United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-83-003e
July 1987

Air

EPA

Evaluation of Rural Air Quality Simulation Models

Addendum E: Graphic Summary of the Performance of Rural Air Quality Models

EPA-450/4-83-003e

Evaluation of Rural Air Quality Simulation Models

Addendum E: Graphic Summary of the Performance of Rural Air Quality Models

By

William M. Cox
Gerald K. Moss
Joseph A. Tikvart

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Radiation
Office of Air Quality Planning and Standards
Research Triangle Park NC 27711

July 1987

This report has been reviewed by the Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

EPA-450/4-83-003e

PREFACE

EPA is conducting an extensive program to evaluate the performance of air quality simulation models using statistical comparisons between measured and predicted concentrations. The operational performance of several rural models has been documented in tabular* and graphical form (Addendum B) for the Clifty Creek power plant and in tabular form (Addendum A, C and D) for three other large midwestern power plants. The purpose of this addendum is to summarize the performance of the rural models in a graphical format using the accumulated information presented earlier for the four plants. The composite behavior of each model is examined from both an operational and scientific perspective with particular emphasis on comparison of differences in performance among the models.

*Londergan, R. J., D. H. Minott, D. J. Wackter, T. Kincaid and D. Bonitata, 1982. Evaluation of Rural Air Quality Simulation Models. EPA Publication No. EPA-450/4-83-003. U. S. Environmental Protection Agency, Research Triangle Park, N. C. (NTIS No. PB 83-182758).

TABLE OF CONTENTS

		Page
	PREFACE	iii
	FIGURES	vi
1.	INTRODUCTION	1
2.	ANALYSIS OF GRAPHIC SUMMARIES	3
3.	SUMMARY AND CONCLUSIONS	15
	REFERENCES	17

APPENDICES

INTRODUCTION TO THE APPENDICES

A. APPENDIX A--KINCAID POWER PLANT
B. APPENDIX B--PARADISE POWER PLANT
C. APPENDIX C--MUSKINGUM RIVER POWER PLANT

FIGURES

Number		Page
1	Fractional Bias By Model Using High 25 Values For Each Data Base	9
2	Composite Fractional Bias By Averaging Period Using High 25 Values	10
3	Composite Net Fractional Bias Using High 25 Values	11
4	Composite Fractional Bias By Model Using High 25 Values For Each Stability Class	12
5	Composite Fractional Bias By Model Using High 25 Values For Each Wind Speed Class	13

Section 1

INTRODUCTION

EPA has an extensive ongoing program to evaluate the performance of air quality simulation models used for regulatory purposes. The performance of several rural models has been documented through extensive tabulations of performance statistics using field data collected around four rural power plants (1,2,3,4). The purpose of this report is to present graphically the composite performance of four rural models using the performance data assembled previously. At each site, SO_2 is the pollutant used for evaluation purposes.

The four rural models evaluted are (1) MPTER developed by EPA (2) MPSDM developed by ERT, (3) PPSP developed by the Martin Marietta Company, and (4) TEM-8A developed by the Texas Air Control Board. This analysis is confined to these rural models, since they essentially span the range of technology represented by the rural models as a group. The data base consists of two years assembled for Clifty Creek (1975 and 1976), two years at Muskingum River (1975 and 1976), one year at Paradise (1976), and essentially one year of data collected at Kincaid (1980 and portions of 1981).

The body of the report focuses on the composite operational performance of each model with emphasis on differences among the models. The Appendices provide a more in-depth analysis of the diagnostic performance components of the four models for three of the data bases. The graphical format used in the Appendices is similar to that used previously in which the performance of these models was presented graphically using the data assembled for the fourth data base at Clifty Creek (5).

SECTION 2

ANALYSIS OF GRAPHIC SUMMARIES

The graphic summaries of model performance (Figures 1-5) were developed from information presented in the Appendices and from an earlier report in which a comprehensive set of graphics were presented for Clifty Creek (5). When possible, the summary statistic is a composite average of the same statistic computed for each data base. In some instances, (e.g. individual stations) composite results among data bases are not meaningful and hence, only general trends are noted.

The performance of the models is displayed using the fractional bias statistic in which model bias is normalized. The general expression for the fractional bias is given by:

$$FB = 2^* (OB - PR)/(OB + PR)$$

The fractional bias of the average is computed using this equation where OB and PR refer to the averages of the observed and predicted values. The same expression is used to calculate the fractional bias of the standard deviation where OB then refers to the standard deviation of the observed values and PR refers to the standard deviation of the predicted values. The fractional bias was selected as the basic measure of performance because it has two desireable features. First, the fractional bias is symmetrical and bounded. Values for the fractional bias range between -2.0 (extreme overprediction) and +2.0 (extreme underprediction). Second, the fractional bias is a dimensionless number which is convenient for comparing the results from studies involving different concentration levels or even different pollutants.

The first figure displays four plots corresponding to the four models being evaluated. Only the outcomes for 1-hour averages are shown; however, the results using 3-hour and 24-hour averages are reasonably similar. In this figure, the fractional bias of the average is plotted vs the fractional bias of the standard deviation using only the 25 highest observed and predicted concentrations, unpaired in space and/or time. For a given model, six points are plotted, one point for each of the six data bases used in the evaluation. The placement of the points relative to the center of the plot (0,0) is a measure of the bias of a particular model. Variation among the data bases is indicated by the degree of spread of the points on each plot. For reference purposes, a "box" is included near the center of each panel which corresponds to performance that is within a factor-of-two for both the average and the standard deviation. Values of the fractional bias (average or standard deviation) that are less than -0.67 overpredict by more than a factor-of-two while values that exceed +0.67 underpredict by more than a factor-of-two. Values that fall entirely within the box correspond with performance that is within a factor-of-two for both performance measures.

For the MPTER model (upper left panel of Figure 1), the bias of both the average and the standard deviation is within a factor-of-two for each of the data bases except one. For Muskingum River (1976) the standard deviation of predicted values is larger than the standard deviation of observed values by over a factor of two (fractional bias less than -0.67).

For MPSDM, there is a greater tendency for overprediction than evidenced for MPTER. For Clifty Creek (1976) both the bias of the average and standard deviation are overpredicted by a factor greater than two. At Clifty Creek, for 1975, only the average is overpredicted by a factor greater than two

while at Muskingum River, for 1976, only the standard deviation is overpredicted by a factor greater than two. For the other three data bases, (Muskingum 1975, Paradise, and Kincaid), both measures of bias are within a factor-of-two.

For PPSP, large overpredictions are evident for each of the six data bases. In each case, both the average and standard deviation of the high 25 values are overpredicted by a factor that exceeds two.

Among the four models, TEM-8A appears to have the smallest average bias as indicated by the closeness of the points to the vertical axis. This contrasts with the bias of the standard deviation, which TEM-8A overpredicts significantly at Clifty Creek (1976) and Muskingum River (1976) and underpredicts significantly (fractional bias greater than +0.67) at Kincaid.

Figure 2 presents the composite performance of the models in which the mean fractional biases of the six data bases are displayed. Separate panels are shown corresponding to 1-hour, 3-hour and 24-hour averages. For 1-hour averages (first panel), TEM-8A exhibits the least average fractional bias followed closely by MPTER which is also relatively unbiased. MPSDM shows a tendency for modest overpredictions while PPSP overpredicts significantly for both the average and the standard deviation. The value plotted for each model in this panel is positioned at the centroid of the values shown on Figure 1 which displayed the performance of each data set for 1-hour averages. As the averaging period increases to 3-hours (top right panel) and to 24-hours (lower panel), there is a noticeable tendency for the points to shift upward and to the right indicating smaller overpredictions (PPSP and MPSDM) and larger underpredictions (MPTER and TEM-8A). For each of the

three averaging periods, PPSP seems to be clearly outperformed by the other three models.

Figure 3 redisplays the information shown in Figure 2 in a different format in which only the three better performing models (MPTER, MPSDM and TEM-8A) are compared. A composite measure of model bias is shown in the form of "buildings" where each building corresponds to one of the three averaging periods across the front axis and one of the three models across the receding axis. The measure of model performance used is the "net fractional bias", which is computed as the square root of the sum of squares of the two measures of bias, i.e., the bias of the average and the bias of the standard deviation. The composite measure of bias, which is the height of each building, is just the net fractional bias averaged across the six data bases.

Two features appear worth noting from this plot. One, there is a noticeable contrast in performance with averaging period. For MPSDM, the composite bias decreases with increasing averaging period while the opposite trend is apparent for TEM-8A. Secondly, MPTER appears to perform marginally better on average than the other two models.

The last two figures display information similar to that shown in Figure 1 except the biases for selected meteorological categories are contrasted.

The data consists of the 25 highest observed and 25 highest predicted values within each of four selected stability categories (Figure 4) and within each of three selected wind speed categories (Figure 5). The stability categories were selected as: very unstable (class A or B); unstable (class C); neutral (class D); and stable (class E, F, or G). The wind speed categories

were selected as: low wind speeds (less than 2.5 m/sec); medium (2.5 to 5 m/sec); and high (greater than 5 m/sec).

In Figure 4, the composite performance for each of the four stability categories is shown for each of the four models. The first panel shows a tendency for MPTER to underpredict the more stable conditions and to overpredict the more unstable conditions. For MPSDM, (top right panel), the spread between stability categories is more apparent; stable conditions are associated with very large underpredictions while neutral, unstable and very unstable conditions are associated with fairly large overpredictions, on average. For PPSP, the composite performance indicates large overpredictions for each stability category. For TEM-8A, the unstable categories are associated with very low bias. However, as the stability class moves from unstable to neutral and stable, an increasing tendency for underprediction becomes obvious.

Figure 5 shows the bias of the four models for three wind speed categories. For PPSP and MPSDM, there is a clear tendency for low winds to be associated with large overpredictions and high winds to be associated with small overpredictions or small underpredictions. For MPTER, no trend is apparent with wind speed class; only the medium winds show any tendency for bias where slight (average) overpredictions are evident. For TEM-8A, the high winds are associated with underpredictions. While the two lower wind speed classes exhibit no noticeable bias. Although composite results are not meaningful for individual stations, there were distinct patterns in the bias of the models as a function of station to source distance (Figures 11-16 in each of the three Appendices and the analogous figures in the Clifty Creek

stations accompanied by relatively large over and underpredictions at the stations nearest the source. For very unstable conditions, the models as a group displayed a tendency for decreasing overprediction as distance between source and receptor increases. For stable conditions, the pattern appeared to reverse (except PPSP), i.e., the models tended to underpredict at most of the stations but the degree of underprediction decreases as the source distance increases.

The fractional bias of the models as a function of hour of the day revealed very distinct patterns that were persistent among the six data bases (See Figures 17-20 in the Appendices and the corresponding figures in the Clifty Creek report). The typical diurnal pattern consisted of marked underpredictions during the early morning and late evening hours and moderate overpredictions during the midday hours.

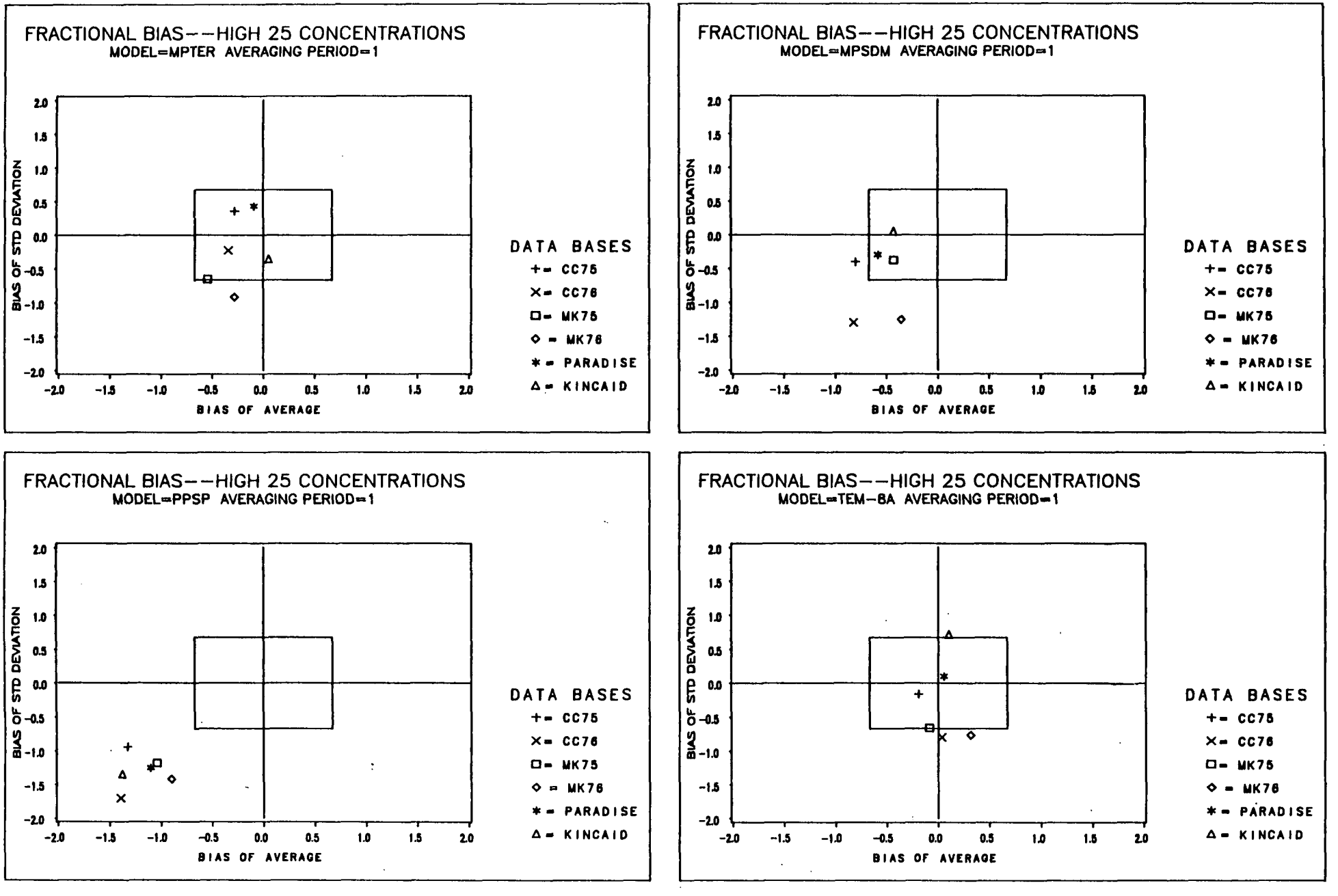


Figure 1. Fractional Bias By Model Using High 25 Values For Each Data Base

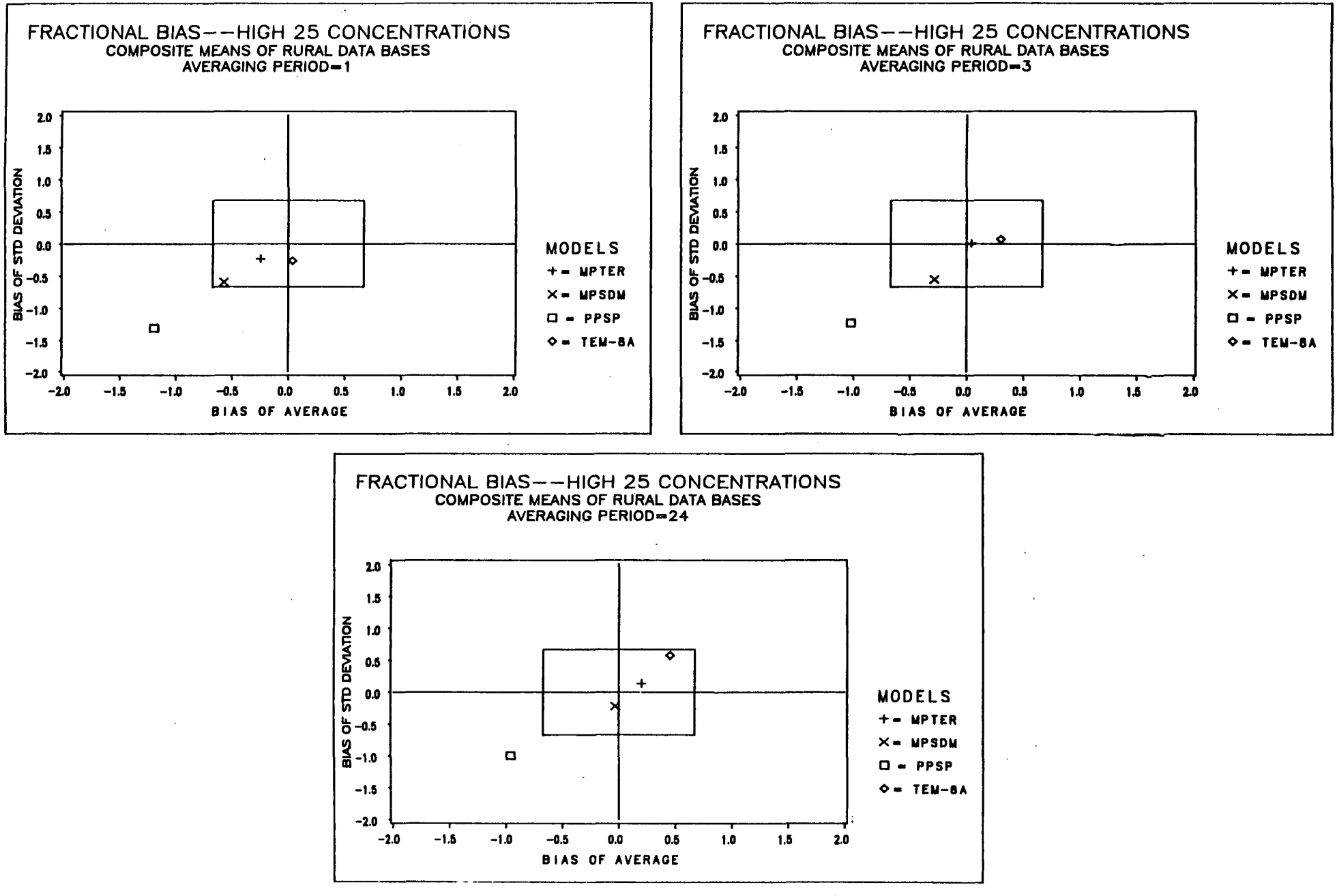


Figure 2. Composite Fractional Bias By Averaging Period Using High 25 Values

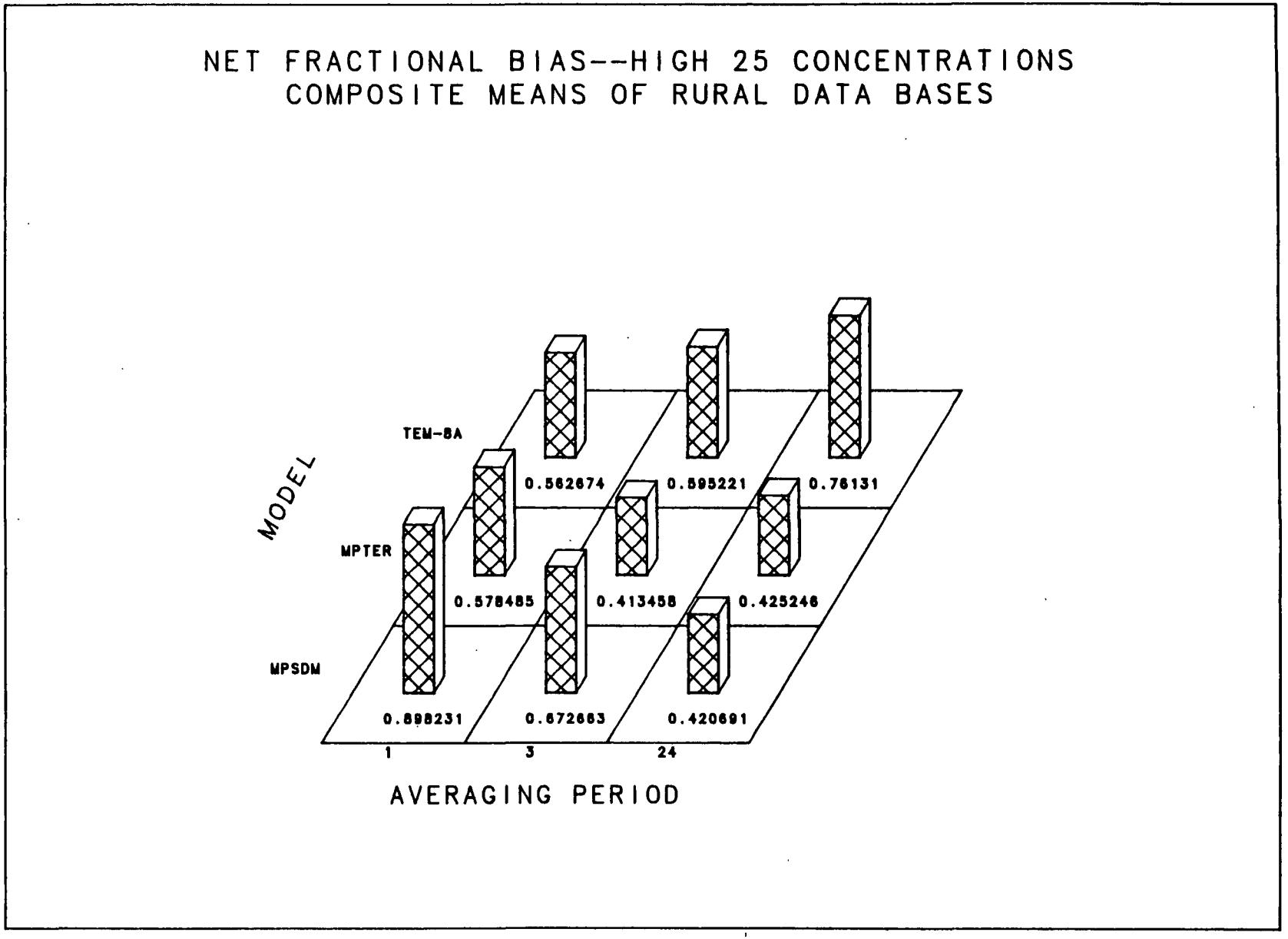


Figure 3. Composite Net Fractional Bias Using High 25 Values

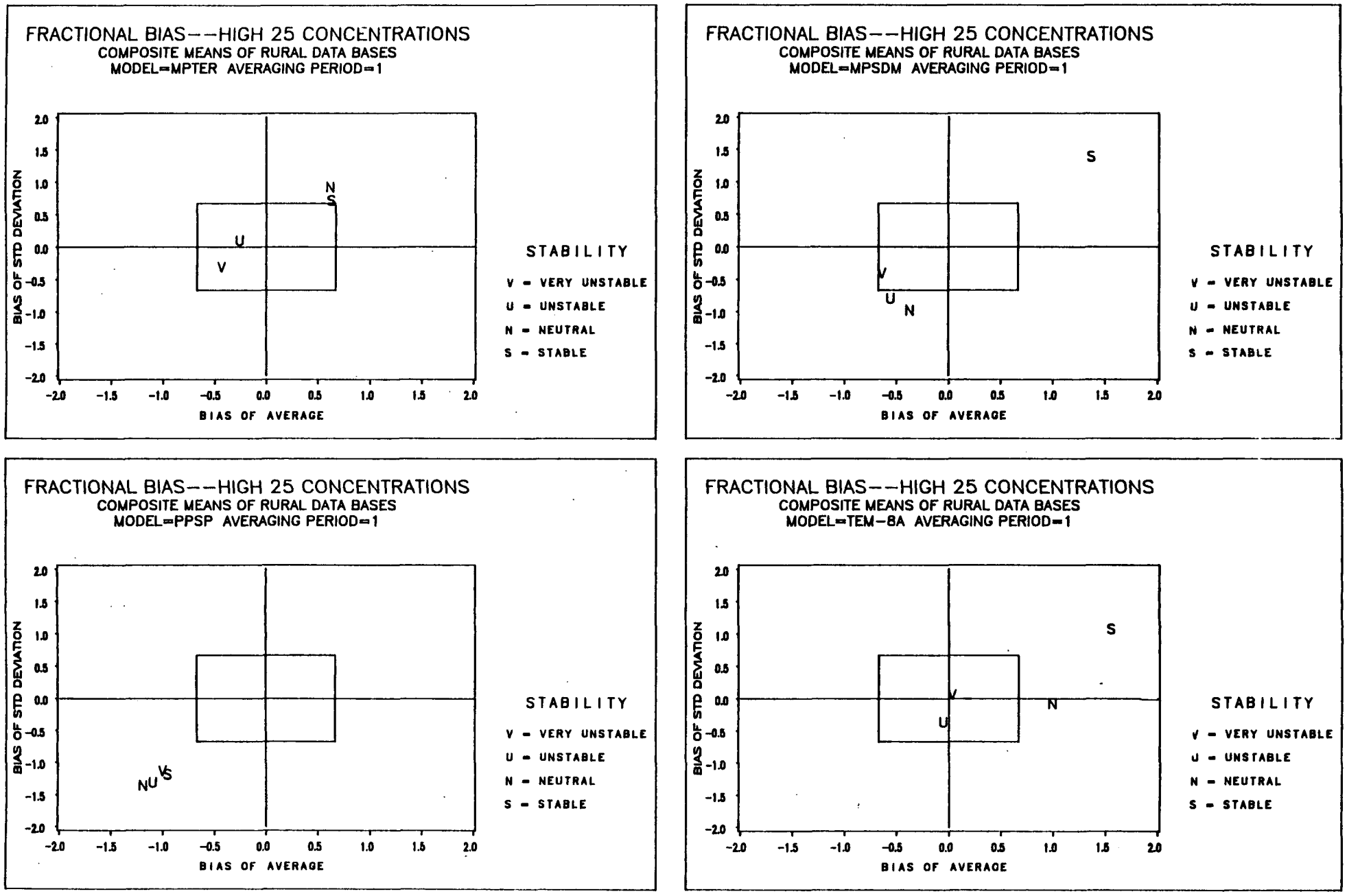


Figure 4. Composite Fractional Bias By Model Using High 25 Values For Each Stability Class

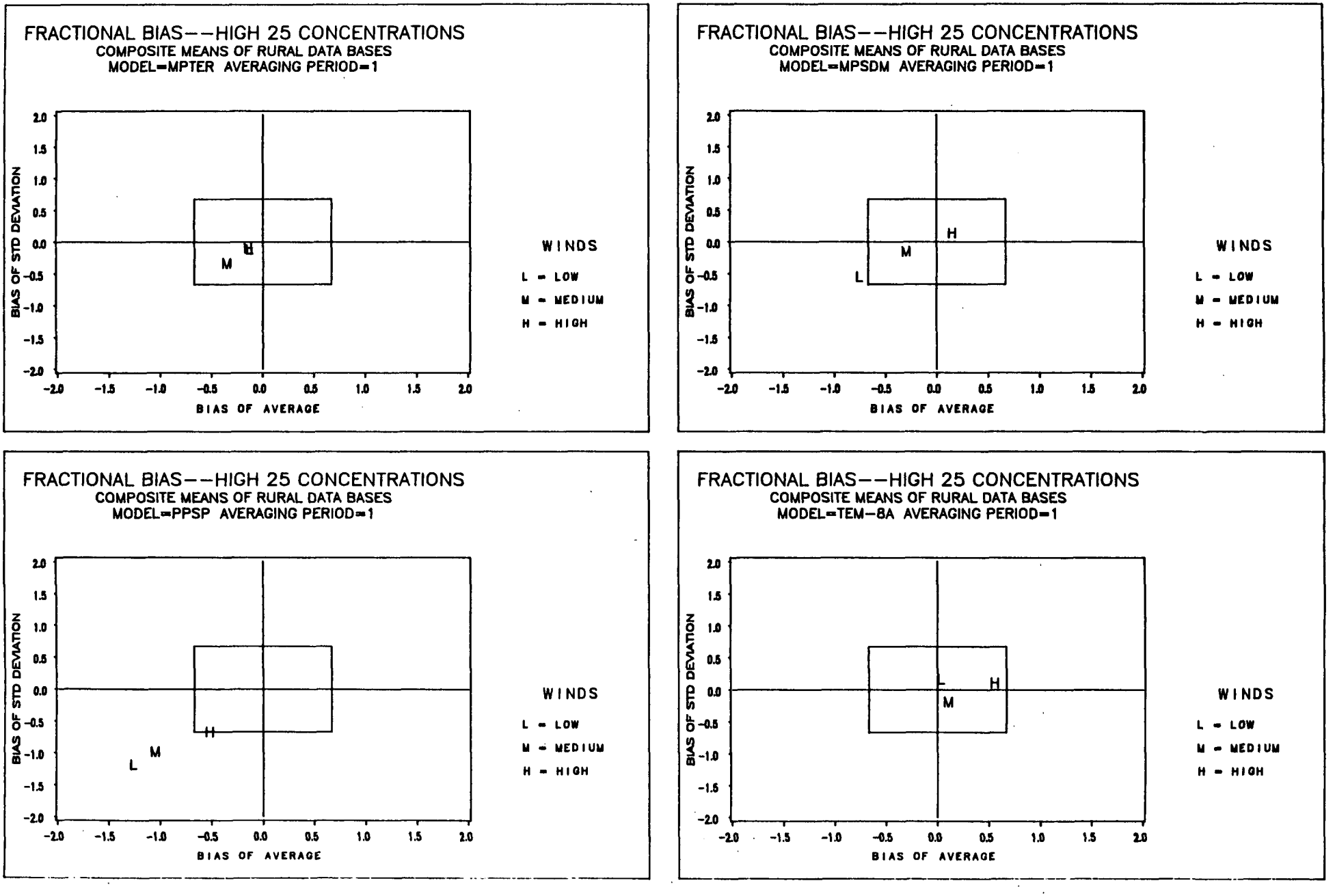


Figure 5. Composite Fractional Bias By Model Using High 25 Values For Each Wind Speed Class

SECTION 3

SUMMARY AND CONCLUSIONS

The performance of four rural air quality simulation models has been presented in the form of comprehensive graphic displays using SO_2 data collected at four power plant sites. Detailed graphic displays of model performance were presented in an earlier report for Clifty Creek, and for three additional data bases in the Appendices to this report. The results from the six data bases (two years of data were available at Muskingum River and Clifty Creek), were used to construct composite measures of model performance and displayed in the form of fractional bias plots using the mean and standard deviation of the highest 25 concentrations. In addition, variations in performance among the data bases, among averaging periods, and among meteorological conditions were examined.

