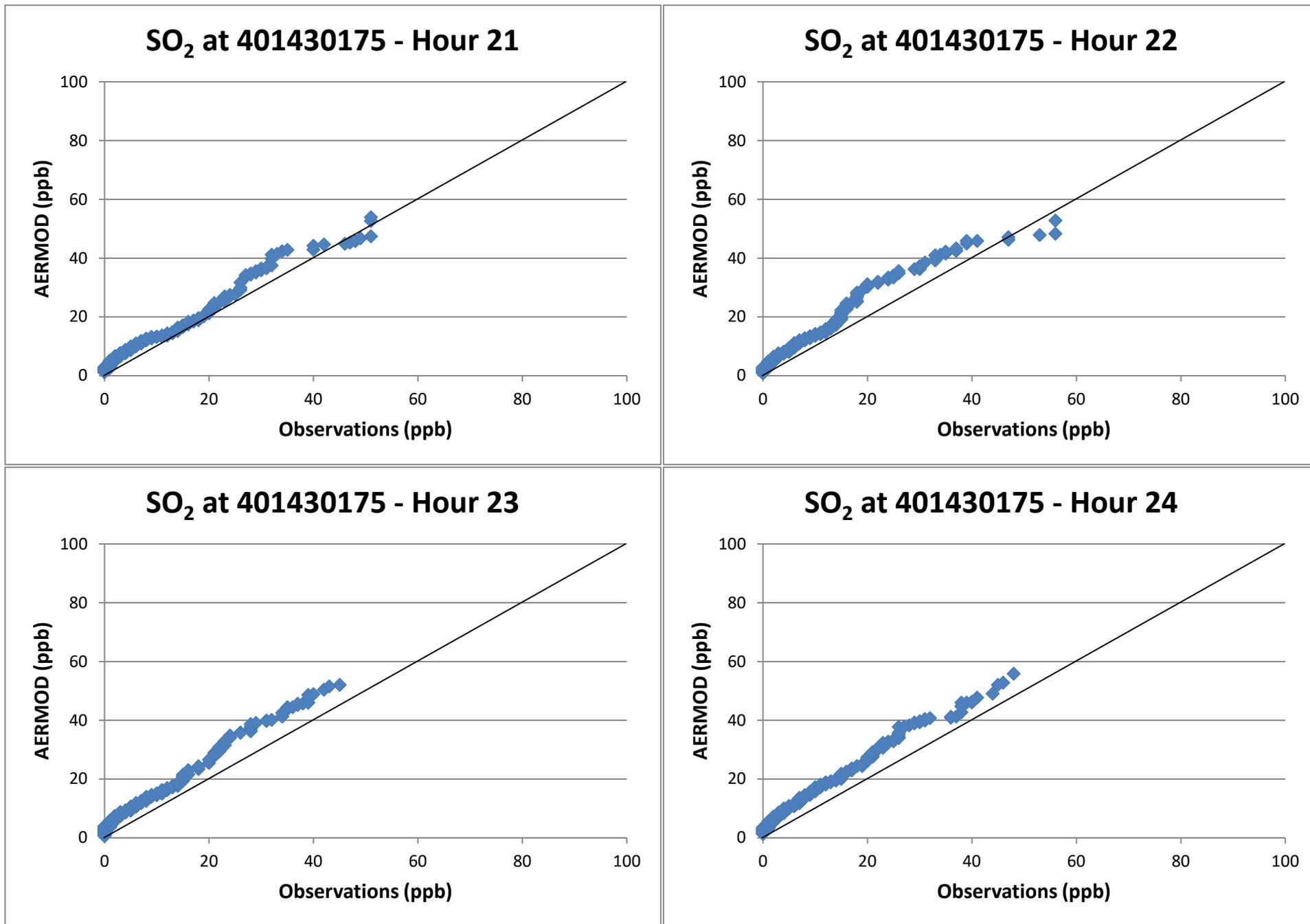
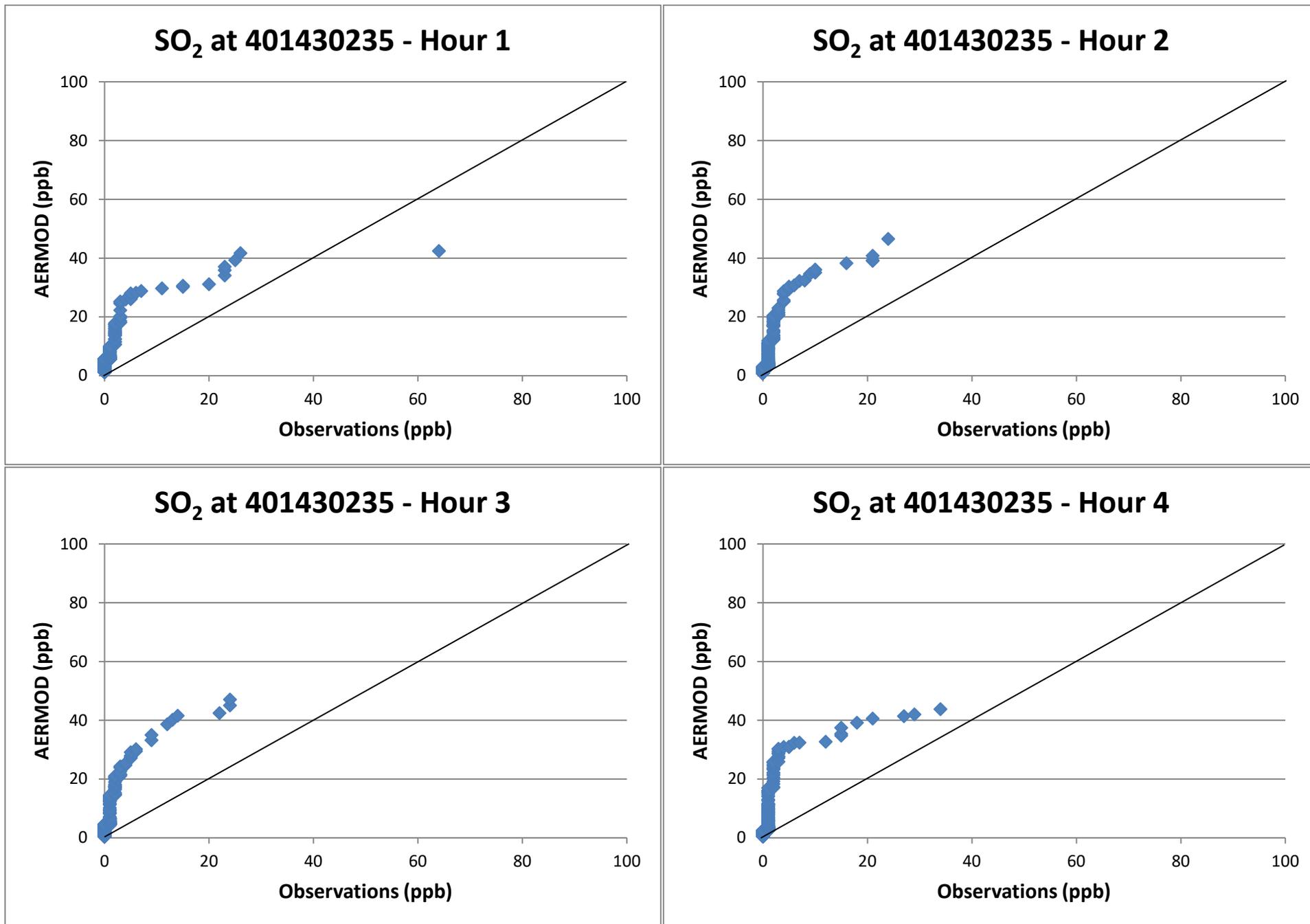


