CRC Report No. A-89-1

METHODS FOR ADJUSTING BIASES IN OZONE MODEL OUTPUTS FOR USE IN ATTAINMENT DEMONSTRATIONS AND EXPOSURE ASSESSMENTS

Deliverable 1B

February 2014

COORDINATING RESEARCH COUNCIL, INC.
5755 NORTH POINT PARKWAY·SUITE 265·ALPHARETTA, GA 30022

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The four main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance, heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); and light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars). CRC's function is to provide the mechanism for joint research conducted by the two industries that will help in determining the optimum combination of petroleum products and automotive equipment. CRC's work is limited to research that is mutually beneficial to the two industries involved, and all information is available to the public.

Methods for Adjusting Biases in Ozone Model Outputs for use in Attainment Demonstrations and Exposure Assessments

P.S. Porter(1), S.T. Rao(2), C. Hogrefe(3), E. Gego(1) and R. Mathur(3)

(1) Porter-Gego, Idaho Falls, Idaho, USA, 83401,
(2) North Carolina State University, Raleigh, NC 27695, USA,
(3) USEPA NERL, ORD, RTP, NC, USA, 27711

I. Introduction

Two uses for photochemical models in the United States are attainment demonstrations and exposure assessments in unmonitored locations. These uses are constrained by the lack of fidelity between model outputs and observations. Despite continual improvement, model predictions are imperfect, having both reducible and irreducible errors. Reducible errors (structural and parametric) are attributable to our incomplete or inadequate understanding of the relevant atmospheric processes and errors in model inputs variables (emissions, meteorology, boundary conditions). Irreducible errors arise from our inability to properly characterize the atmosphere with correct initial conditions. In addition, model estimates are grid-averages while observations are mostly point measurements from fixed surface monitoring sites.

It is desirable to bring statistical properties of model outputs into closer harmony with pollutant concentrations observed at monitoring sites so that models can more effectively be used to (1) predict future air quality conditions for a given set of inputs (attainment demonstration), and (2) predict pollutant concentrations at unmonitored locations for use in exposure estimates. The statistical properties of interest are extreme values (the design value being the 4th highest 8 hour daily maximum ozone, 8H4), and the means and standard deviations of time series.

Attainment demonstrations use photochemical models to design emission control strategies for meeting future air quality standards. The demonstration consists of a *base case* (current or past conditions) plus an estimate of future conditions (*prediction case*). Future conditions are derived from model experiments using hypothetical emission control strategies and past meteorology. The USEPA-recommended use of models in an attainment demonstration *(USEPA, 2007)* makes use of the assumption that the ratio of future to base case ozone design values is the same for model predictions and observations:

$$\boldsymbol{RRF = \frac{MF}{MB} = \frac{OF}{OB}} \qquad \textbf{(1)}$$

RRF= <u>R</u>elative <u>R</u>eduction <u>F</u>actor
MF = <u>M</u>odel prediction for <u>F</u>uture (future 8H4)
MB = <u>M</u>odel prediction for <u>B</u>ase case (past 8H4)
OB = <u>O</u>bserved <u>B</u>ase case design value (present 8H4)
OF = <u>O</u>bserved in the <u>F</u>uture (future 8H4)

$$\boldsymbol{DVF = RRF \bullet OB} \qquad \textbf{(2)}$$

DVF = <u>D</u>esign <u>V</u>alue for <u>F</u>uture

RRF is a bias adjustment that assumes model predictions can be used in a relative sense, that is, even though model outputs are biased they can be used to predict change from a given condition.

Model predictions for the future are usually based on a combination of past meteorology and future emission scenarios designed to reduce ozone concentrations in non-attainment areas. This combination of meteorology and emissions usually can't be verified because it did not occur (there are no comparable observations). Verification requires the use of emission inventories estimated from past conditions, a scenario that may not be of interest to policy-makers.

Ozone model predictions are also used to estimate human and ecosystem exposure in unmonitored locations. One objective of an exposure study is to link human health outcomes to ozone concentrations via a mathematical model. For example, the Cox proportional hazards model links survival time to pollutant exposure as follows:

$$\boldsymbol{\gamma = \frac{1}{\lambda_o \exp\left[\beta\,(O3 - \overline{O3})\right]}} \qquad \textbf{(3)}$$

γ = expected survival time (years) of population members
O3 = exposure at a particular location
$\overline{O3}$ = mean regional exposure
λ_o = baseline incidence rate
β = relative risk

It is desirable to be able to use model predictions at unmonitored locations to estimate the relative risk (**β**) parameter of equation (3). However, unbiased estimates of **β** require unbiased ***standard deviation*** estimates from the model (Porter etal, 2012).

In this paper we assess four new methods for using regional photochemical models in attainment demonstrations and verify them with a long-term model experiment. Included are adjustments to time series (matching the mean and variance of model time series to observations) and CDF matching (matching sample CDFs of modeled and observed concentrations, disregarding time sequence). We also estimate the ability of these methods to produce contemporaneous predictions, that is, same-year values. We compare the new approaches with the RRF and with unadjusted model values. Performance metrics are provided for all.

II. Data and Methods

A. Model Setup

CMAQ (Northeastern US, 1988-2005, source Hogrefe et al, 2009)(Figure 1)

Model:	MM5v3.7.2 CMAQv4.5.1, CB4, aero3, years 1988 –2005.
Emissions:	NEI1990, 1996-2001, OTC2002, OTC2009 (SMOKE).
Domain:	Northeastern U.S., 36 km / 12 km (details in Hogrefe et al, 2009)

B. Observations

Hourly, 1996-2005
USEPA Air Quality System (AQS) site:
http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm

Sites in the model domain that were utilized are 80% complete for a given ozone season (ozone Season: 1 May – 30 September, 153 days). There are approximately 250 monitoring sites in the model domain for any given year.

C. Adjustment Methods

1. Mean and variance matching

For each ozone season, adjust the mean and variance of the modeled 8H ozone time series so that they match those observed:

$$\boldsymbol{future\ prediction} = \boldsymbol{CMAQ\ future} + \tag{4}$$

$$\{\boldsymbol{CMAQ\ future} - \boldsymbol{mean\ CMAQ\ future}\} \bullet \frac{\boldsymbol{\sigma(OBSERVED\ base\ year)}}{\boldsymbol{\sigma(CMAQ\ base\ year)}}$$

$$+ \ \boldsymbol{mean}\,\{\boldsymbol{OBSERVED\ base\ year} - \boldsymbol{CMAQ\ base\ year}\}$$

Base year adjustment factors (ratio of standard deviations and the mean bias) computed for each ozone season (153 days) are applied to other ozone seasons (***prediction*** years). The ***prediction*** year 4th highest 8H value is then compared with that observed. It is implicitly assumed that ratio of standard deviations and the bias computed for the base year will be valid for prediction years.

2. Mean and variance matching with temporal components

Mean and variance matching are applied to temporal components of model and observed values. Temporal components in this case consist of high- and low-frequency components (*'HF'* and *'LF'*, respectively) defined by FIR filters:

$$\boldsymbol{OBSERVED\ (base\ year) = OBSERVED\ (LF, base\ year) + OBSERVED(HF, base\ year)} \quad \textbf{(5)}$$

$$\boldsymbol{CMAQ \quad = CMAQ(LF) + CMAQ(HF)} \quad \textbf{(6)}$$

The principal behind this approach is that low- and high- frequency processes are driven by different phenomenon, with high-frequency variation attributable mostly to synoptic scale weather, and low-frequency variation driven by seasonal emissions. It follows that components with different driving forces should have different adjustments. Therefore, adjust the mean and variance of each component of the modeled 8H ozone time series so that they match that observed:

$$\boldsymbol{future\ (LF) = CMAQ\ future\ (LF) +} \quad \textbf{(7)}$$

$$\boldsymbol{\{CMAQ\ future\ LF - mean\ CMAQ\ future\ LF\} \bullet \frac{\sigma(OBSERVED\ base\ year\ LF)}{\sigma(CMAQ\ base\ yearLF)}}$$

$$\boldsymbol{+\ mean\ \{OBSERVED\ base\ year\ LF -\ CMAQ\ base\ year\ LF\}}$$

$$\boldsymbol{future\ (HF) = CMAQ\ future\ (HF) +} \quad \textbf{(8)}$$

$$\boldsymbol{\{CMAQ\ future\ HF - mean\ CMAQ\ future\ HF\} \bullet \frac{\sigma(OBSERVED\ base\ year\ HF)}{\sigma(CMAQ\ base\ yearHF)}}$$

$$\boldsymbol{+\ mean\ \{OBSERVED\ base\ year\ HF -\ CMAQ\ base\ year\ HF\}}$$

$$\boldsymbol{future = future\ (LF) + future\ (HF)} \quad \textbf{(9)}$$

As with method 1, method 2 adjusts the time series (not just 4^{th} highest) and assumes base year *variance ratios* and *mean bias* will be the same in future years for both high- and low- frequency components.

3. Cumulative Distribution Function matching (CDF matching 1)

Observations and model estimates are sorted and ranked to establish sample CDF's. The **CMAQ** CDF is modified so that it will match the **OBSERVED** CDF as closely as possible. The linear coefficients identified for the base year are then used to modify the CDF's of all other years.

$$\textbf{\textit{OBSERVED base year QUANTILES}} = \textbf{\textit{Ko}} + \textbf{\textit{K1}} \bullet \textbf{\textit{CMAQ base year QUANTILES}} \qquad (10)$$

$$\textbf{\textit{future QUANTILES}} = \textbf{\textit{Ko}} + \textbf{\textit{K1}} \bullet \textbf{\textit{CMAQ future QUANTILES}}$$

The parameters Ko and K1 rotate and displace the CMAQ CDF such that the root mean squared distance between modeled and observed quantiles are minimized. The original time order of **CMAQ** and observations is lost. As with the other methods, the values of Ko and K1 estimated for the base year are applied to future years.

4. Cumulative Distribution Function matching (CDF matching 2)

$$\textbf{\textit{OBSERVED base year QUANTILES}} = \textbf{\textit{Ko}} + \textbf{\textit{K1}} \bullet \textbf{\textit{CMAQ base year QUANTILES}} \qquad (11)$$

$$\textbf{\textit{future prediction 4}}^{th} \textbf{\textit{ highest}} = \{\textbf{\textit{CMAQ future 4}}^{th} \textbf{\textit{ highest}} - \textbf{\textit{CMAQ base year 4}}^{th}\textbf{\textit{highest}}\}$$

$$\textbf{\textit{Ko}} + \textbf{\textit{K1}} \bullet \textbf{\textit{CMAQ base year QUANTILES}}$$

As with Method 3, the time order is lost. In contrast with Method 3, the parameters Ko and K1 are applied only to the base year.

D. Performance metrics

1. Mean bias (MB) and relative (or fractional) mean bias (RMB)

Mean bias averages the deviations from observed:

$$\boldsymbol{MB = \sum \frac{\{ESTIMATE - OBSERVED\}}{n}} \qquad \textbf{(12)}$$

Positive deviations cancel negative deviations, so it is possible to have a small MB but also have large errors. In calculating the relative mean bias, also referred to as fractional mean bias, divide each term in equation (12) by the **OBSERVED** value.

2. Mean absolute bias (MAB) ***and relative (or fractional) mean absolute bias (RMAB)***

Mean absolute bias averages the absolute deviations:

$$\boldsymbol{MAB = \sum \frac{|ESTIMATE - OBSERVED|}{n}} \quad \textbf{(13)}$$

Positive and negative deviations do not cancel but systematic positive or negative biases cannot be discerned. In calculating the relative mean absolute bias divide each term in equation (13) by the **OBSERVED** value.

3. Root mean squared error (RMSE) and relative (or fractional) root mean square error (RRMSE)

The root mean squared error averages the squares of deviations about the observed values:

$$\boldsymbol{RMSE = \left[\sum \frac{\{ESTIMATE - OBSERVED\}^2}{n}\right]^{1/2}} \quad \textbf{(14)}$$

The mean squared error can be thought of as a sum of variances of the estimates and observations plus the squared differences in the means of the estimates and observations. Together with bias, errors can be assigned to the spread of the deviations and differences in mean values.

4. Correlation

Pearson's (linear) correlation coefficient is given by:

$$\mathbf{R} = \boldsymbol{\frac{1}{n-1} \sum \frac{\{x - \bar{x}\} \bullet \{y - \bar{y}\}}{\sigma_x \bullet \sigma_y}} \quad \textbf{(15)}$$

Where ***x*** and ***y*** are estimates and observations, respectively, and **R** is an indication of pattern agreement between two variables (usually time series in this report) (bias in the mean and/or standard deviation plays no role in **R**).

5. Fraction of cases within 10% of observed

The fraction sites with adjusted ozone values within 10% of the observed value were calculated.

III. Application of Adjustment Methods: Examples for a Single Site

In this section, each method is illustrated using a single site (site 360130005, northwestern Pennsylvania) with base and future years 2000 and 2005, respectively. These examples consider only one site and are not necessarily representative of all sites. It will be shown in section IV that results considering all sites are highly variable.

A. CMAQ Outputs

Time series plots of **CMAQ** and **OBSERVED** base and future years are shown in Figures 2 and 3, respectively. The future (2005) 8H4 **CMAQ** value is 74.2, while the **OBSERVED** 8H4 is 77.8.

B. RRF

The RRF is given by (8H4):

$$\boldsymbol{RRF = \frac{74.2}{73.6} = 0.991 = \frac{OF}{83.2}} \qquad \textbf{(16)}$$

RRF= Relative Reduction Factor

MF = Model prediction for Future (future design value, 8H4)
MB = Model prediction for Base case (past design value 8H4)

OB = Observed Base case design value (present design value 8H4)
OF = Observed in the Future (future design value 8H4)

$$\boldsymbol{DVF = 0.991 \bullet 83.2 = 82.4} \qquad \textbf{(17)}$$

The **OBSERVED** future (2005) design value is 77.8 leaving an error of +4.6 ppb (5.9%).

Figure 4 shows the RRF (future CMAQ/base year CMAQ) plotted vs. the observed ratio (future observation/base year observation) for all sites for the 2000 base year and 2005 future year. In this pair of years the assumption that changes in the model are matched by changes in the observations doesn't hold up very well.

C. Mean and variance matching

Adjusted CMAQ is shown along with the original CMAQ and observations for the base year in Figure 5. Notice that the mean and variance of 'adjusted CMAQ' match those observed (see the caption and Table 1), and that the standard deviation adjustment is greater than 1 (14.4/8.9). Adjusted values are close to some of the higher observed values (for example the value above 90 ppb about day 40), but not others (for example the observed value greater than 80 ppb on about day 140). What one is really after for this application is the 4th highest value, which is more easily seen in a QQ plot (Figure 6). Notice that the adjusted CMAQ 8H4 (82.5) is closer to OBSERVED (83.2) than is CMAQ (73.5). The error is -0.7 ppb, -0.8%.

The same variables are shown for the prediction year of 2005 (Figure 7). Notice (in the caption and Table 1) that this time the mean and variance of the adjusted CMAQ do not match those observed. A QQ plot (Figure 8) shows that the 8H4 adjusted CMAQ (84.4) is not closer to observed (77.8) than CMAQ (74.2). The error is 6.6 ppb, 8.5%.

D. Mean and variance matching of temporal components

Base year low- and high- frequency components are shown in Figure 9 and 10, respectively. The standard deviations of both CMAQ components are smaller than observed (see figure caption). Notice that the adjusted components are closer to **OBSERVED** than is **CMAQ**. It is evident that the sum of the adjusted low- and high-frequency components are closer to the observations than is the original **CMAQ** output (Figure 11) for this same-year comparison. From the QQ plots of low- and high-frequency components and their sum (Figures 12 - 14, respectively), one can see that adjusting the variance upward has the effect of rotating the component **CMAQ** QQ plots counter-clockwise until they match what is **OBSERVED.**

Prediction-year (2005) comparisons of components can be found in Figures 15-17. As in the year 2000, the adjustment to the low-frequency component brings the standard deviation closer to what is **OBSERVED.** However, the mean is over-corrected and is farther from the **OBSERVED** mean than is **CMAQ.** With respect to the high-frequency component, the variance is over-adjusted. Mean values for this component are always near zero for **OBSERVED** and **CMAQ.** The 8H4 for the sum of components, the metric we are really after, has moved closer to what is **OBSERVED** than **CMAQ** (Table 1). Quantile plots for 2005 components and their sum are shown in Figures 18-20. The 8H4 error is 1.5 ppb or 1.9%.

E. CDF matching 1

CDF matching is simpler than temporal components. The principle behind CDF matching is that a summer season of 8H ozone measurements is a random sample from a particular CDF, and the 8H4 is an extreme value from this distribution. Therefore, matching the **CMAQ** and **OBSERVED** CDF's will lead to agreement with the 8H4. The base year QQ plot (Figure 21) shows that the adjusted CMAQ QQ plot lies directly on top of that observed (by design); in this case the adjustment brings the adjusted 8H4 closer to what is **OBSERVED**. The same adjustment is applied to the 2005 prediction year (Figure 22). In the case of 2005, the 8H4 is over-adjusted so that it is not closer to **OBSERVED** than **CMAQ**.

F. CDF matching 2

'CDF matching 2' adds the base-year adjusted CMAQ CDF to the difference in 8H4 between prediction- and base- year 8H4 values:

$$\textit{prediction 8H4} = \textit{adjusted CMAQ base year 8H4} + \{\textit{CMAQ prediction 8H4} - \textit{CMAQ base year 8H4}\} \qquad (18)$$

Adjustment factors are not used for prediction years. For this example, the adjusted 8H4 is given by (see Table 1):

$$\begin{aligned} \textit{8H4(2005)} &= \textit{8H4(adjusted 2000)} + \{\textit{CMAQ 8H4 2005} - \textit{CMAQ 8H4 2000}\} \qquad (19) \\ &= 82.3 + \{74.2 - 73.5\} \\ &= 83.0 \end{aligned}$$

The error is 5.2 ppb or 6.7%.

Table 1. Summary of examples

	method	mean	std. dev.	8H4
base year 2000	observed	50.0	14.4	83.2
	CMAQ	53.4	8.91	73.5
	RRF	NA	NA	83.2
	MEAN VAR. MATCH	50.0	14.4	82.5
	TEMP. COMP.	50.0	14.3	82.3
	CDF MATCH 1	50.0	14.3	82.3
	CDF MATCH 2	NA	NA	82.3
future year 2005	observed	52.0	13.4	77.8
	CMAQ	52.2	10.5	74.2
	RRF	NA	NA	82.4
	MEAN VAR. MATCH	48.8	17.0	84.4
	TEMP. COMP.	48.7	14.5	79.3
	CDF MATCH 1	48.0	16.9	83.4
	CDF MATCH 2	NA	NA	83.0

IV. Results for all Sites

A. Predicting design values

For every pair of years from 1996-2005, one year served as a base year and the other a prediction year. There are a total of 90 pairs where the two years are different. For each base/prediction year pair there are roughly 200 sites from which performance metrics (described in section II) were computed.

Figure (23) shows the RMSE for base year 1996 and prediction years 1997-2005. RMSE values range from about 6 to 15 ppb. No one method is the best for all the prediction years, but ***CDF MATCH 2*** is the best most often (6 years). ***TEMP.COMP*** RMSE values are much higher than for the other methods for this base year (results for all base year/prediction year pairs are summarized below).

Poor performance for all methods occurred for prediction year 2002, which happens also to be a year with high ozone (Figure 24). Notice also from Figure 24 that CMAQ ozone has a decreasing temporal trend whereas OBSERVED ozone does not. Model performance tends to improve with increasing ozone concentrations (Figure 25). Since this is a period of decreasing emissions (Figure 26), decreasing trends in ozone accompanied by decreasing performance in CMAQ might be expected. Relative RMSE eliminates ozone level as a factor explaining performance (Figure 27) and leaves the RMSE of the ***TEMP.COMP*** method with an upward temporal RMSE trend.

MAB for the same set of years is shown in Figure 28. Again, no one method is the best for all the prediction years, though ***CDF MATCH 2*** is the best most often (6 years). ***TEMP. COMP*** no longer seems much worse than the other methods. The year 2002 is the poorest performing for all methods. Temporal patterns in RMAB (Figure 29) appear to be similar to those for MAB, and this time there is no temporal trend in ***TEMP.COMP*** performance.

Correlations (R) for same set of years are shown in Figure 30 (all are significant at the 5% level). As before, no one method is the best for all the prediction years (***CDF MATCH 2*** is best for 6 years). All methods have an apparent downward temporal trend (best correlation for prediction year 1997). Domain-wide relative correlation is shown in Figure 31.

From a spatial image of the RMSE for all base year/prediction year pairs (Figure 32) there seems to be poorer performance for all methods along sea and lake coastal areas. For CMAQ there also seems to be poorer performance in Ohio.

Table 2 shows 'best' results for all 90 base/prediction year pairs, which is a count of the number of times a given method had the best metric. Using this criterion the best method is ***CDF MATCH 2***, which has the best RMSE, MAB and R^2 nearly half the time. The ***TEMP.COMP*** method is second, being the best about 16-20% of the time. ***RRF*** is best only about 2 percent of the time, while the occurrence of best ***CMAQ*** and ***MEAN/VAR MATCH*** range from 11.1 to 15.6% of the time.