From an operational performance standpoint (Figures 1-3), PPSP performed poorly across the six data bases, overpredicting the mean and the spread of the highest 25 values by well over a factor-of-two. MPSDM also tended to overpredict the highest concentrations but not by the degree found for PPSP. Both MPTER and TEM-8A appeared to have relatively little bias, especially for 1-hour and 3-hour averages. TEM-8A demonstrated a noticeble tendancy for underprediction of the highest 24-hour averages.

Breakdown of performance by stability category (Figure 4) indicated a strong tendency for underprediction of the highest concentrations for stable conditions (except for PPSP) and overpredictions for the more unstable conditions. It is particularly noteworthy for very unstable and unstable conditions

that the composite bias for TEM-8A across the six inventories is virtually nill. Variations in model bias among wind speed categories (Figure 5) was relatively small showing a tendency for slight underpredictions for high wind speeds and little bias for the medium and low wind speed categories. Exceptions to this pattern were obvious for PPSP, where overpredictions were evident for each wind category, and, for MPTER, where the differences in performance among wind categories were very small.

Although composite results are not meaningful for individual stations, the trend in the fractional bias with distance between the stations and the source were similar among the data sets. There appeared to be a consistent tendency for decreasing overprediction with distance for unstable conditions and, conversely, a tendency for decreasing underprediction with distance for the more stable conditions.

Diurnal patterns among the data bases were also reasonably consistent. In most cases, early morning and late evening hours are associated with large underpredictions while the mid-day hours are associated with small to moderate overpredictions.

It is generally concluded that the graphic formats used to assess the operational and diagnostic components of model performance provide insight into model behavior not readily obtained from tabular presentations of data. Hopefully a framework has been established that can be used to develop a standardized graphics approach for displaying and intercomparing the performance of air quality models.

REFERENCES

1. Londergan R. J., Minott D. H., Wackter D. J. and Bonitta D., Evaluation of Rural Air Quality Simulation Models, EPA-450/4-83-003, Environmental Protection Agency, Research Triangle Park, N. C., October 1982.

2. Cox, W. M., and Moss G. K., Evaluation of Rural Air Quality Simulation Models, Addendum A: Muskingum River Data Base, EPA-450/83-003a, Environmental Protection Agency, Research Triangle Park, N. C., June 1985.

3. Cox, W. M., Rorex H. W., and Moss G. K., Evaluation of Rural Air Quality Simulation Models, Addendum C: Kincaid SO2 Data Base, EPA-450/83-003c, Environmental Protection Agency, Research Triangle Park, N. C., March 1986.

4. Cox, W. M., Rorex H. W., Moss G. K., and Baldridge E. W., Evaluation of Rural Air Quality Simulation Models, Addendum D: Paradise SO2 Data Base, EPA-450/83-003d, Environmental Protection Agency, Research Triangle Park, N. C., January 1987.

5. Cox, W. M., G. K. Moss, J. A. Tikvart and K. W. Baldridge, Evaluation of Rural Air Quality Simulation Models, Addendum B: Graphical Display of Model Performance Using the Clifty Creek Data Base, EPA/483-003b, Environmental Protection Agency, Research Triangle Park, N. C., August 1985.

APPENDICES

INTRODUCTION TO THE APPENDICES

APPENDIX A--KINCAID POWER PLANT
APPENDIX B--PARADISE POWER PLANT
APPENDIX C--MUSKINGUM RIVER POWER PLANT

INTRODUCTION TO THE APPENDICES

The performance of four rural models (MPTER, MPSDM, PPSP, and TEM-8A) is presented graphically in a format similar to that used in the earlier report for Clifty Creek. Results for the three inventories are reported separately as follows: Appendix A--Kincaid power plant, Appendix B--Paradise power plant and Appendix C--Muskingum River power plant, 1975 and 1976. The purpose of the Appendices is to provide complete documentation regarding how each model performs both from an operational and diagnostic viewpoint. Also they provide the basis for material presented in the body of the report that analyzes composite performance of the models across the inventories (including Clifty Creek) and for evaluating differences in model performance among the inventories. Throughout both the operational and diagnostic phases of the discussion in these appendices, graphics are displayed for two basic data sets (1) the 25 highest observed and predicted concentrations unpaired in space and time and (2) all observed and predicted concentrations completely paired in space and time.

The figures used to display the performance of the models are presented in the same sequence for each of the three inventories. Figure 1 in each of the three appendicies presents a map that displays the location of the monitoring stations relative to the stack(s) including information regarding terrain. Figures 2-4 presents information oriented to the operational performance of the models. Graphs are provided by averaging period and data set using fractional bias plots, quantile-quantile plots and frequency distributions. Performance is given only for all stations and meteorological conditions combined.

A portion of the diagnostic information is provided in Figures 5-10 in the form of fractional bias plots. These plots display the fractional bias of the standard deviation vs the fractional bias of the average for individual stations, and for individual wind speed and stability categories. Further diagnostic information is provided in Figures 11-16 in the form of bias of the average as a function of station to source distance by stability and wind speed category. Finally, Figures 17-20 present comparisons of model performance in the form of diurnal plots in which model bias of the average is displayed for each hour of the day for all stations combined and as a function of station groupings by distance from the source. For further information regarding interpretation of the various types of graphs, the reader is referred to the Clifty Creek graphics report(5).

APPENDIX A--KINCAID POWER PLANT

Figure A.1 shows the location of each of the 30 SO_2 monitoring stations located around the Kincaid plant. The stations are located at distances ranging between approximately 2 km and 20 km. Since the area surrounding the plant is essentially flat, no terrain features are shown.

Figure A.2 summarizes the direction and relative magnitude of the fractional bias of the average and the fractional bias of the standard deviation for each of the four models. The left hand column displays three panels corresponding to the three averaging periods for the 25 highest observed and predicted concentrations. PPSP obviously overpredicts for each averaging period while the other three models are relatively unbiased. There is a tendency for TEM-8A and MPTER to underpredict the scatter as evidenced by the consistently positive values for fractional bias of the standard deviation. Figure A.3, which displays Q-Q plots of the 25 highest observed and predicted concentrations, supplements the information shown in Figure A.2. The magnitude of overpredictions by PPSP is clear from the distance between the points and the diagonal line of "perfect" fit. Also evident is the somewhat nonlinear relationship between the rank ordered observed and predicted values, especially for PPSP and MPSDM.

The right hand side of Figure A.2 presents fractional bias information using all data pairs. There is a clear shift of points to the right in each panel which indicates a tendency for smaller overpredictions (PPSP) and larger underpredictions (MPTER). Figure A.4, which presents the same information in the form of cumulative frequency distributions, provides additional insight into the nature of the disparity between observed and

predicted values. For example, the disagreement between observed and predicted values again appears to be a function of the magnitude of observed values; the difference is especially evident for 1-hour and 3-hour averages.

Figure A.5 displays the bias at each of the 30 monitoring stations, by model, using the 25 highest concentrations at each station. Consistent with results shown in Figure A.2, PPSP overpredicts at the majority of stations for both the average and the "scatter". Although there is a rather large range in the magnitude of the fractional bias, a more clearly defined pattern will be evident in subsequent displays involving distance between source and monitoring stations. Figure A.6 displays the bias for each of the four stability categories using the 25 highest concentrations for each stability category. Except for PPSP, there is a clear separation in performance among stability categories with large underpredictions for stable conditions and relatively modest overpredictions for the more unstable conditions. Figure A.7 displays similar results for wind speed categories. Compared with the spread among the data points seen for stations and stability conditions, the range in bias among wind speed categories is quite small. For MPTER, MPSDM and TEM-8A, the biases are all within a factor of two. Figures A.8-A.10 are analogous to Figures A.5-A.7 except that all data paired in time and space are used to compute the fractional biases. MPSDM and TEM-8A appear to exhibit the least overall bias as indicated by the clustering of points inside the framed area that corresponds to a factor of two. While PPSP continues to overpredict for most data categories, the degree of overprediction is not as severe as with the 25 highest concentrations.

Figure A.11 displays the fractional bias of the average as a function of distance between the source and each monitoring station. For the more distant stations, each model exhibits a relatively low bias. The biases at the closest stations (less than 4 miles) are obviously larger, ranging from overpredictions by PPSP (less than -1) to underpredictions by MPTER (greater than +1). Figure A.12 and A.13 contain displays identical to Figure A.11 except they are subdivided by stability category (Figure A.12) and wind speed category (Figure A.13). While the pattern for very unstable conditions is similar to that shown in Figure A.11, the range in the bias as a function of distance is somewhat larger for the unstable and neutral conditions. For stable conditions each of the models severely underpredicts at all distances except for PPSP where large overpredictions close to the source become slight underpredictions at the most remote station. Figures A.14-A.16 are identical to the previous three figures except that all paired data are used. Results are reasonably similar to those for the highest 25 values except there is a greater tendency toward underpredictions.

Figure A.17 displays fractional bias of the average as a function of hour of the day using the 25 highest values. For all models, except PPSP, large underpredictions occur early in the morning and again later in the day. Midday hours are associated with relatively little bias. PPSP displays large overpredictions for all hours. Figure A.18 presents the same information except subdivided by stations grouped by distance from the plant. The stations are grouped as follows: Group A -- less than 5 km; Group B -- between 5 and 10 km; Group C -- between 10 and 15 km; Group D -- greater than 15 km. The diurnal patterns for the more distant stations (Group D) contrast with those at the closer stations (Group A) in that the

range in bias is smaller for the more distant stations, especially for PPSP. Finally, Figures A.19 and A.20 display information in an identical format to Figures A.17 and A.18 except all paired data are used. Patterns are quite similar; again, the tendency toward greater underprediction is apparent.

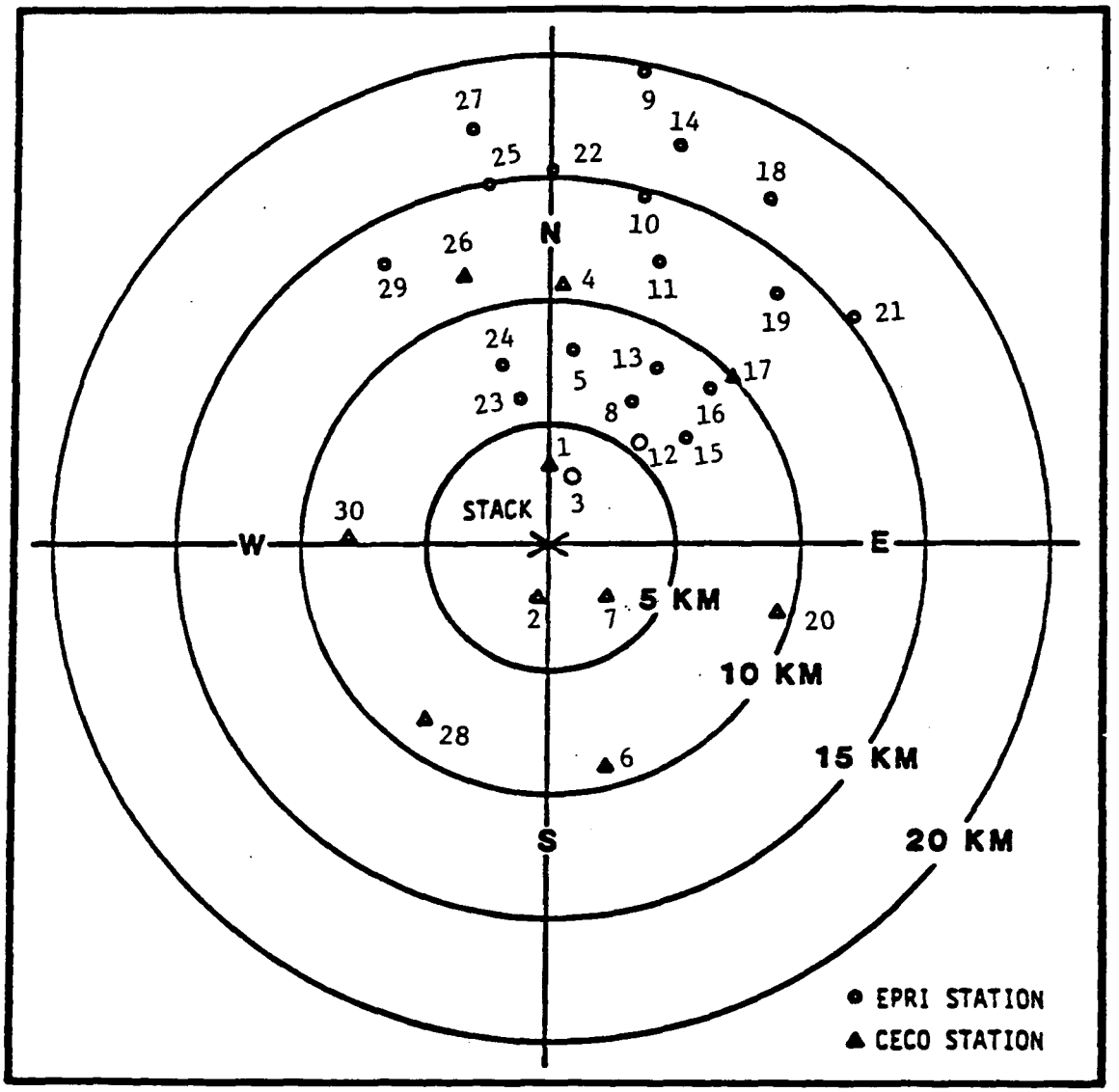


Figure A.1 Field Monitoring Network Near The Kincaid Power Plant

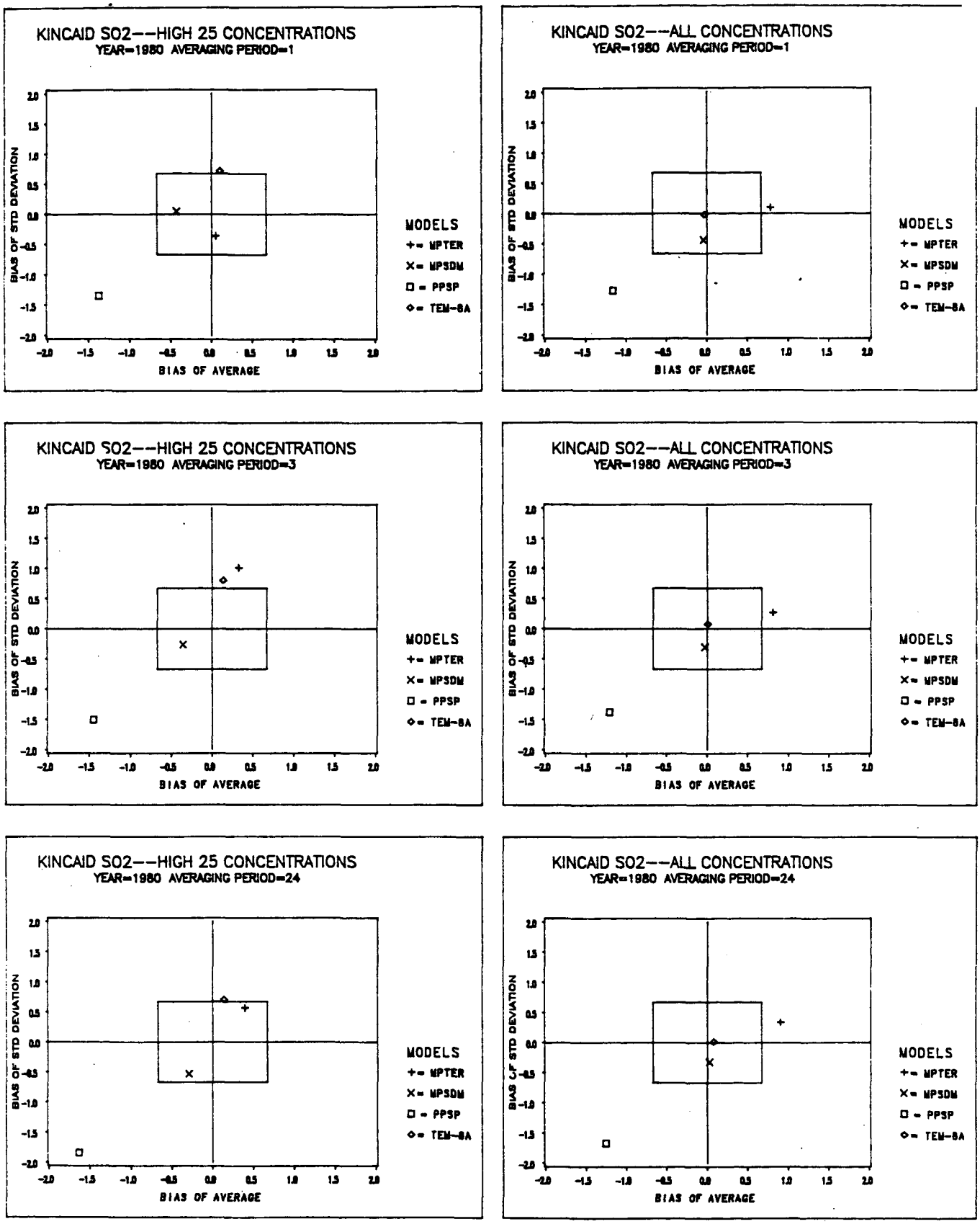
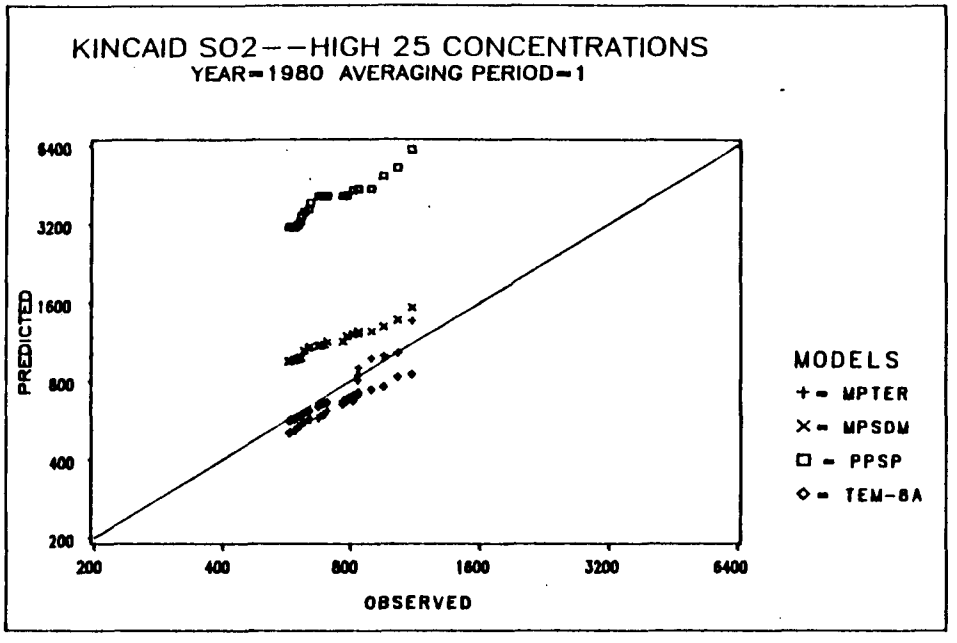
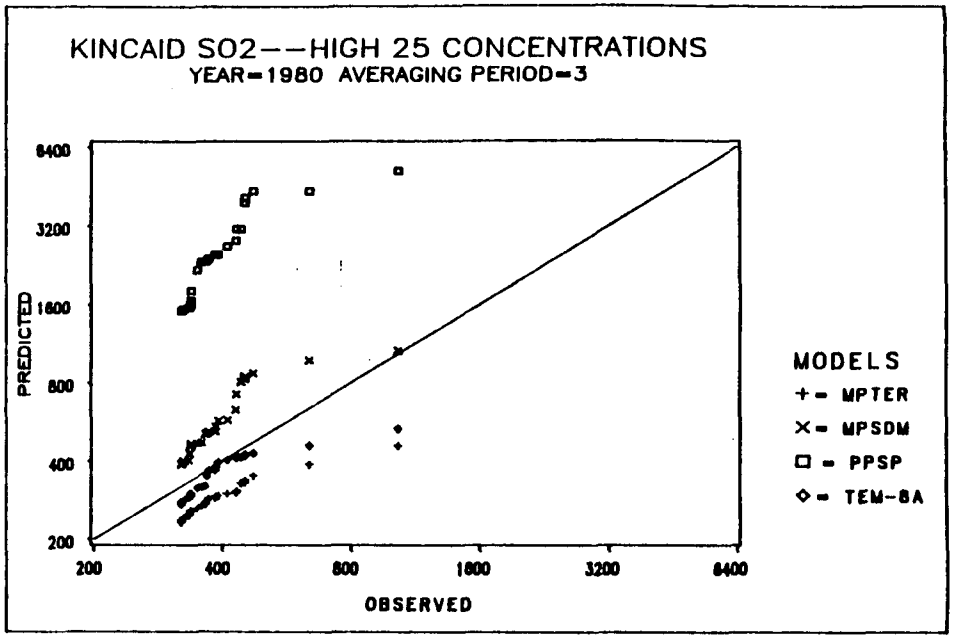


Figure A.2 Fractional Bias By Averaging Period Using High 25 Values And All Paired Values