Figure 40. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430175).



**Figure 41.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).

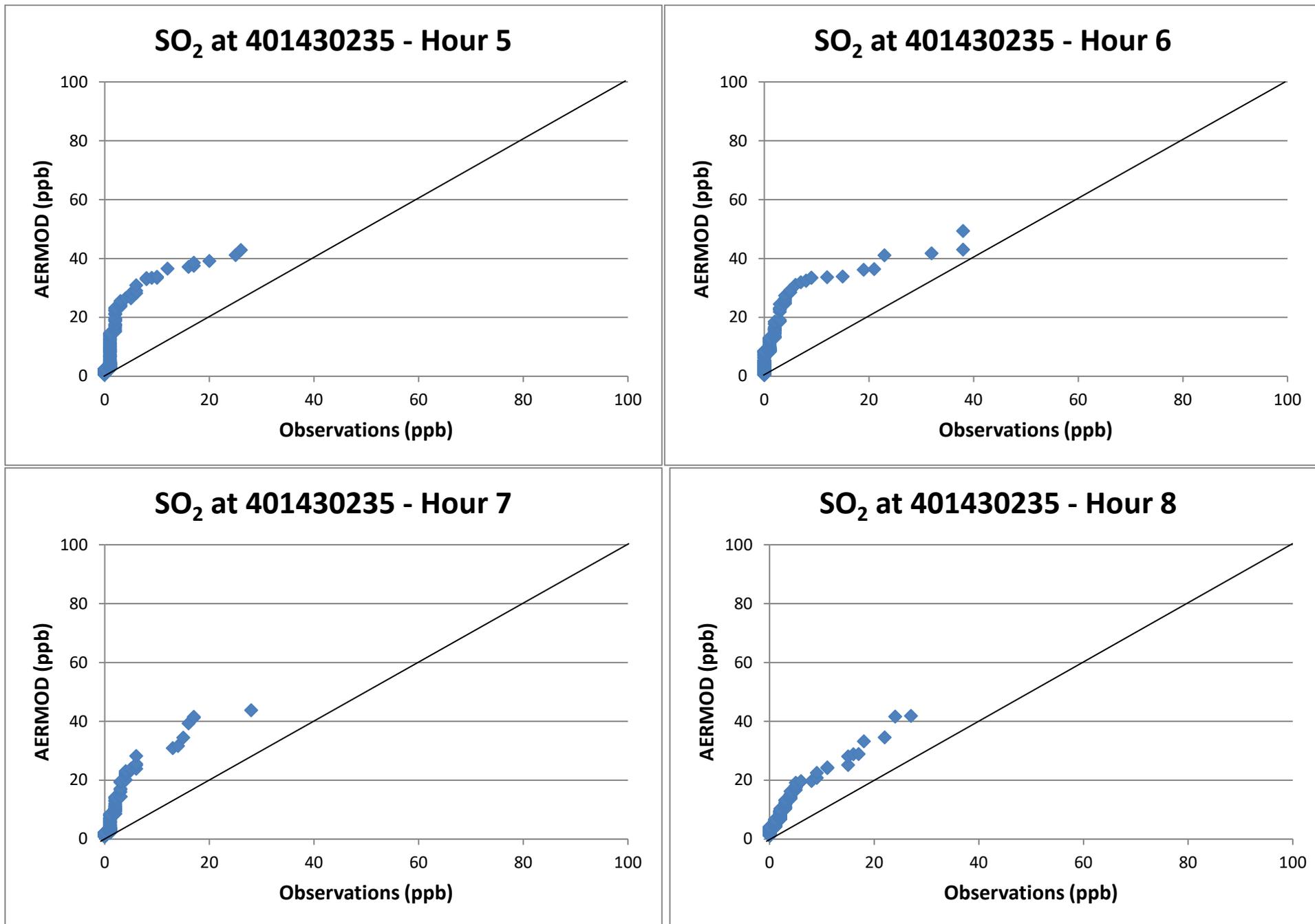