Metrics for all base/prediction year pairs are shown in Table 3. Here we see that while some methods were not often the best, differences in performance are often small. For example, ***RRF*** has an RMSE of 9.4 ppb compared with ***CDF MATCH 2*** of 8.6 ppb. The best methods for MAB, RMSE and R^2 are ***CDF MATCH 2***, ***CDF MATCH 2***, and ***MEAN&VAR MATCH***, respectively.

Table 4 shows the percentage sites where the estimate and observed values differ by less than 10%. CMAQ is the winner here, being within 10% of observations at 92.6% of the sites. The rate for the other methods range from 81% to 86.5%.

Table 2. Percent of base- target-year pairs with best result

Metric	**Method:** RRF	CMAQ	MEAN&VAR MATCH	TEMP. COMP.	CDF MATCH 1	CDF MATCH 2
MB	**11.1**	**13.3**	**25.6**	**34.4**	**7.8**	**7.8**
MAB	**2.2**	**12.2**	**13.3**	**17.8**	**5.6**	**48.9**
RMSE	**2.2**	**15.6**	**11.1**	**20.0**	**9.9**	**42.2**
RMB	**11.1**	**14.4**	**26.7**	**33.3**	**6.7**	**7.8**
RMAB	**3.3**	**12.2**	**11.1**	**16.7**	**4.4**	**52.2**
RRMSE	**2.2**	**11.1**	**15.6**	**16.7**	**6.7**	**47.8**
R2	**1.1**	**12.2**	**13.3**	**16.7**	**8.9**	**47.8**

Table 3. Mean metric of all base- prediction- year pairs

Metric	Method: RRF	CMAQ	MEAN&VAR MATCH	TEMP. COMP.	CDF MATCH 1	CDF MATCH 2
MB	0.067	-4.7	-1.2	0.47	0.0046	-0.068
MAB	7.5	8.0	7.2	7.1	7.0	6.9
RMSE	9.4	9.8	7.0	10.0	8.8	8.6
RMB	0.0061	-0.051	-0.0125	0.016	0.0039	0.0042
RMAB	0.087	0.092	0.084	0.082	0.081	0.081
RRMSE	0.111	0.114	0.106	0.115	0.105	0.103
R2	0.33	0.36	0.37	0.36	0.39	0.42

Table 4. Percent of Sites within 10% of OBSERVED, all Years

Metric	Method: RRF	CMAQ	MEAN&VAR MATCH	TEMP. COMP.	CDF MATCH 1	CDF MATCH 2
MAB	81.0	92.6	86.5	81.4	81.0	83.4

B. Same-year performance

Same-year performance compares base-year time series of adjusted 8H ozone with base-year observed time series (i.e., 8H4 comparisons aren't made and there are no 'future' predictions of 8H4). Same-year comparisons are useful for contemporaneous exposure assessments. Adjusted model values would be spatially interpolated to the entire domain. The domain-wide MAB, RMSE and correlation (Figures 33–35, respectively) show ***TEMP.COMP*** and ***CDF MATCH 1*** outperforming the other methods (***RRF*** and ***CDF MATCH 2*** aren't shown because they adjust only 8H4 values, not the entire time series). As noted above, an important criteria for fitting exposure models (i.e., equation 3), is that the standard deviation of the adjusted values matches those observed. The same year standard deviations for **CMAQ**, ***CDF MATCH 1***, and **OBSERVED** are shown in Figure 36 (the **MEAN/VAR MATCH** and **TEMP.COMP.** methods automatically have the same variance as observed). Conclusions from these results are that raw CMAQ values shouldn't be used for exposure studies, and that ***CDF MATCH 1*** has, on average, standard deviations very close to what are observed.

V. Summary

This report evaluates methods for reducing bias in CMAQ ozone predictions so that they can be more effectively used in attainment demonstrations and exposure studies. Attainment demonstrations need unbiased estimates of design values (8H4) for future years. Exposure assessments require contemporaneous unbiased standard deviation estimates.

Four methods were developed and tested along with raw CMAQ and RRF values, using a long-term simulation study covering the northeastern US. Three of the methods project base-year CMAQ bias metrics to prediction-years. The projected metrics are (1) mean and variance, (2) mean and variance of temporal components, and (3) regression parameters of a QQ plot. The fourth method uses QQ regression parameters only for the base year.

Adjustment of modeled values so that they more closely resemble observations usually improves performance in attainment demonstrations. Conclusions about performance depend on the metric used. Simple counts of best base year/prediction year (Table 2) show that methods using base year parameters (***MEAN/VAR MATCH*** and ***TEMP.COMP***) to adjust prediction year model values are often less effective than methods that do not (***CDF MATCH 2***). This means that the characteristics of CMAQ bias are not stationary and likely depend on emissions and meteorology.

Domain-wide mean metrics (Table 3) indicate ***CDF MATCH 2*** has the smallest MAB, RMAB, and RRMSE, and highest R^2. ***CDF MATCH 1*** has the smallest MB and RMB, while ***MEAN/VAR MATCH*** has the smallest RMSE. Differences among the methods for this set of metrics tend to be small. For all methods estimated values are within +/- 10% of the observed value in more than 80% of cases.

Same-year performance is best for ***TEMP.COMP*** and ***CDF MATCH 2*** (Figures 33-35). Raw CMAQ outputs should not be used for exposure modeling because they will produce biased parameter estimates.

VI. Literature Cited

Hogrefe, C., B. Lynn, R. Goldberg, C. Rosenzweig, E. Zalewsky, W. Hao, P. Doraswamy, K. Civerolo, J. Ku, G. Sistla, P. Kinney (2009) A combined model-observation approach to estimate historic gridded fields of PM2.5 mass and species concentrations, *Atmospheric Environment* **43**:2561-2570.

Porter, P.S., Gego, E., Garcia, V. and S.T. Rao. 2012. Benefits of using enhanced air quality information in human health studies. Air Pollution Modeling and its Application XXII, Steyn, D. and Castelli, S.T., eds, Elsevier, The Netherlands

USEPA. 2007. Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze. EPA -454/B-07-002, April

FIGURES

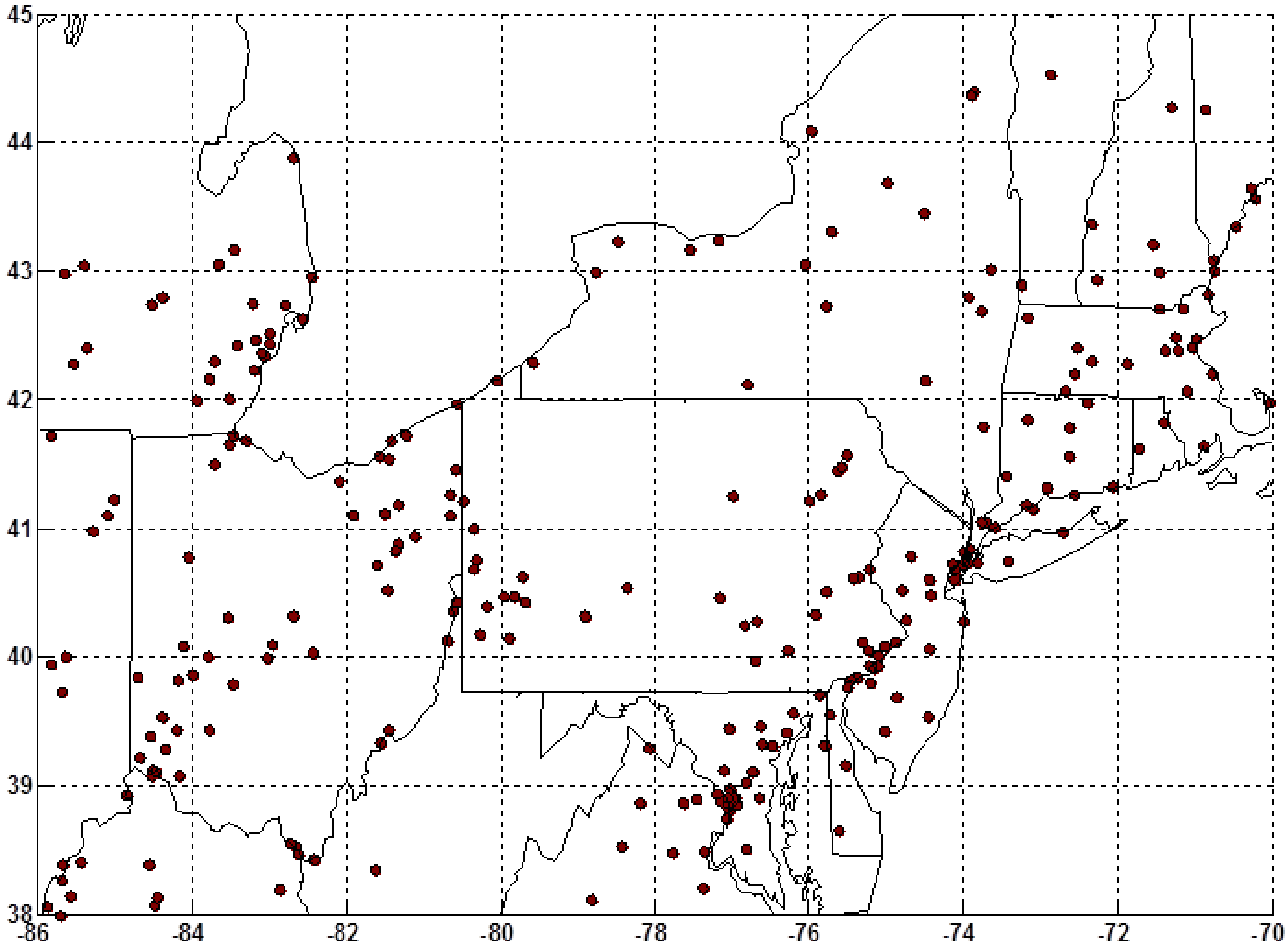


Figure 1. Model domain and monitoring sites

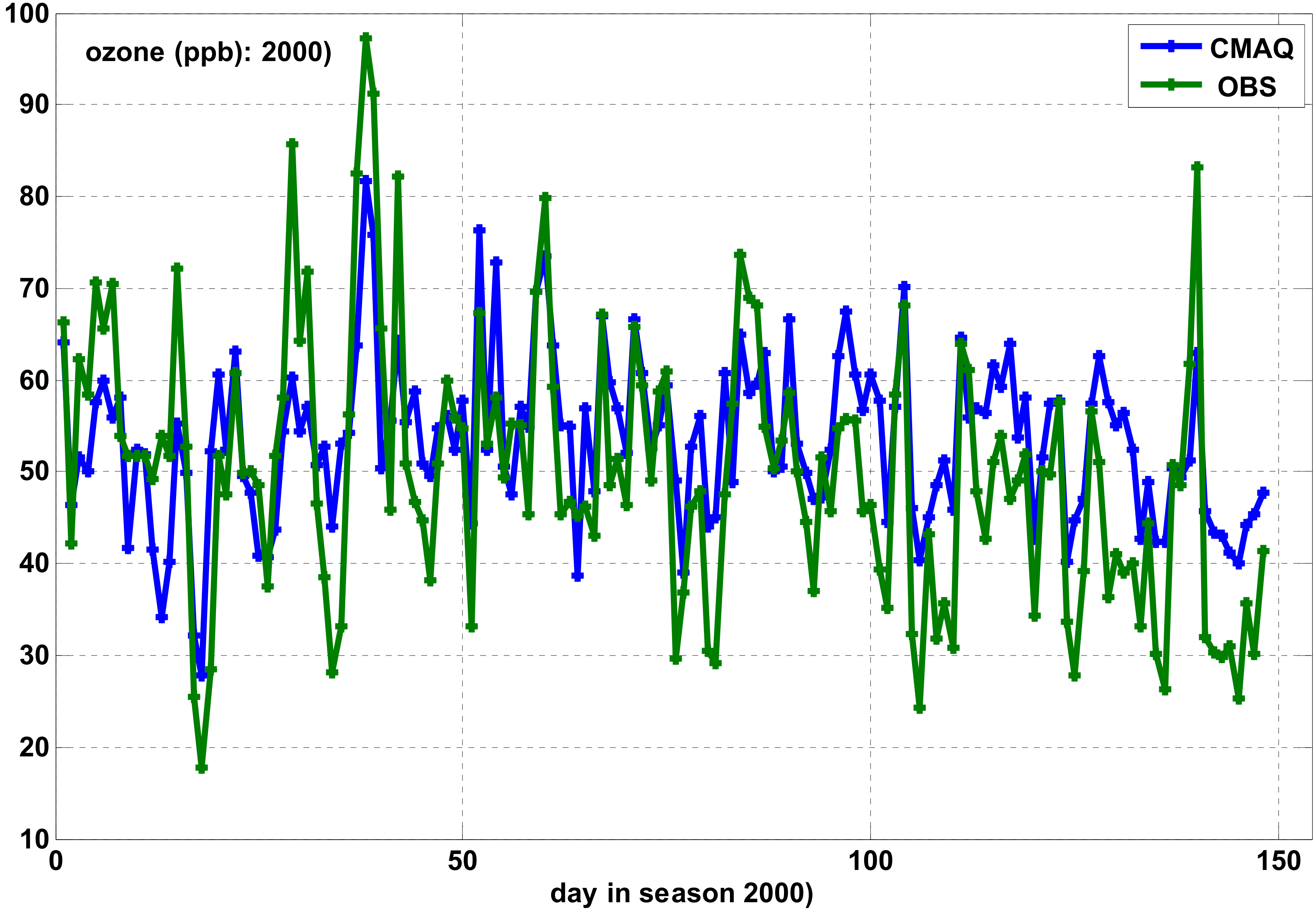


Figure 2. Base year (2000) ozone

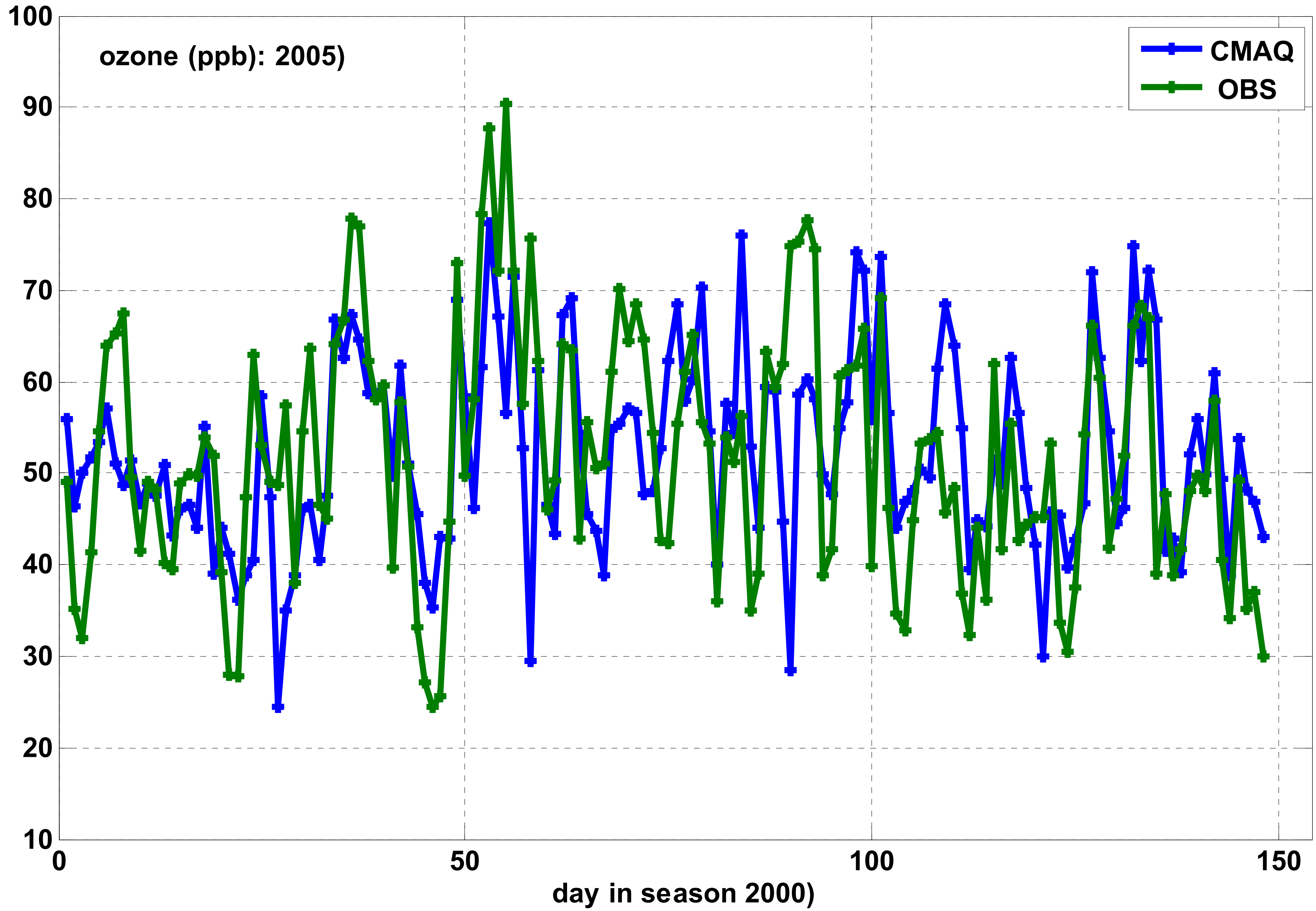


Figure 3. Prediction year ozone

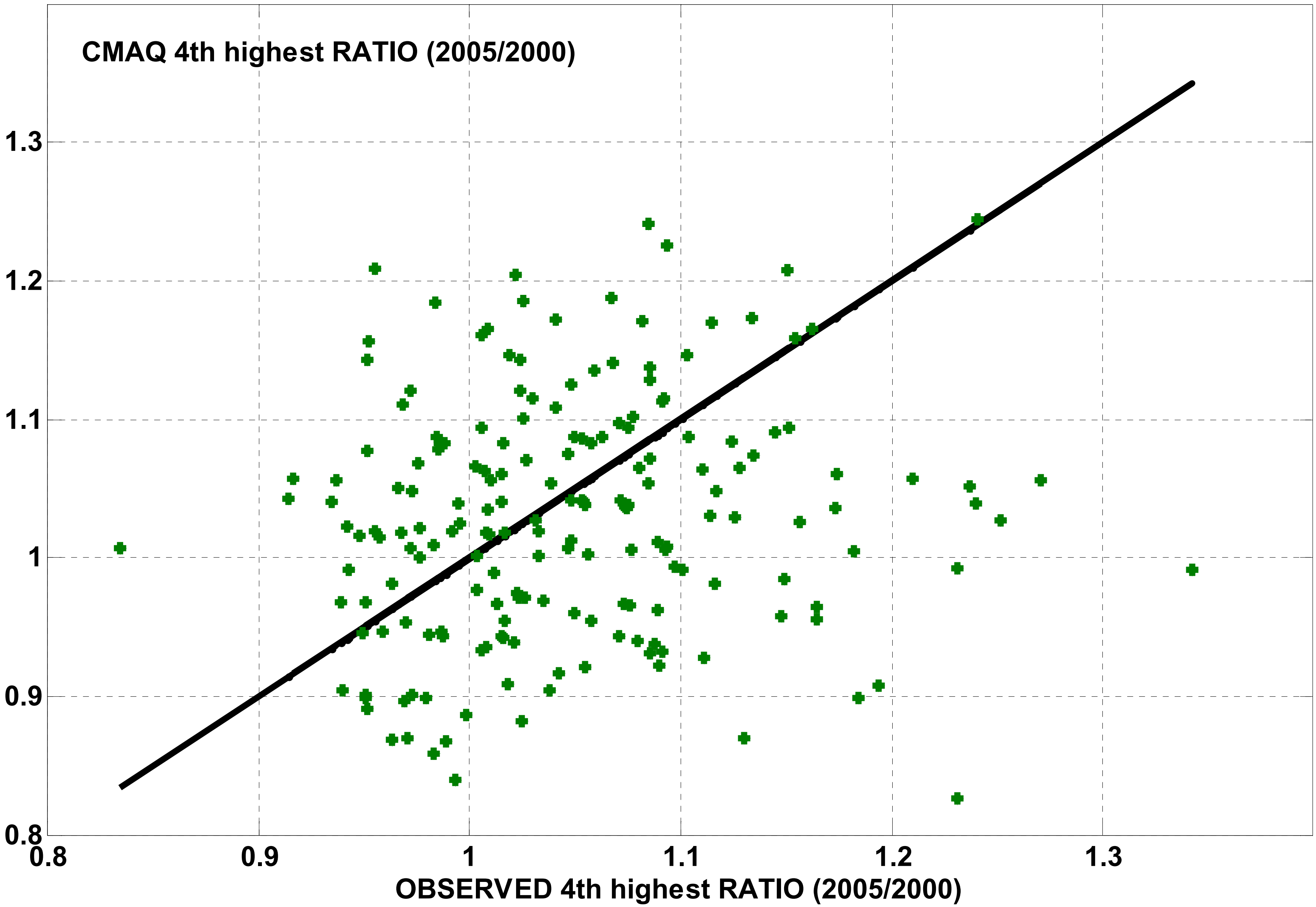


Figure 4. RRF (4th highest values, future CMAQ/base year CMAQ) plotted against the ratio for OBSERVED (future / base year)