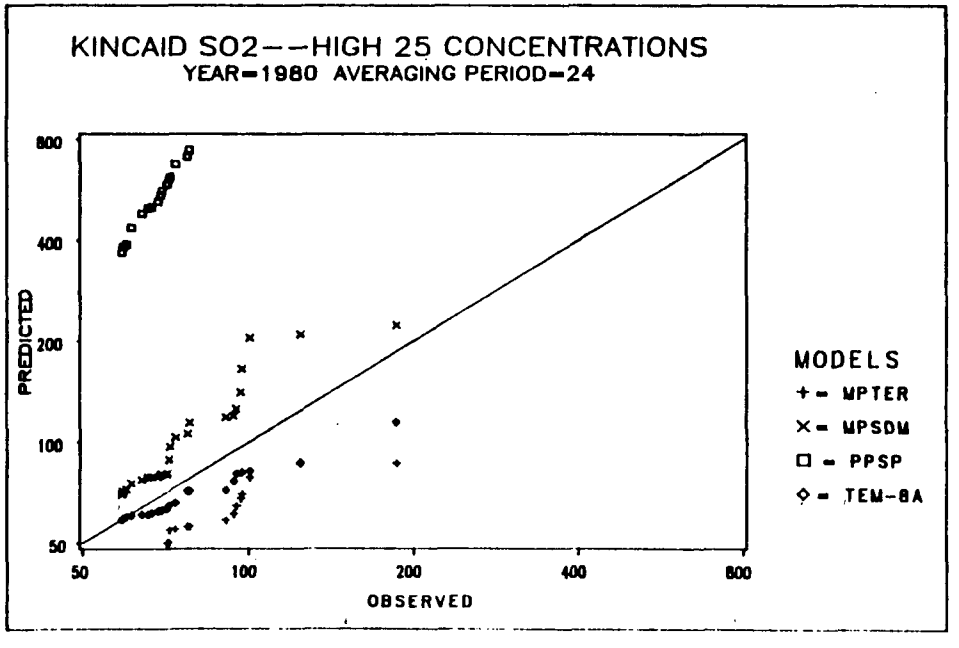
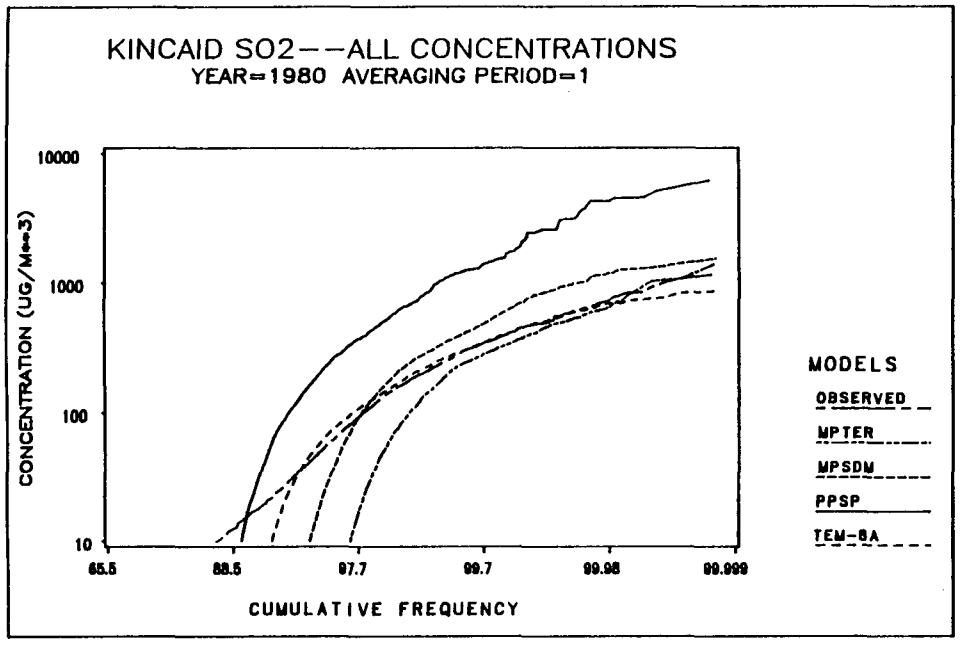
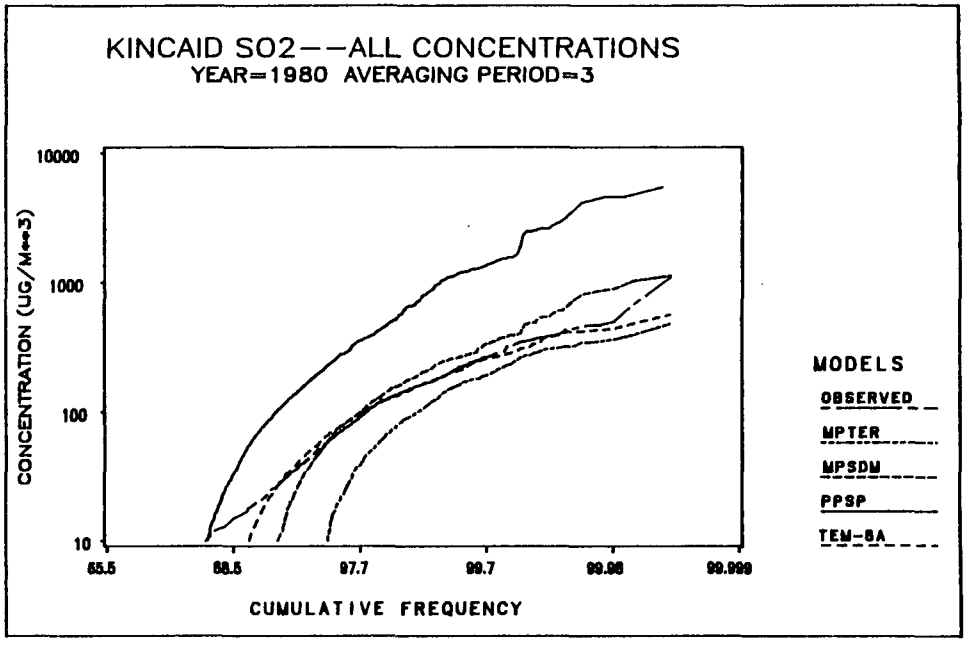


Figure A.3 Predicted Quantiles Vs Observed Quantiles By Averaging Period Using High 25 Values





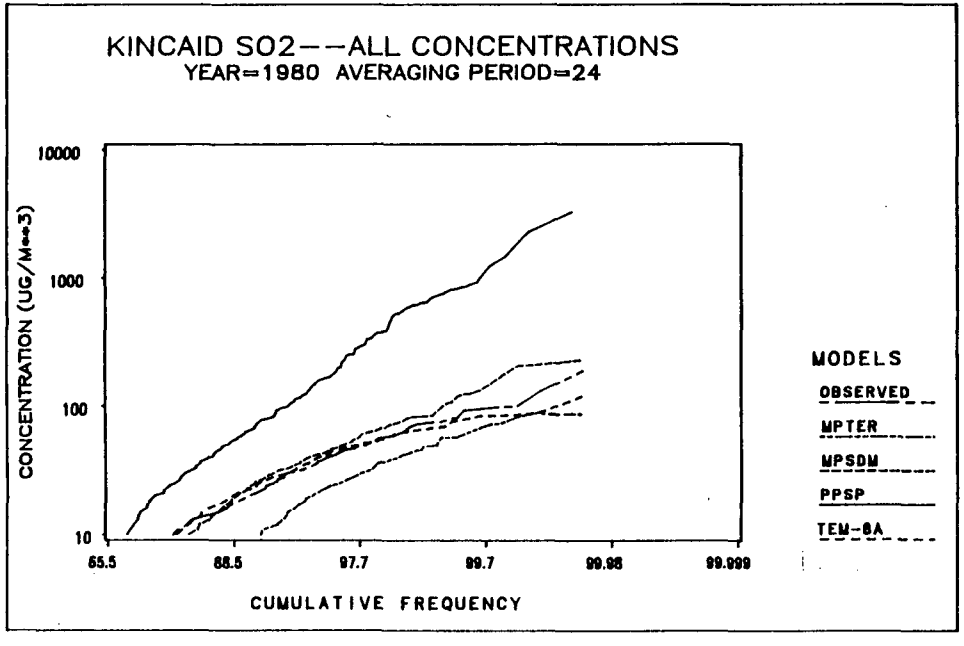


Figure A.4 Cumulative Frequency Distributions By Averaging Period Using All Paired Values

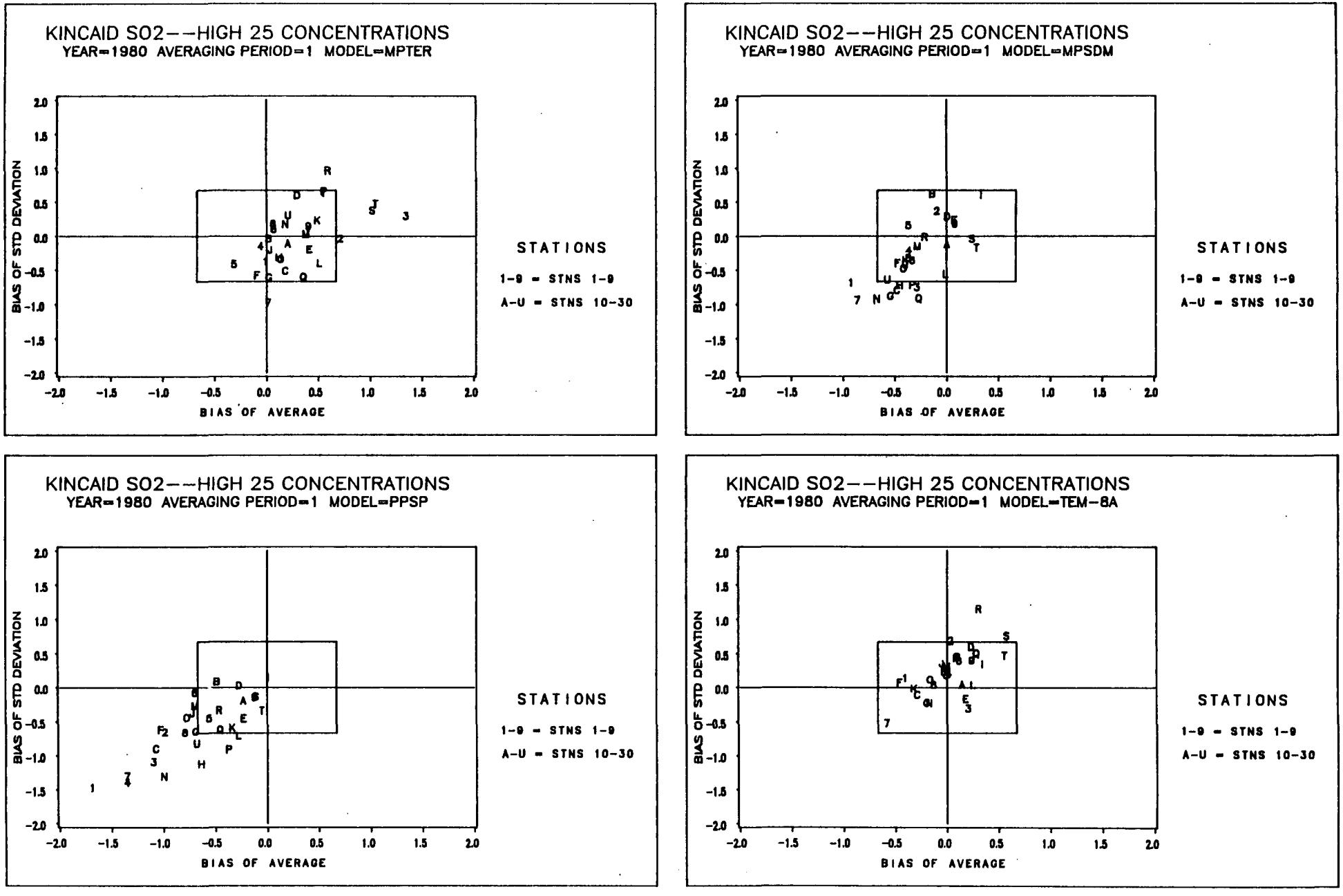


Figure A.5 Fractional Bias By Model Using 25 High Values For Each Station

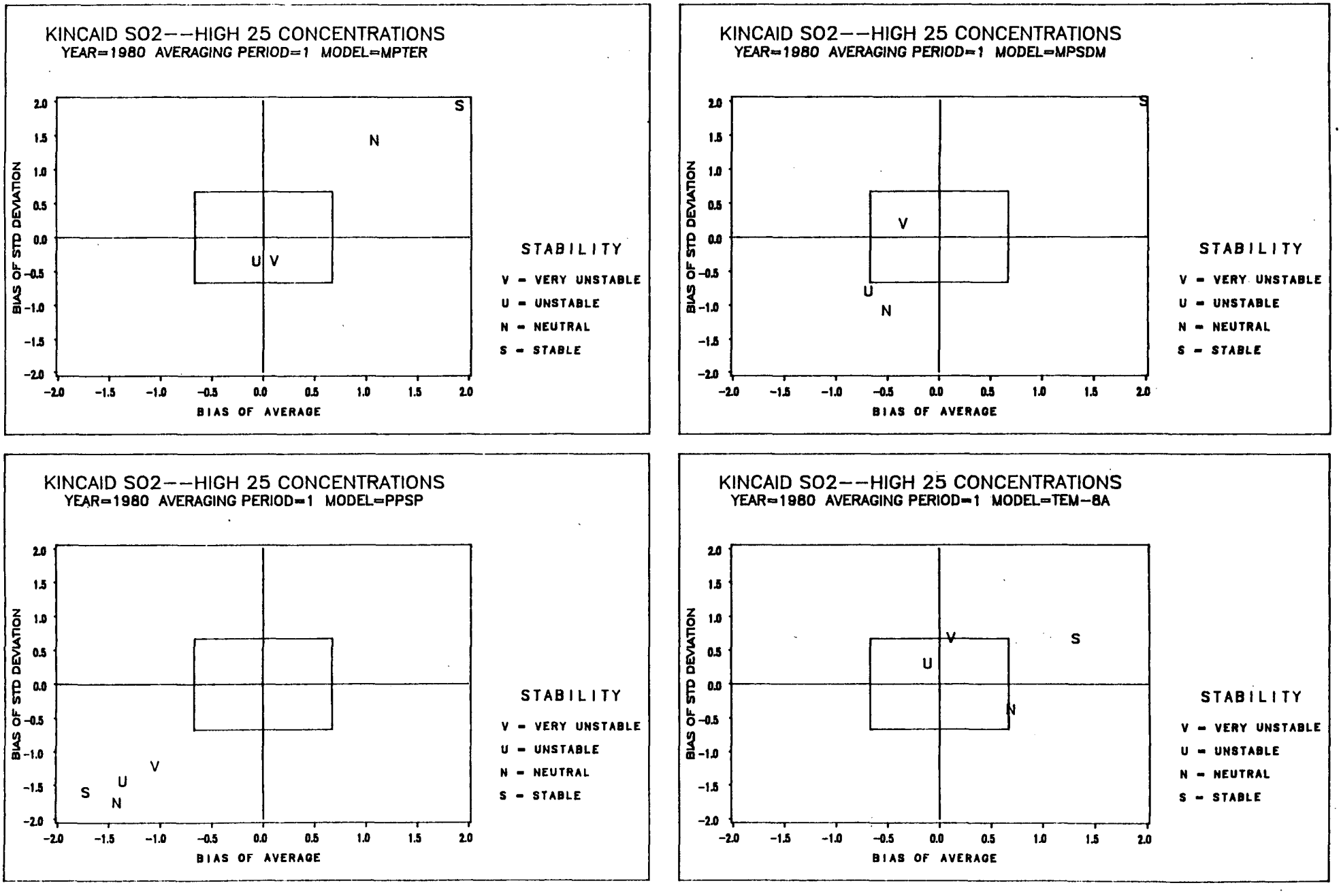


Figure A.6 Fractional Bias By Model Using 25 High Values For Each Stability Class

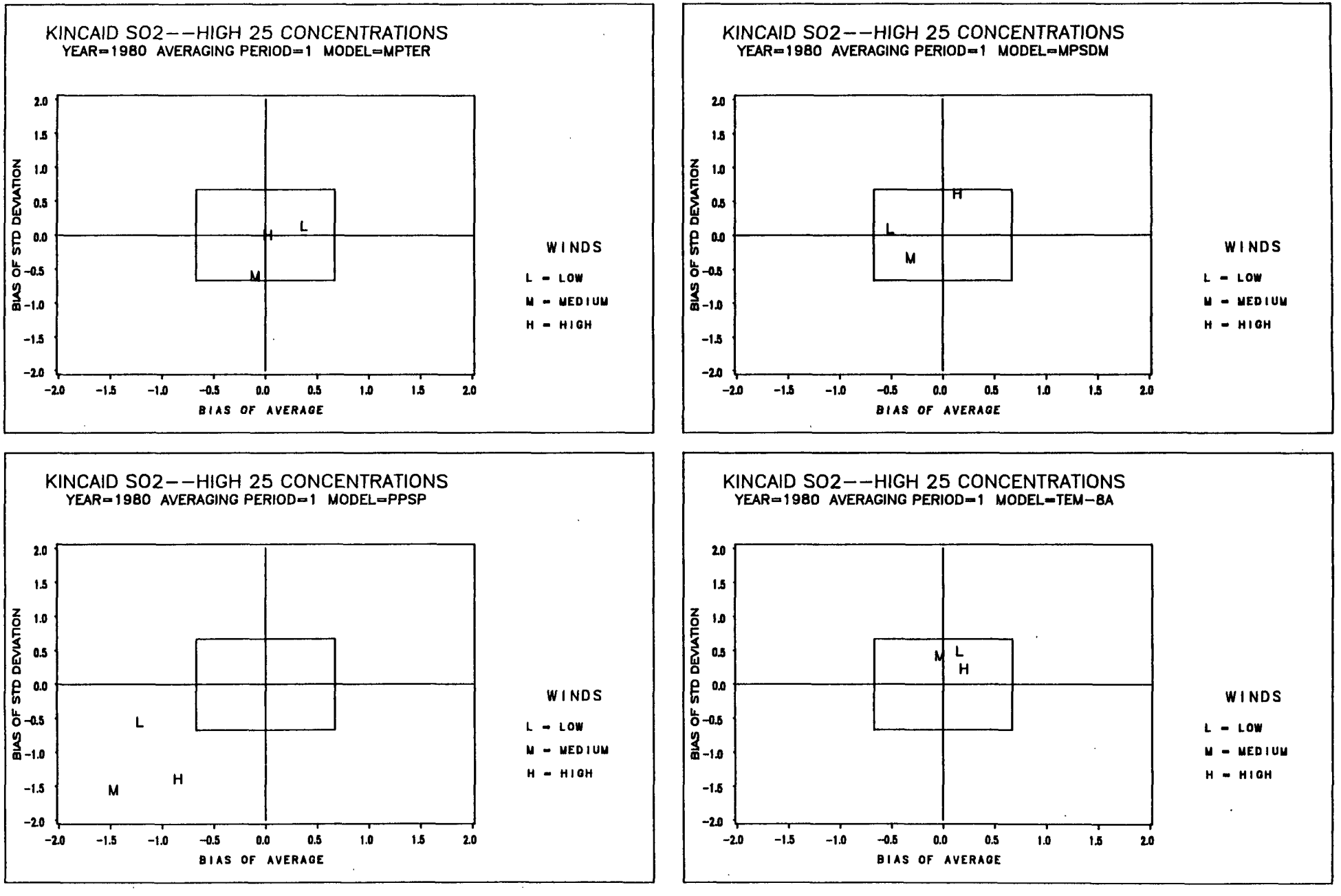


Figure A.7 Fractional Bias By Model Using 25 High Values For Each Wind Speed Class

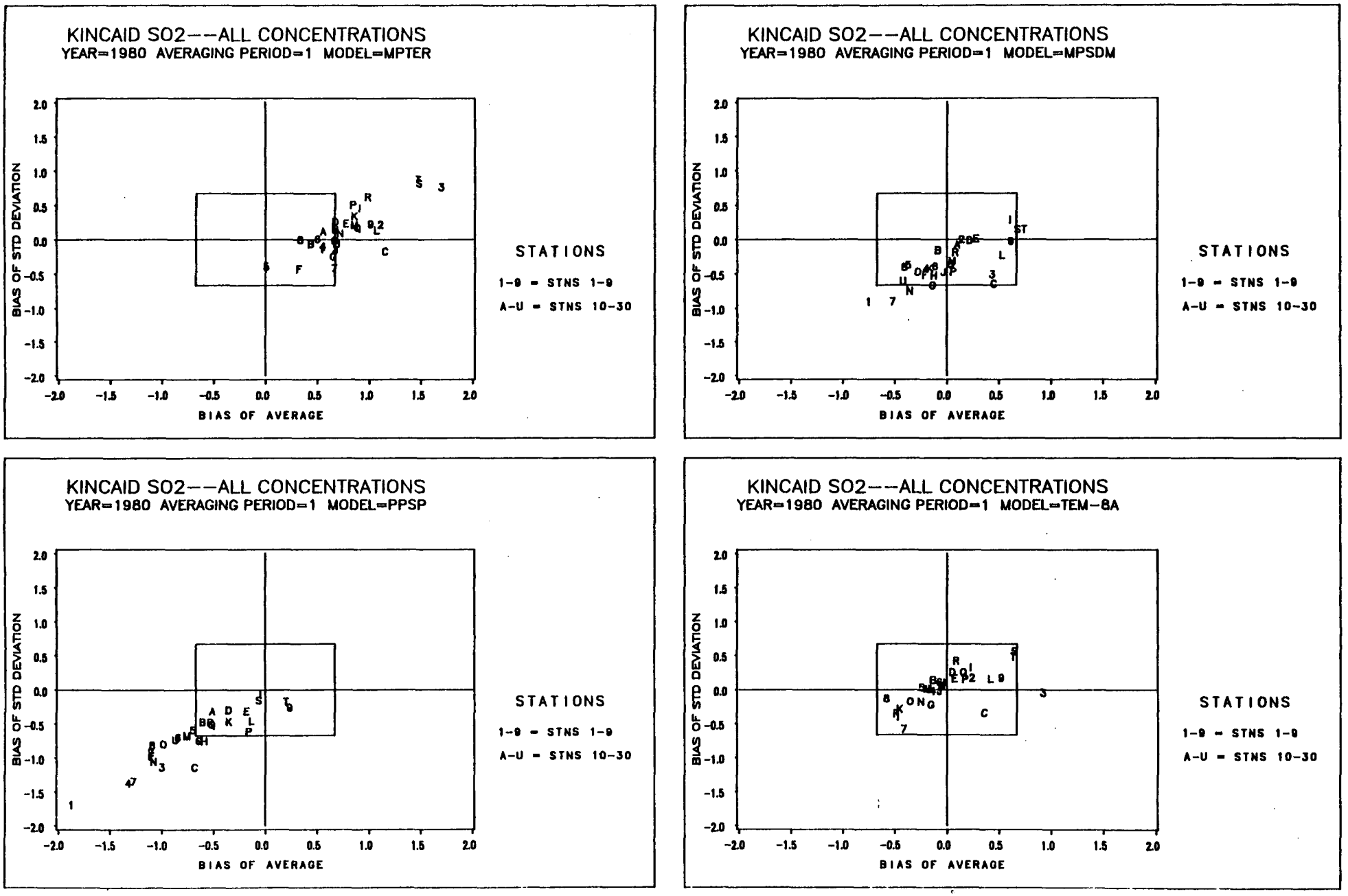


Figure A.8 Fractional Bias By Model Using All Paired Values For Each Station

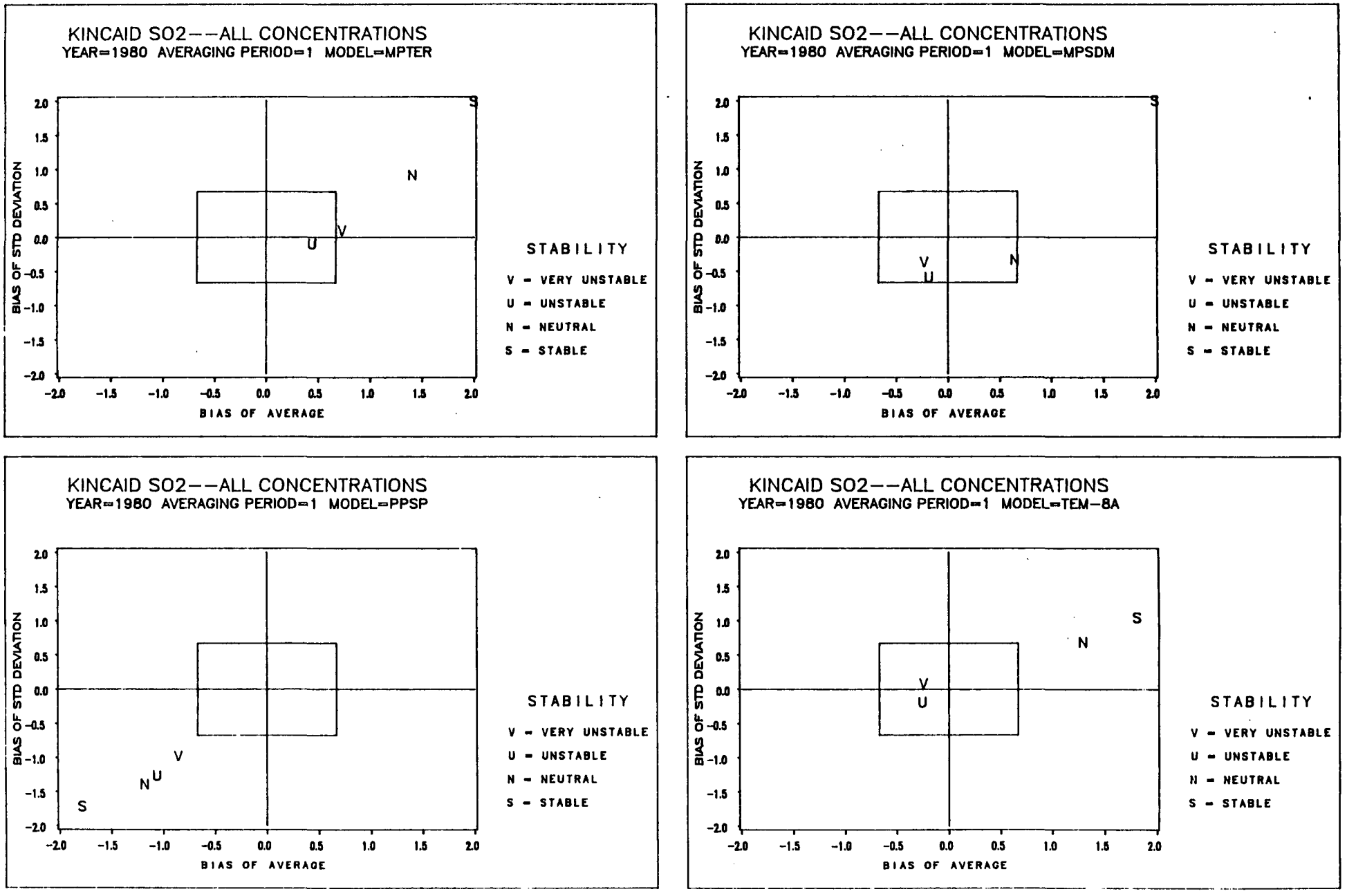


Figure A.9 Fractional Bias By Model Using All Paired Values For Each Stability Class

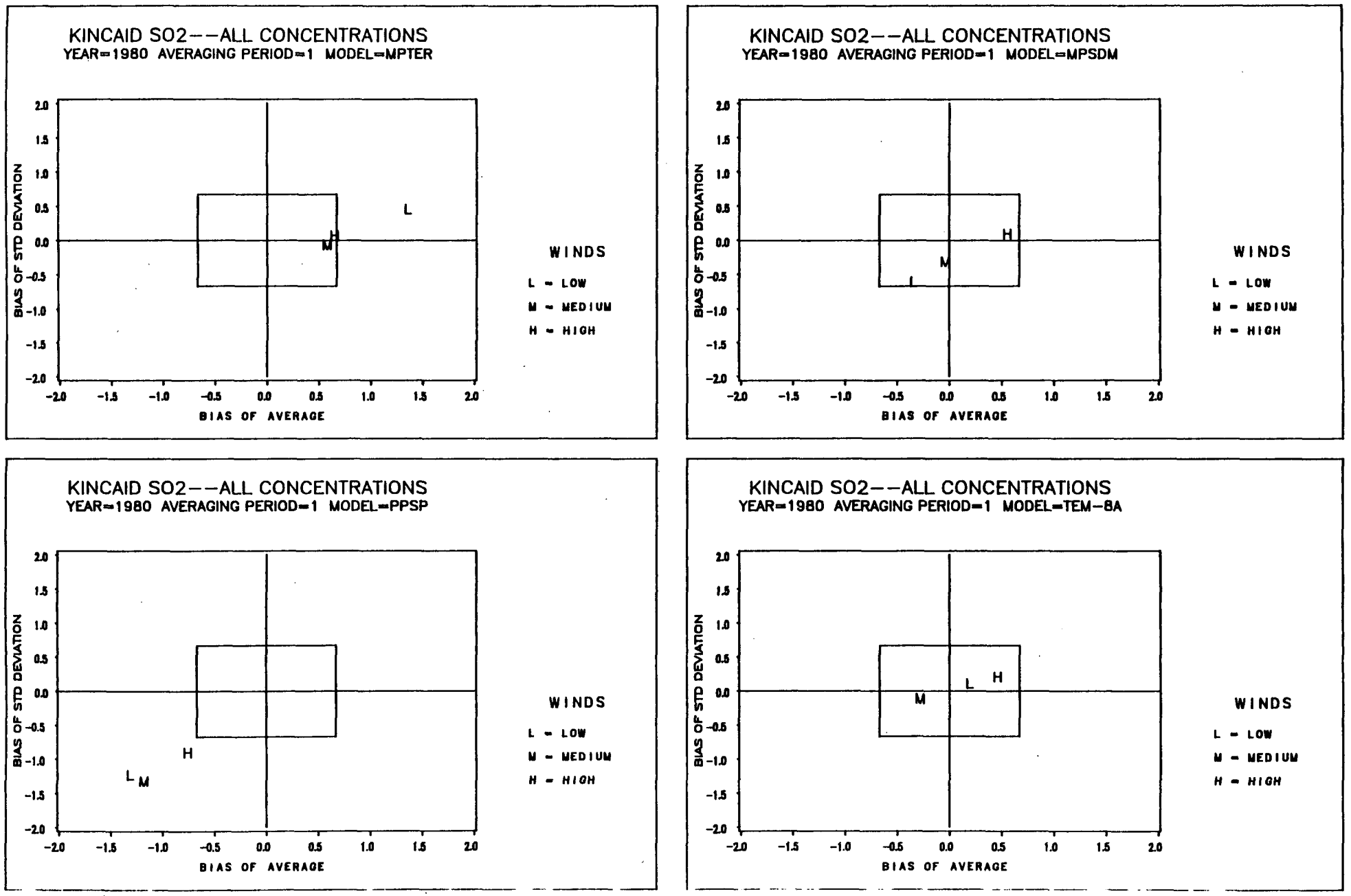


Figure A.10 Fractional Bias By Model Using All Paired Values For Each Wind Speed Class