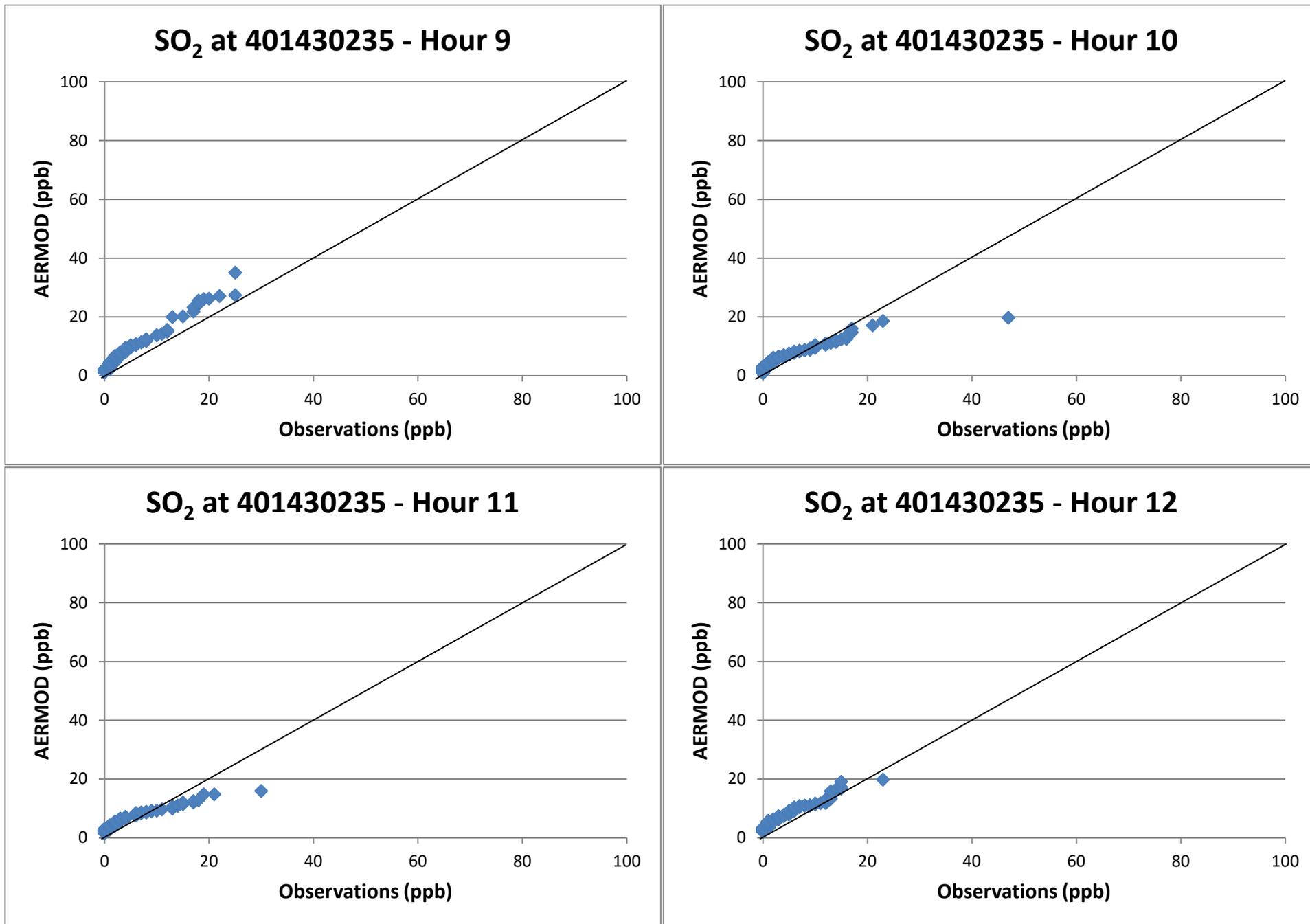
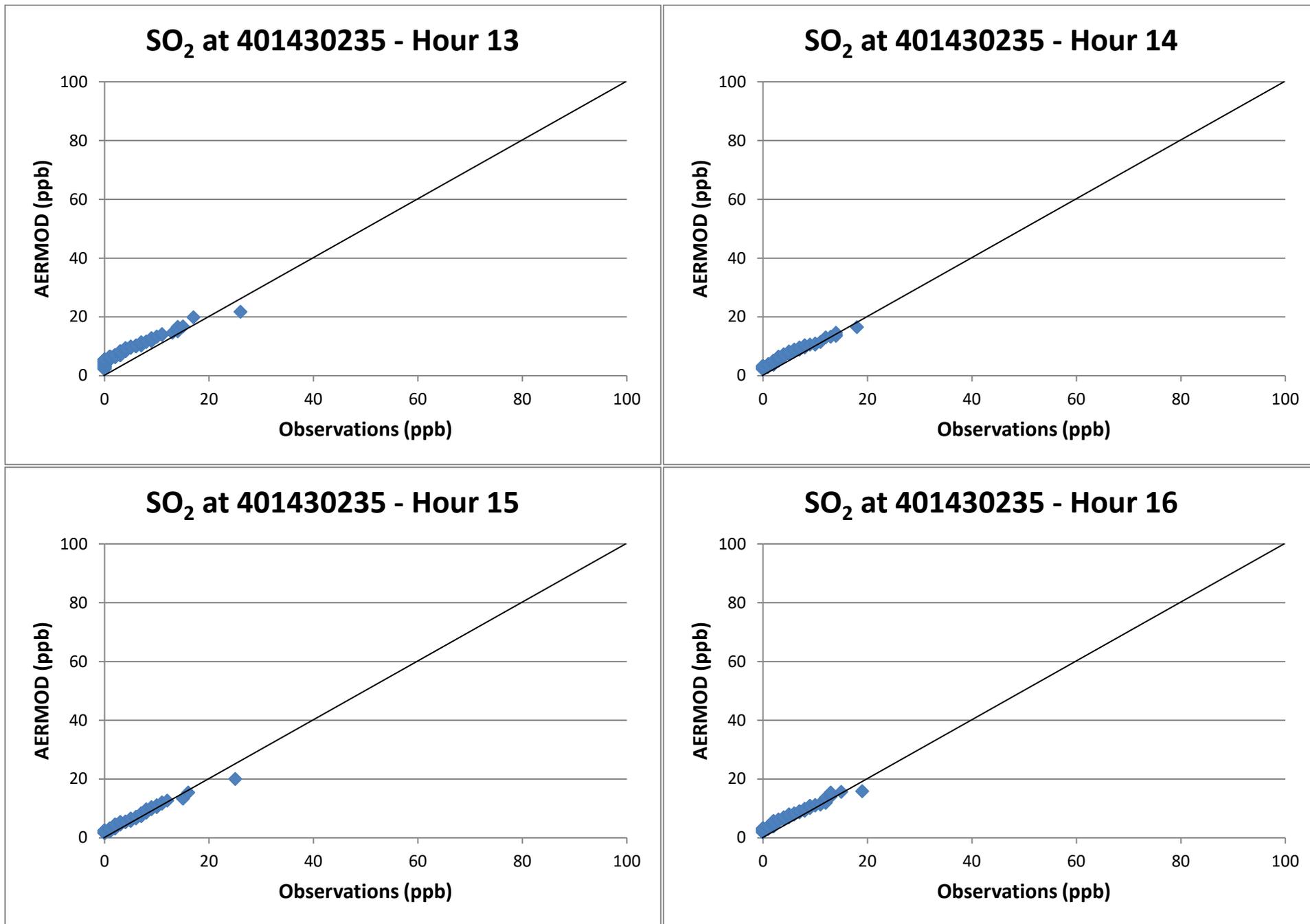


Figure 42. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).