EXAMPLES: SINGLE SITE

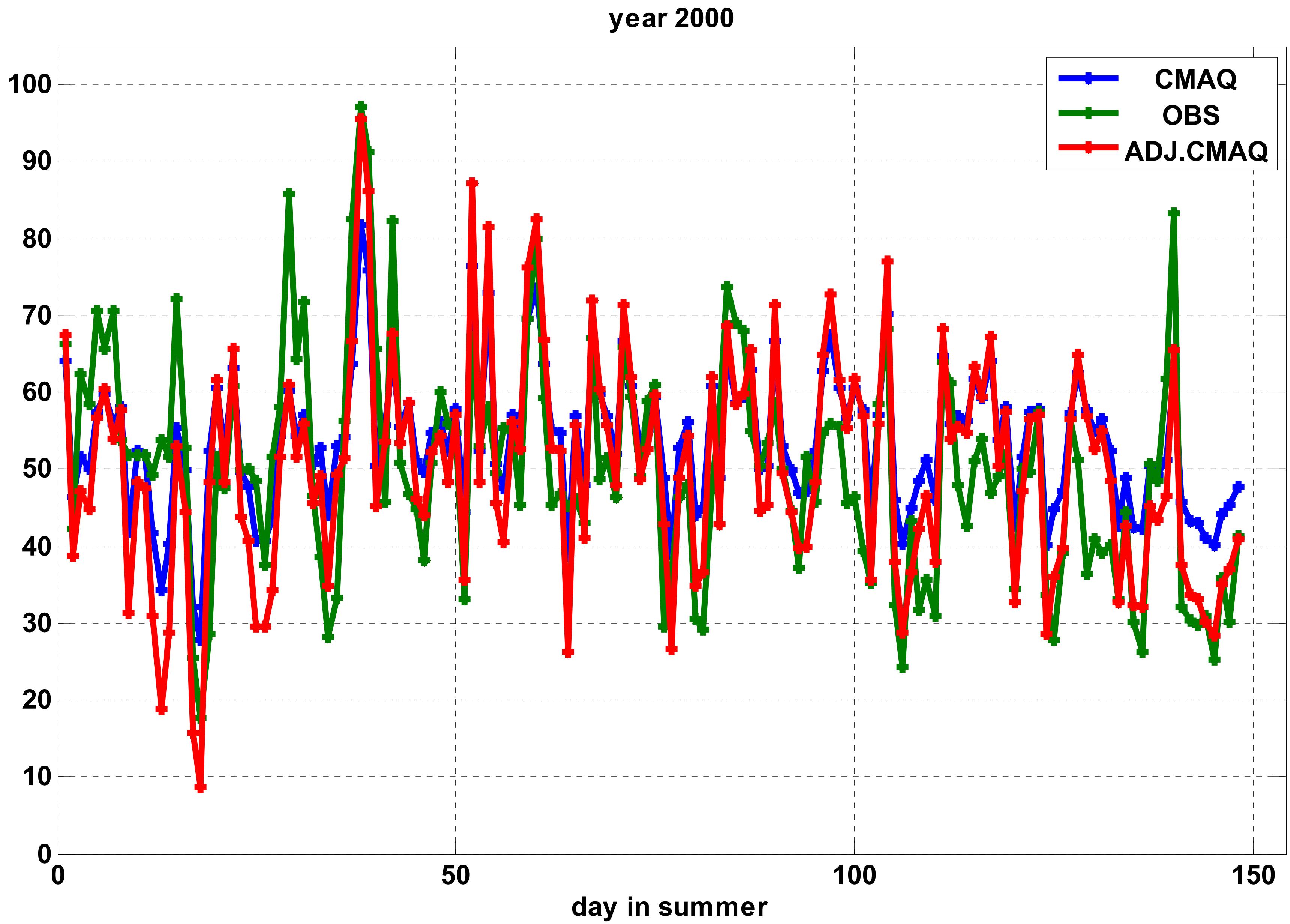


Figure 5. Ozone base year 2000: CMAQ (mean 53.4, standard deviation 8.9, 4TH,73.5), OBSERVED (mean 50.0, standard deviation 14.4, 4TH. 83.2), and ADJUSTED CMAQ (mean 50.0, standard deviation 14.4, 4TH, 82.5).

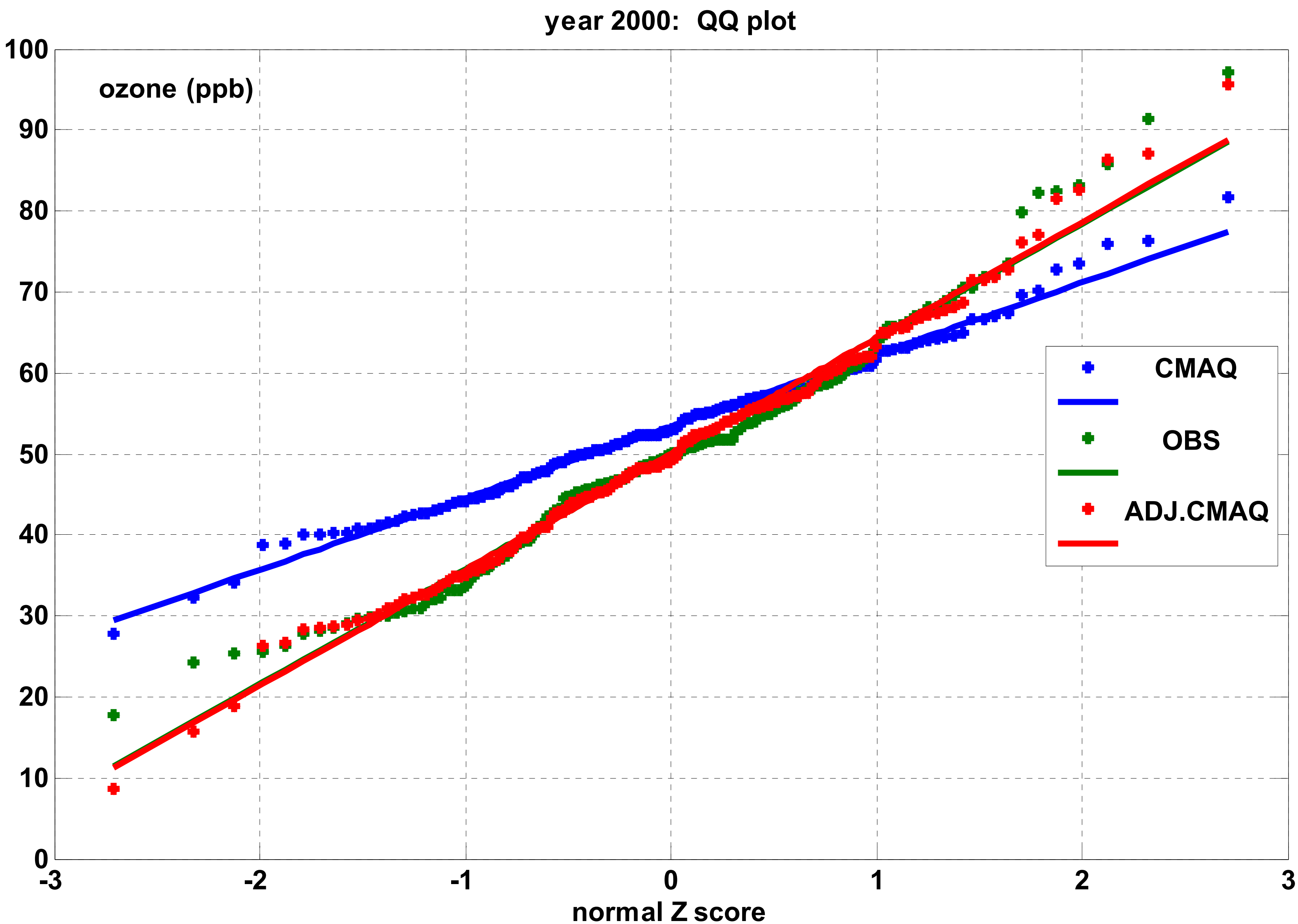


Figure 6. QQ plot of Ozone base year 2000: CMAQ (mean 53.4, standard deviation 8.9, 4^{TH} 73.5), OBSERVED (mean 50.0, standard deviation 14.4, 4^{TH} 83.2), and ADJUSTED CMAQ (mean 50.0, standard deviation 14.4, 4^{TH} 82.5).

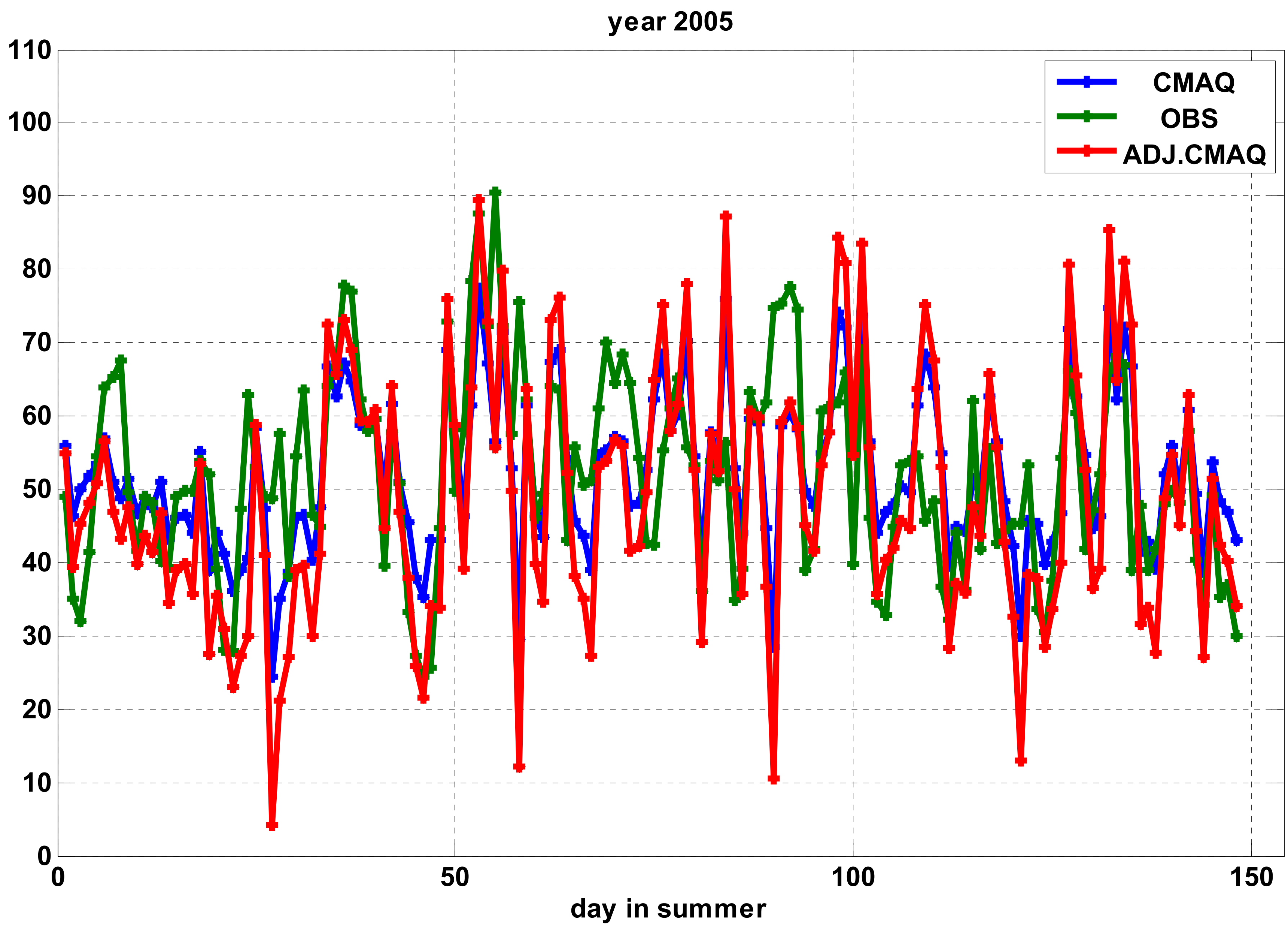


Figure 7. Ozone prediction year 2005: CMAQ (mean 52.2, standard deviation 10.5, 4^{TH} 74.2),
OBSERVED (mean 52.2, standard deviation 13.4, 4^{TH} 77.8),
and ADJUSTED CMAQ (mean 48.8, standard deviation 16.9, 4^{TH} 84.4).

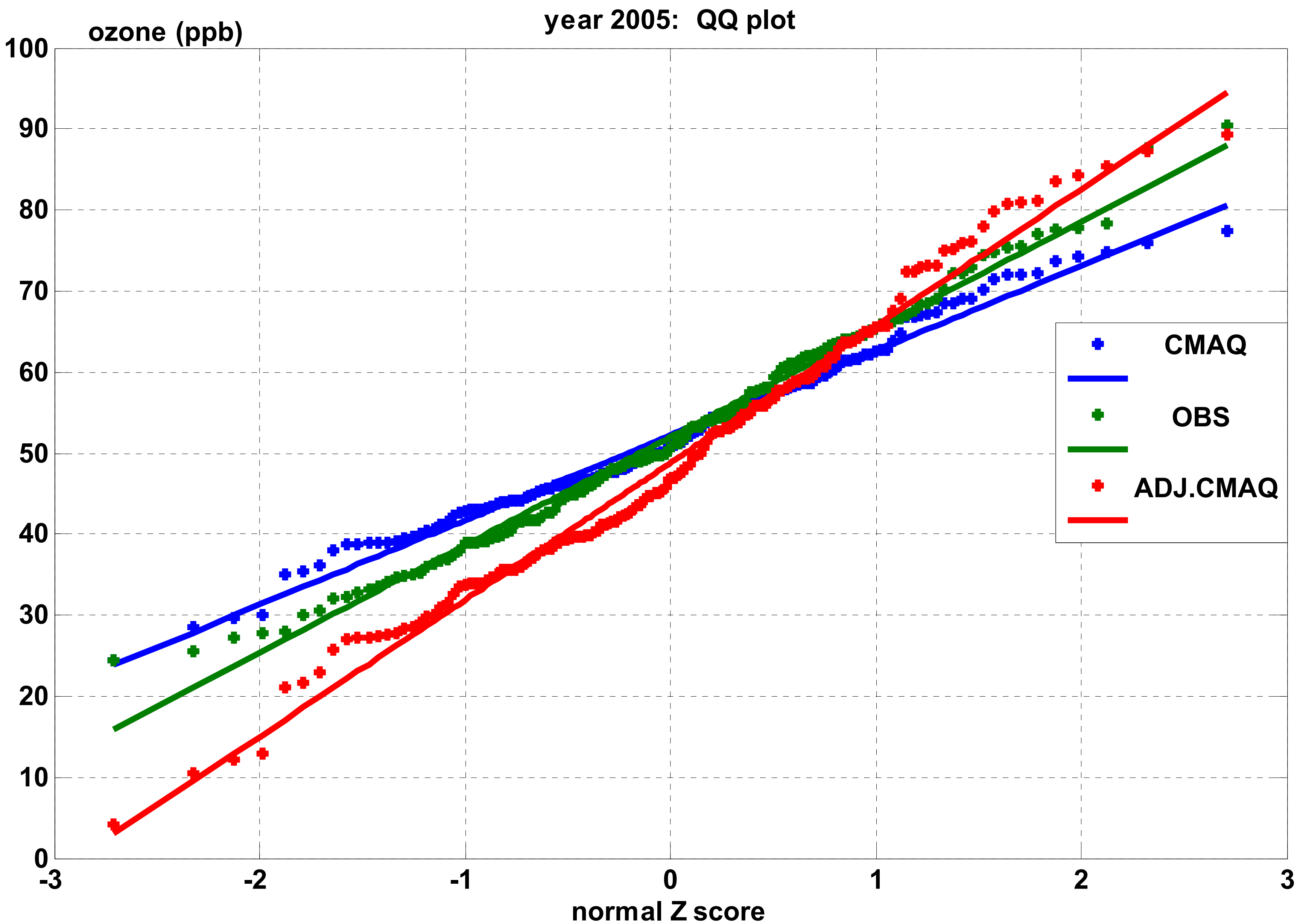


Figure 8. QQ Ozone prediction year 2005: CMAQ (mean 52.2, standard deviation 10.5, 4^{TH} 74.2), OBSERVED (mean 52.2, standard deviation 13.4, 4^{TH} 77.8), and ADJUSTED CMAQ (mean 48.8, standard deviation 17.0, 4^{TH} 84.4).

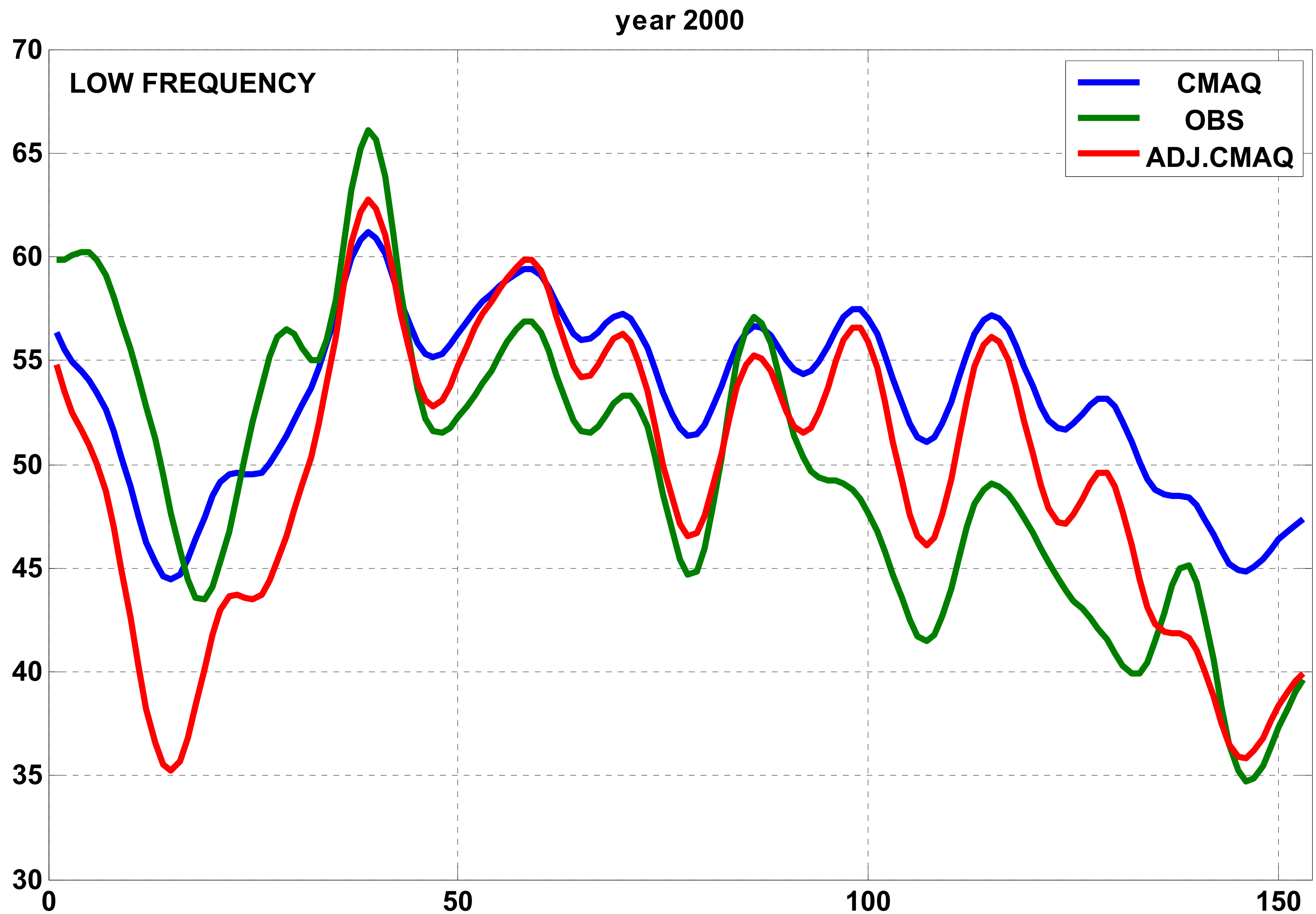


Figure 9. Low-frequency Ozone base year 2000: CMAQ (mean 53.2, standard deviation 4.21)
OBSERVED (mean 49.6, standard deviation 6.93)
and ADJUSTED CMAQ (mean 49.6, standard deviation 6.93)