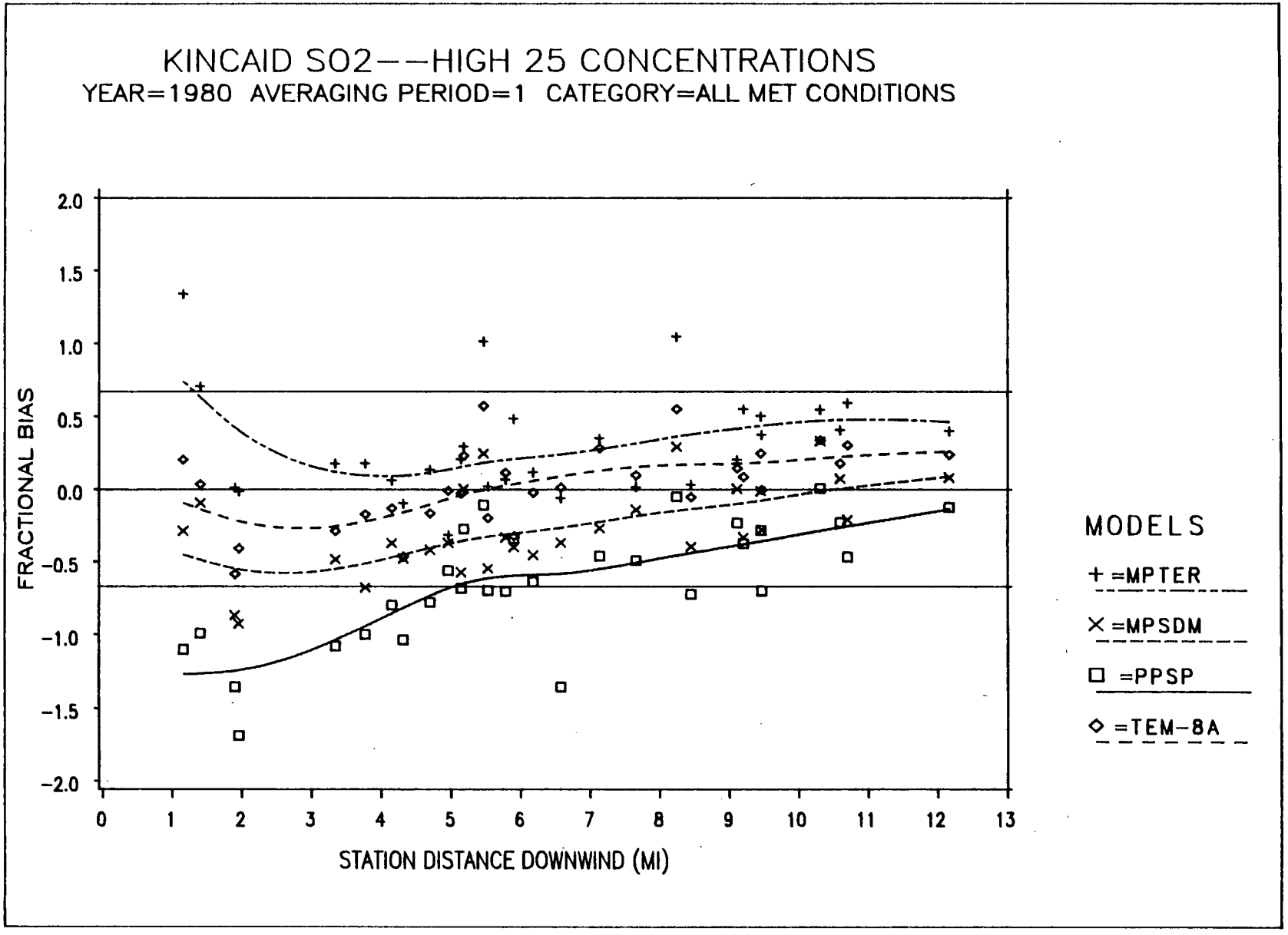


Figure A.11 Fractional Bias Of The Average Vs Station Distance Using High 25 Values

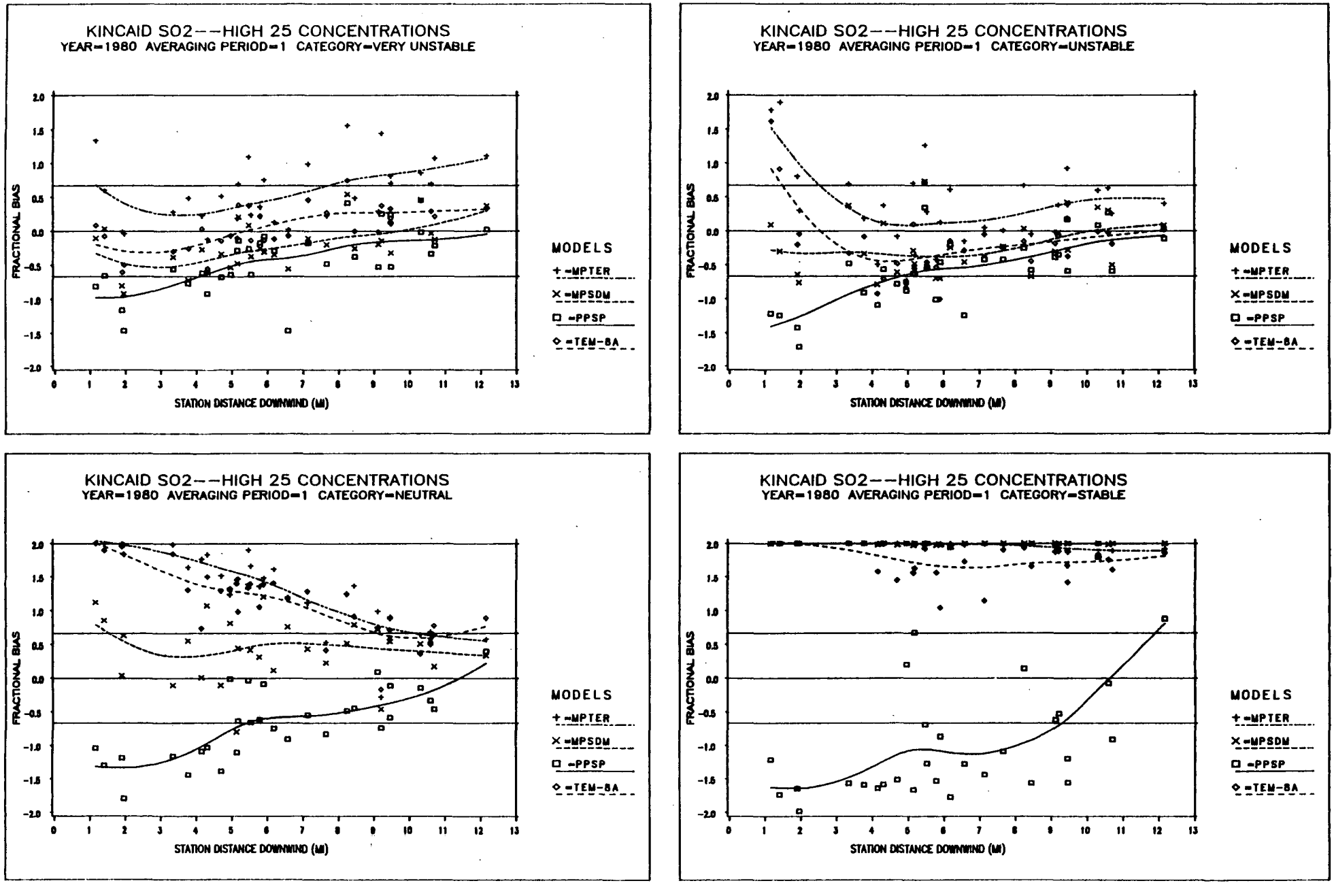
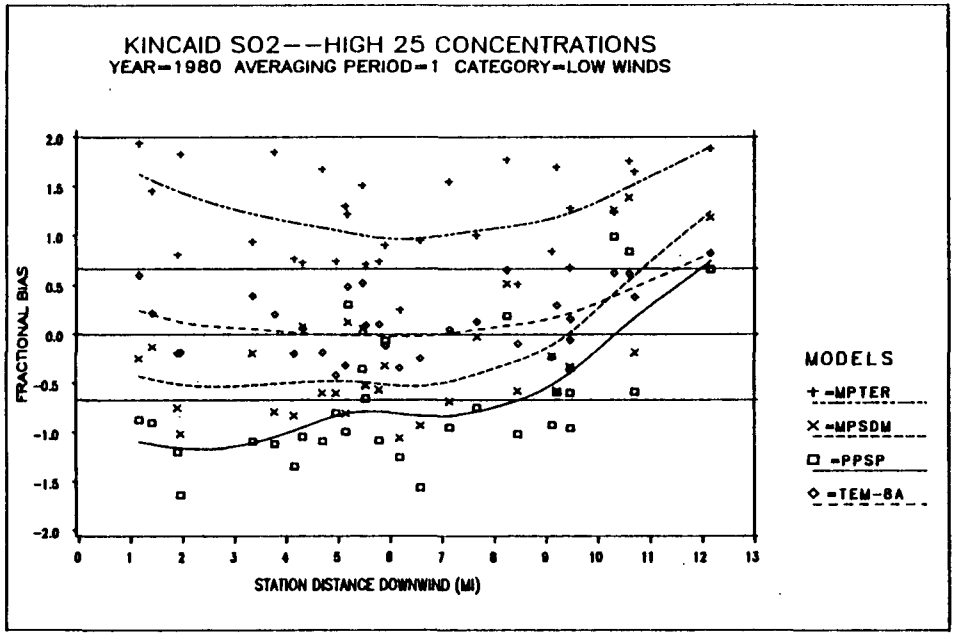
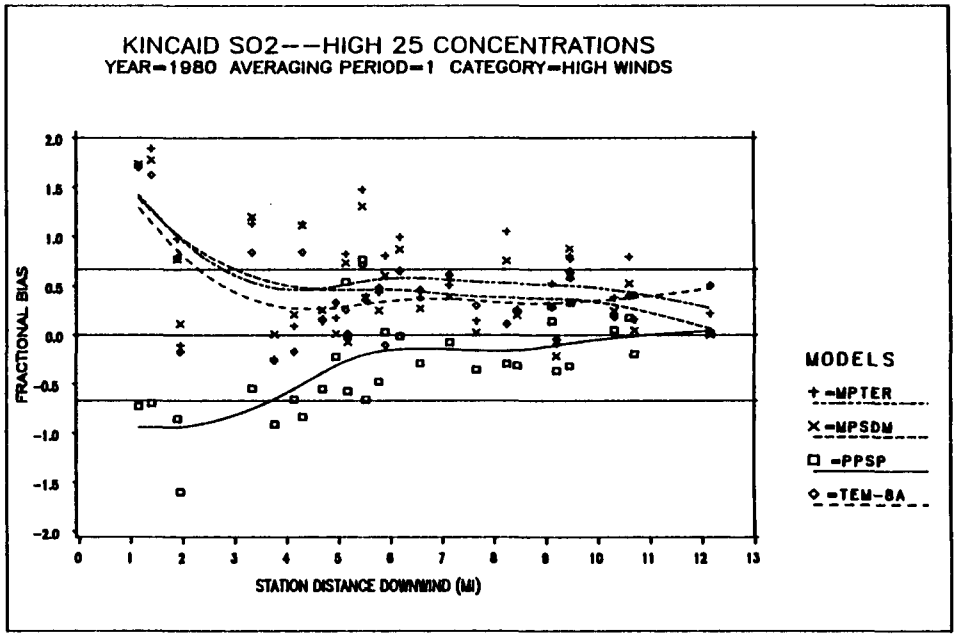


Figure A.12 Fractional Bias Of The Average Vs Station Distance By Stability Class Using High 25 Values





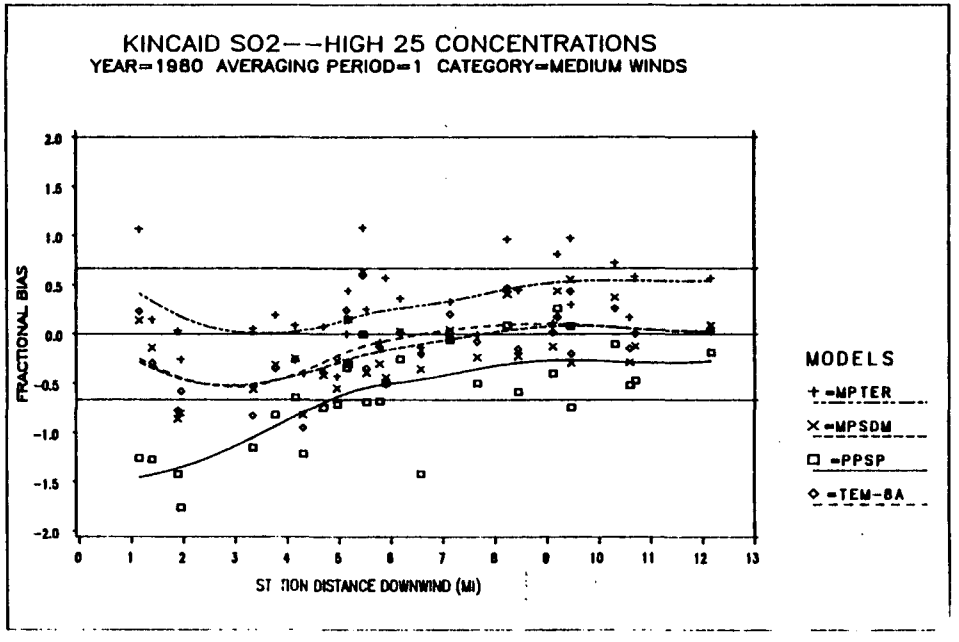


Figure A.13 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using High 25 Values

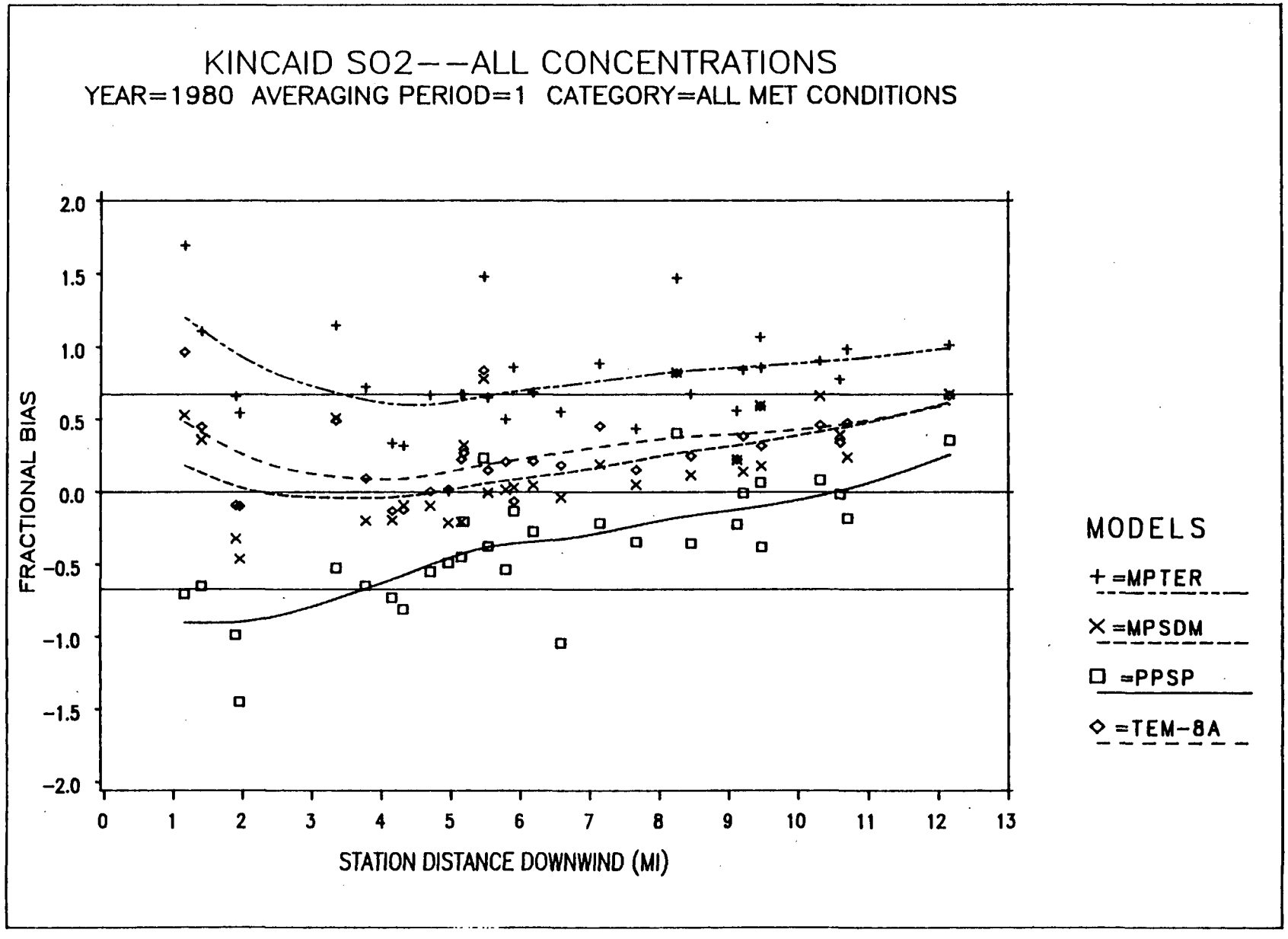


Figure A.14 Fractional Bias Of The Average Vs Station Distance Using All Paired Values

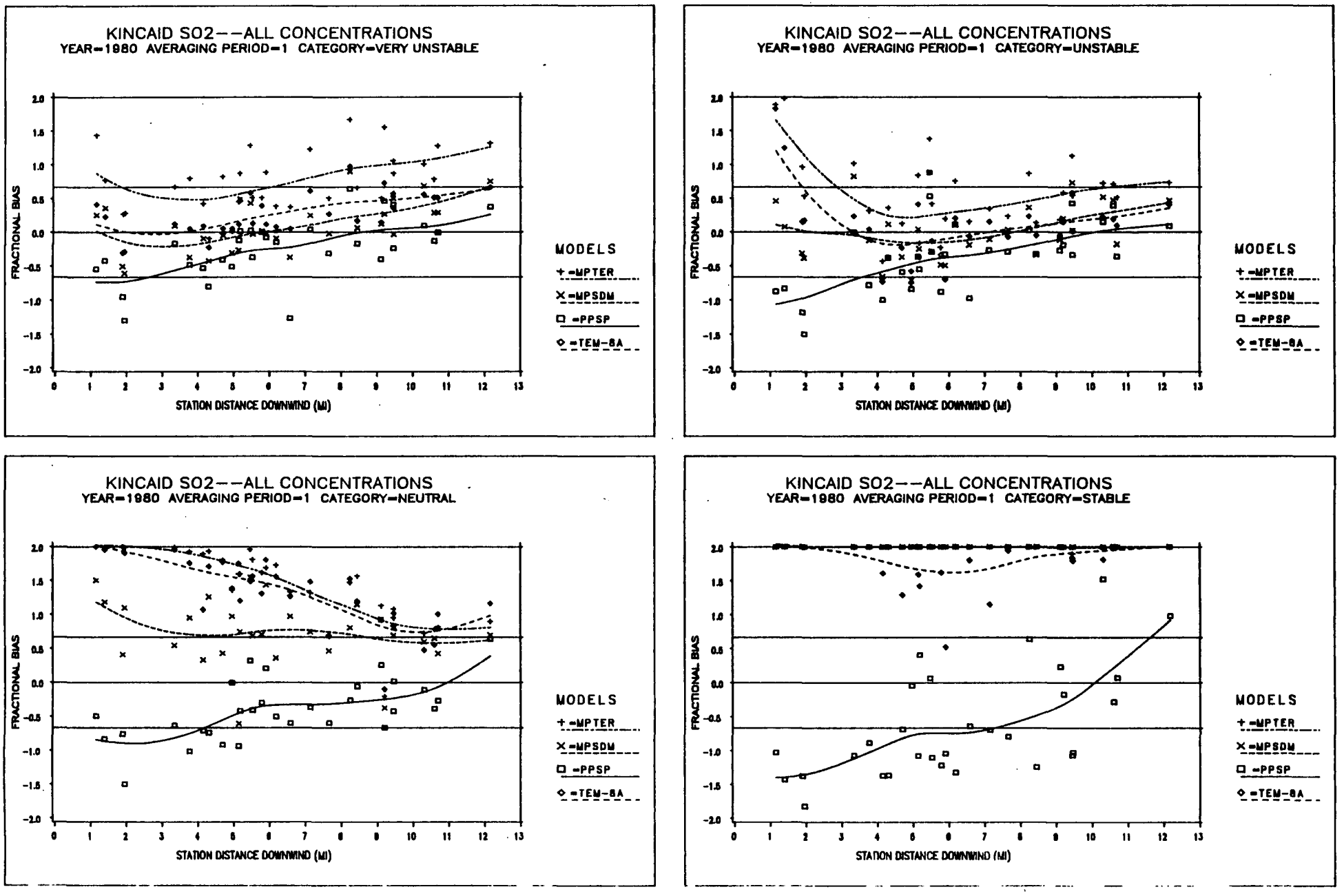


Figure A.15 Fractional Bias Of The Average Vs Station Distance By Stability Class Using All Paired Values

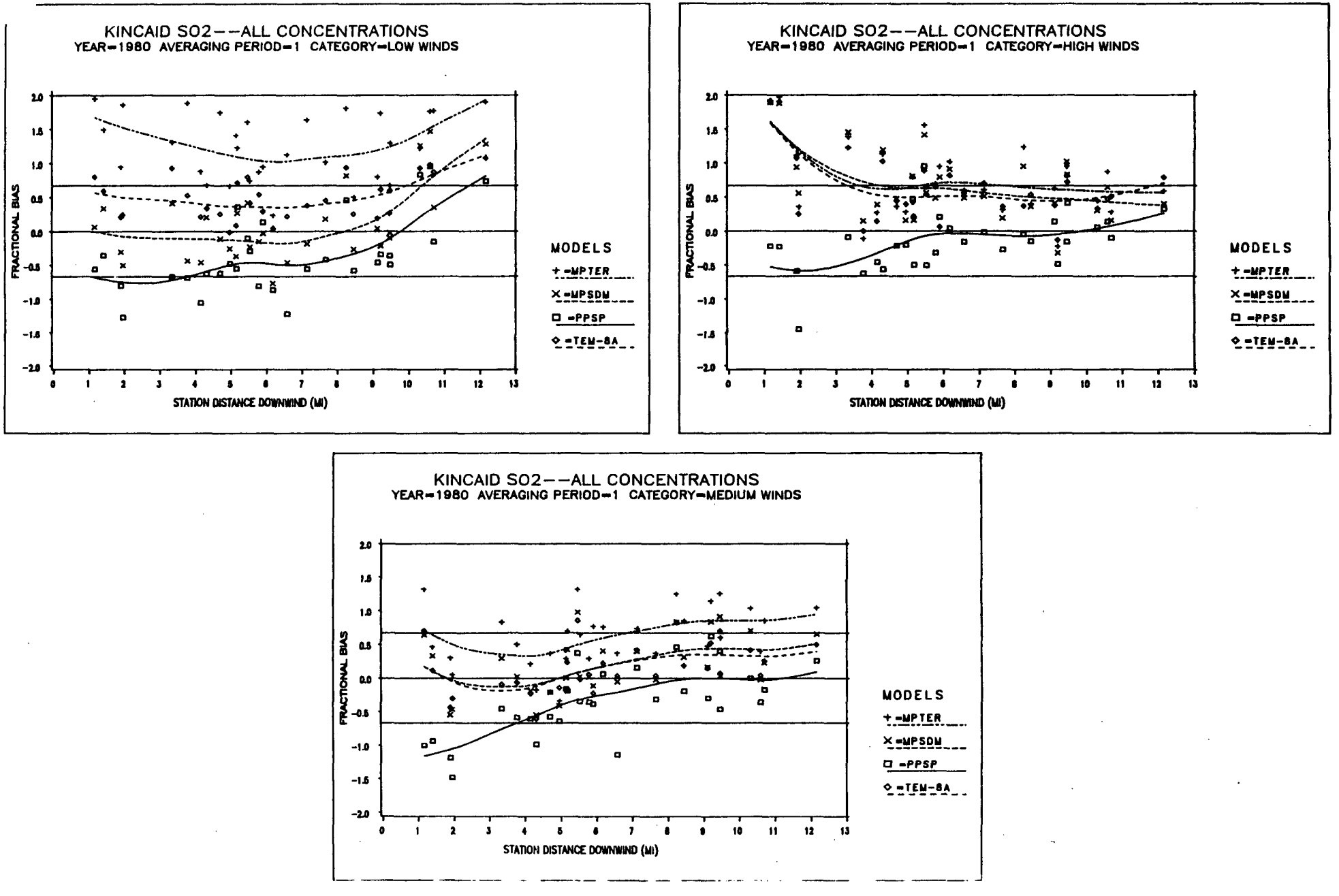


Figure A.16 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using All Paired Values