**Figure 43.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).



**Figure 44.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).

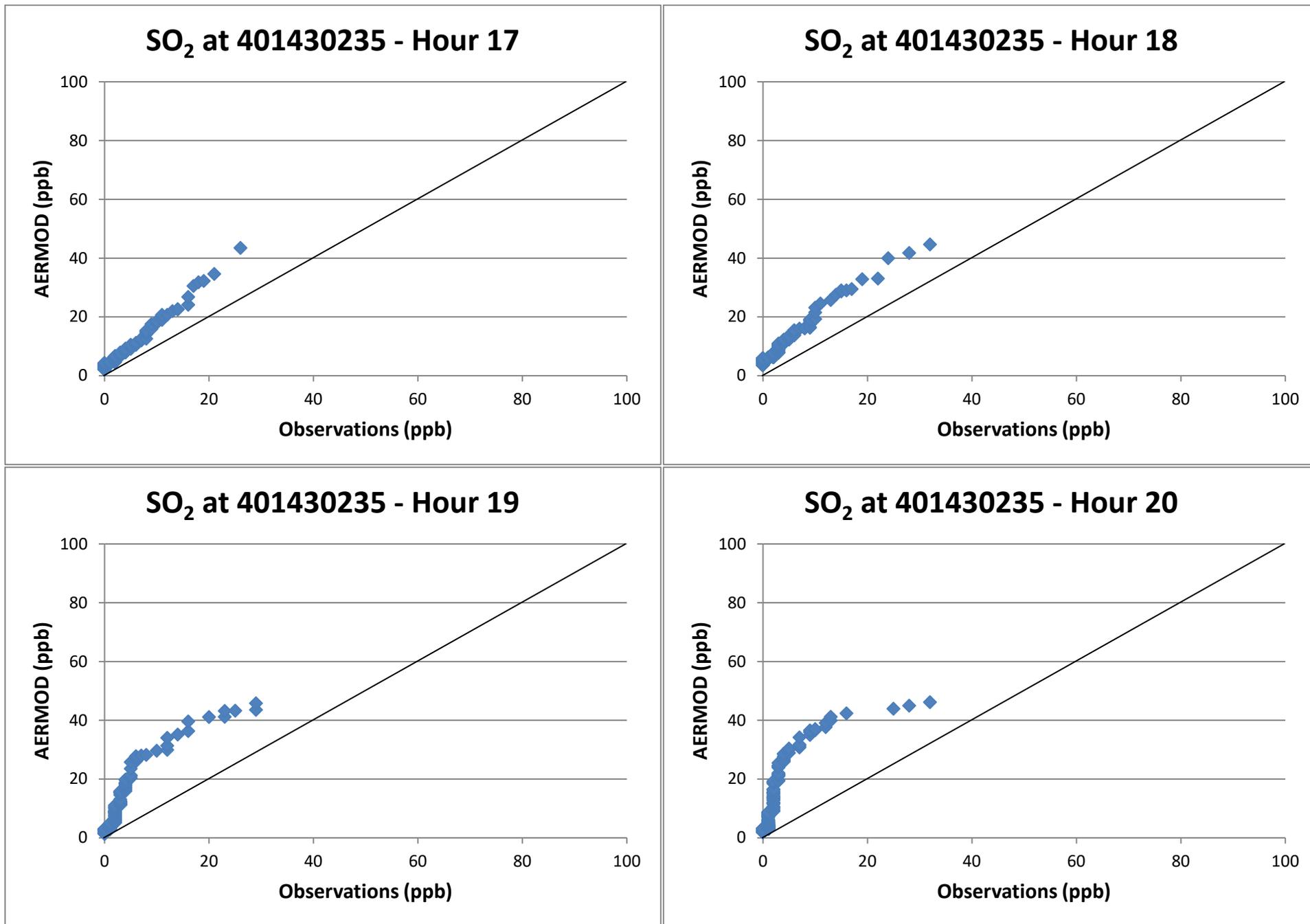


Figure 45. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).