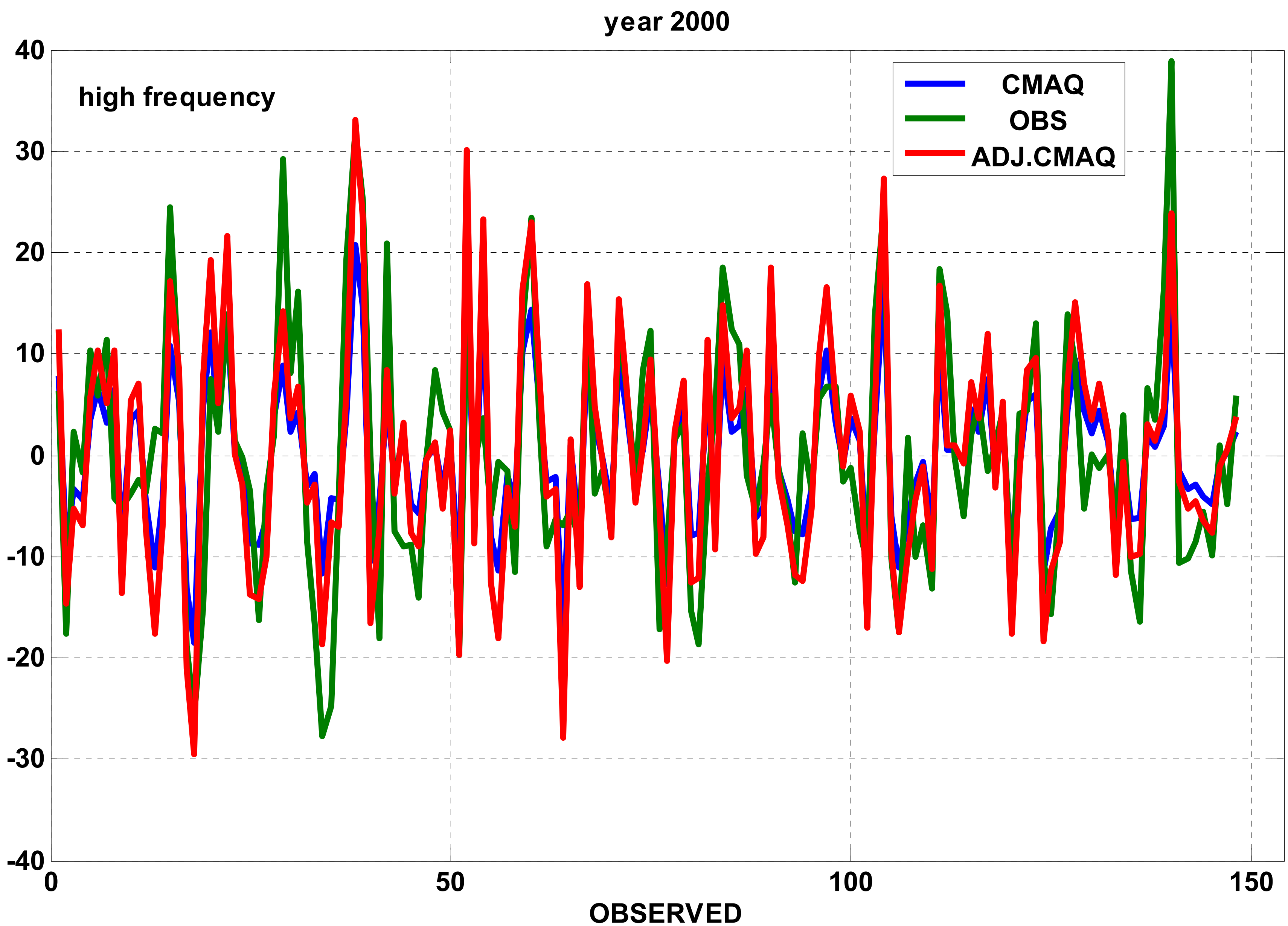


Figure 10. High-frequency Ozone base year 2000: CMAQ (mean -0.0345, standard deviation 7.34), OBSERVED (mean -0.0280, standard deviation 11.7), and ADJUSTED CMAQ (mean -0.0280,, standard deviation 11.7).

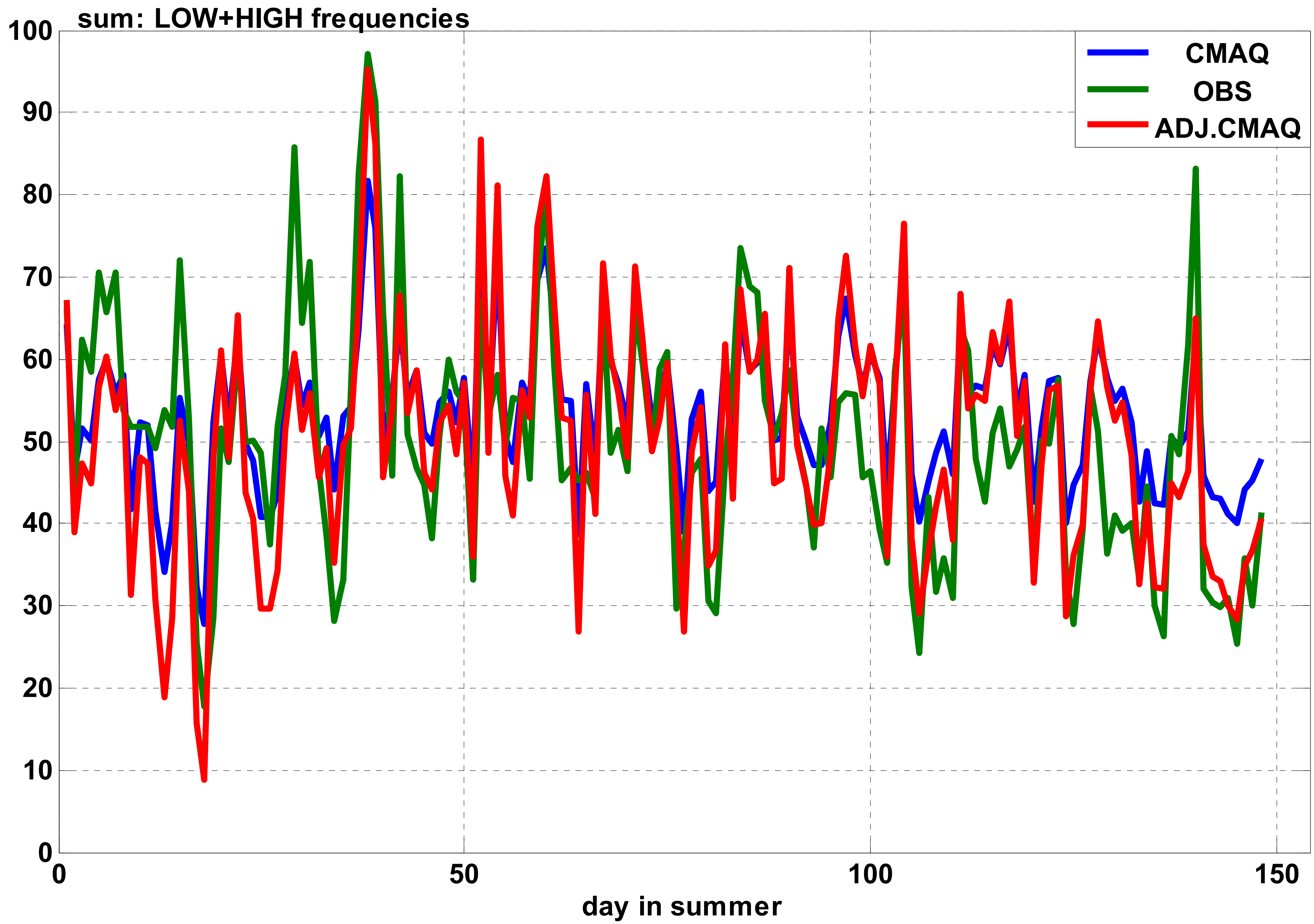


Figure 11. Sum of high- and low-frequency Ozone base year 2000: CMAQ (mean 53.4, standard deviation 8.91, 4TH 73.5), OBSERVED (mean 50.0, standard deviation 14.4, 4TH 83.2), and ADJUSTED CMAQ (mean 50.0, standard deviation 14.3, 4TH 82.3.

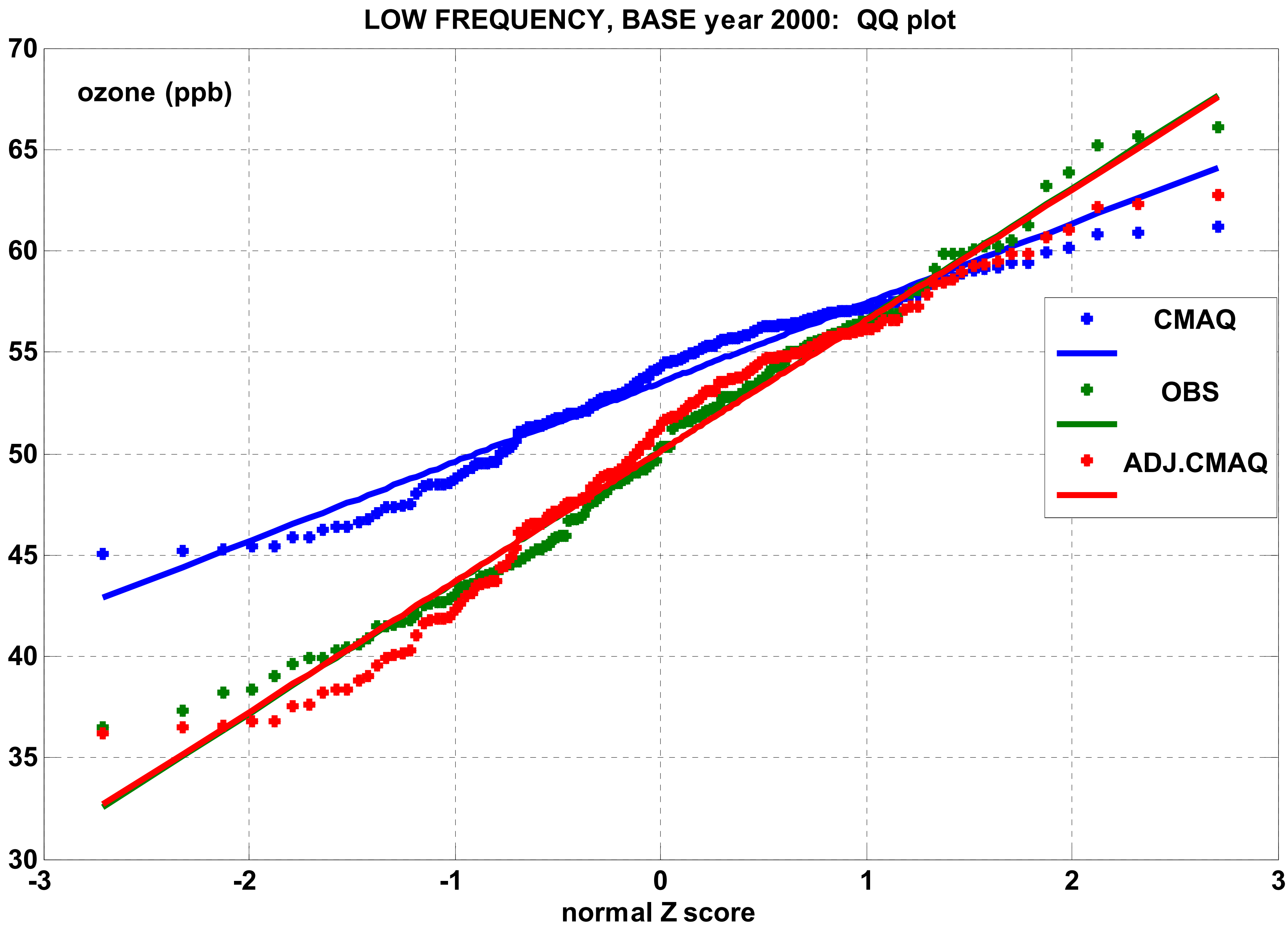


Figure 12. QQ of Low-frequency Ozone base year 2000: CMAQ (mean 53.2, standard deviation 4.21),
OBSERVED (mean 49.6, standard deviation 6.93),
and ADJUSTED CMAQ (mean 49.6, standard deviation 6.93).

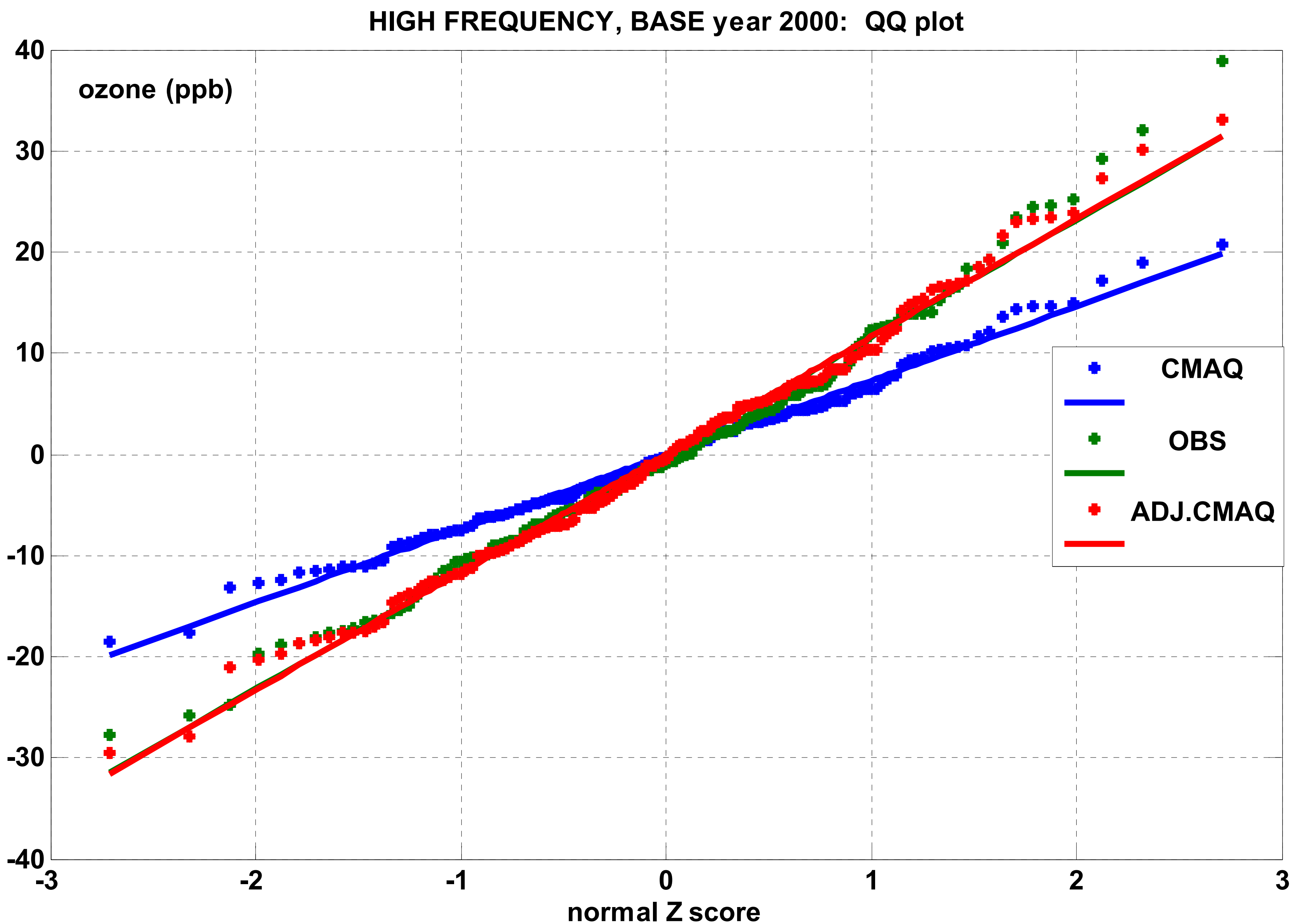


Figure 13. QQ of High-frequency Ozone base year 2000: CMAQ (mean -0.0345, standard deviation 7.34), OBSERVED (mean -0.0280, standard deviation 11.7), and ADJUSTED CMAQ (mean -0.0280, standard deviation 11.7).

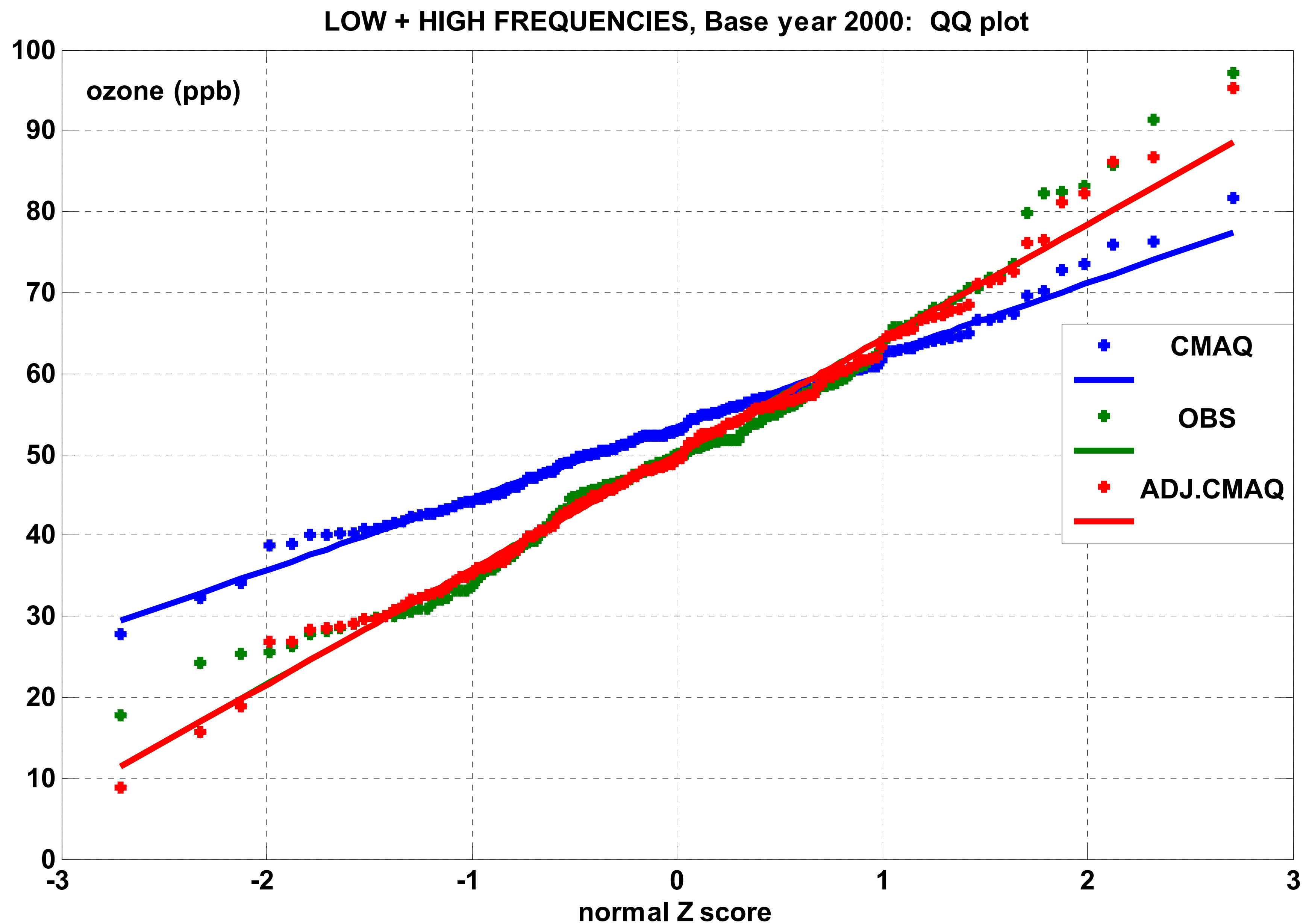


Figure 14. QQ, sum of High- and LOW frequencies Ozone base year 2000:
CMAQ (mean 53.4, standard deviation 8.9, 4^{TH} 73.5),
OBSERVED (mean 50.0, standard deviation 14.4, 4^{TH} 83.2),
and ADJUSTED CMAQ (mean 50.0, standard deviation 14.3, 4^{TH} 82.3).