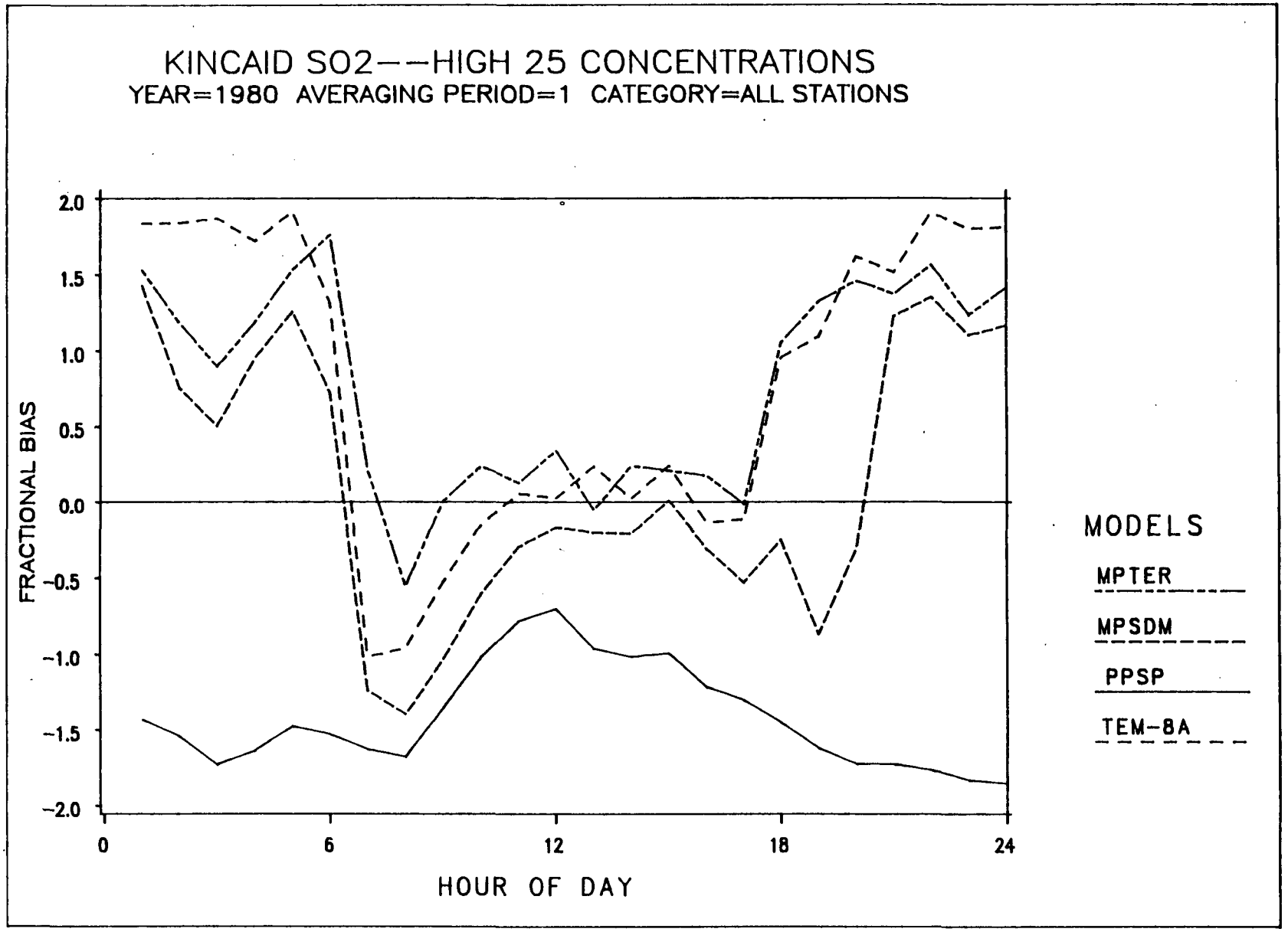


Figure A.17 Fractional Bias Of The Average Vs Hour Of The Day Using High 25 Values

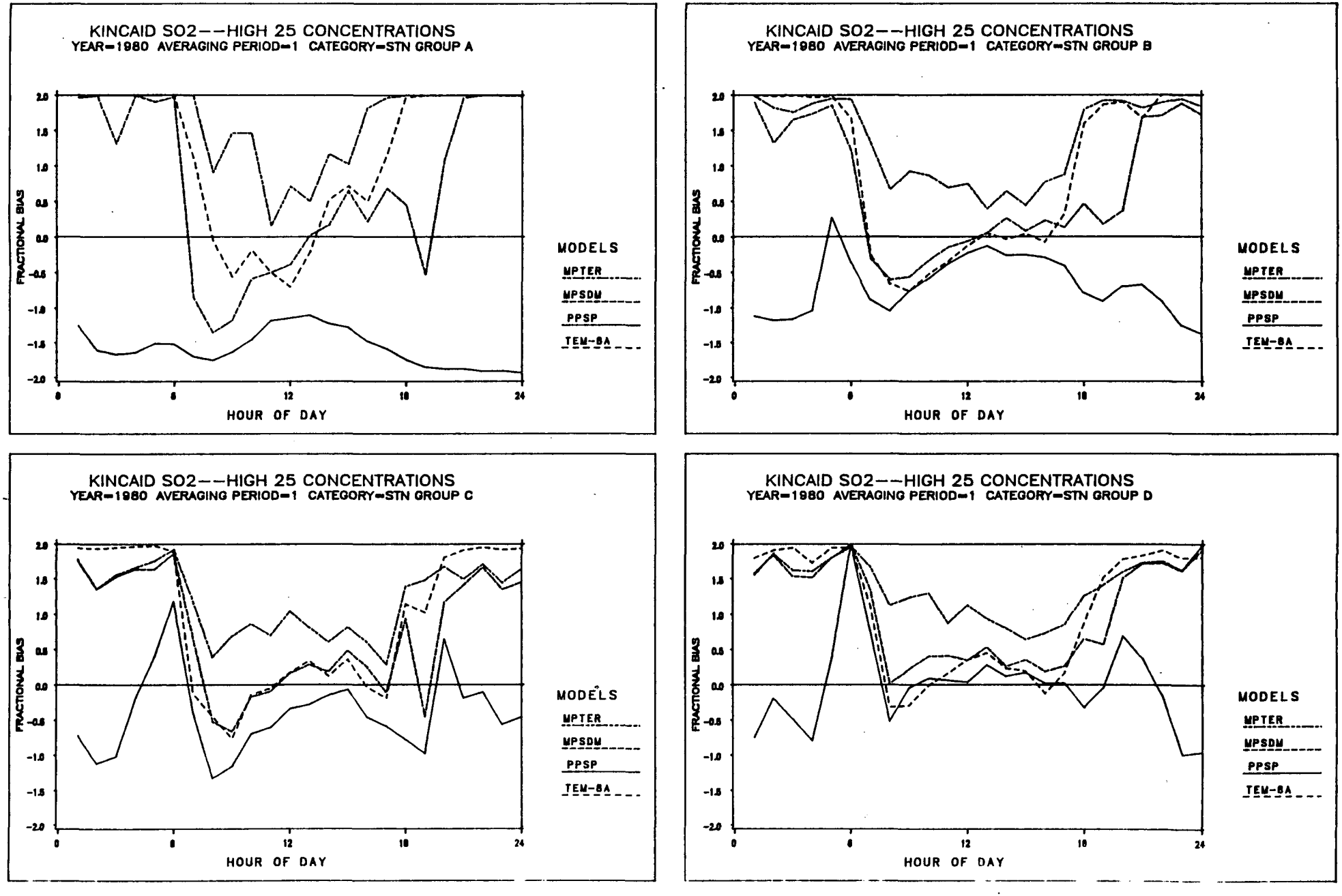


Figure A.18 Fractional Bias Of The Average Vs Hour Of The Day By Station Grouping Using High 25 Values

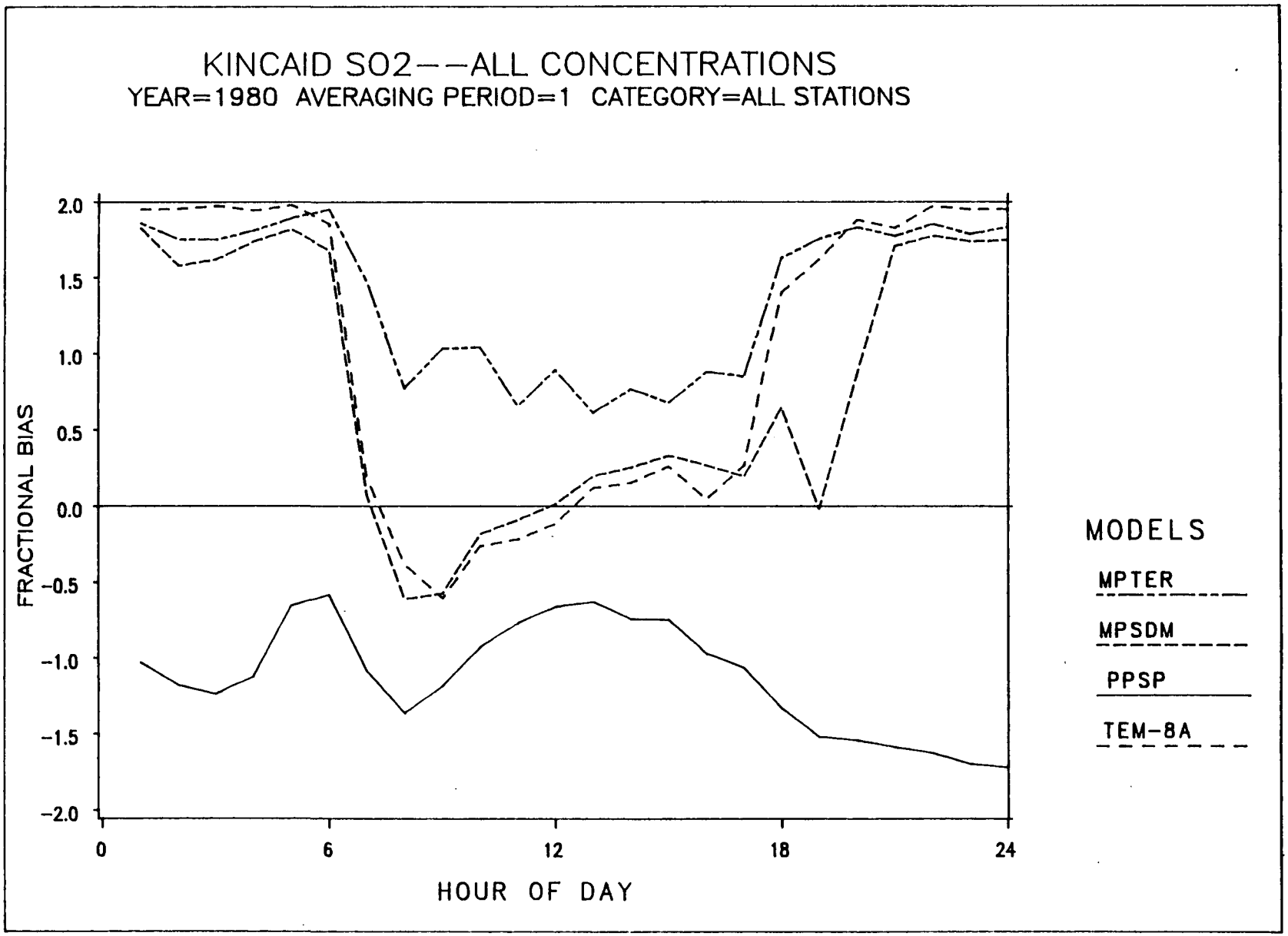


Figure A.19 Fractional Bias Of The Average Vs Hour Of The Day Using All Paired Values

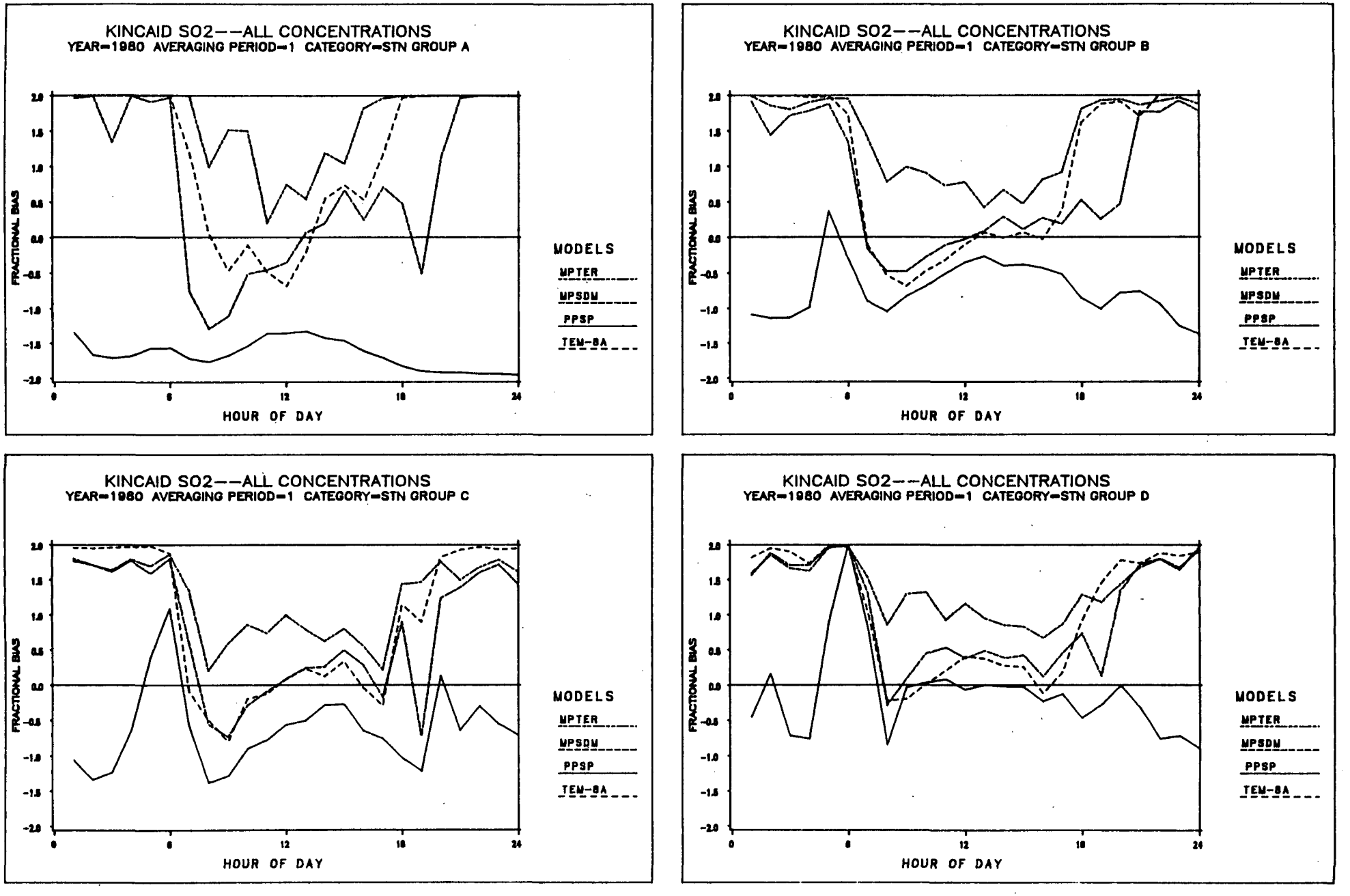


Figure A.20 Fractional Bias Of The Average Vs Hour Of The Day By Station Grouping Using All Paired Values

APPENDIX B--PARADISE POWER PLANT

Figure B.1 shows the location of the 12 monitoring stations surrounding the Paradise Power plant. The distance between the stations and source ranges from approximately 3 km up to 17 km from the plant site. Although some terrain features are shown on the figure, the elevation of the monitors relative to plant grade are unimportant (less than 40 meters). When reviewing the graphics for this plant, the reader should be aware that stations 1, 2, 4 and 5 reported slightly less than 40 percent of the possible hourly SO_2 concentrations. This becomes particularly important for data sets that contain only a few observations per data category, e.g., data subsets at individual stations by meteorological conditions or hour of the day.

Figure B.2 presents the fractional bias of the average vs the fractional bias of the standard deviation for each of the four models. The information shown in the left hand side of this figure and from the companion Q-Q plot (Figure B.3) indicate large overpredictions by PPSP for all three averaging periods which contrast sharply with the relatively low bias for the other three models. Also there is a general tendency for increasing underprediction as averaging period increases. Overpredictions of 1-hour averages by MPSDM tend to vanish for 24-hour averages while a lack of bias for MPTER and TEM-8A for 1-hour averages become underpredictions for 24-hour averages. Similar trends are apparent from examination of the all data categories for fractional bias (right hand side of Figure B.2) and the cumulative frequency plots (Figure B.4).

Diagnostic information in the form of fractional bias plots for stations, stability categories, and wind speed categories are shown in Figures B.5-B.7

for the highest 25 values. The fractional bias of 1-hour averages indicates that MPTER, MPSDM and TEM-8A are relatively unbiased as evidenced by the preponderance of the 12 stations found within the factor of two frame in the center of each panel. Of these three models, MPSDM indicates a tendency for overpredictions at stations 1, 8 and 12. Fractional bias for stability categories (Figure B.6) shows a familiar pattern; stable conditions are associated with large underpredictions while the more unstable conditions are associated with low bias and/or overpredictions. PPSP overpredicts for each of the stability categories including stable conditions. Fractional bias for the wind speed categories (Figure B.7) shows little difference in the bias among the three wind speed categories. MPTER and TEM-A both perform very well as indicated by the close proximity of the three points to the center of the fractional bias panel (0,0). As expected, the all concentration data sets (Figures B.8-B.10) indicate a tendency for greater underprediction. For wind speed categories (Figure B.10), a somewhat greater separation in performance among the wind speed categories is evident compared with that for the 25 highest values (Figure B.7).

Fractional bias vs station distance (Figures B.11-B.16) show patterns that may be affected somewhat by the missing data noted above for stations 1,2,4 and 5. The effect may be most pronounced for station 1 and station 2 since they are the closest and farthest stations, respectively. With this caveat in mind, the patterns of over/under-predictions for the four models appear to be a function of stability (Figure B.12) and wind speed (Figure B.13) categories. For very unstable conditions, the models as a group shift from overprediction toward slight underprediction as station distance increases while the models take on different patterns as the stability increases.

Neutral and stable conditions appear to be associated with decreasing underprediction with distance for each of the models except PPSP. Patterns for wind speed categories are also obvious; large variations in bias near the plant for low winds become somewhat smaller for medium winds.

Finally, the fractional bias for all stations (Figure B.17--high 25 and Figure B.19--all data) and for each of the 12 stations (Figure B.18 and Figure B.20) is presented. The stations were not grouped by station distance from the source to prevent possible masking of the effects of missing data. Obviously the plots for stations having large amounts of missing data (stations 1, 2, 4 and 5) are much "noisier" than for the other stations. This noisiness is probably a consequence of both the smaller quantity of data and possible seasonal effects (monitoring did not begin at these stations until mid-August). For most of the stations, familiar diurnal patterns do appear; underpredictions occur for early morning and late evening hours accompanied by relatively low biases near mid-day.

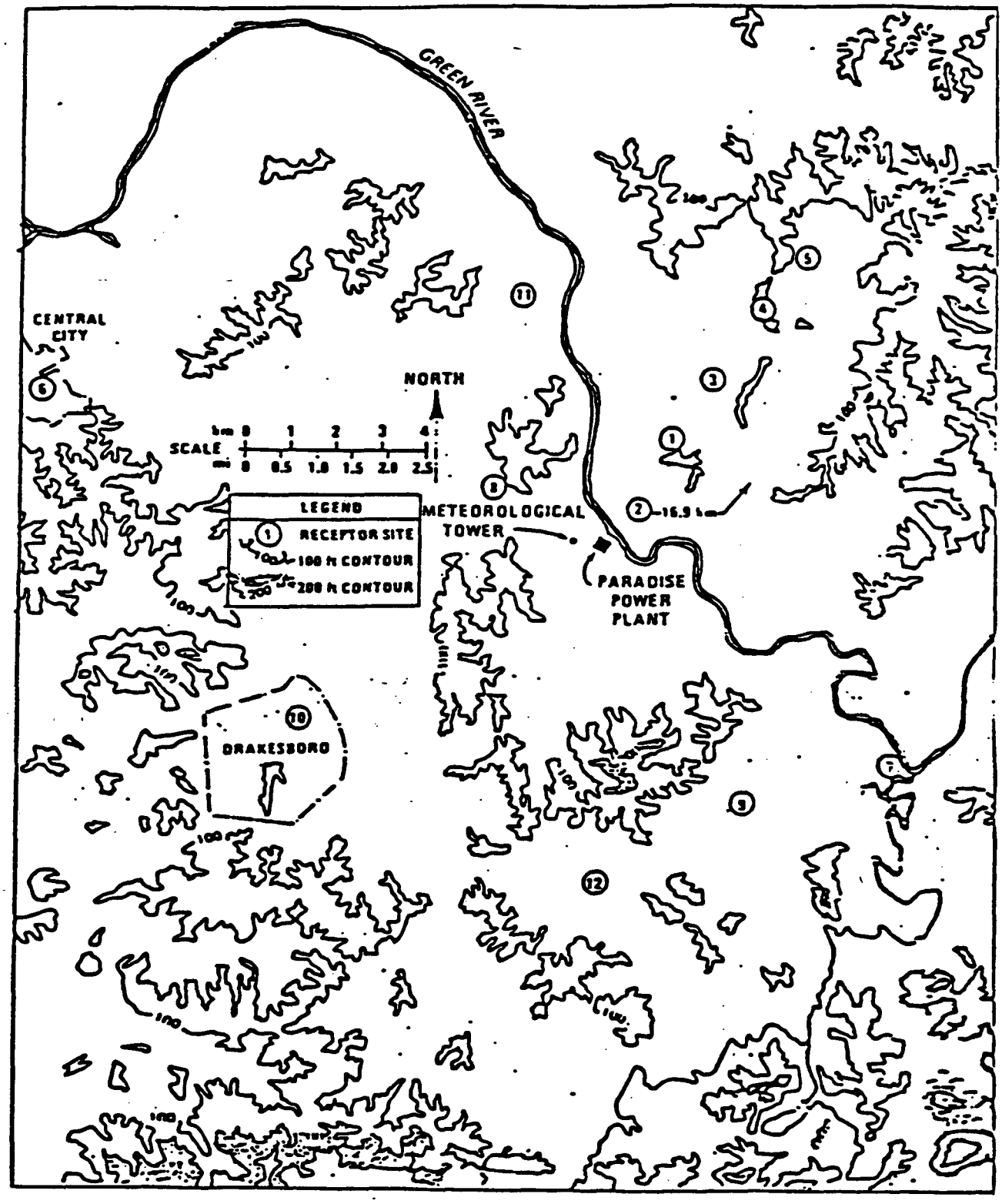


Figure B.1 Field Monitoring Network Near The Paradise Power Plant

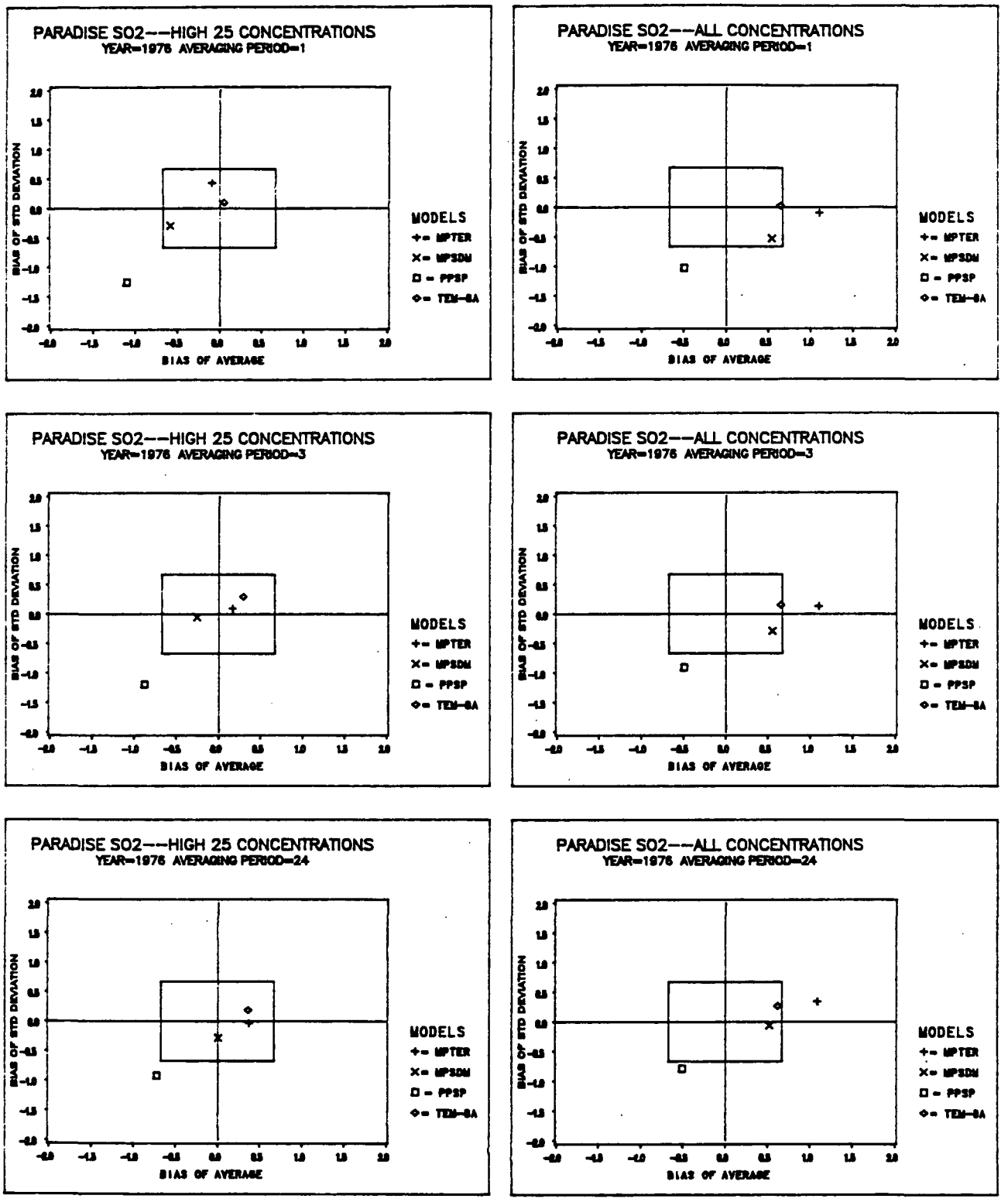


Figure B.2 Fractional Bias By Averaging Period Using High 25 Values and All Paired Values