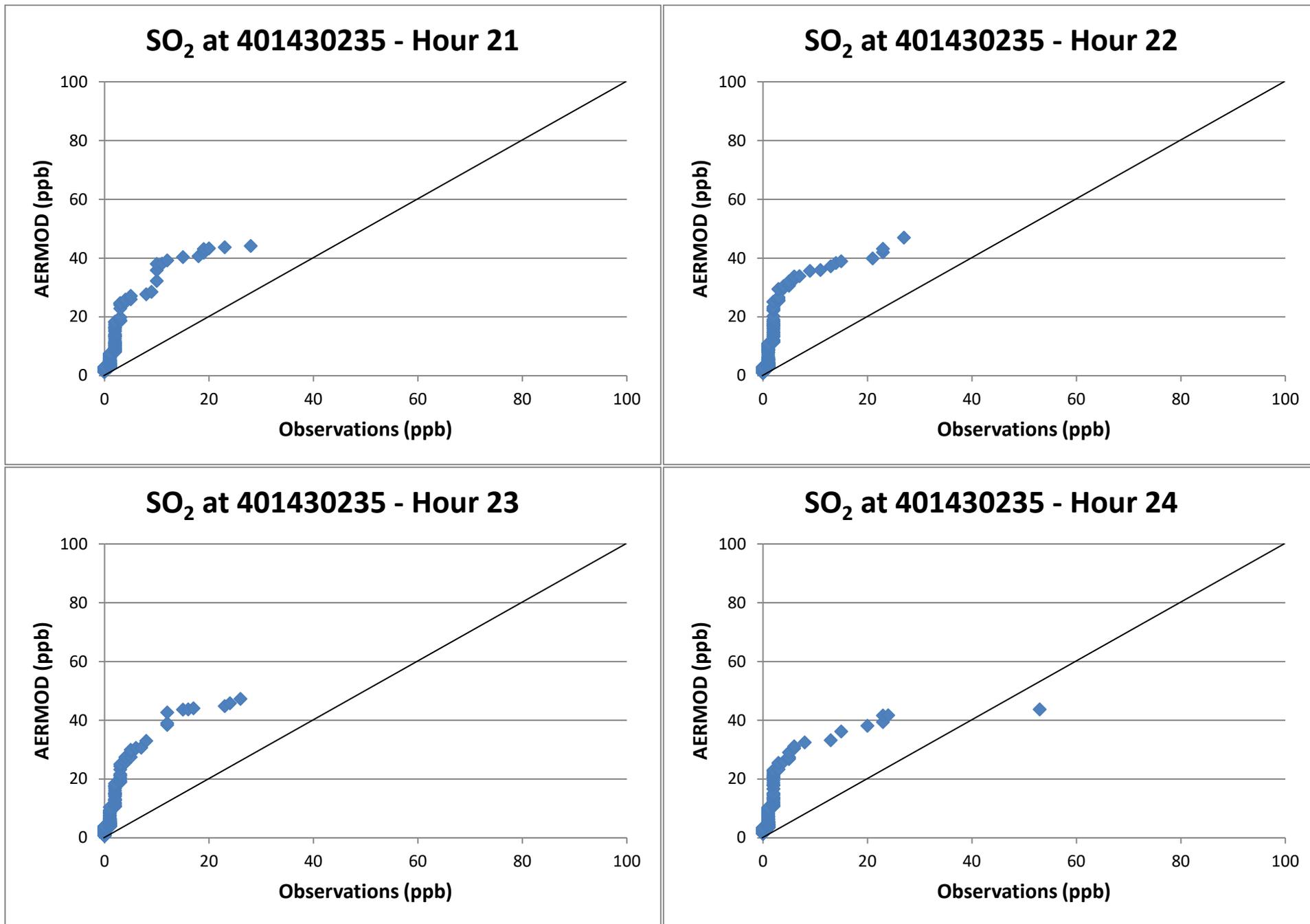


Figure 46. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401430235).

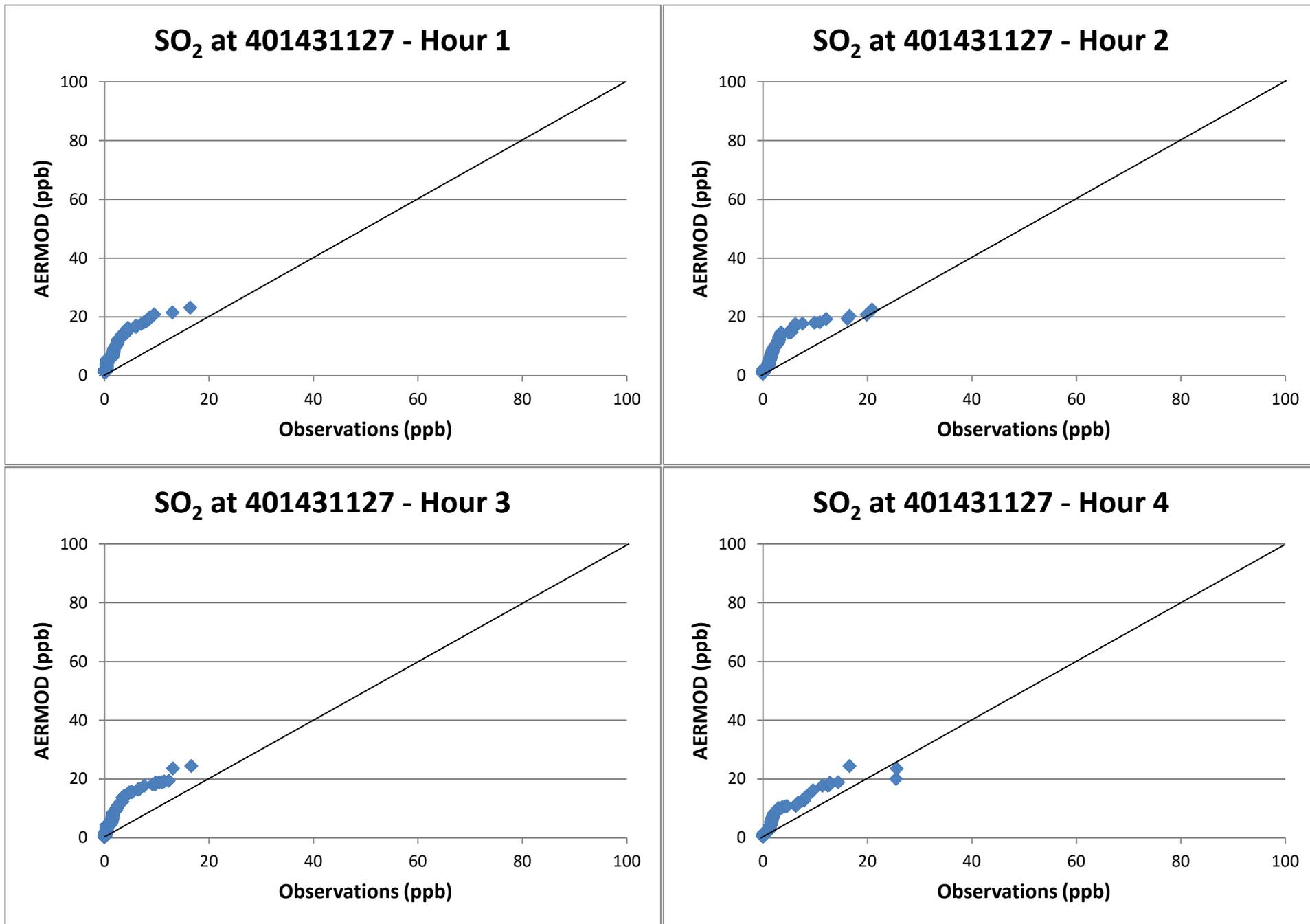


Figure 47. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).