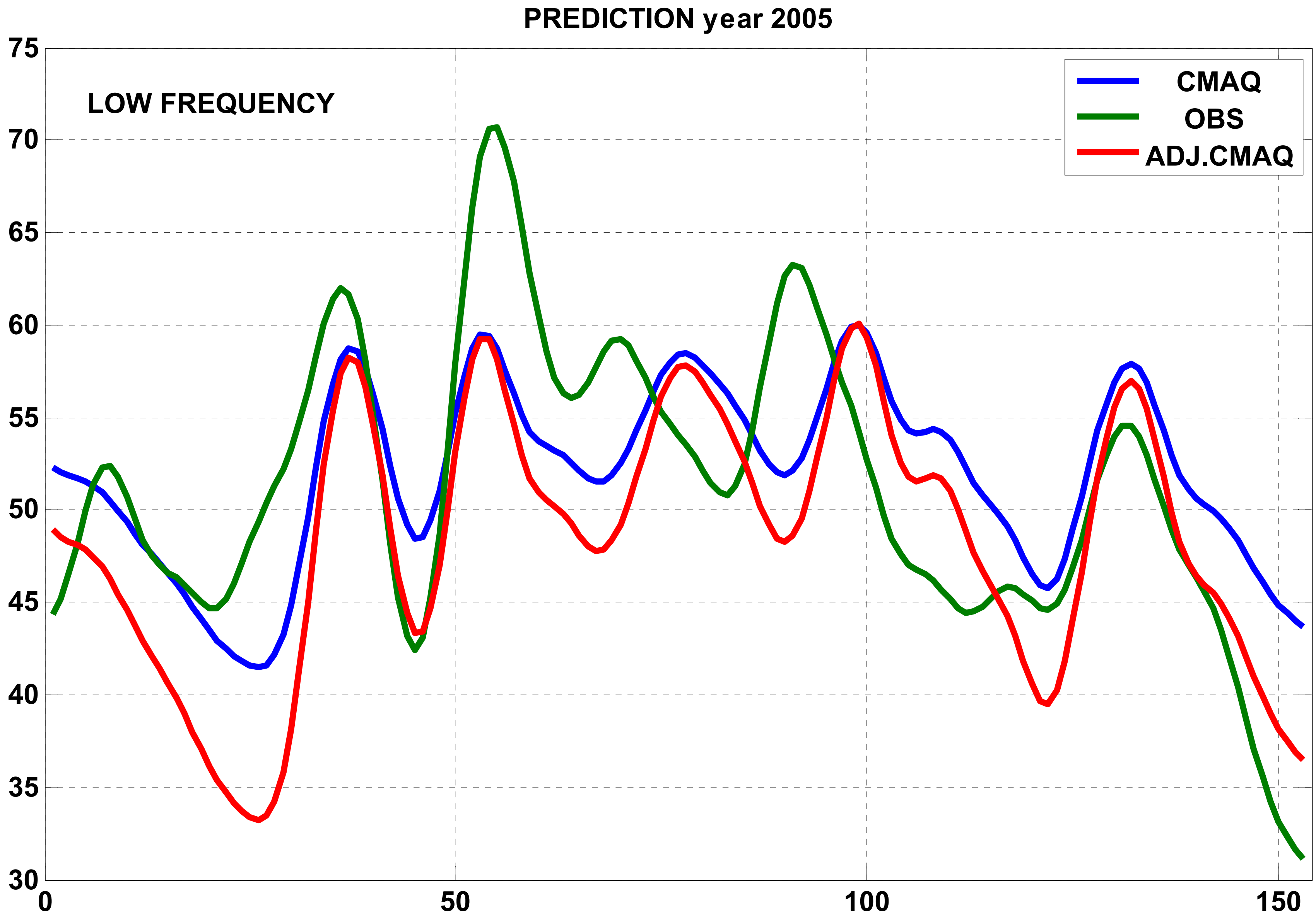


Figure 15. Low-frequency Ozone for prediction year 2005: CMAQ (mean 51.9, standard deviation 4.80), OBSERVED (mean 51.3, standard deviation 7.74), and ADJUSTED CMAQ (mean 48.4, standard deviation 6.90,).

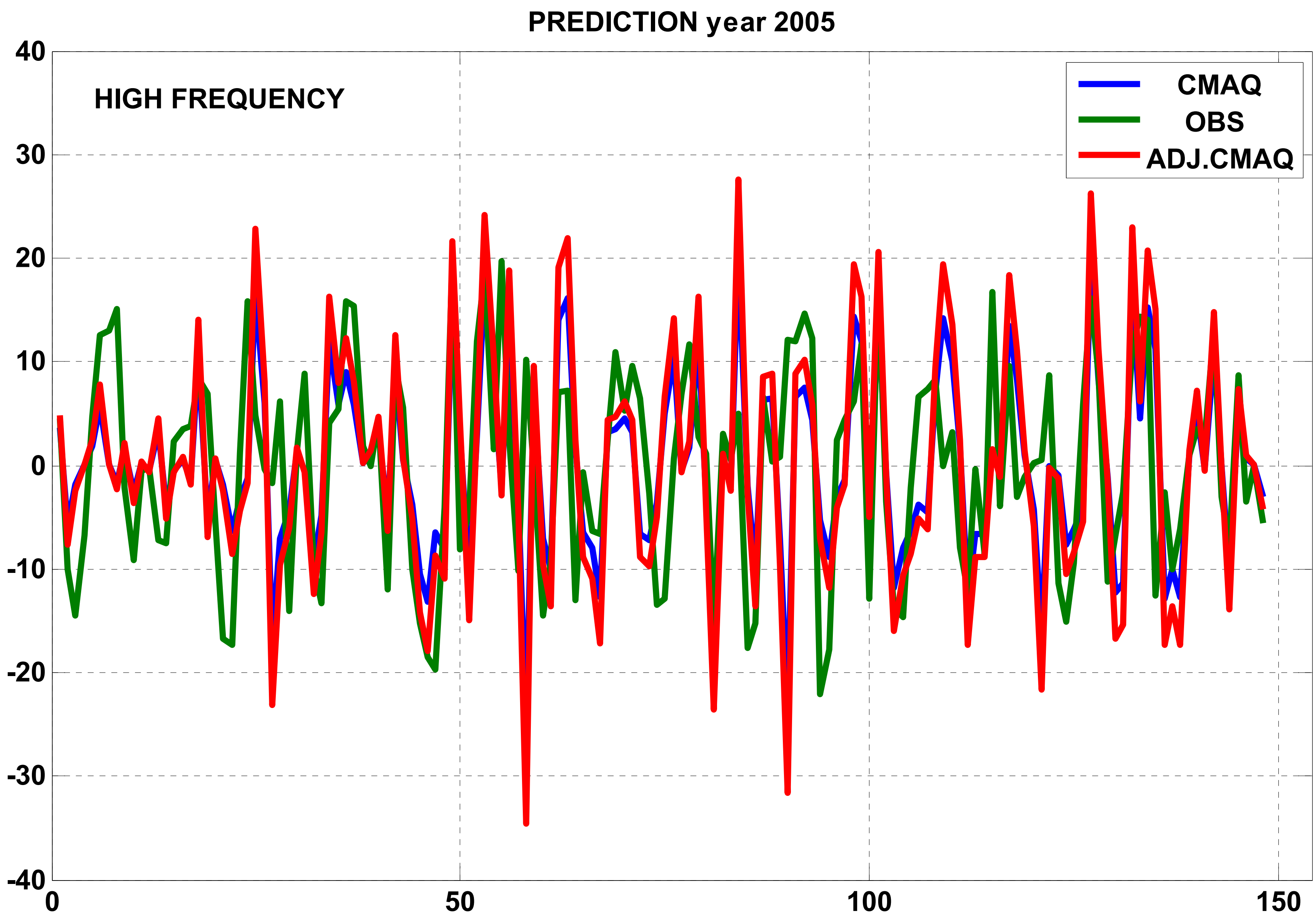


Figure 16. High-frequency Ozone for prediction year 2005: CMAQ (mean 0.0043, standard deviation 8.62), OBSERVED (mean 0.0208, standard deviation 9.87), and ADJUSTED CMAQ (mean 0.0108, standard deviation 11.68).

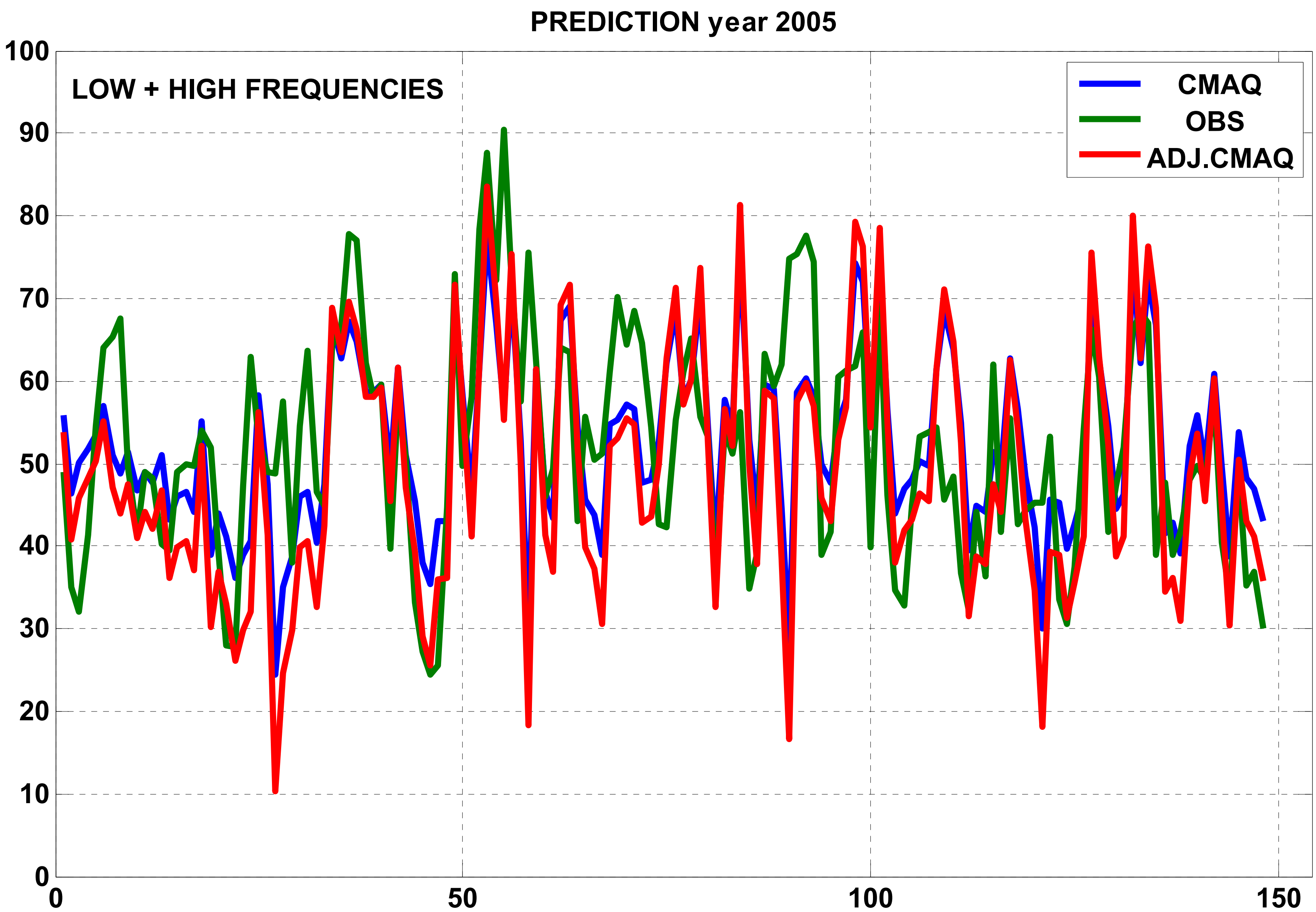


Figure 17. Sum of High- and Low-frequency Ozone for prediction year 2005: CMAQ (mean 52.2, standard deviation 10.5, 4^{TH} 74.2), OBSERVED (mean 52.0, standard deviation 13.4, 4^{TH} 77.8), and ADJUSTED CMAQ (mean 48.7, standard deviation 14.5, 4^{TH} 79.3).

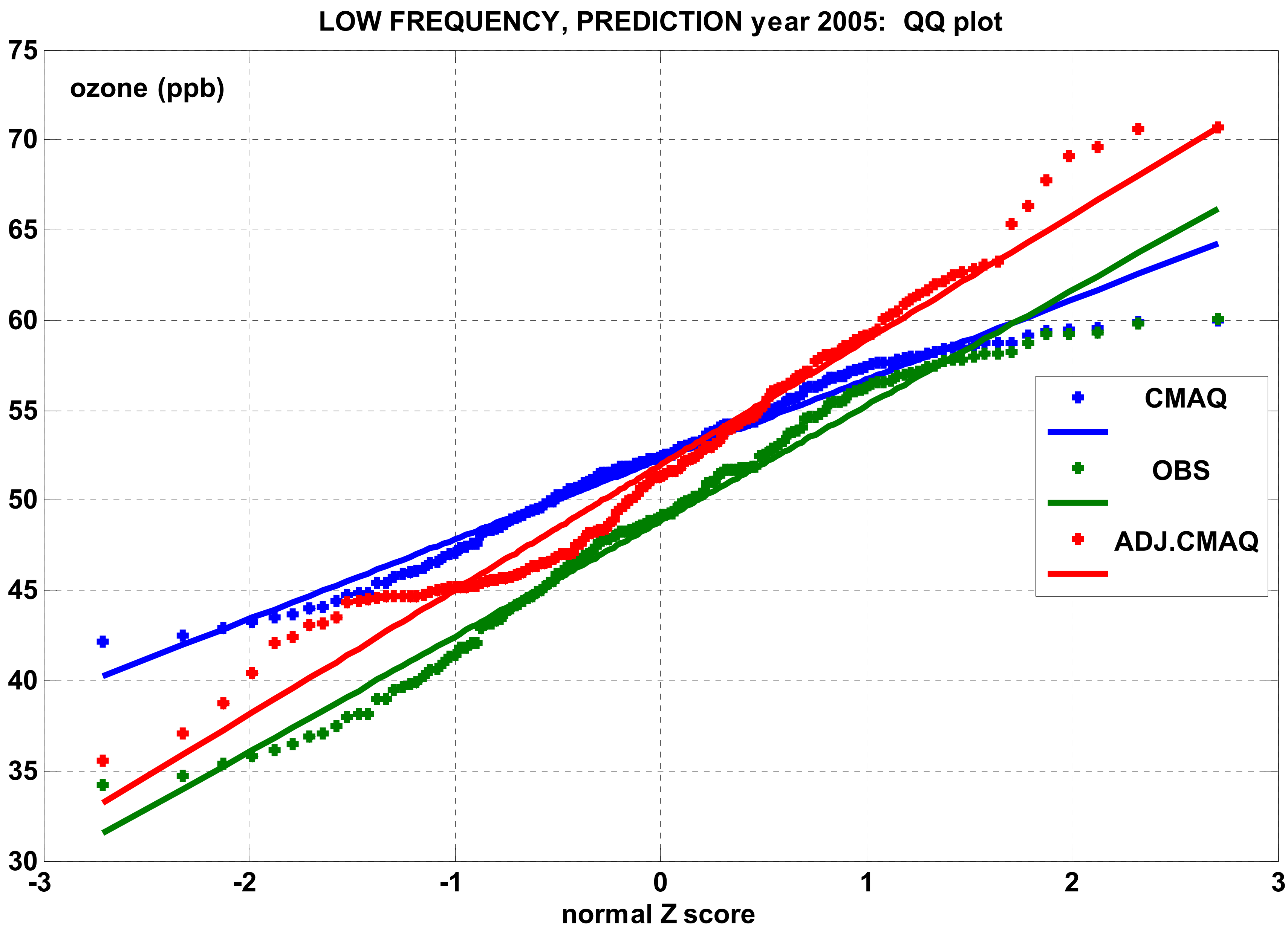


Figure 18. QQ of Low-frequency Ozone for prediction year 2005: CMAQ (mean 51.9, standard deviation 4.80), OBSERVED (mean 51.3, standard deviation 7.74), and ADJUSTED CMAQ (mean 48.4, standard deviation 6.93).

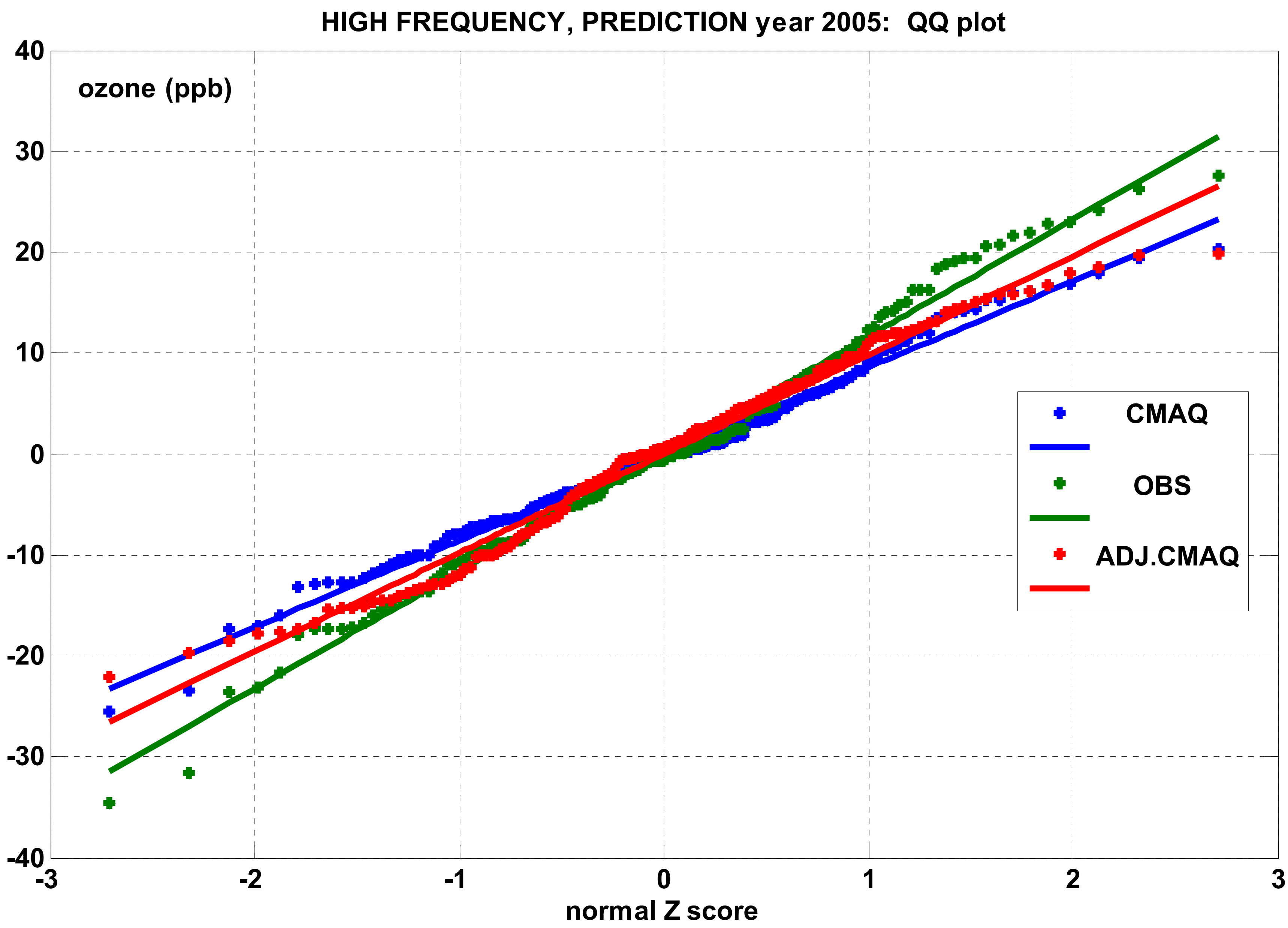


Figure 19. QQ of High-frequency Ozone for prediction year 2005: CMAQ (mean 0.0043, standard deviation 8.62), OBSERVED (mean 0.0208, standard deviation 9.87), and ADJUSTED CMAQ (mean 0.0108, standard deviation 11.68).

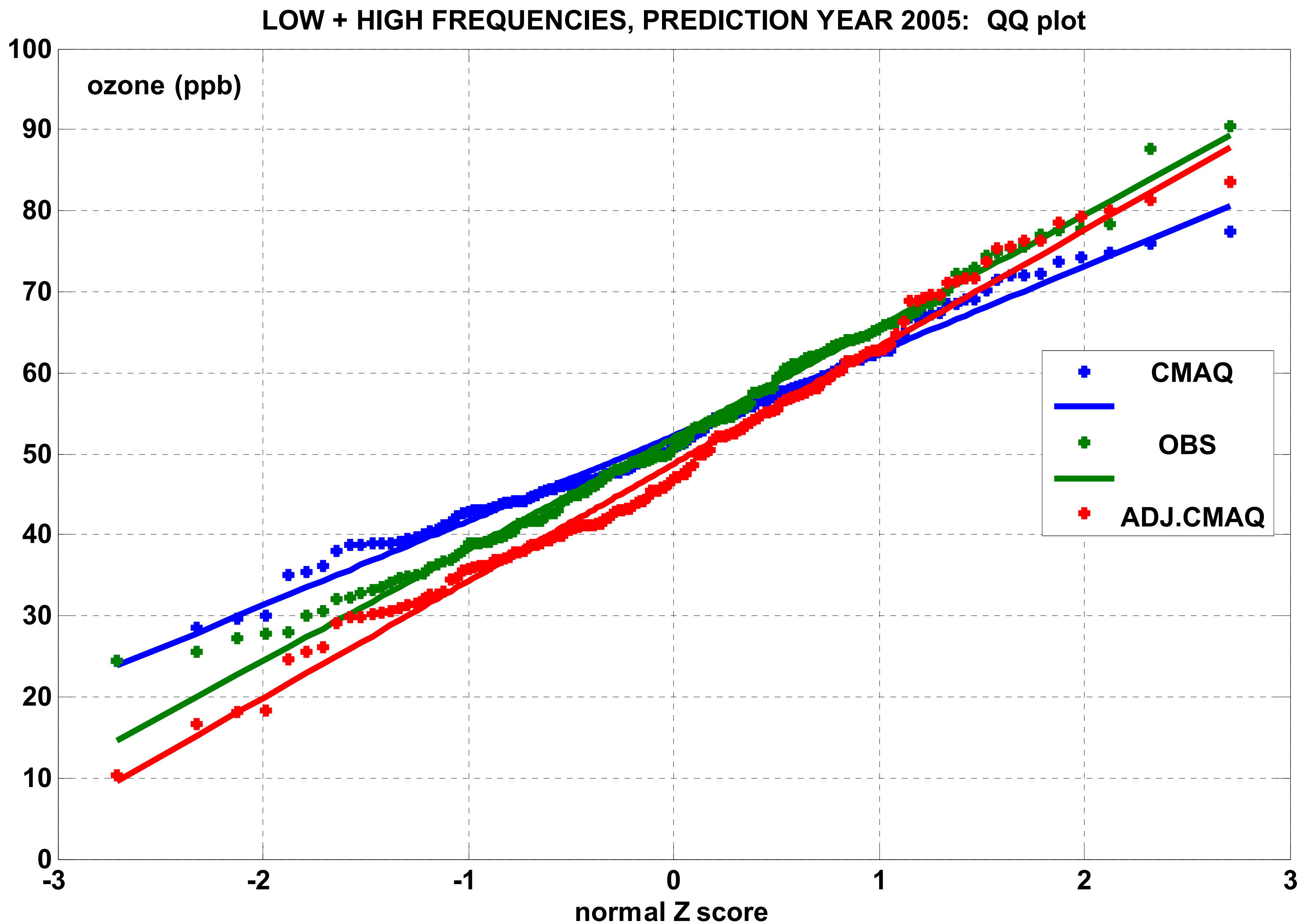


Figure 20. QQ of sum of High- and Low-frequency Ozone for prediction year 2005:
CMAQ (mean 52.2, standard deviation 10.5, 4TH 74.2),
OBSERVED (mean 52.0, standard deviation 13.4, 4TH 77.8),
and ADJUSTED CMAQ (mean 48.7, standard deviation 14.5, 4TH 79.3).