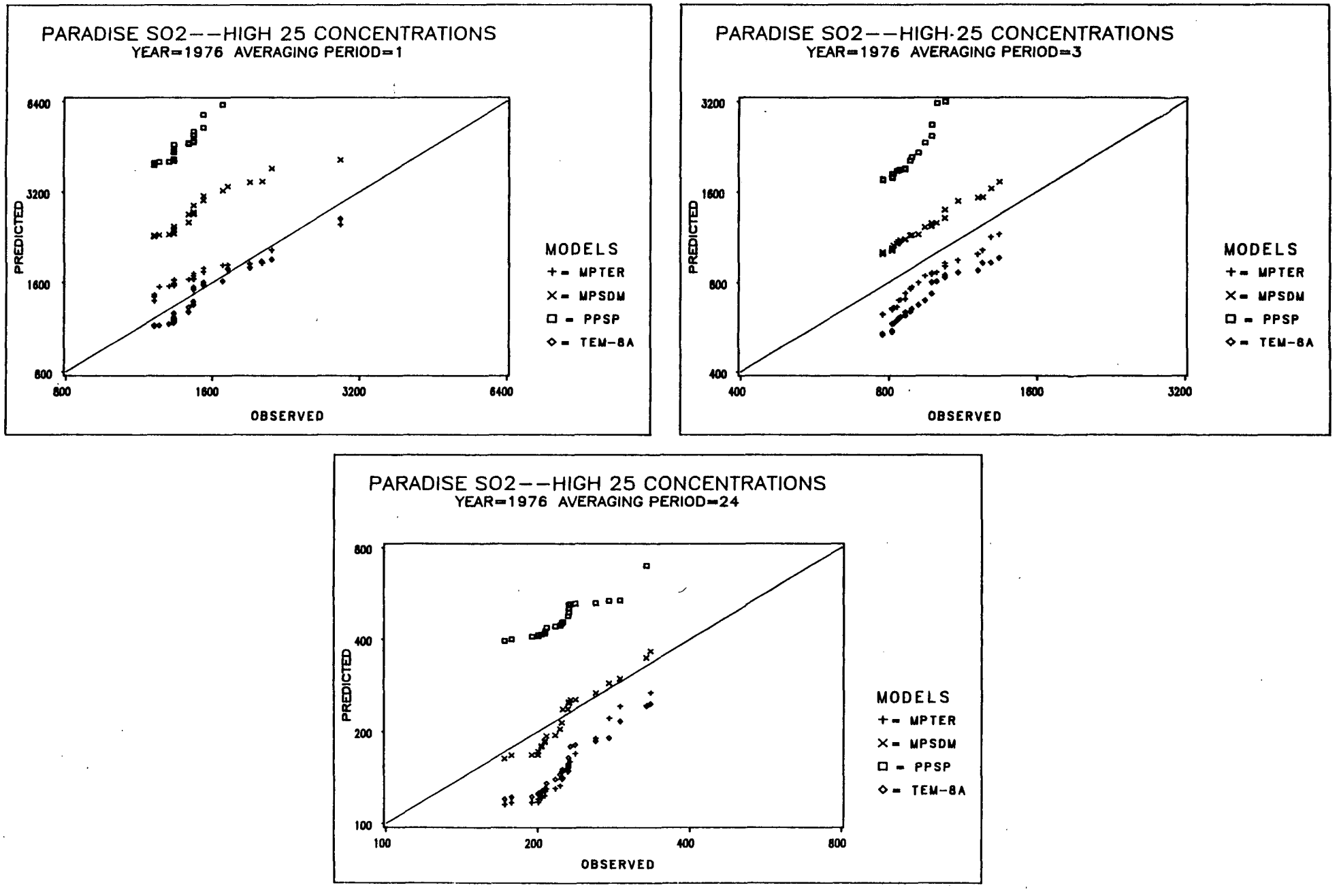
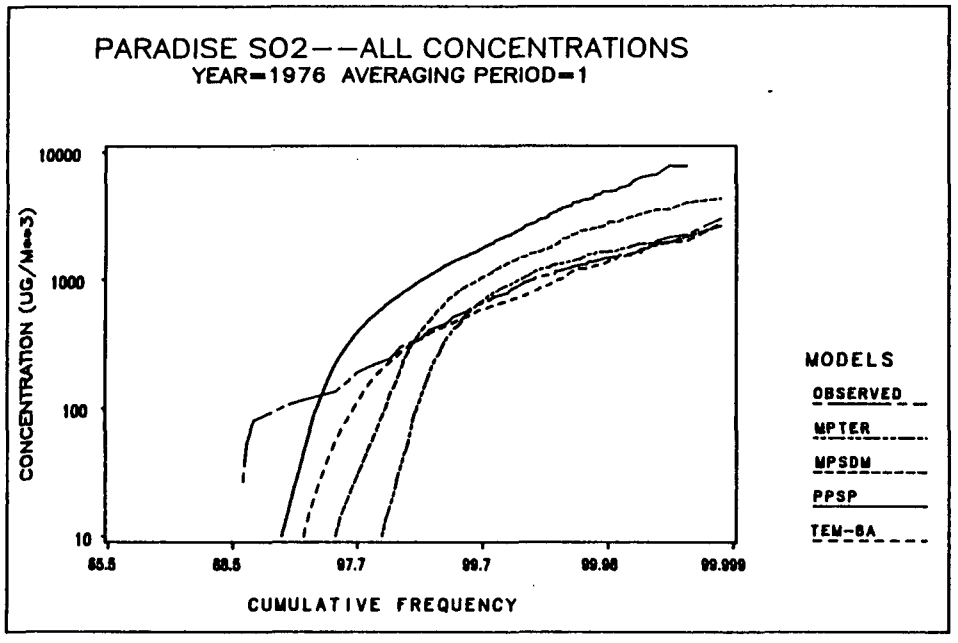
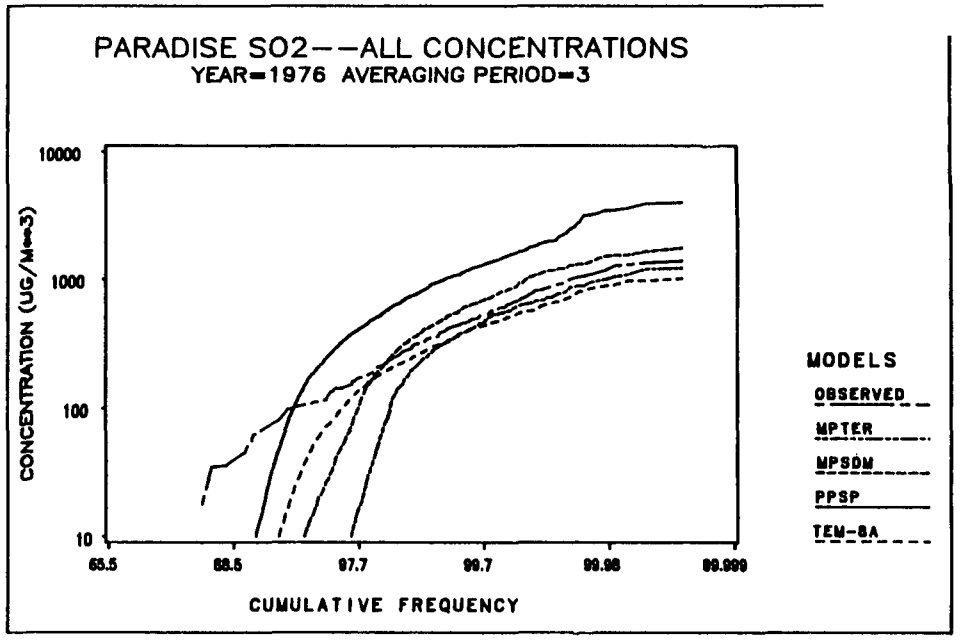


Figure B.3 Predicted Quantiles Vs Observed Quantiles By Averaging Period Using High 25 Values





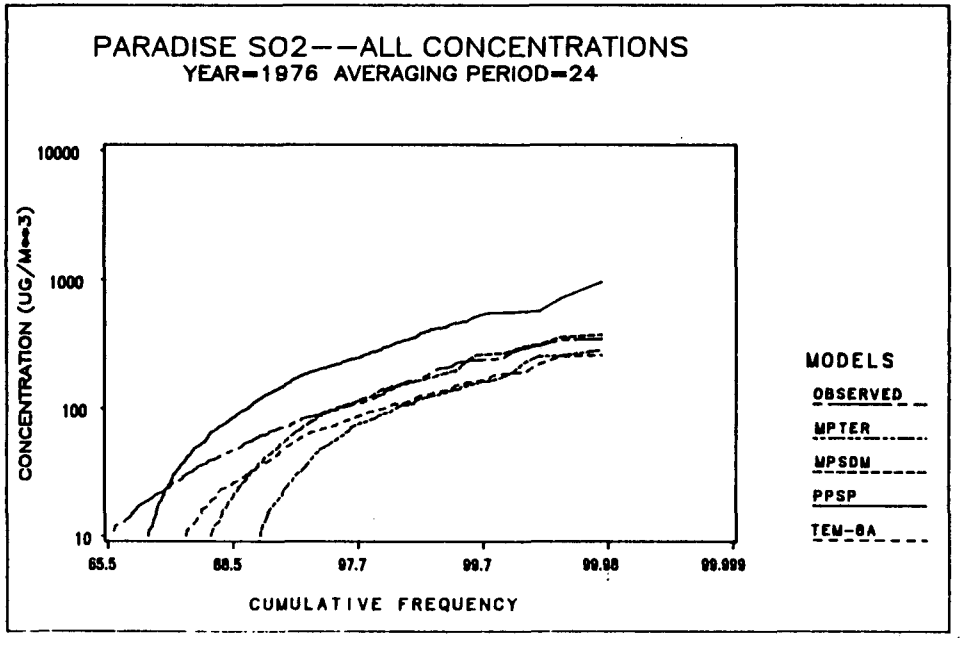


Figure B.4 Cumulative Frequency Distributions By Averaging Period Using All Paired Values

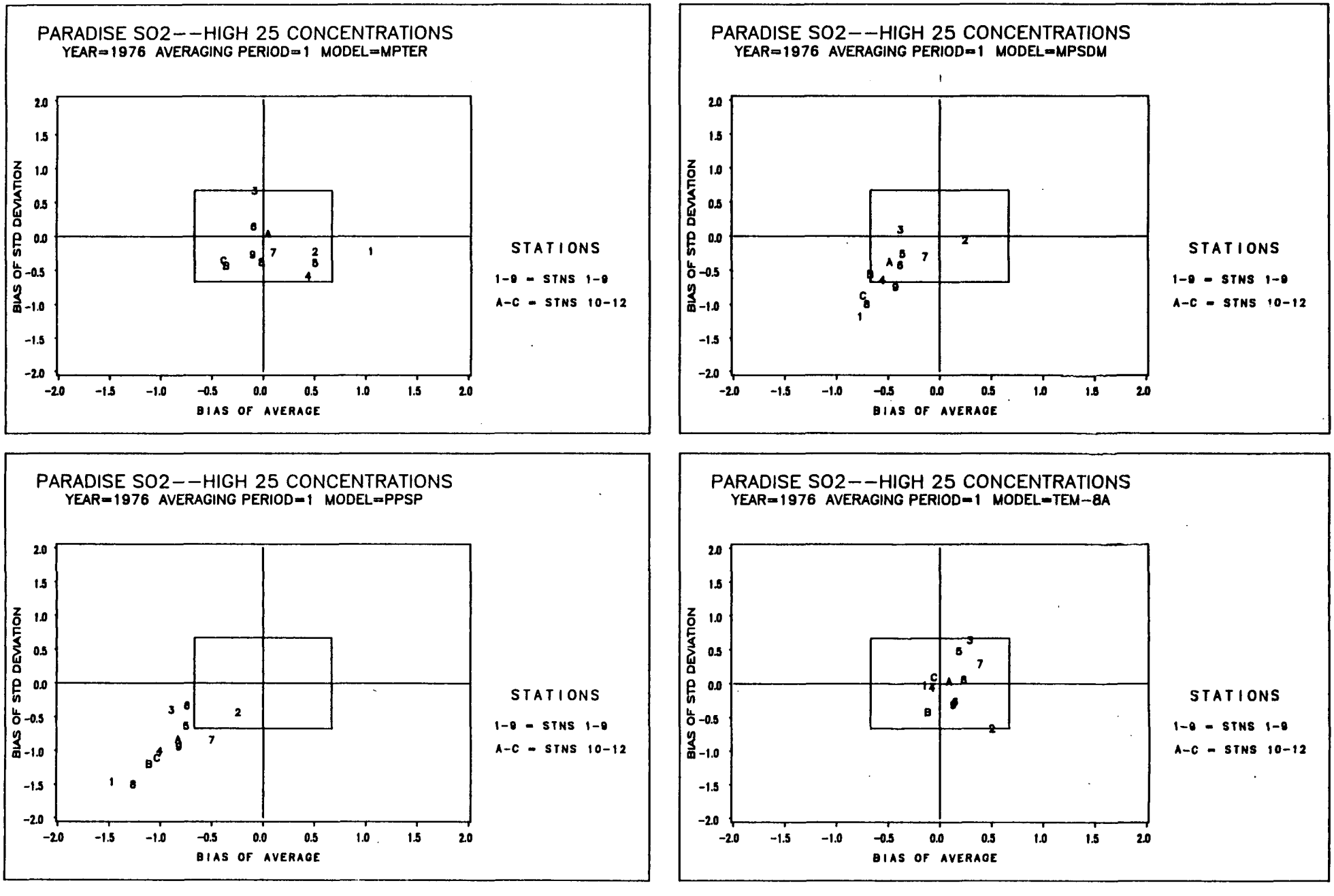


Figure B.5 Fractional Bias By Model Using 25 High Values For Each Station

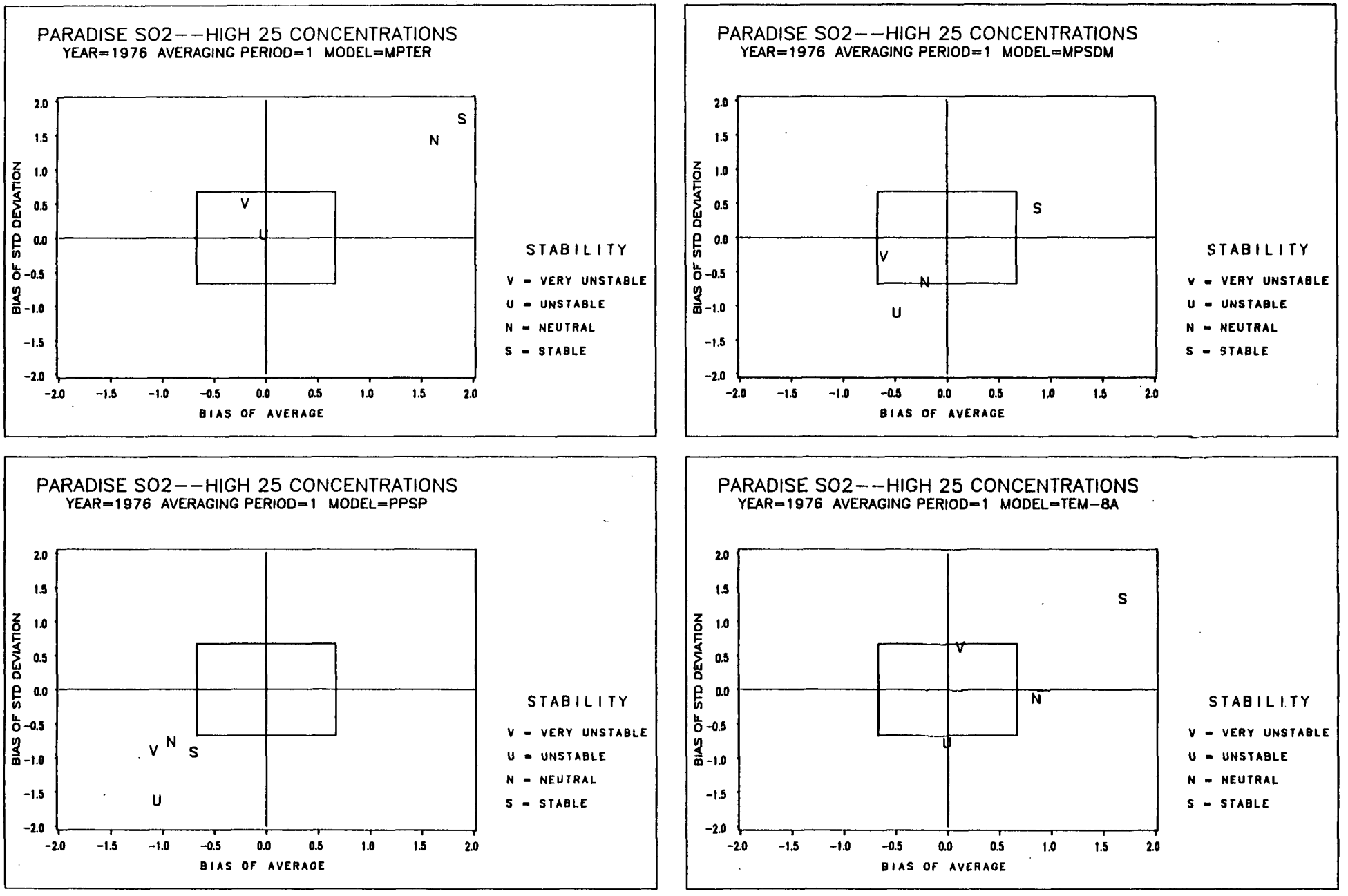


Figure B.6 Fractional Bias By Model Using 25 High Values For Each Stability Class

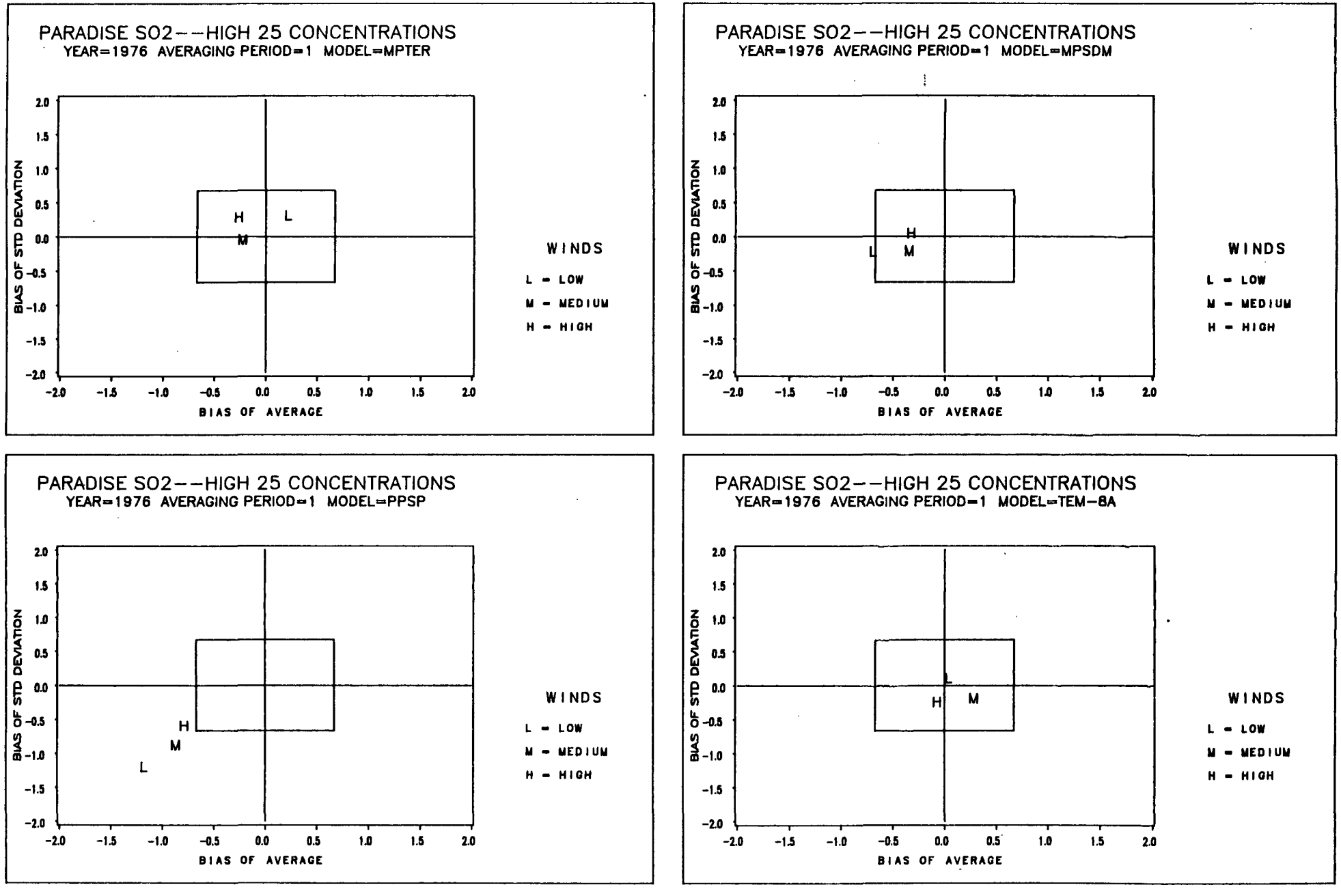


Figure B.7 Fractional Bias By Model Using 25 High Values For Each Wind Speed Class

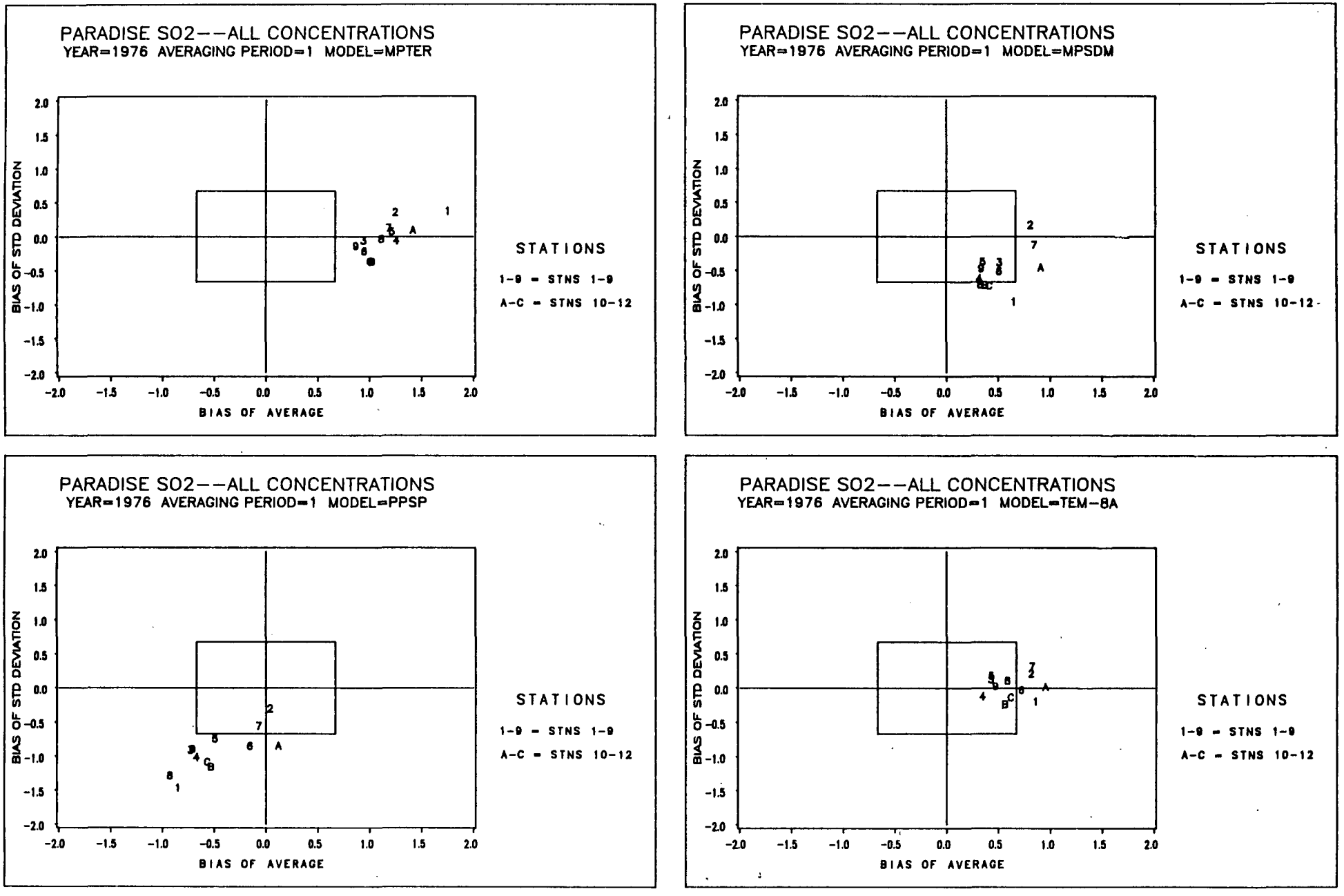


Figure B.8 Fractional Bias By Model Using All Paired Values For Each Station

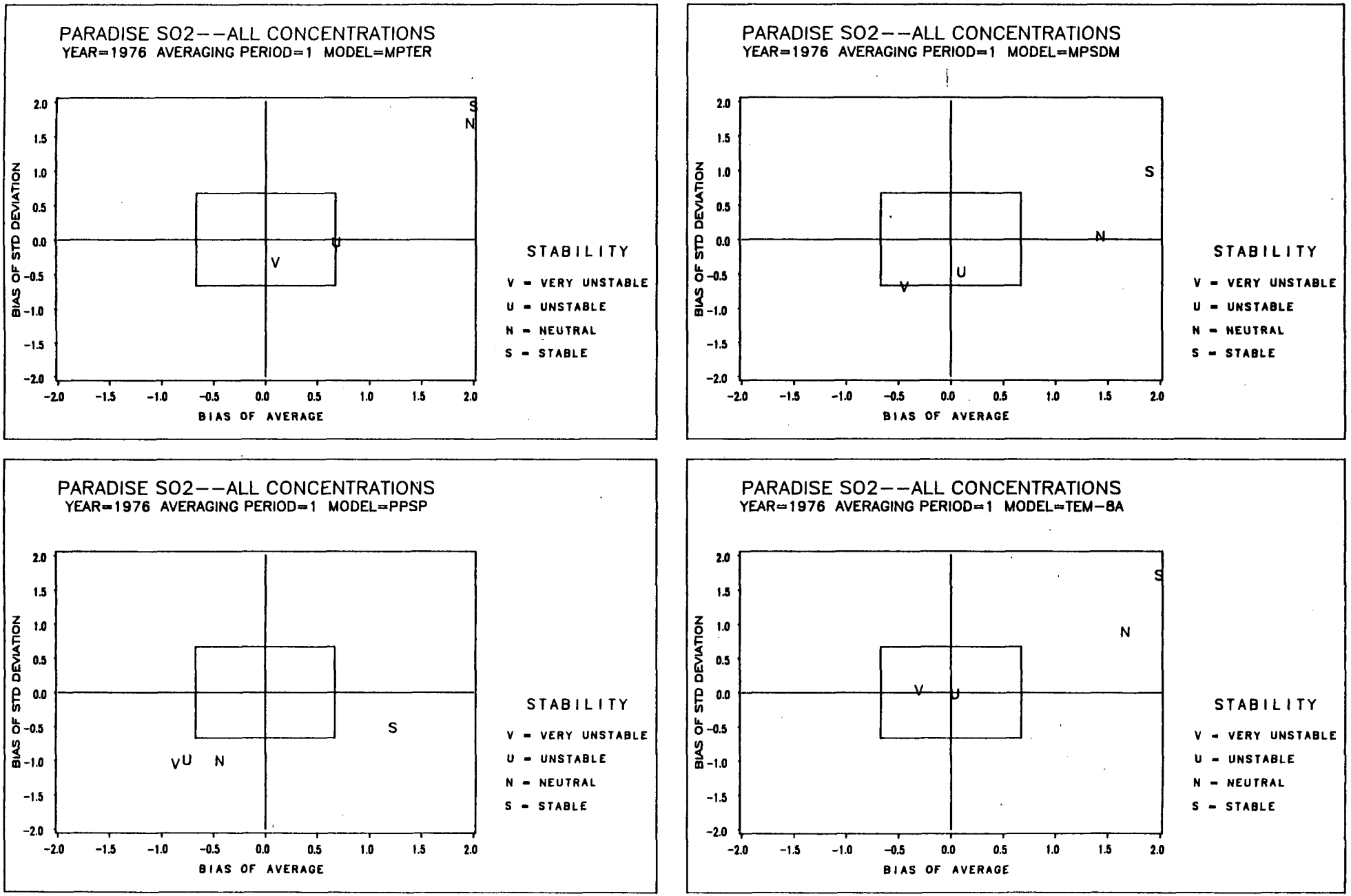


Figure B.9 Fractional Bias By Model Using All Paired Values For Each Stability Class