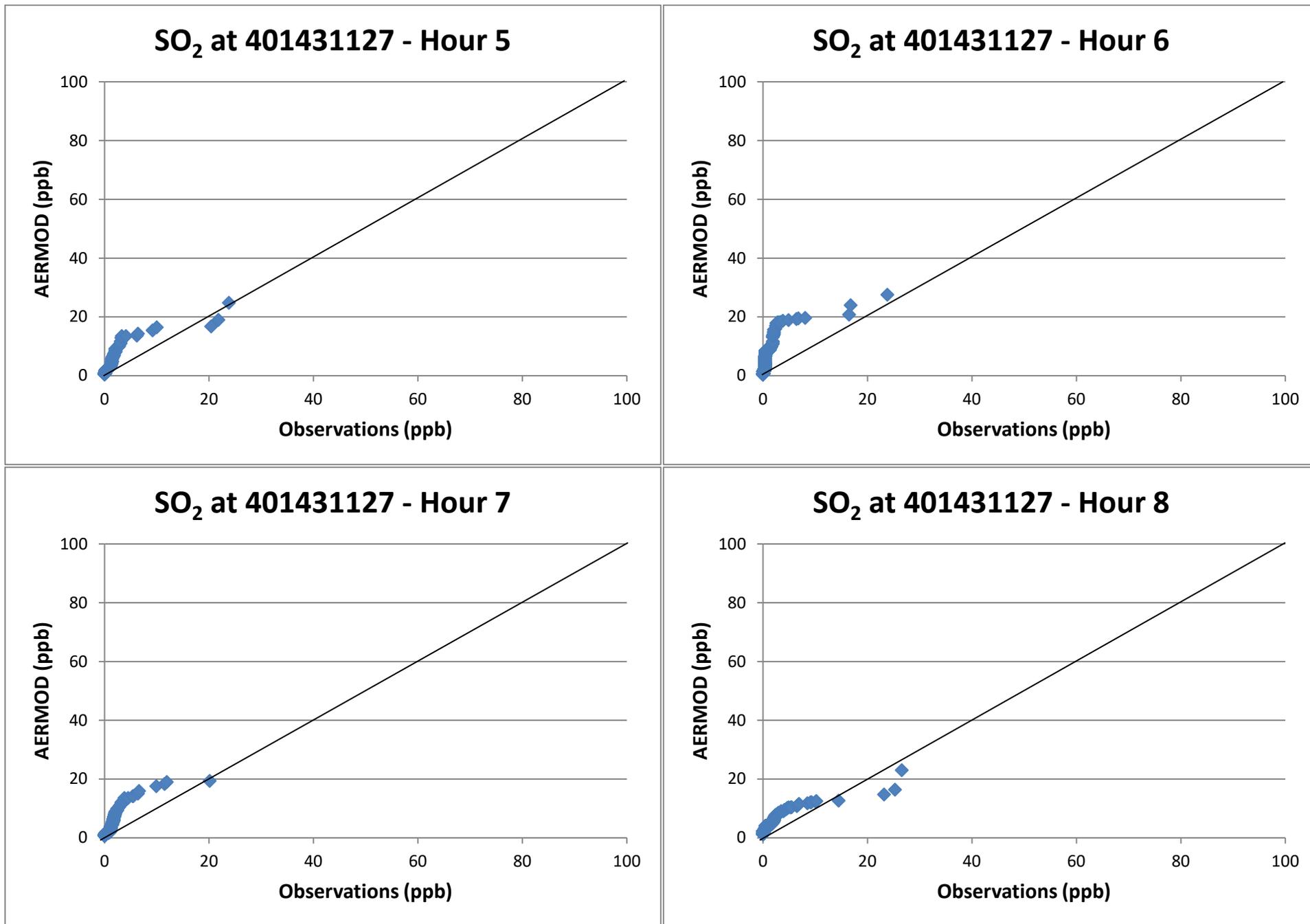
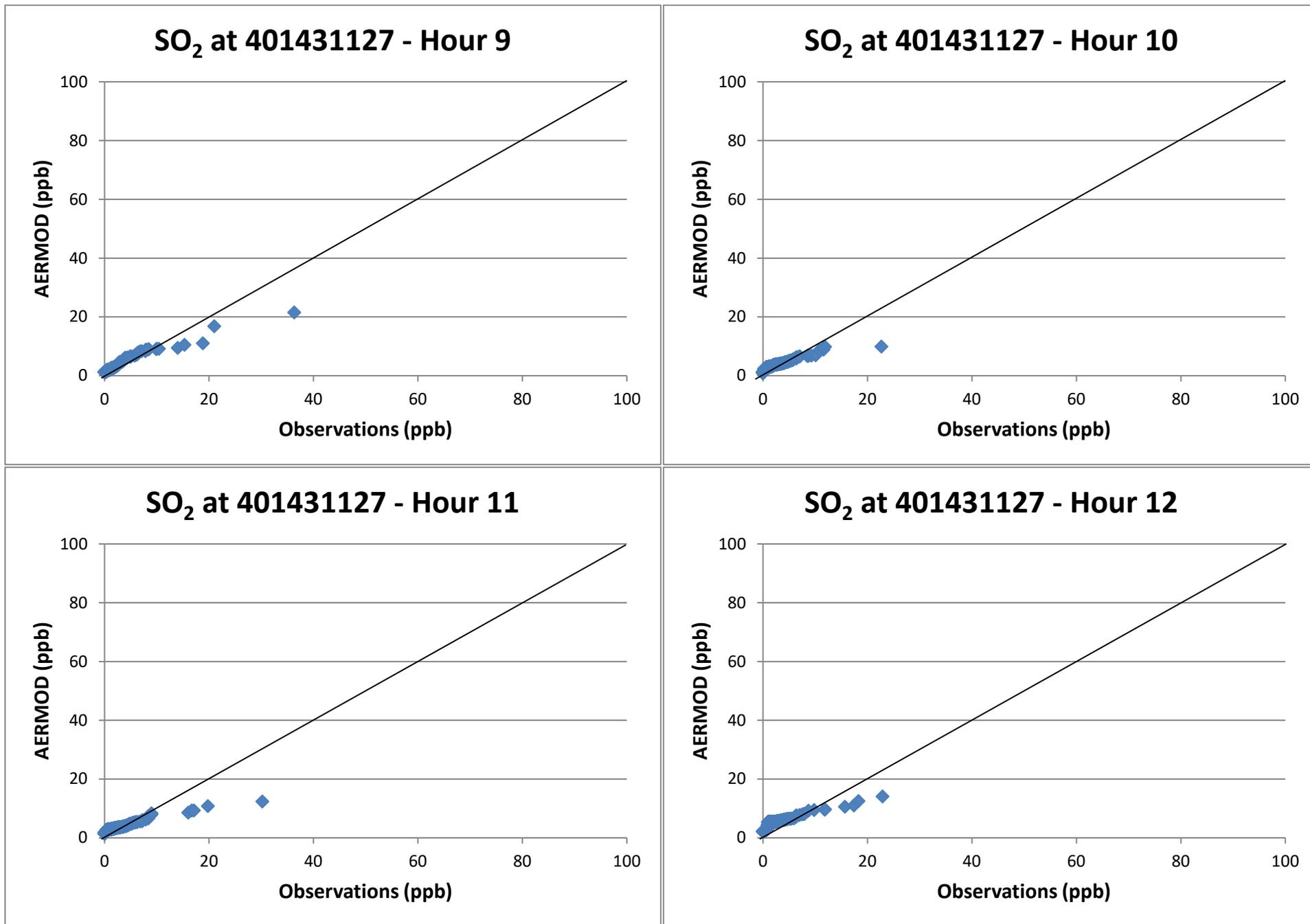
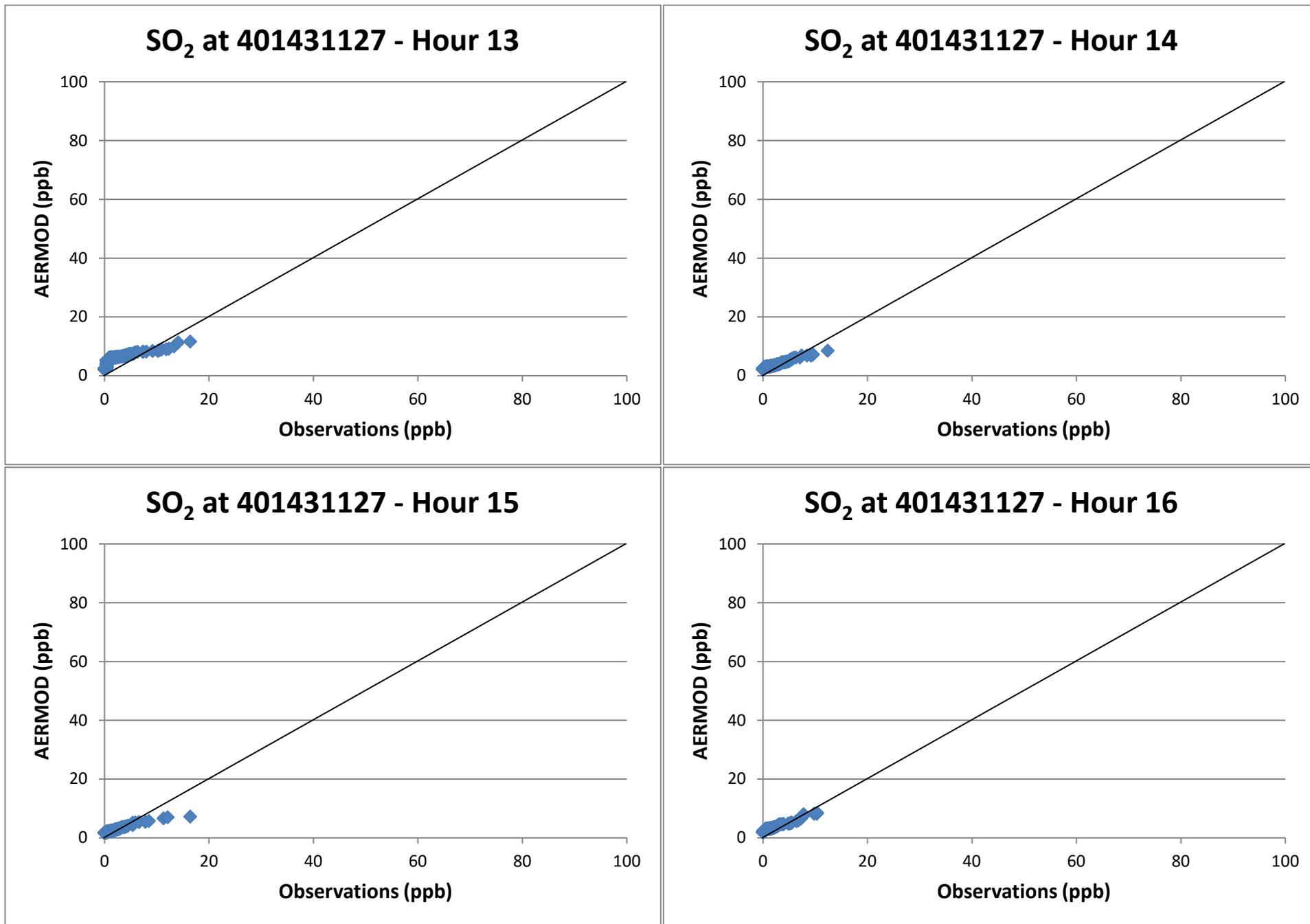