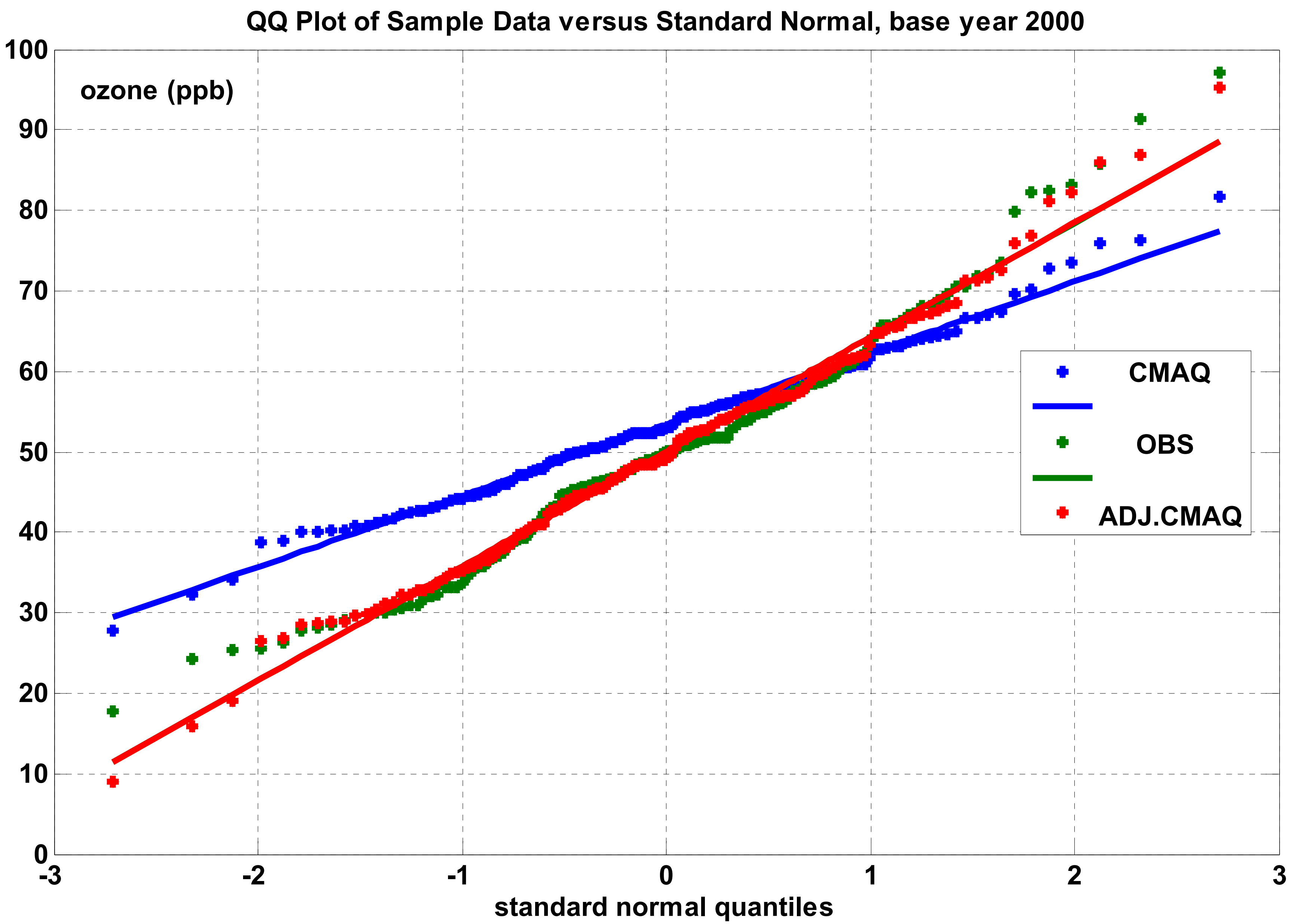


Figure 21. QQ of sum Ozone for base year 2000: CMAQ (mean 53.4, standard deviation 8.91, 4^{TH} 73.5),
OBSERVED (mean 50.0, standard deviation 14.39, 4^{TH} 83.2),
and ADJUSTED CMAQ (mean 50.0, standard deviation 14.29, 4^{TH} 82.3).

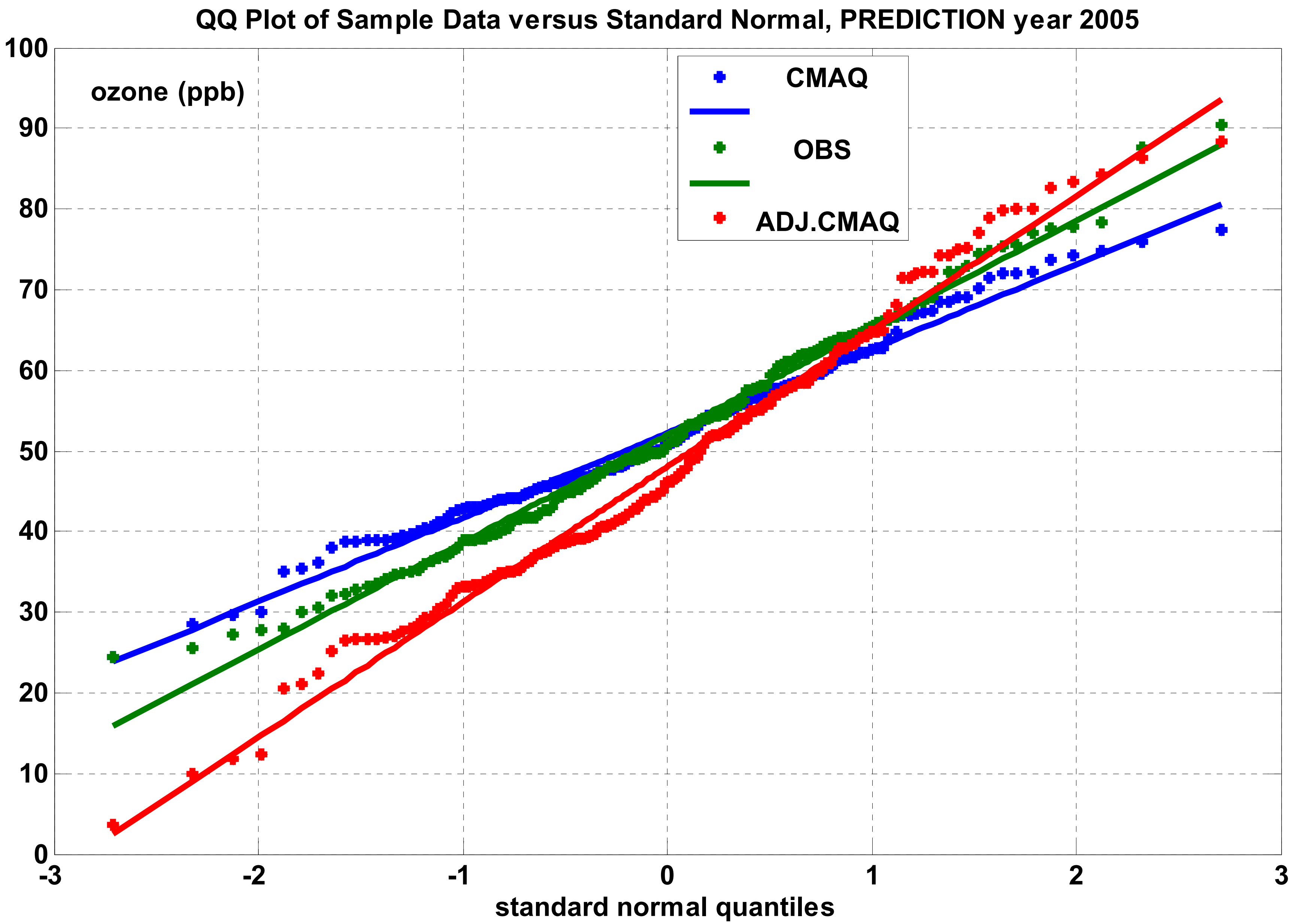


Figure 22. QQ of sum Ozone for prediction year 2005: CMAQ (mean 52.2, standard deviation 10.5, 4^{TH} 74.2), OBSERVED (mean 52.0, standard deviation 13.4, 4^{TH} 77.8), and ADJUSTED CMAQ (mean 48.0, standard deviation 16.9, 4^{TH} 83.4).

RESULTS
FOR ALL SITES

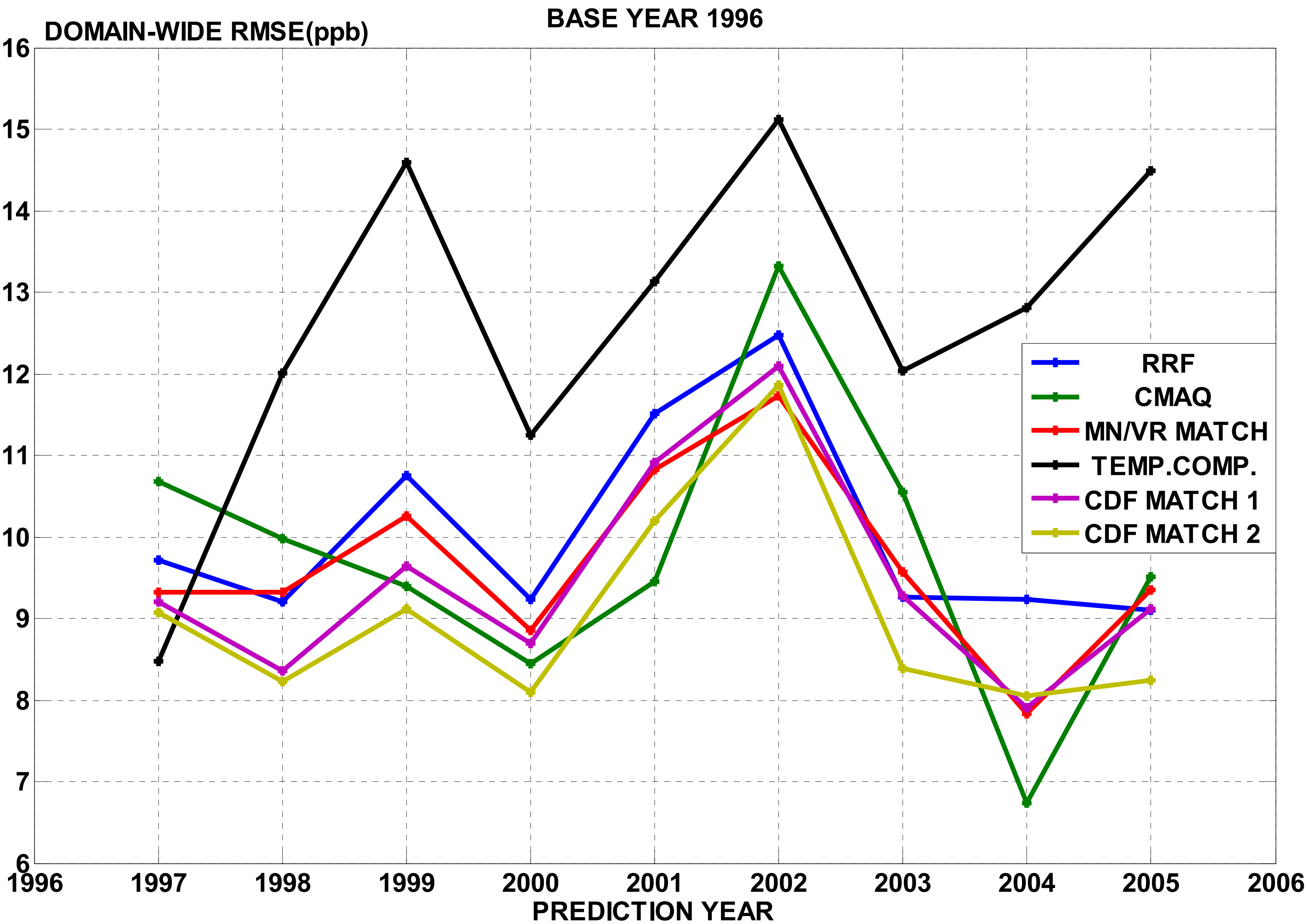


Figure 23. RMSE, domain-wide for base year 1996

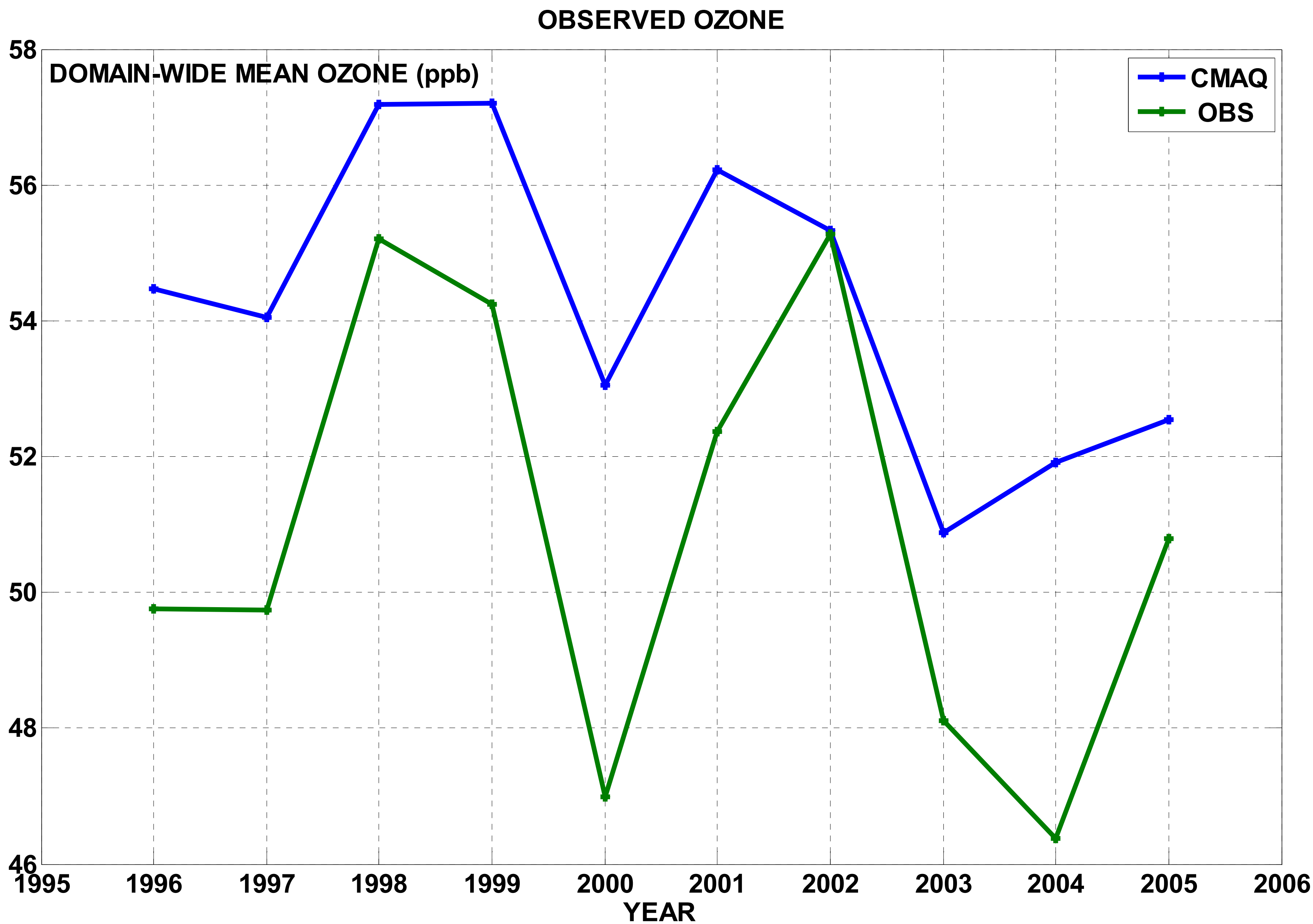


Figure 24. Domain-wide mean OBSERVED ozone for all years

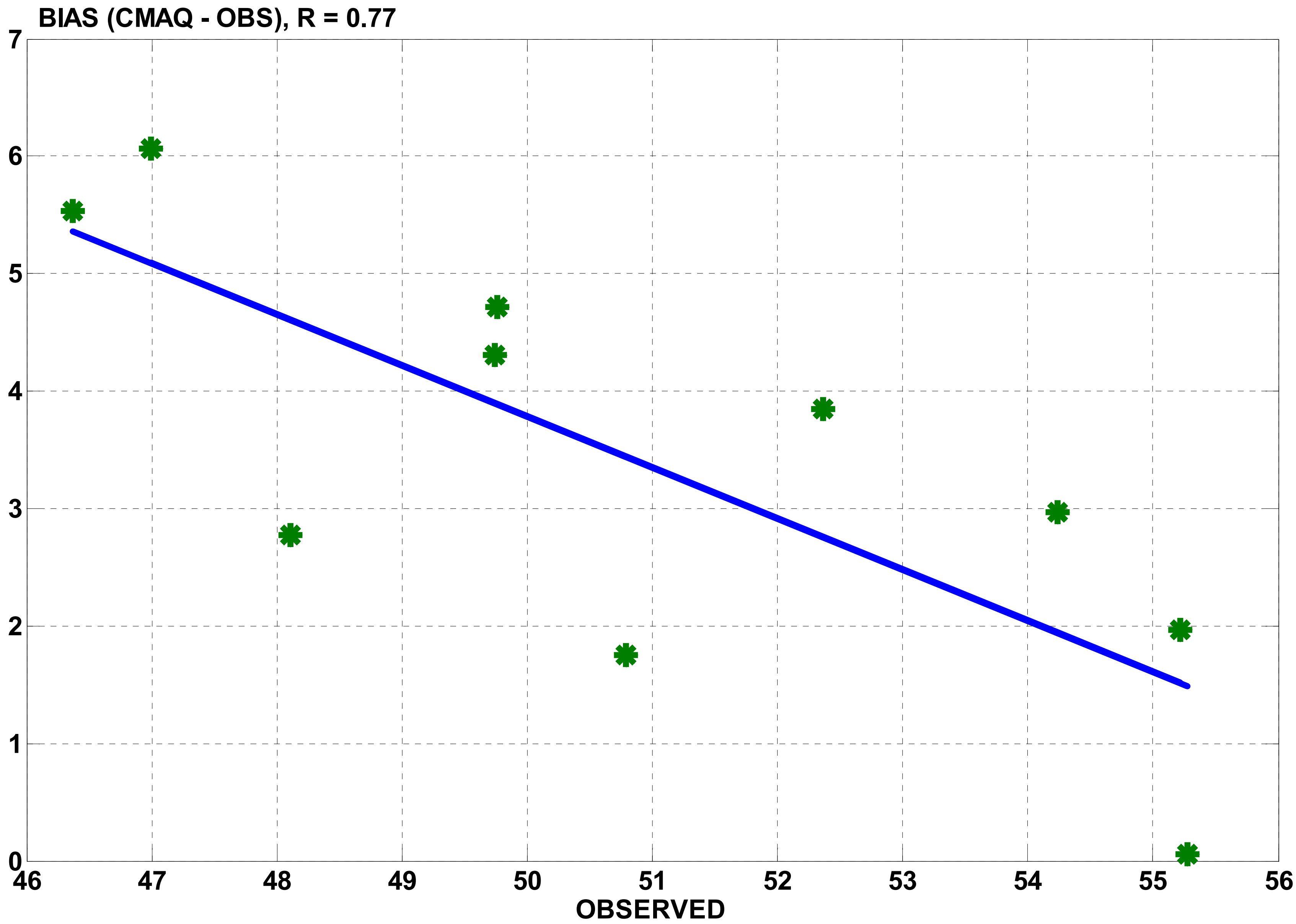


Figure 25. Domain-wide mean BIAS (CMAQ – OBSERVED) as a function of OBSERVED ozone for all years