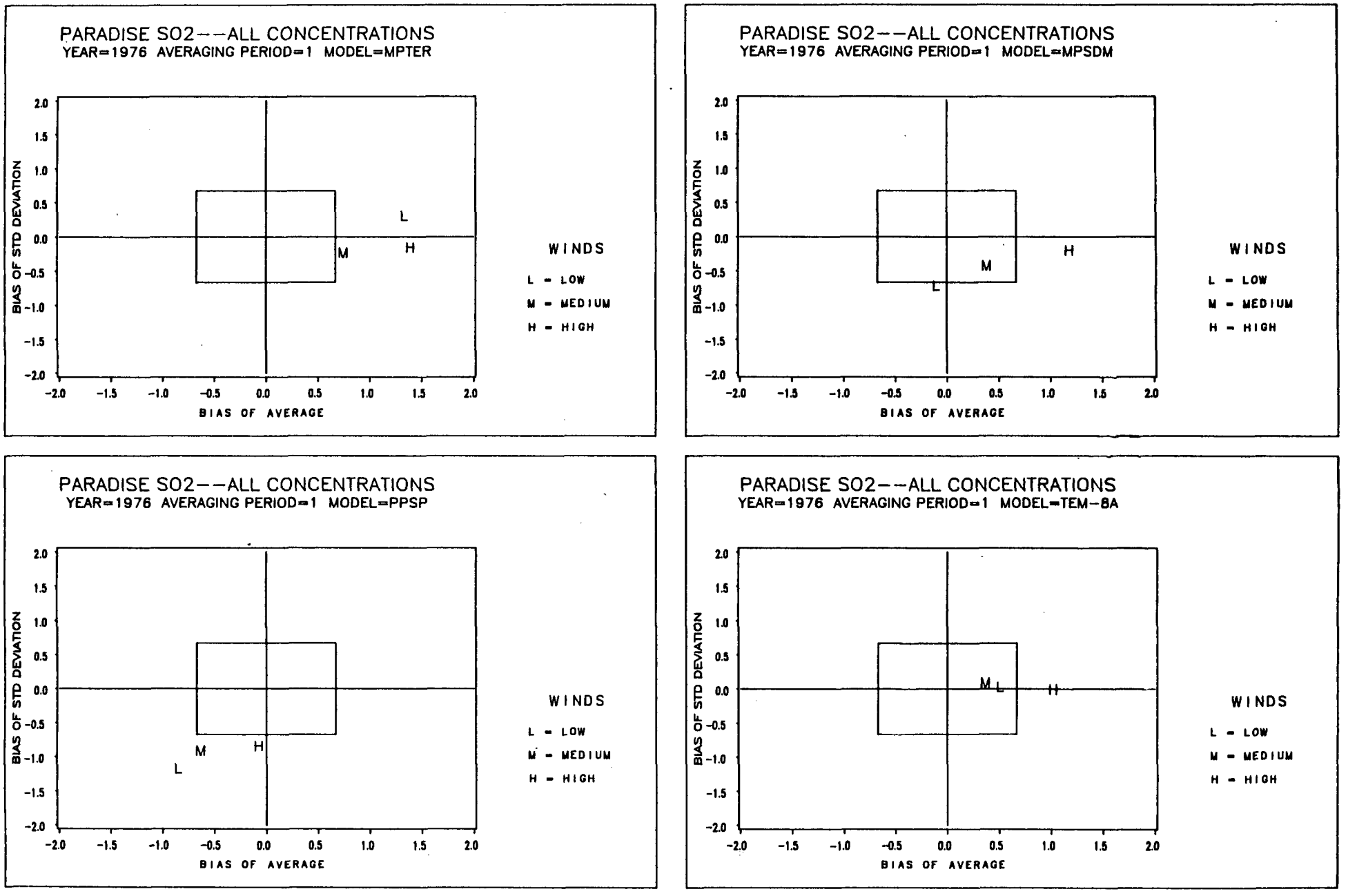


Figure B.10 Fractional Bias By Model Using All Paired Values For Each Wind Speed Class

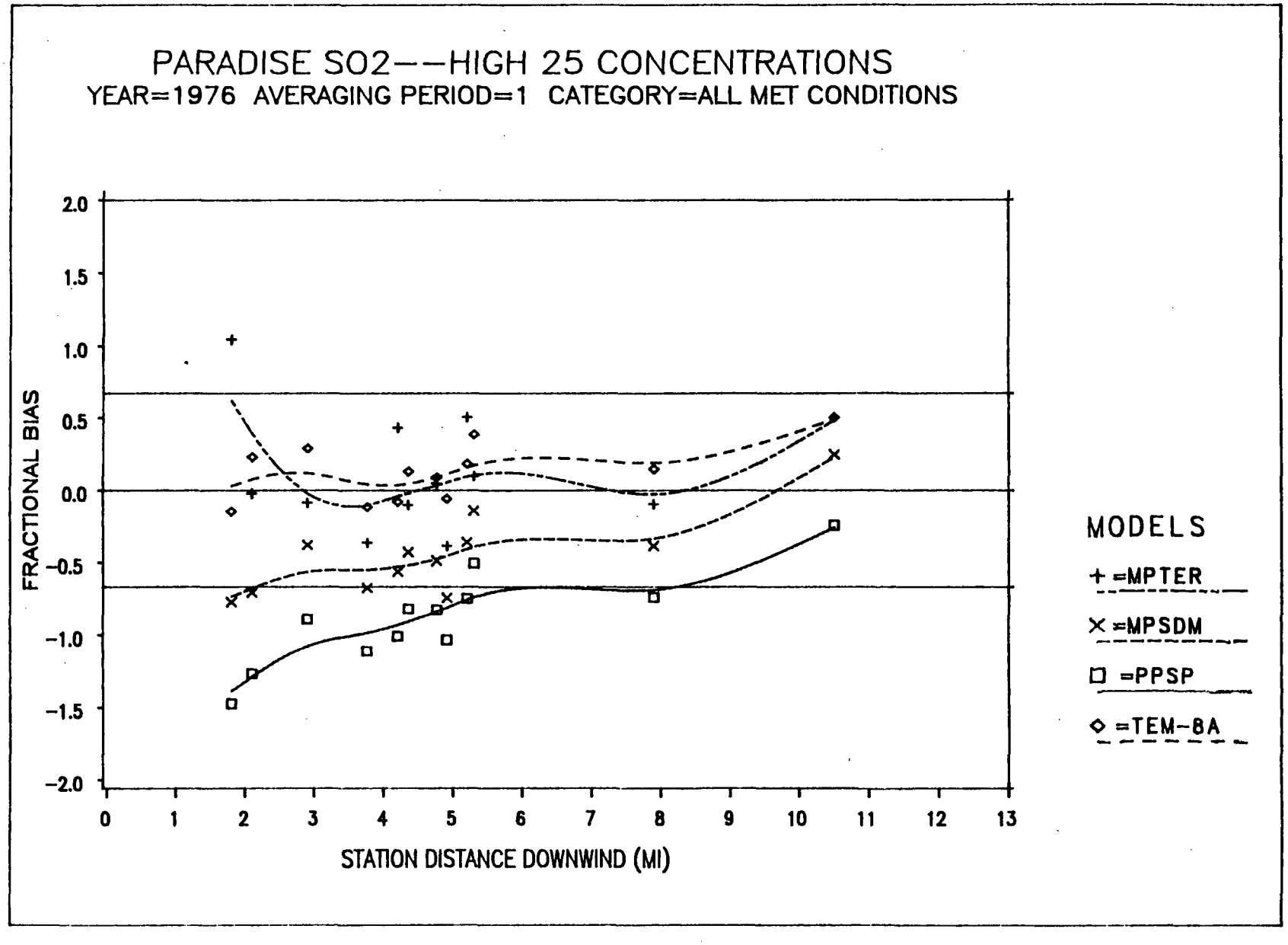


Figure B.11 Fractional Bias Of The Average Vs Station Distance Using High 25 Values

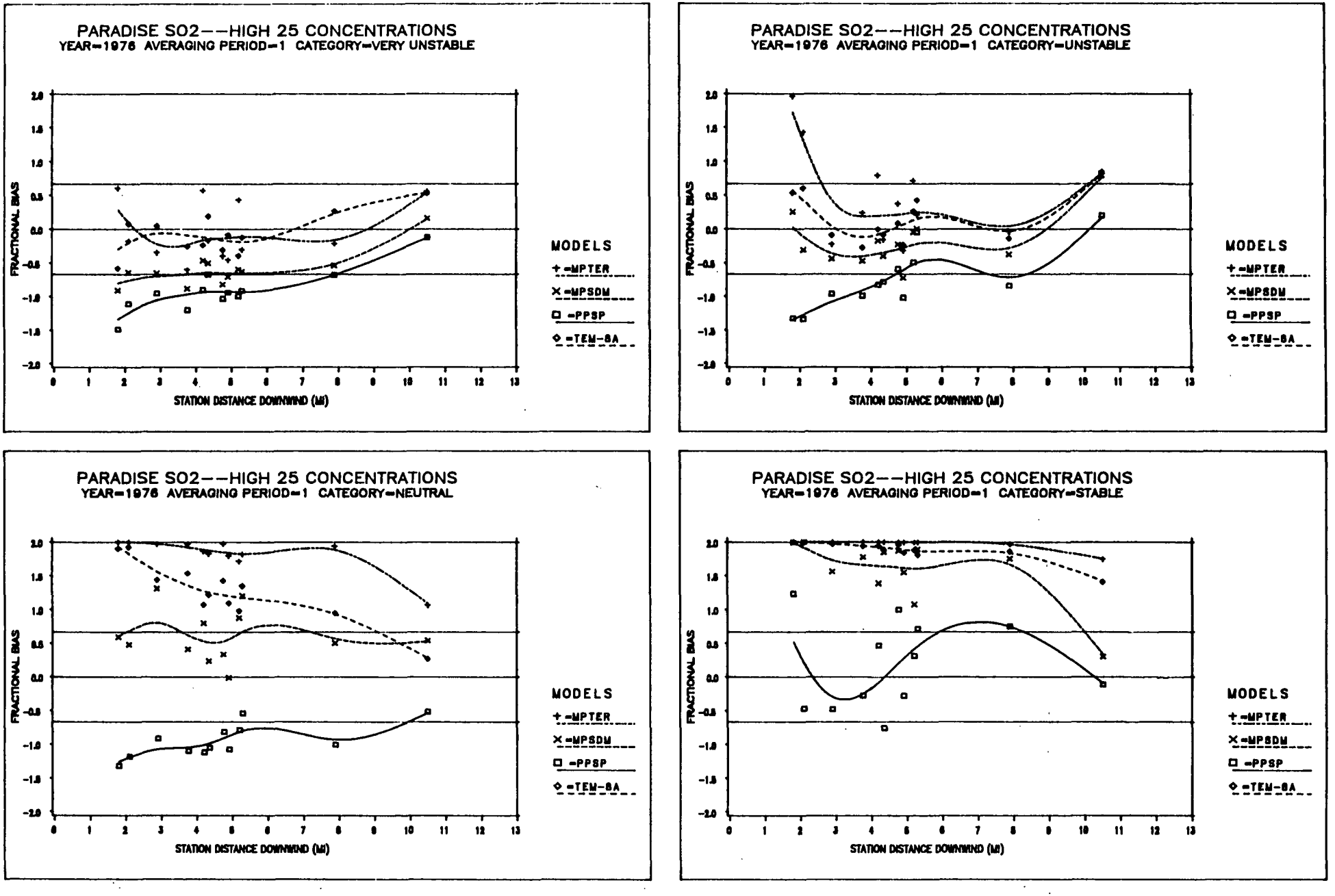
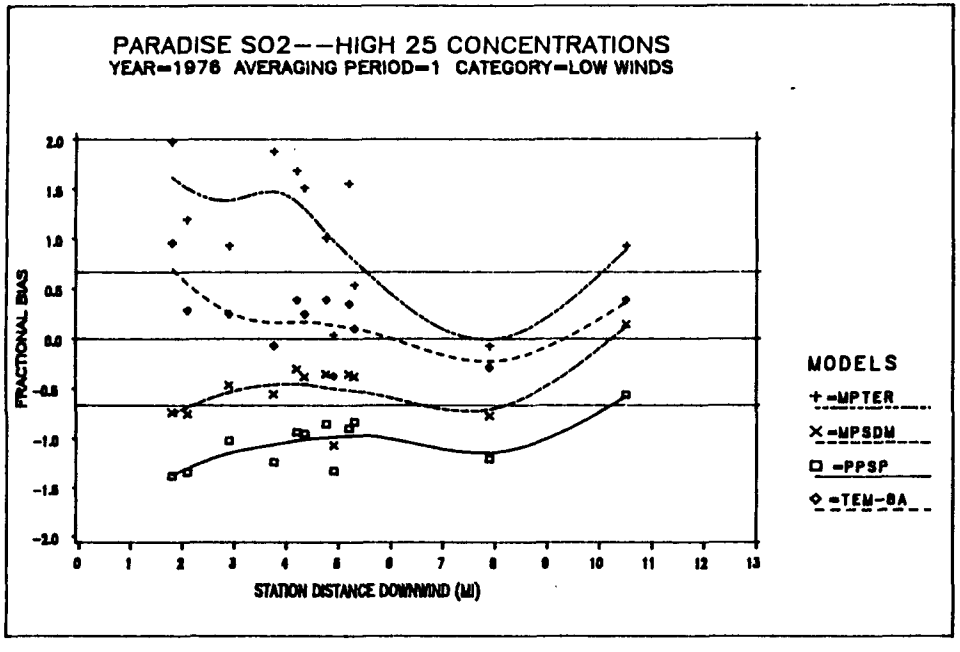
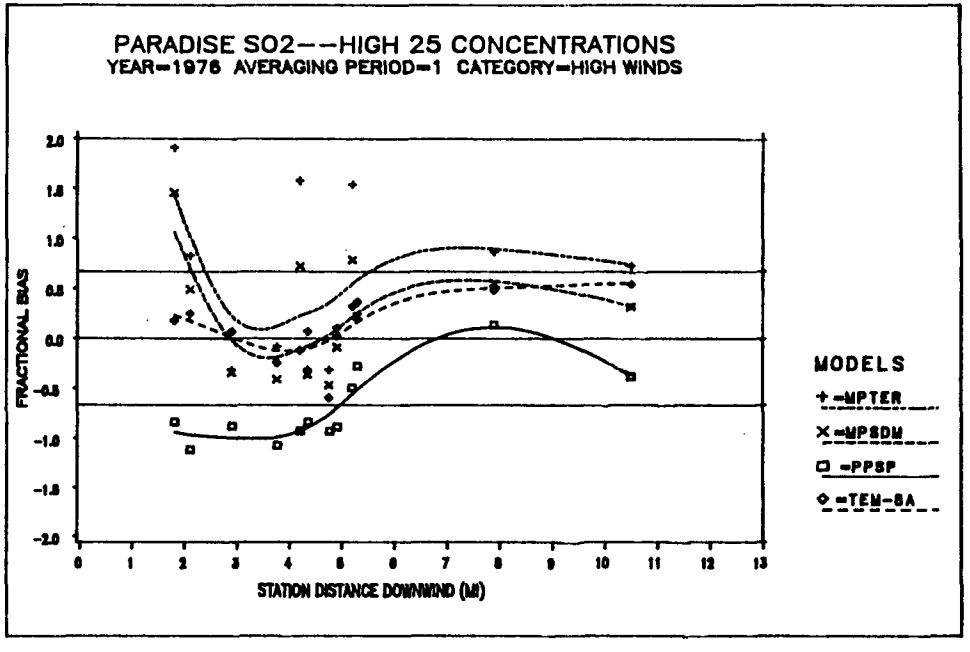


Figure B.12 Fractional Bias Of The Average Vs Station Distance By Stability Class Using High 25 Values





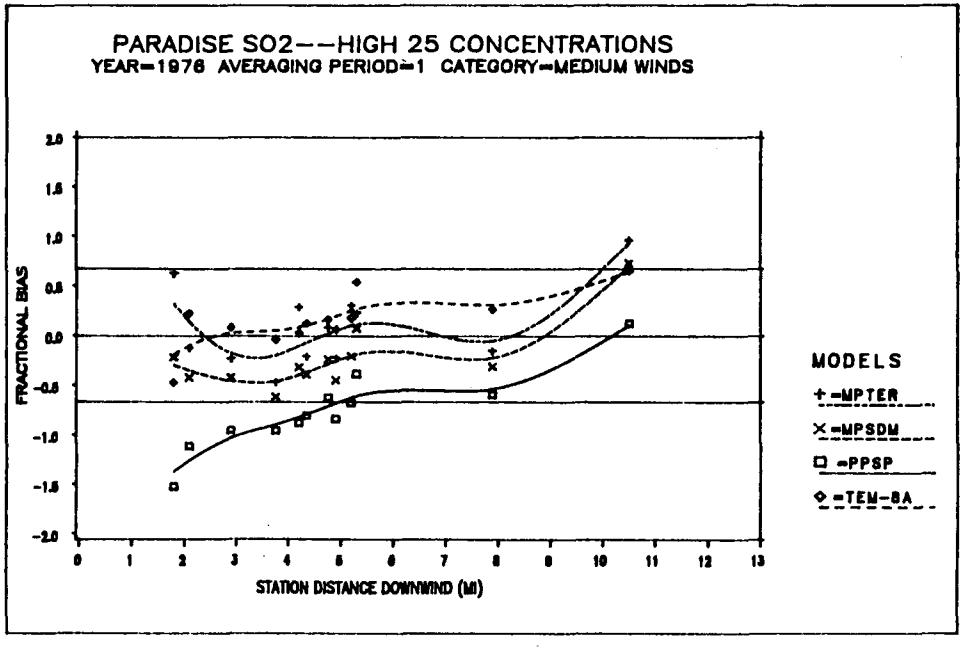


Figure B.13 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using High 25 Values

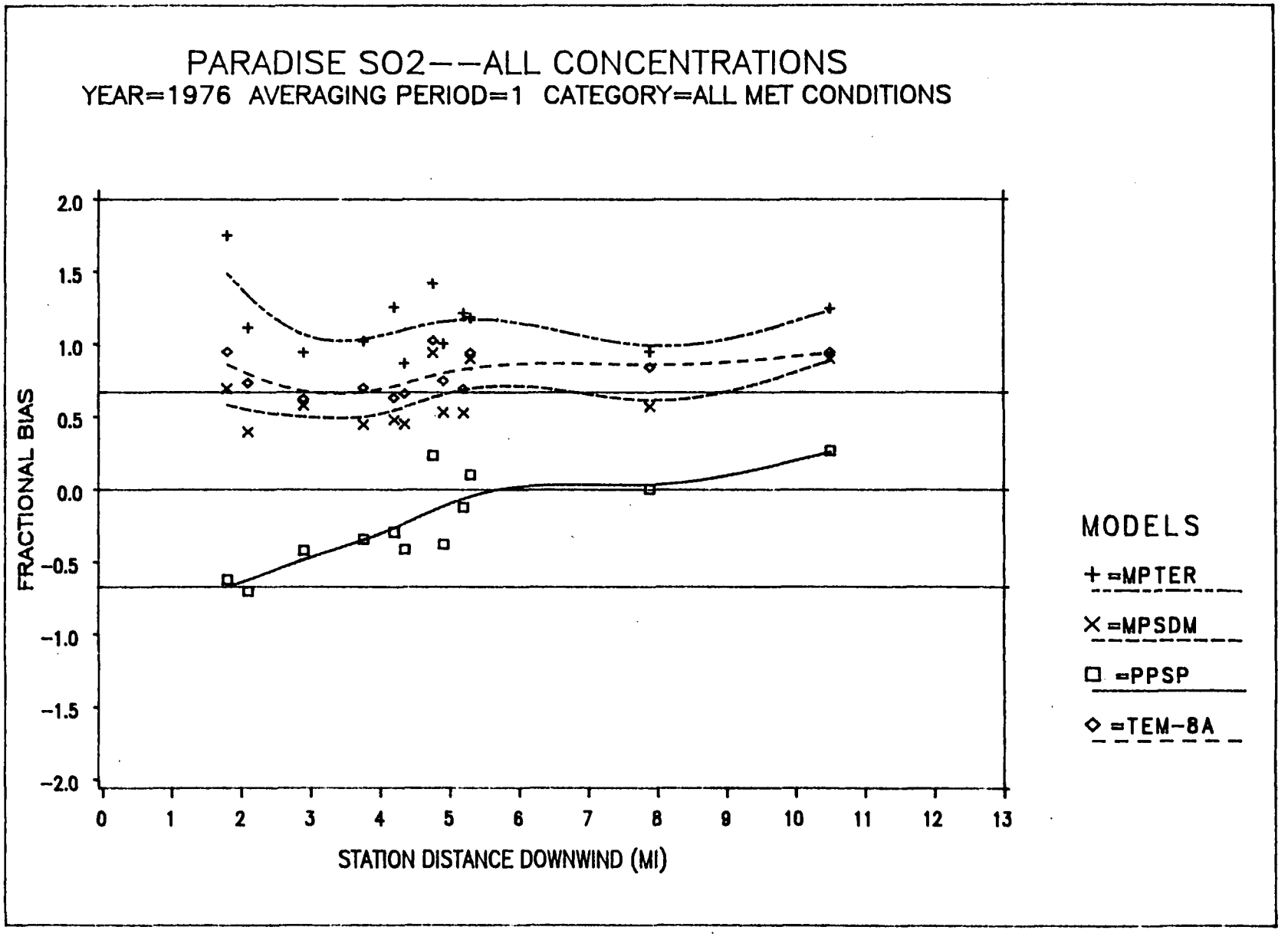


Figure B.14 Fractional Bias Of The Average Vs Station Distance Using All Paired Values

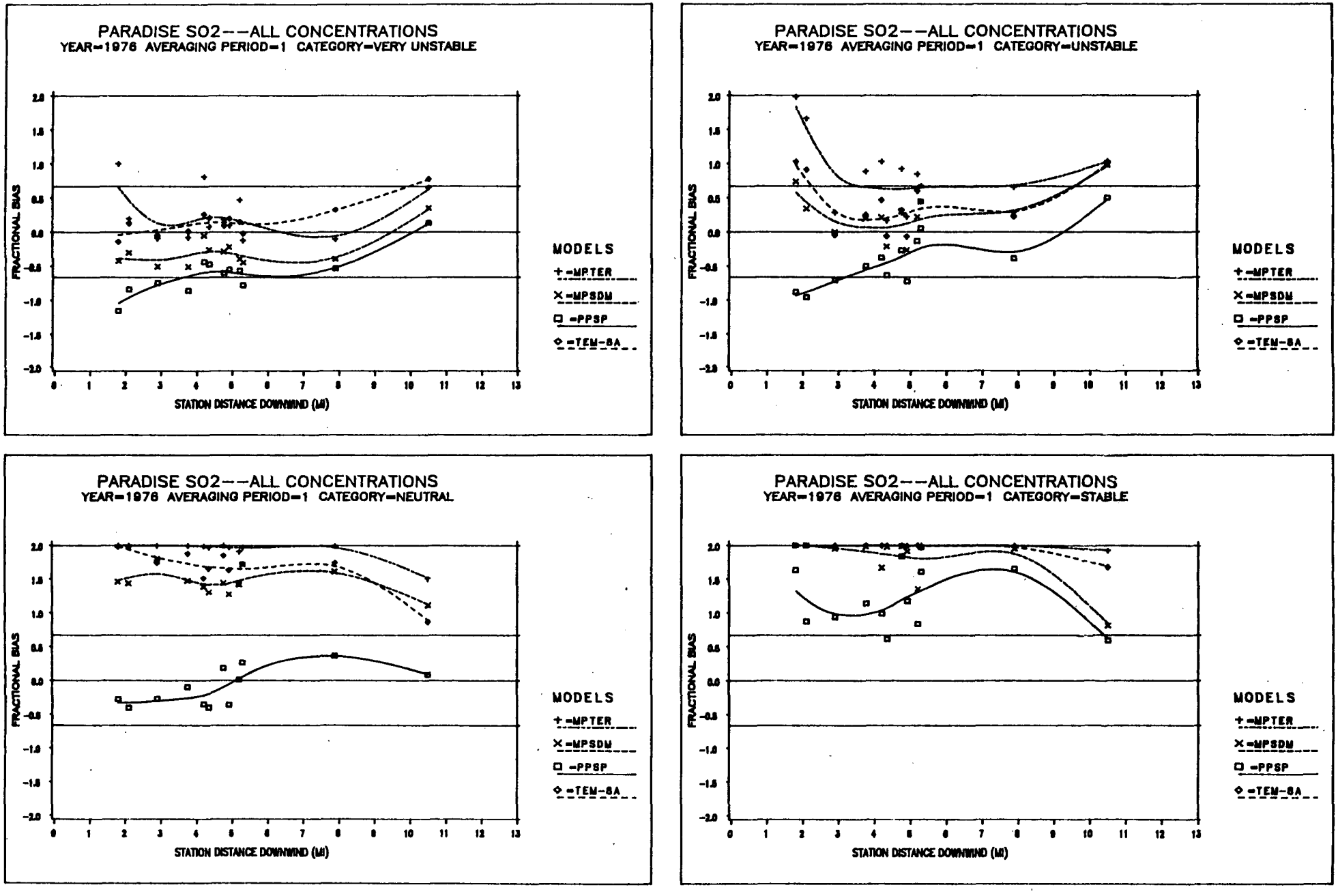
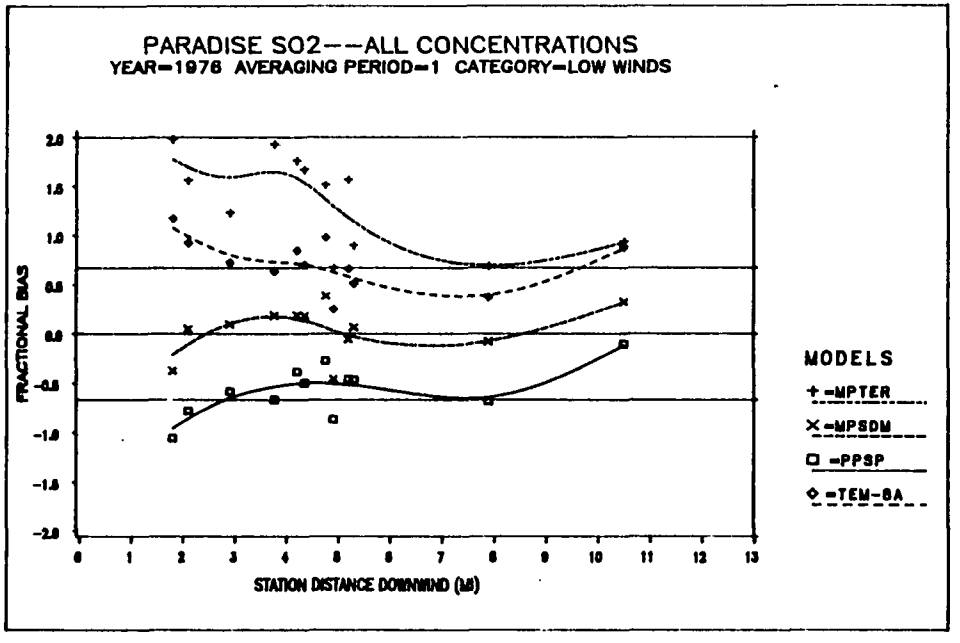
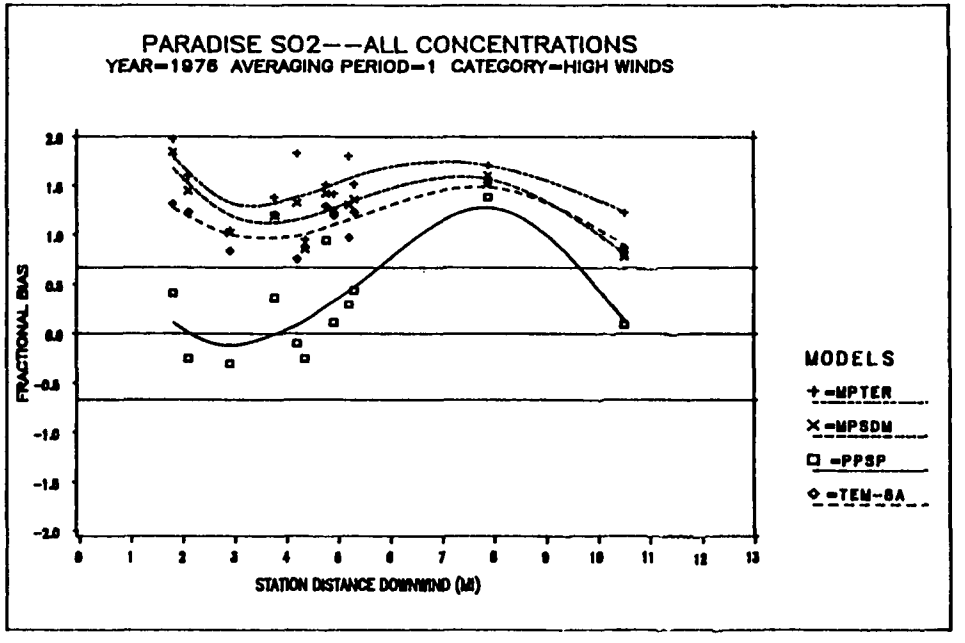