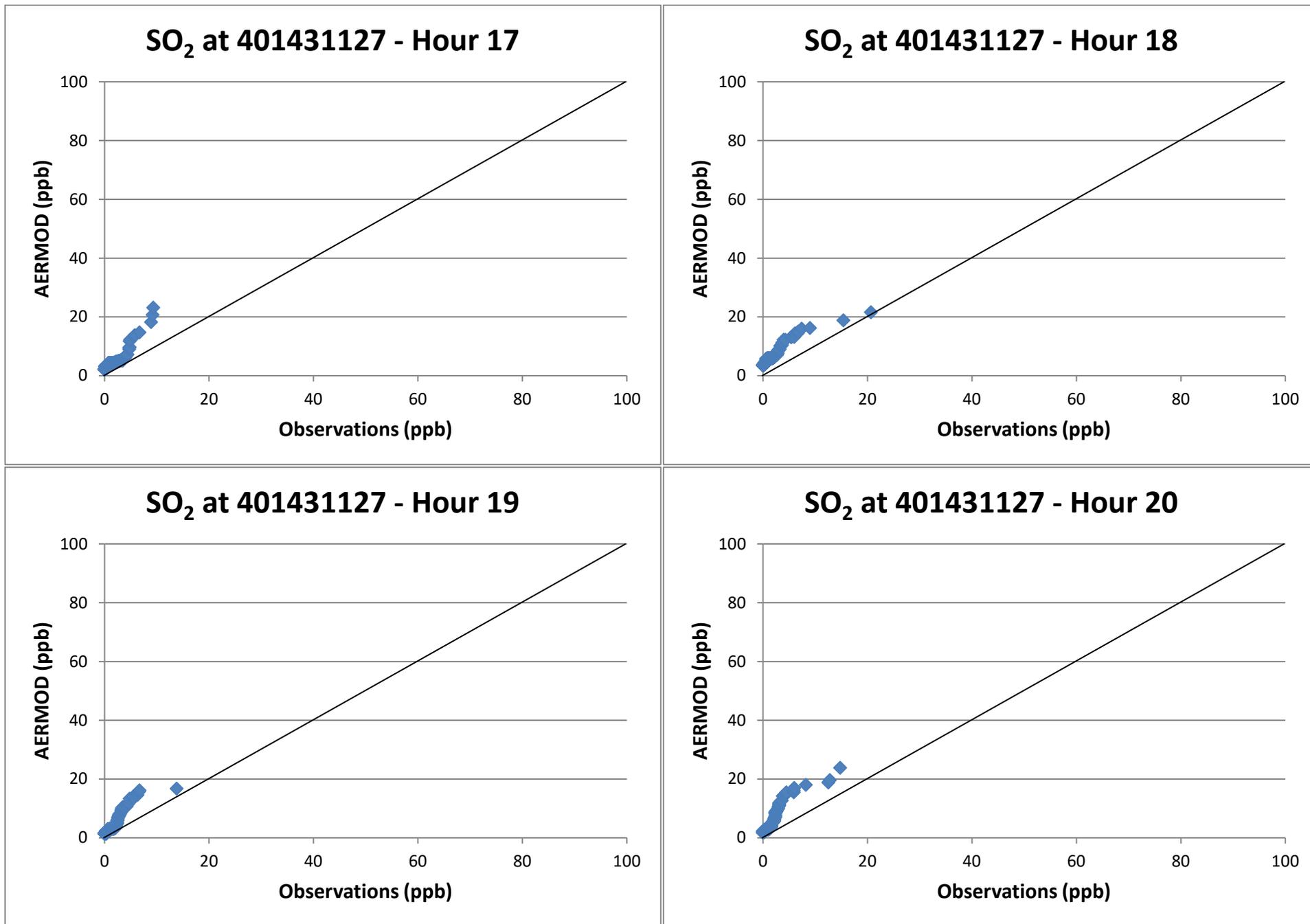
Figure 48. QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).