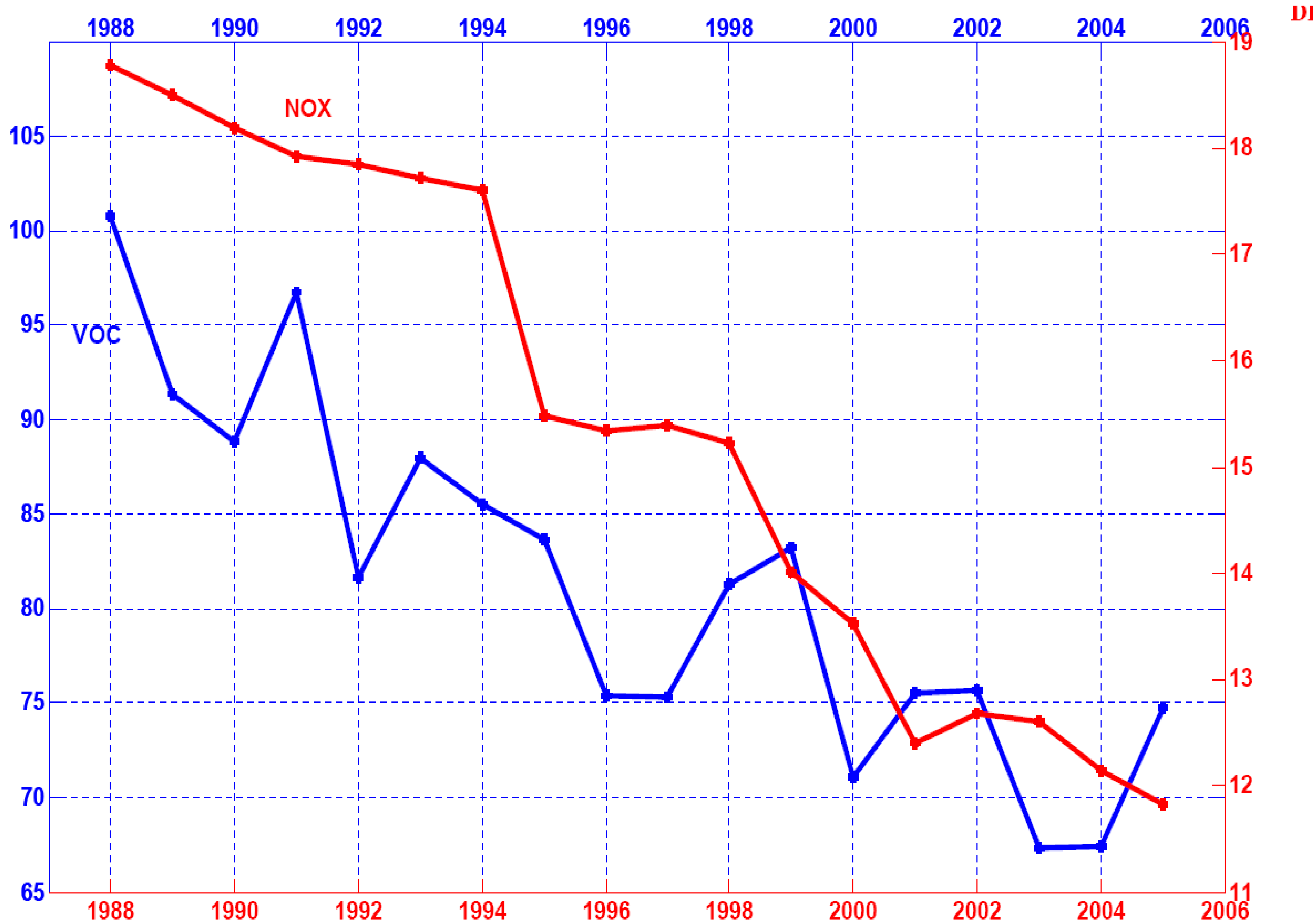


Figure 26. Domain-wide NOX and VOC emissions (normalized)

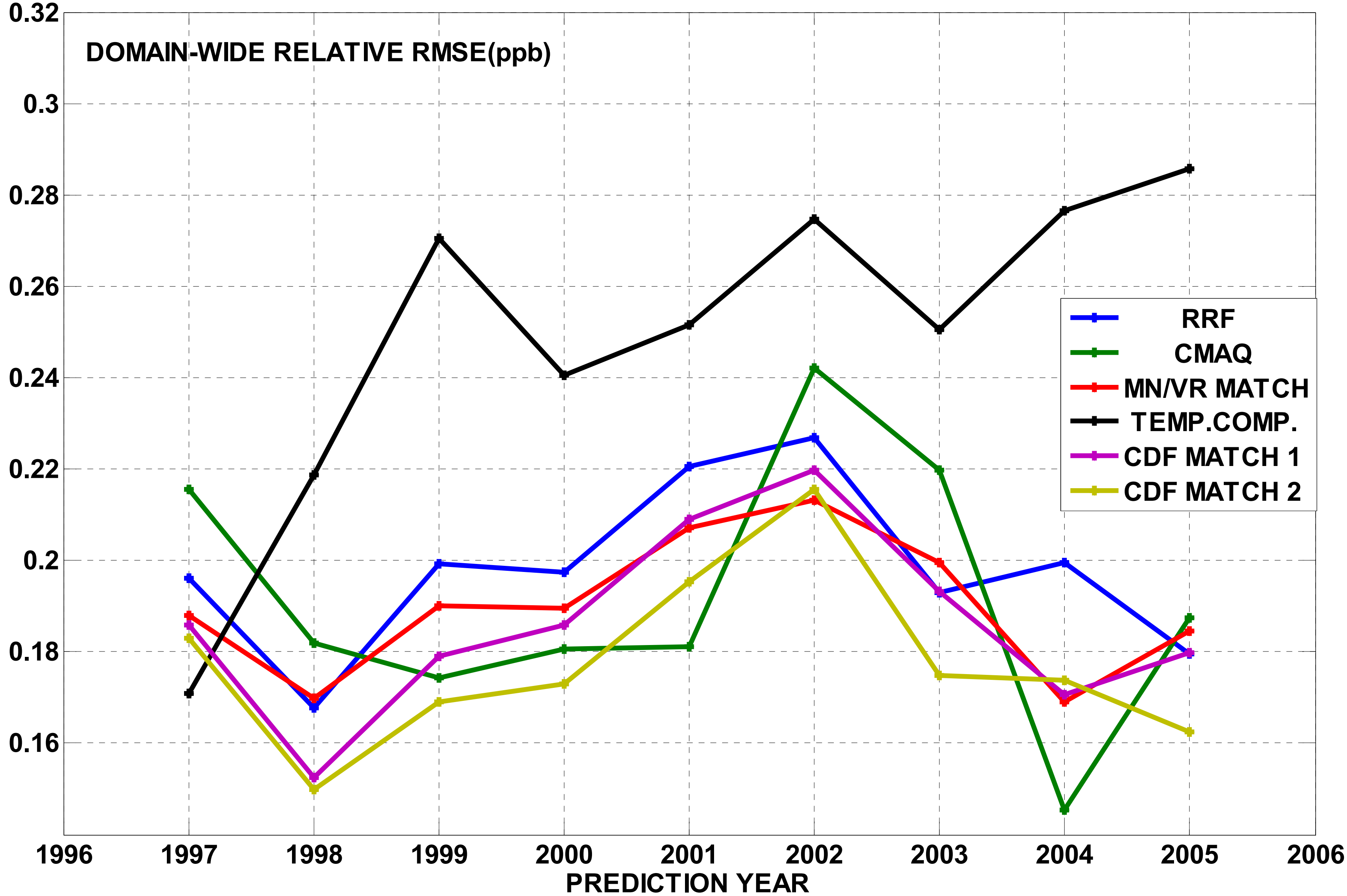


Figure 27. RELATIVE RMSE, domain-wide for base year 1996

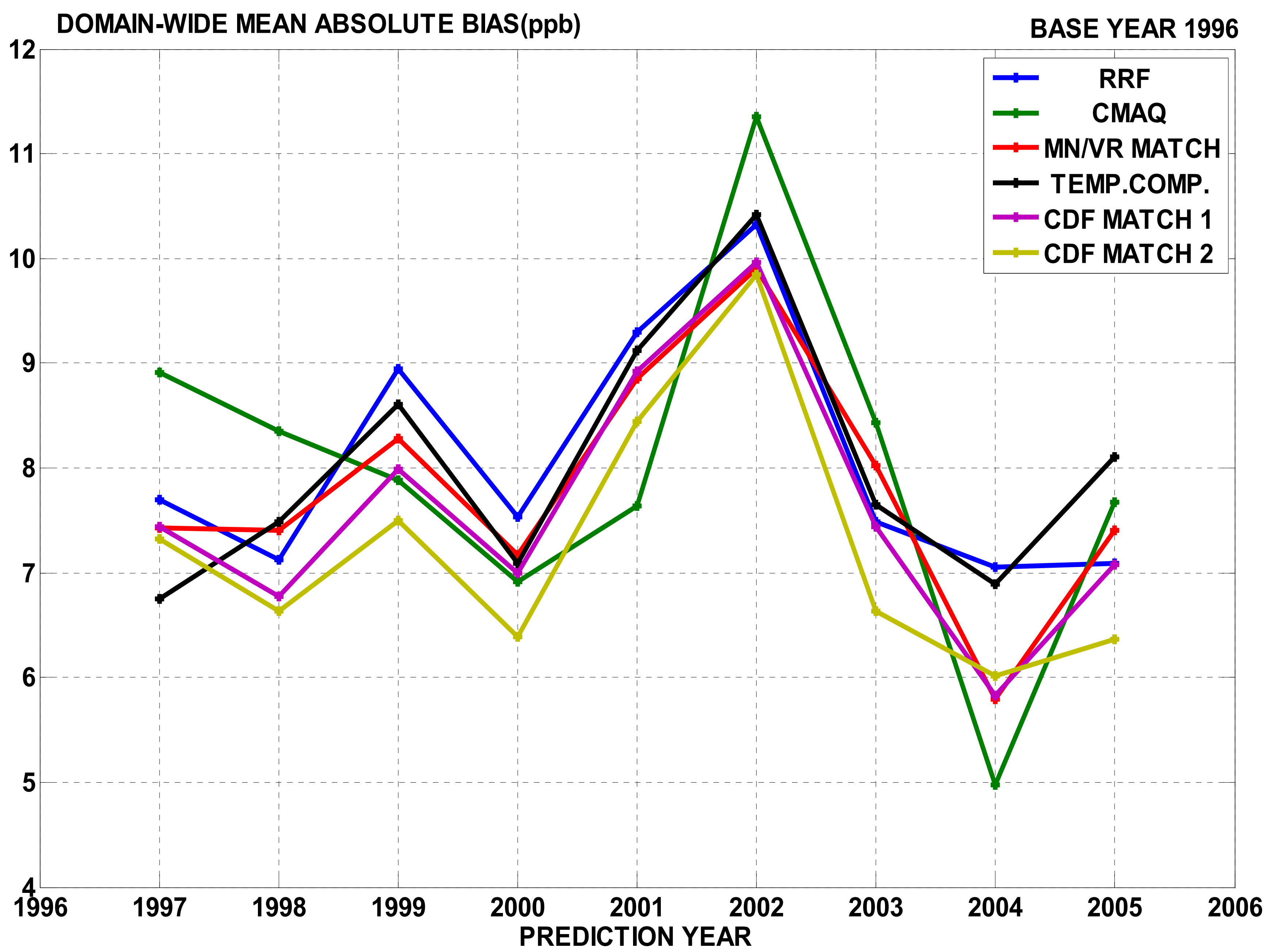


Figure 28. Domain-wide MAB, for base year 1996

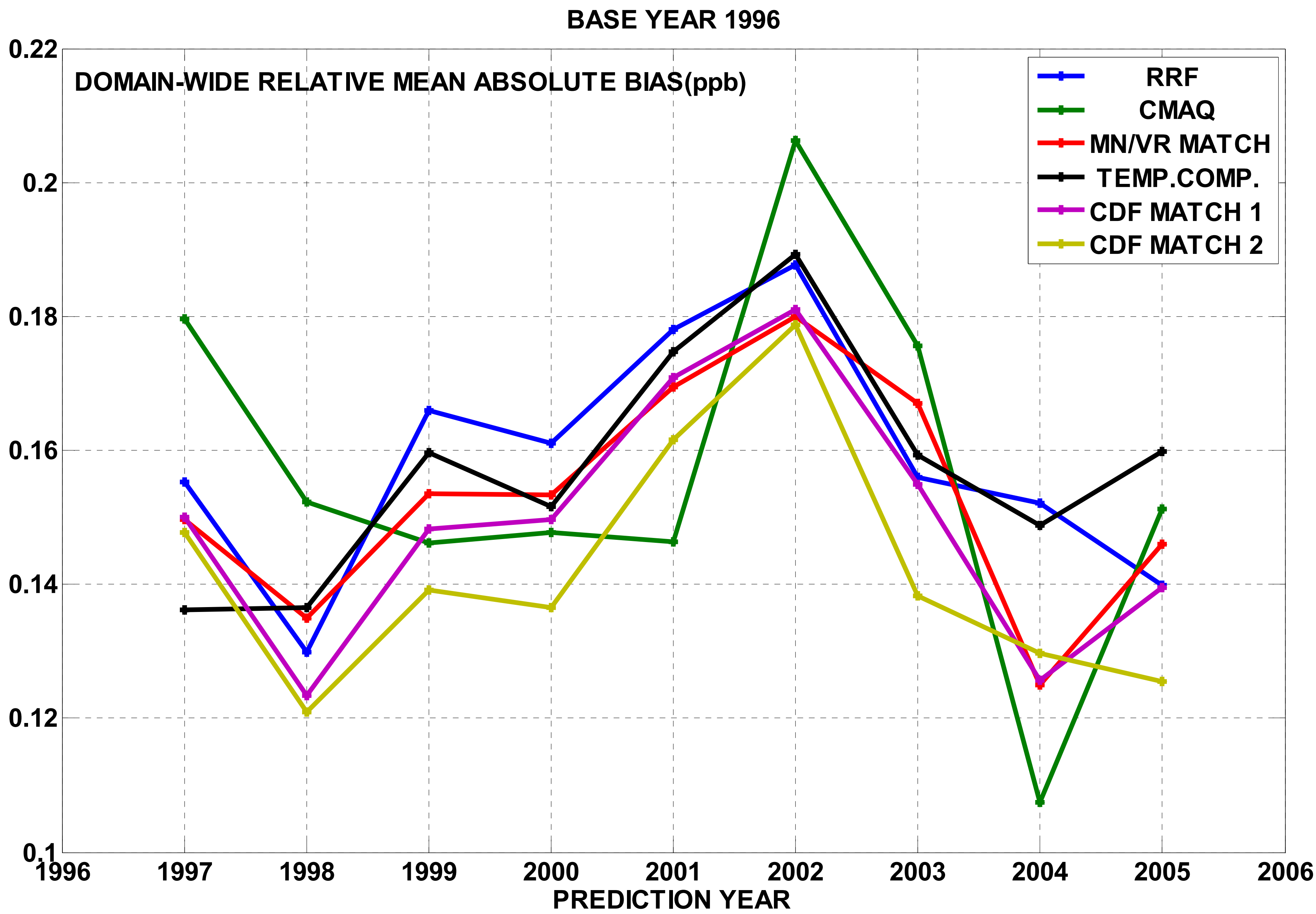


Figure 29. Domain-wide RELATIVE MAB, for base year 1996

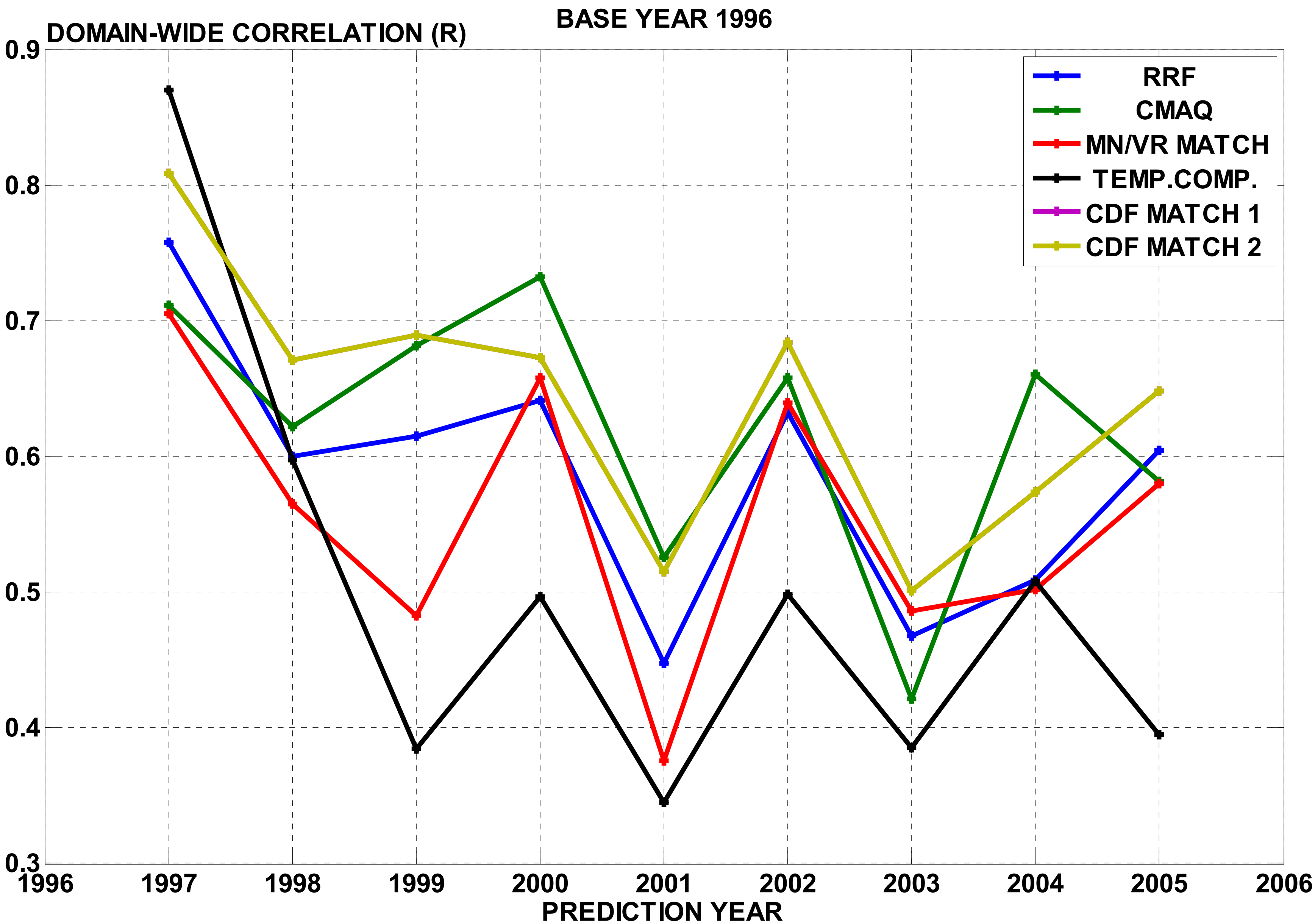


Figure 30. Domain-wide CORRELATION (R), for base year 1996

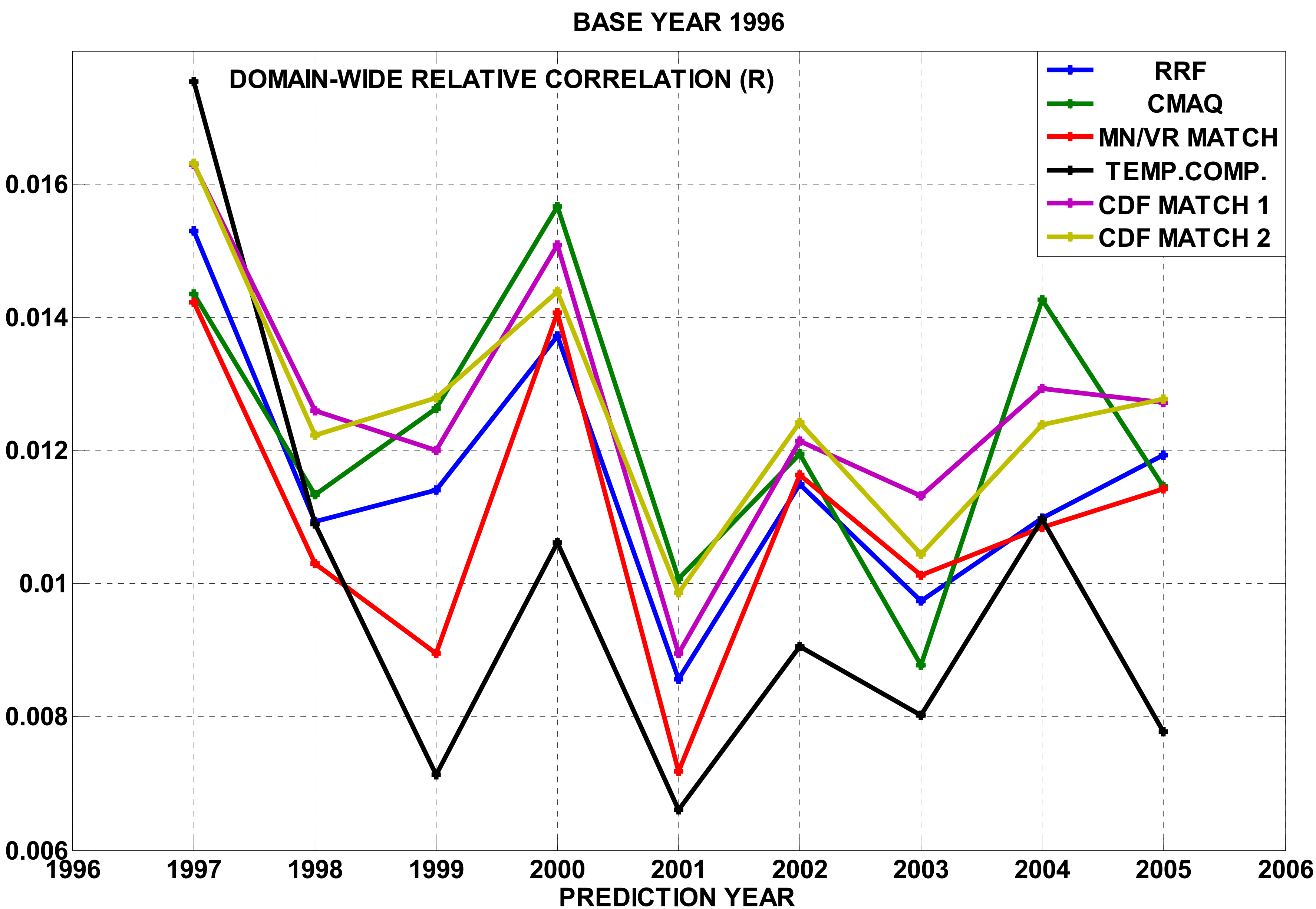


Figure 31. Domain-wide RELATIVE CORRELATION (R), for base year 1996

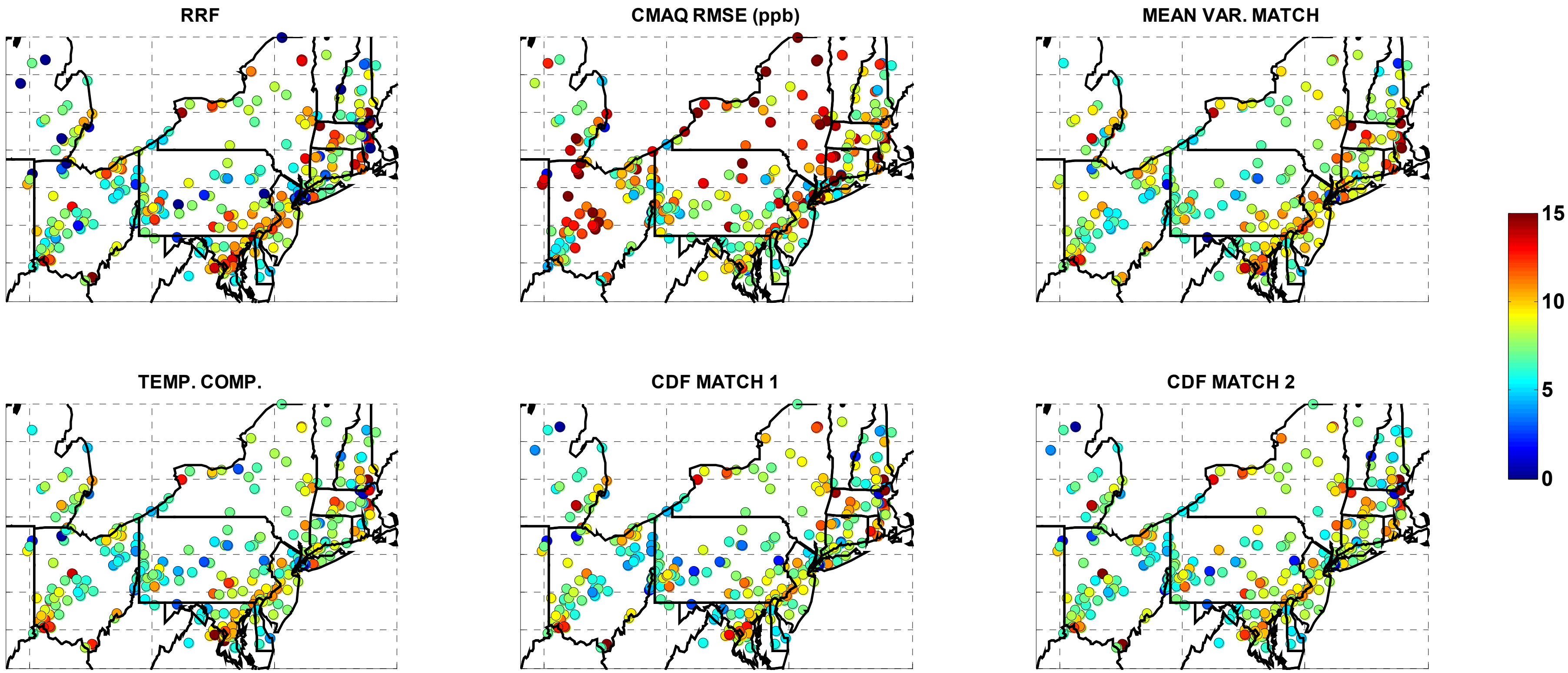


Figure 32. Spatial image of RMSE, all prediction years for each method

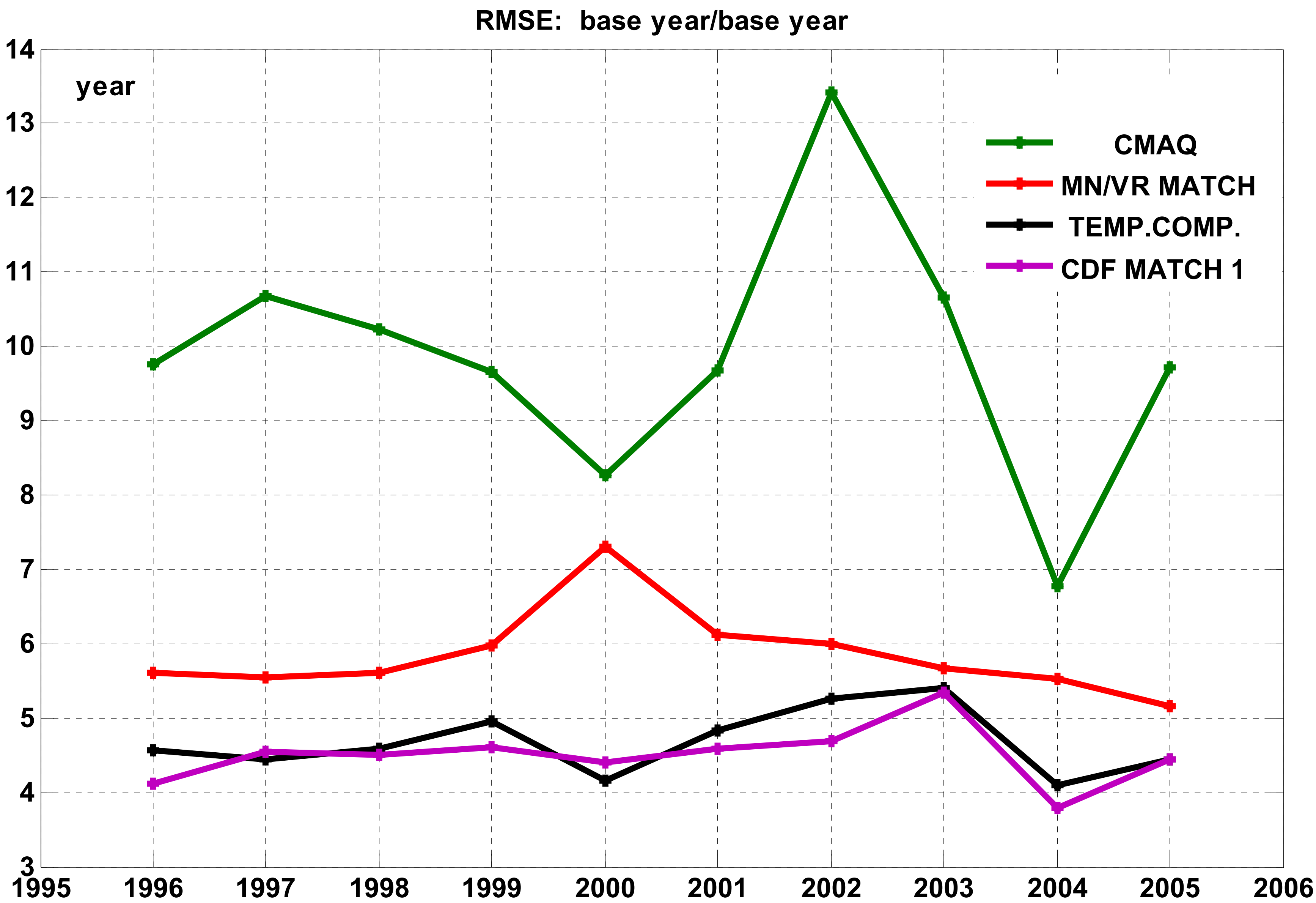


Figure 33. Same-year domain-wide RMSE for time series of 8H

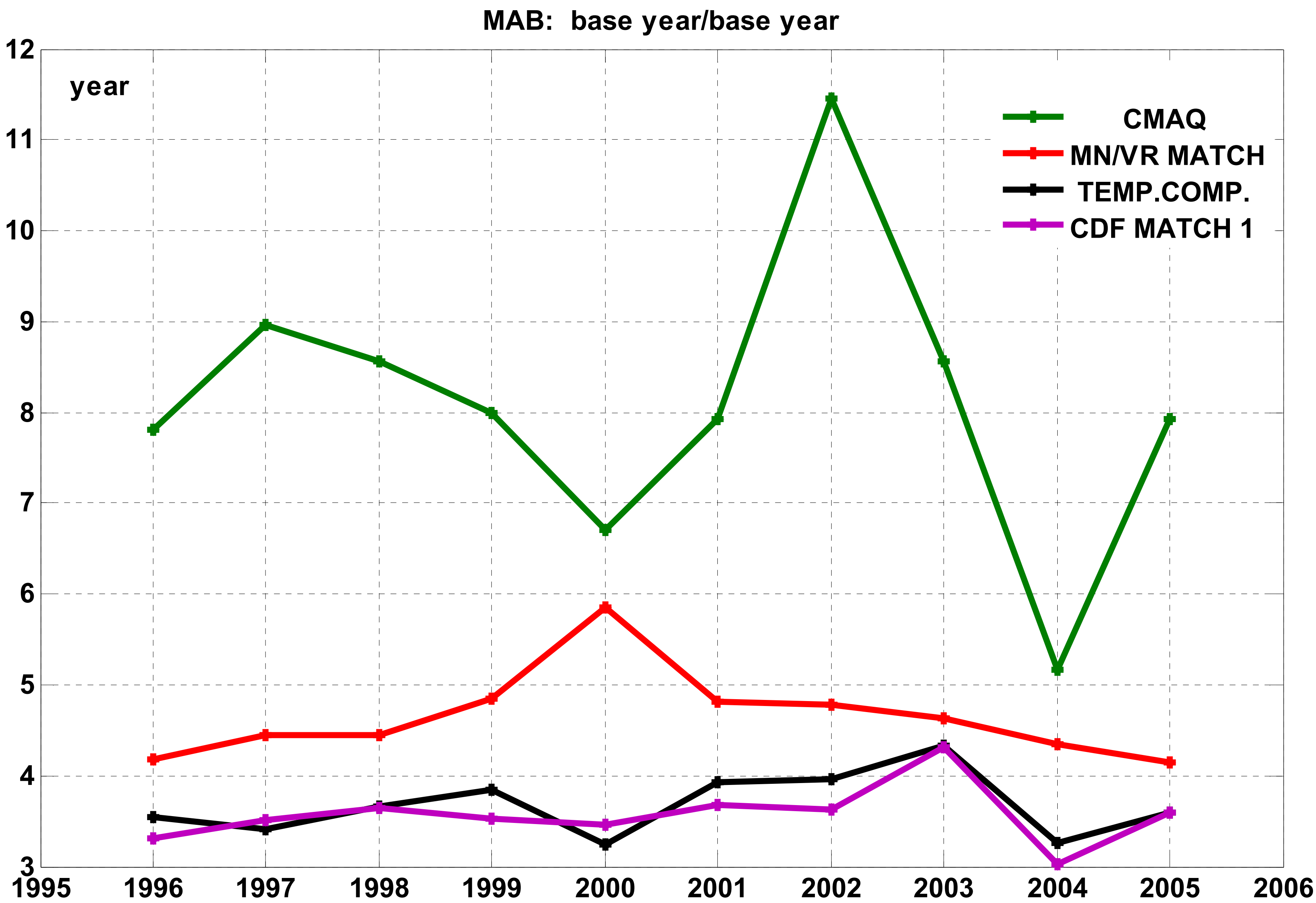
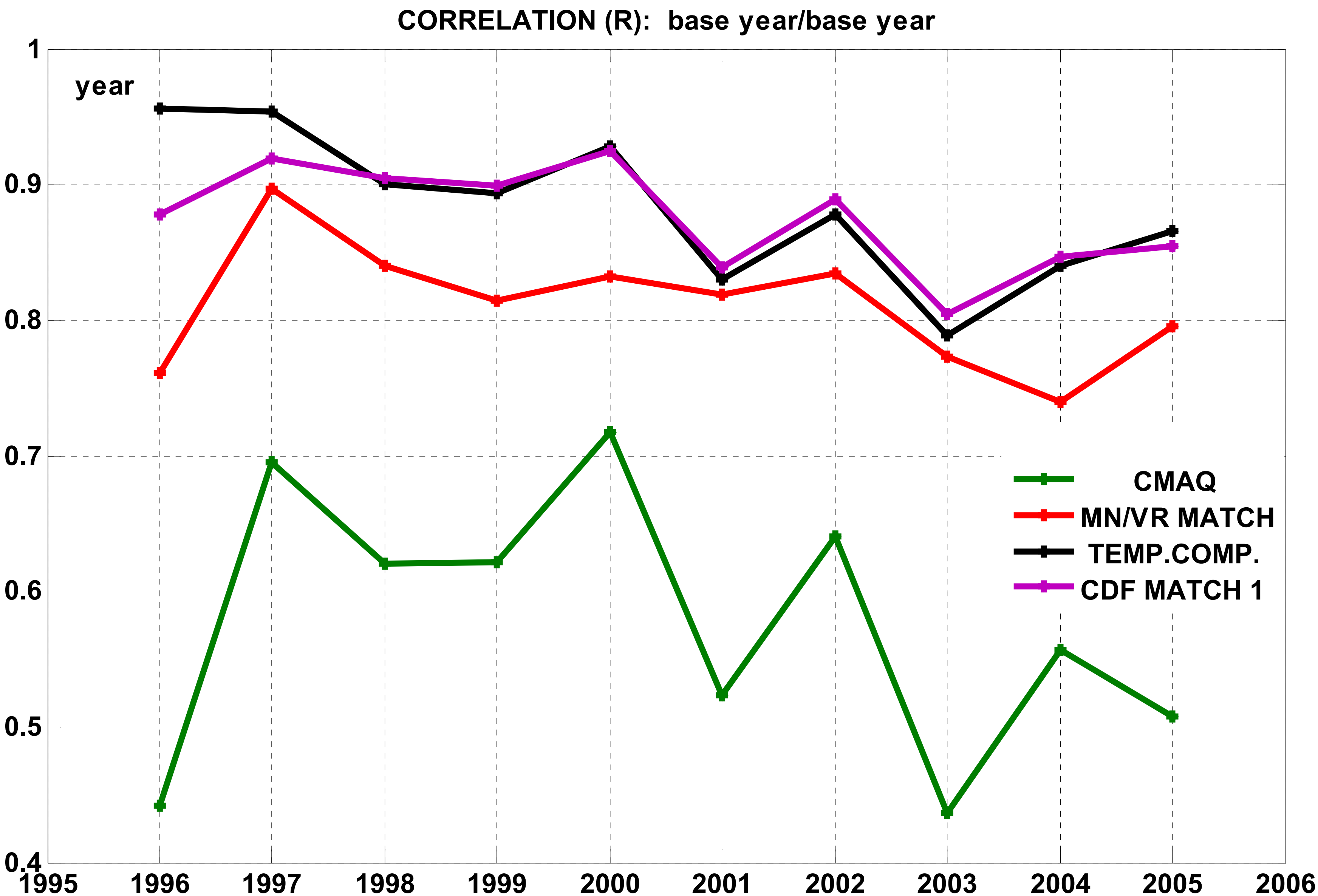


Figure 34. Same-year domain-wide for time series of 8H MAB



53

Figure 35. Same-year domain-wide CORRELATION (R) for time series of 8H.

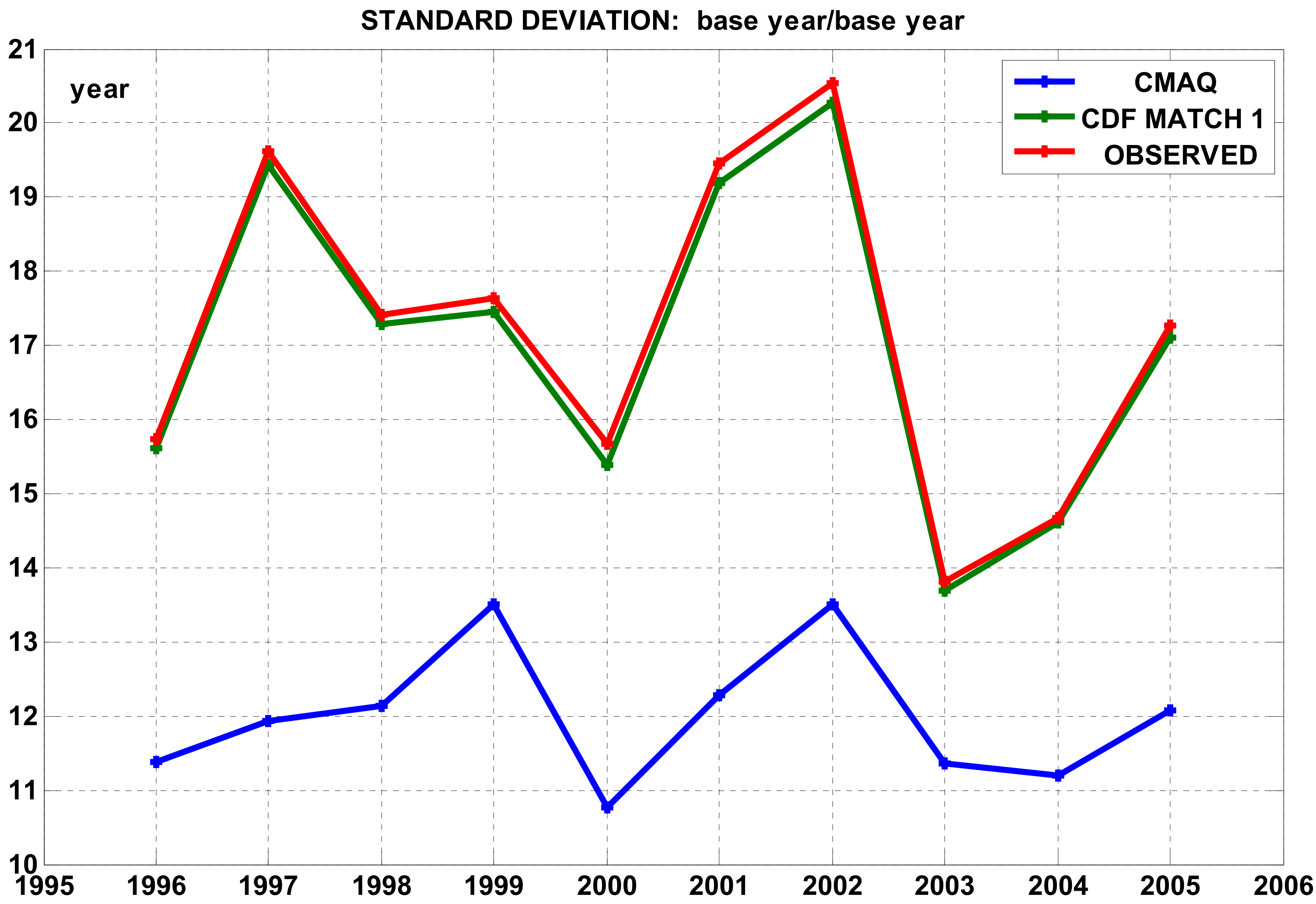


Figure 36. Same-year domain-wide STANDARD DEVIATION for time series of 8H.