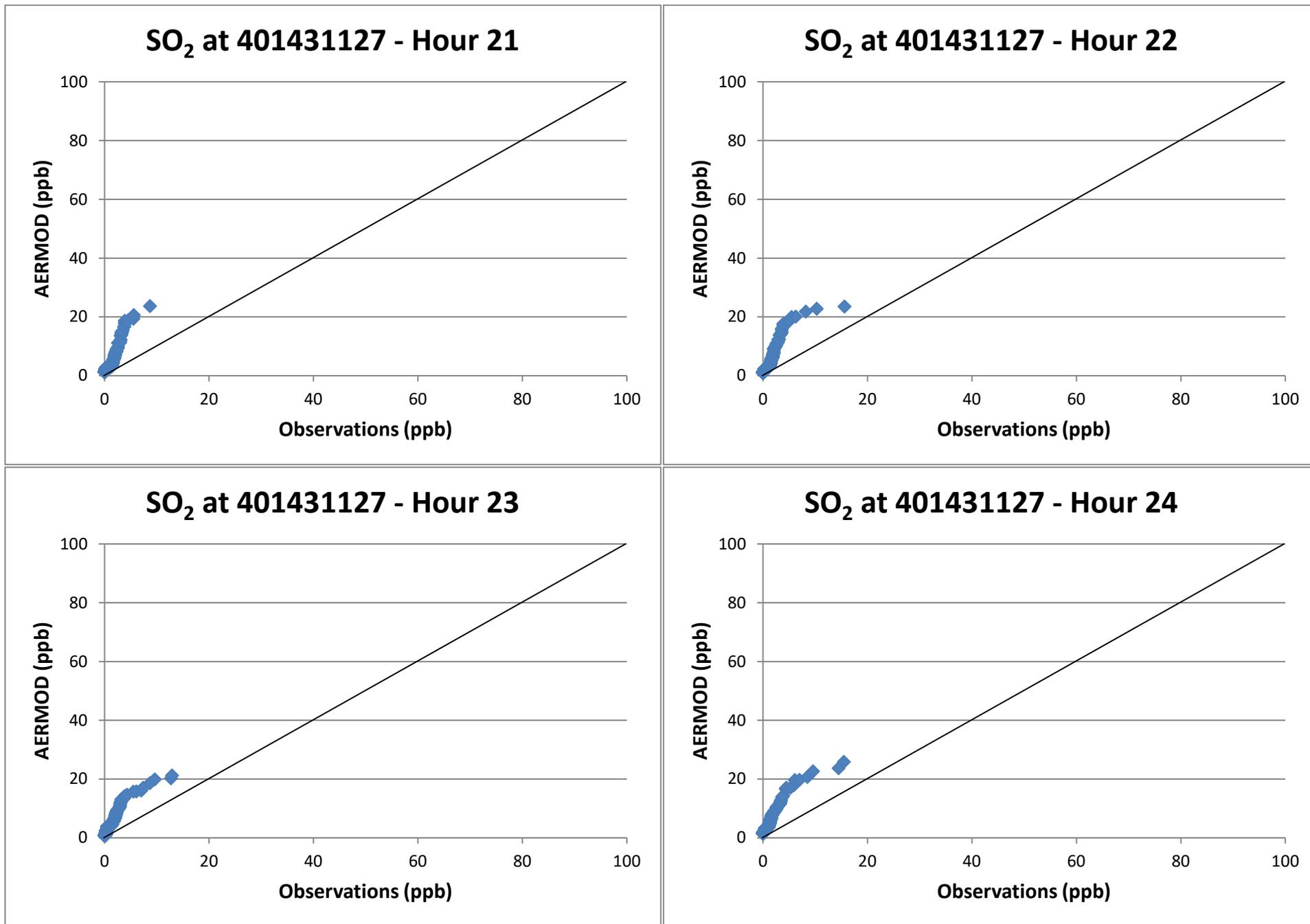
**Figure 49.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).



**Figure 50.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).



**Figure 51.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).



**Figure 52.** QQ-plots for measured 1-hour SO<sub>2</sub> concentration vs. modeled 1-hour SO<sub>2</sub> concentration by hour-of-day for Tulsa, OK (401431127).