Figure B.15 Fractional Bias Of The Average Vs Station Distance By Stability Class Using All Paired Values





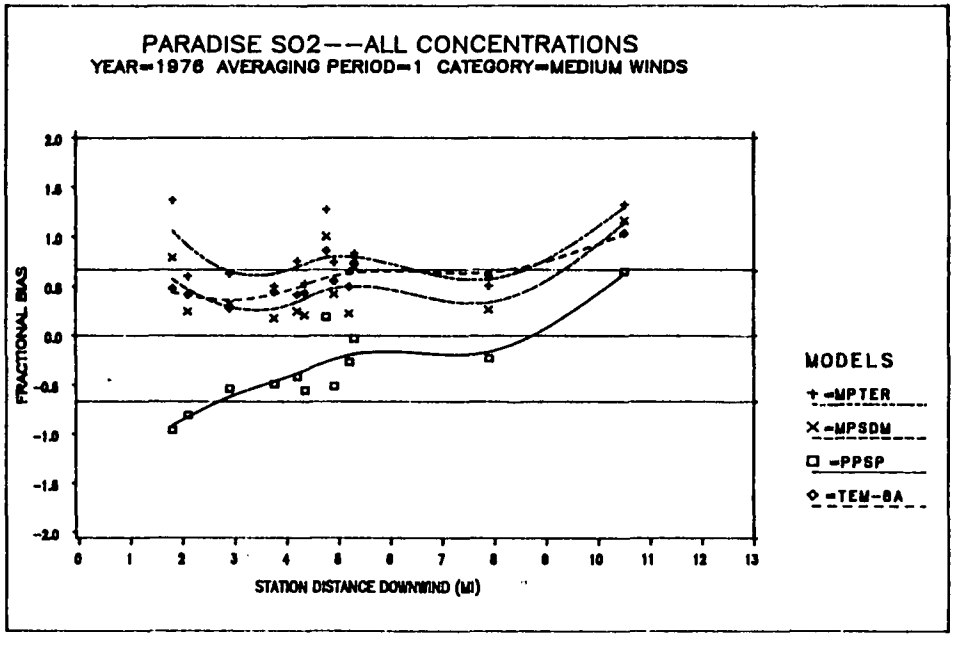


Figure B.16 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using All Paired Values

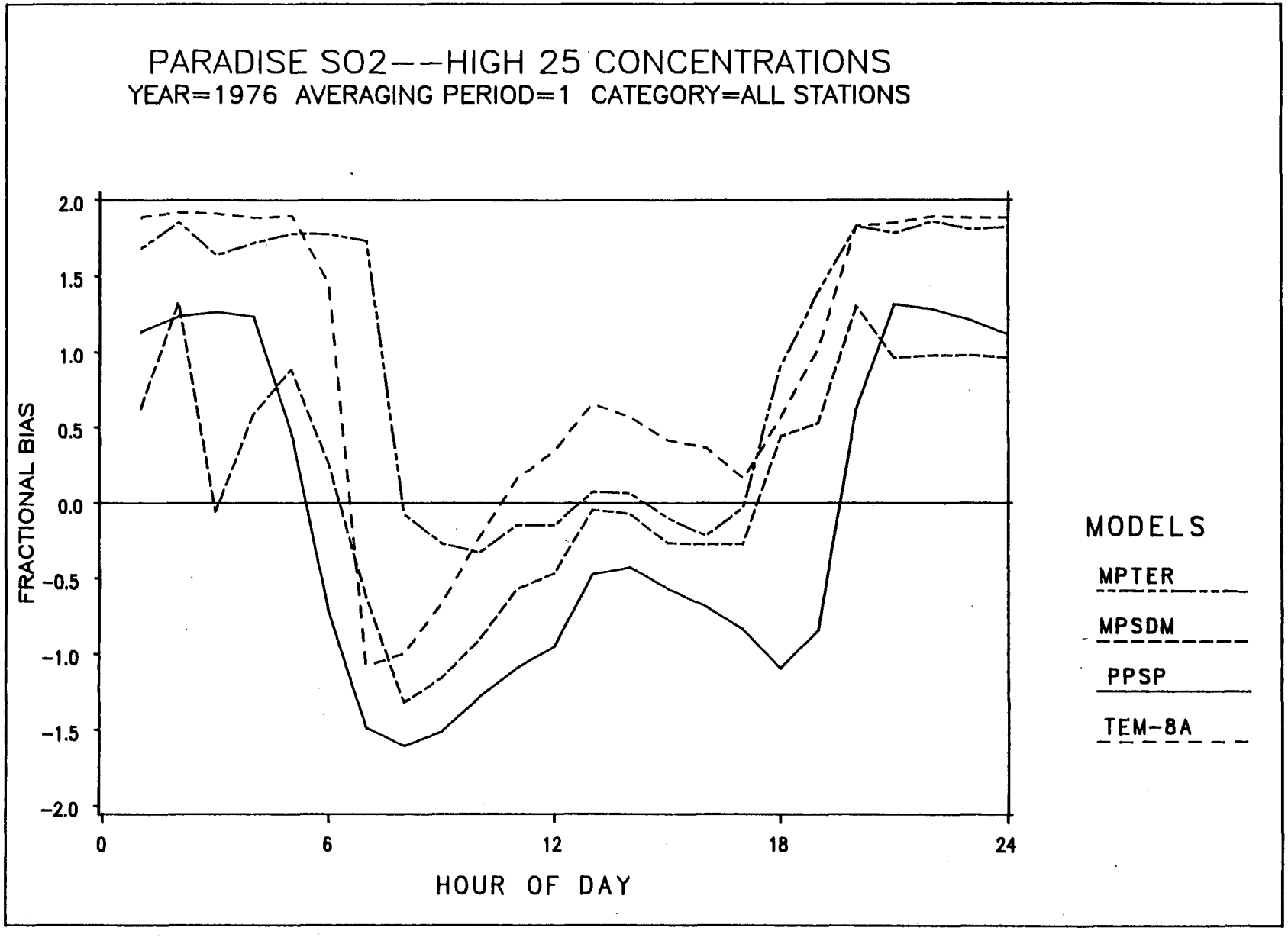


Figure B.17 Fractional Bias Of The Average Vs Hour Of The Day Using High 25 Values

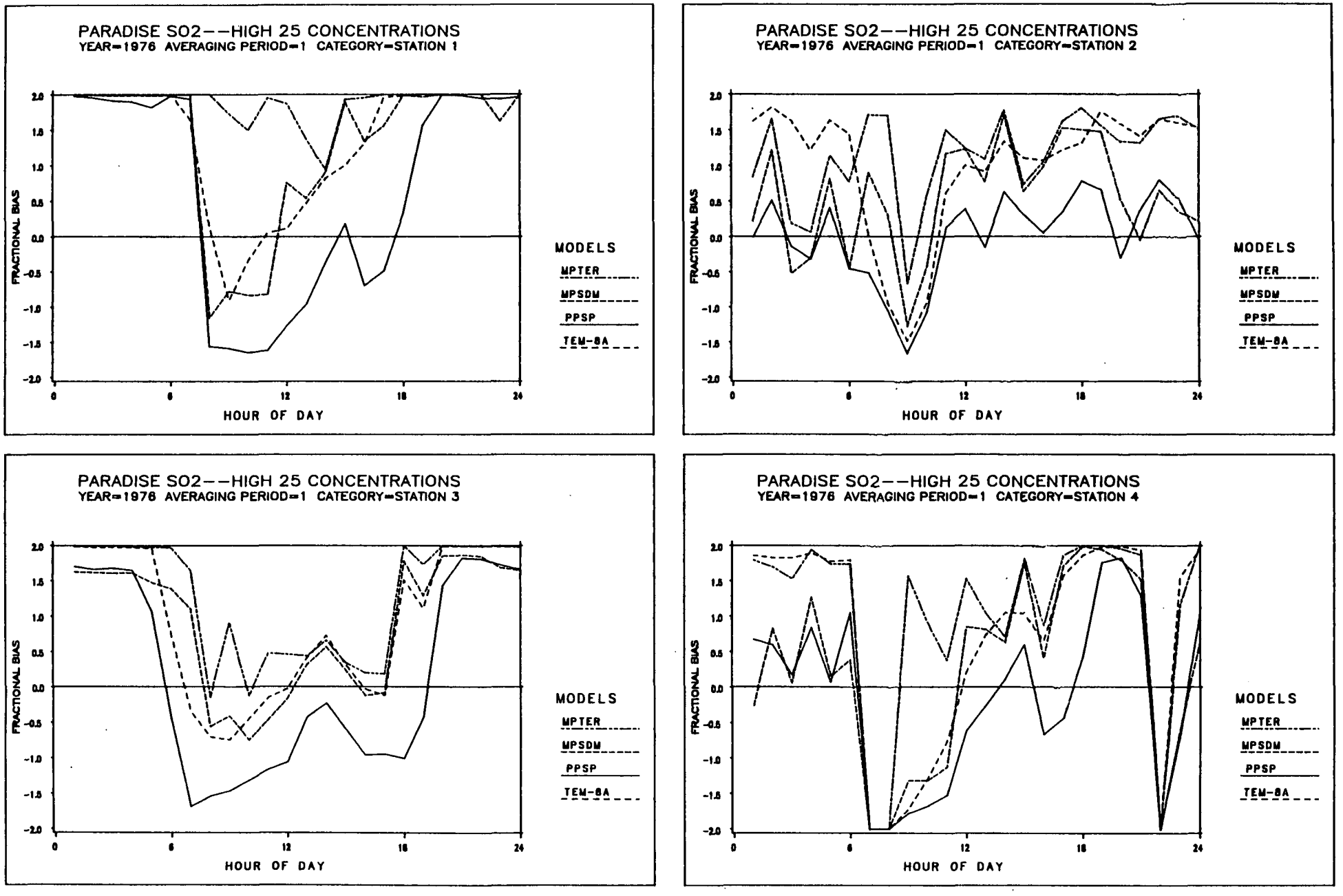
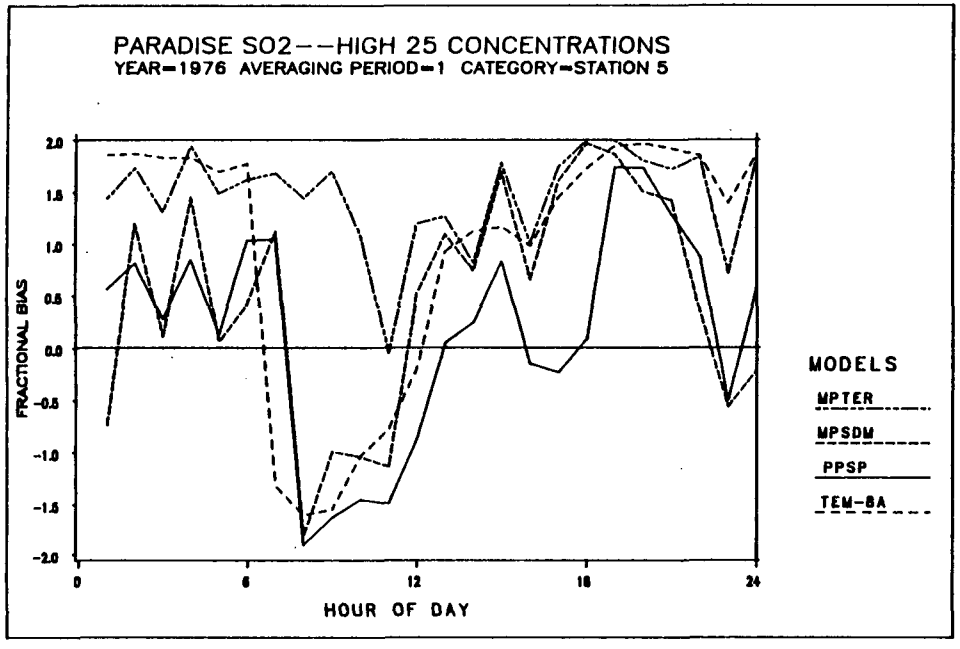
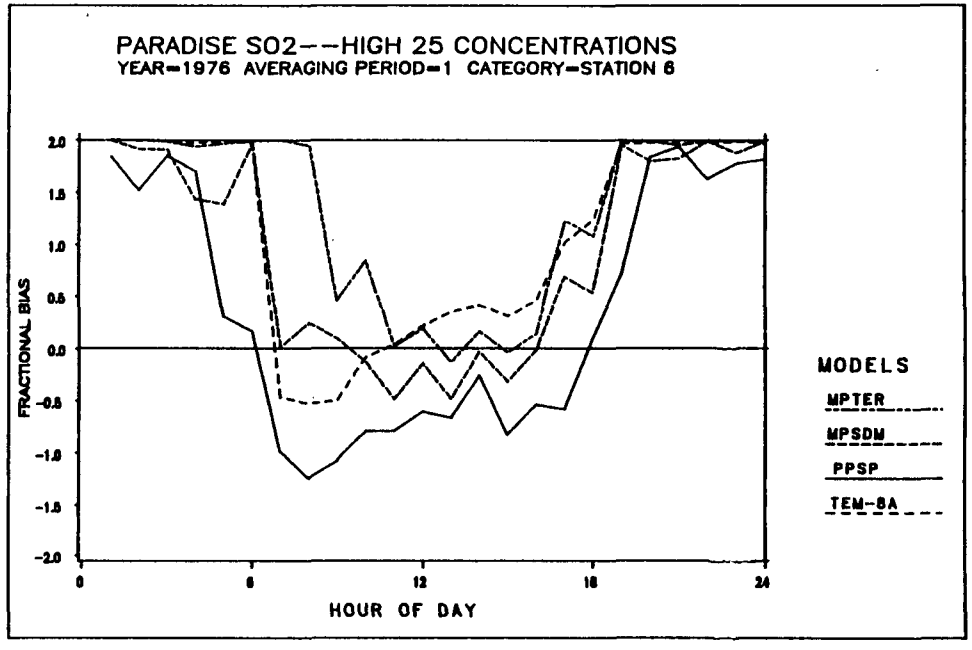
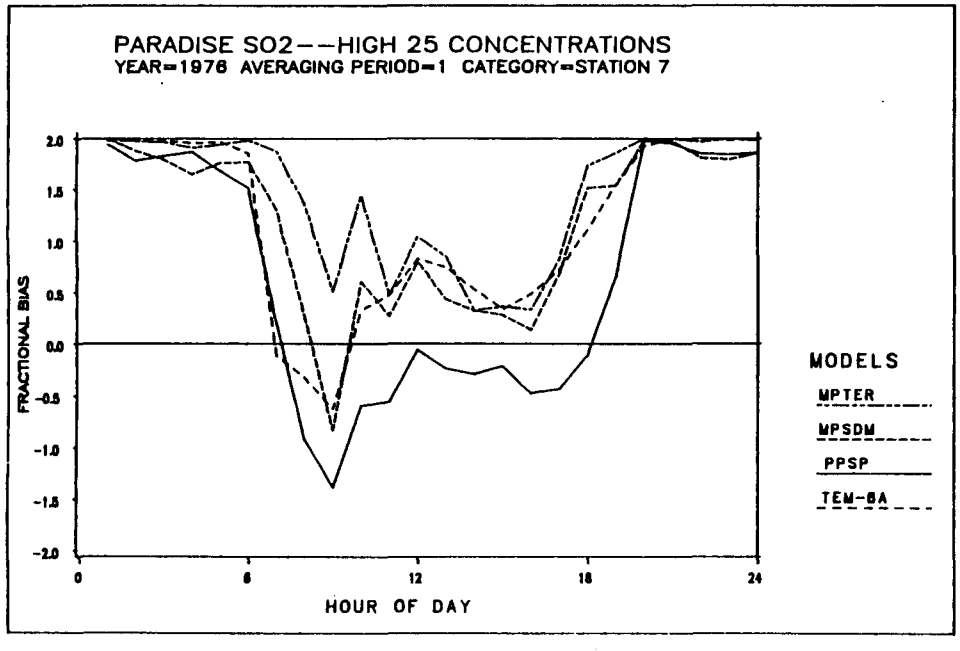


Figure B.18 Fractional Bias Of The Average Vs Hour Of The Day By Station Using 25 High Values







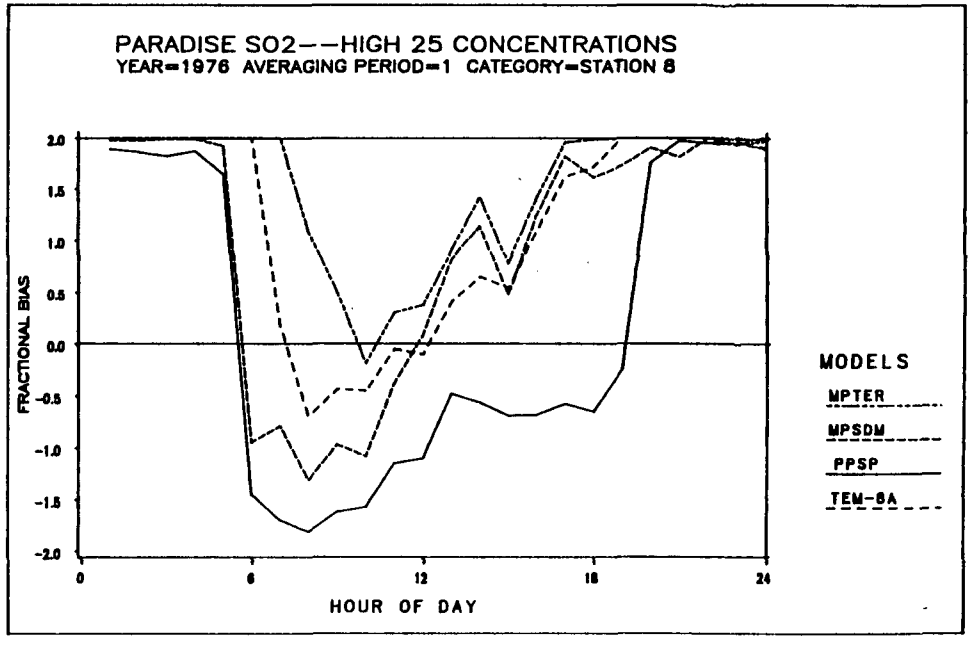
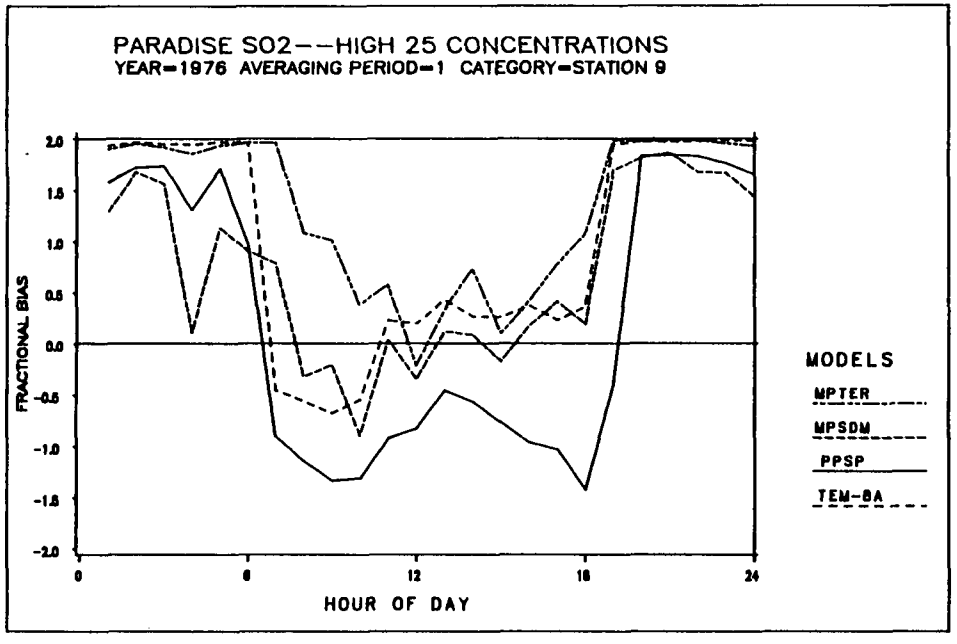
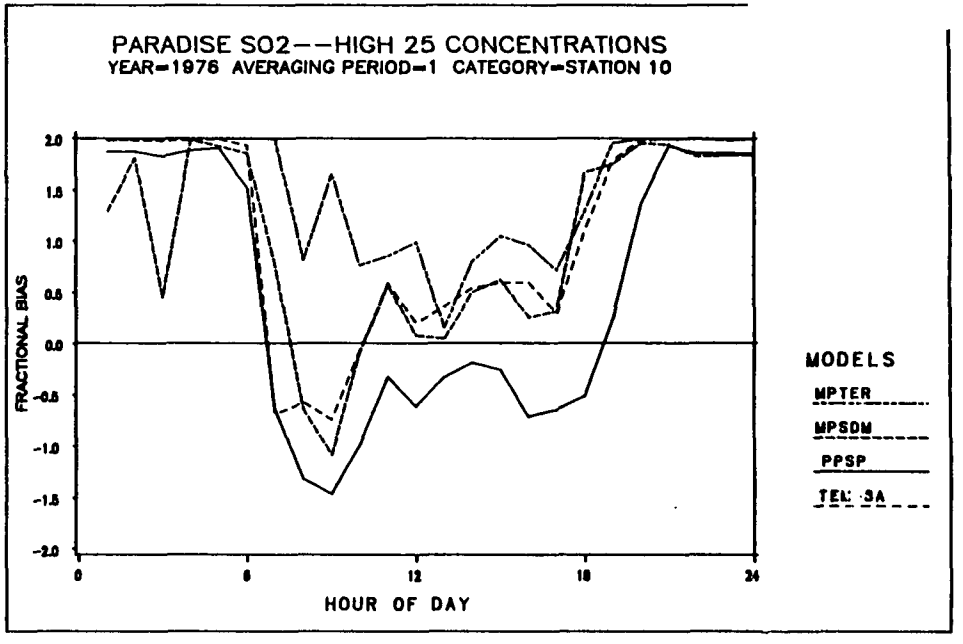
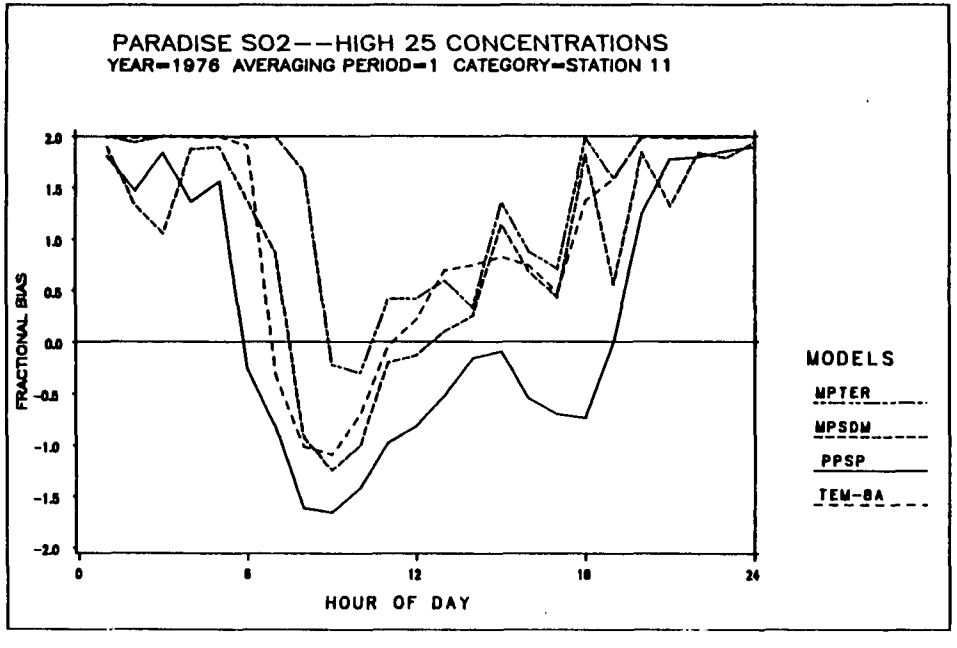


Figure B.18 (Continued)







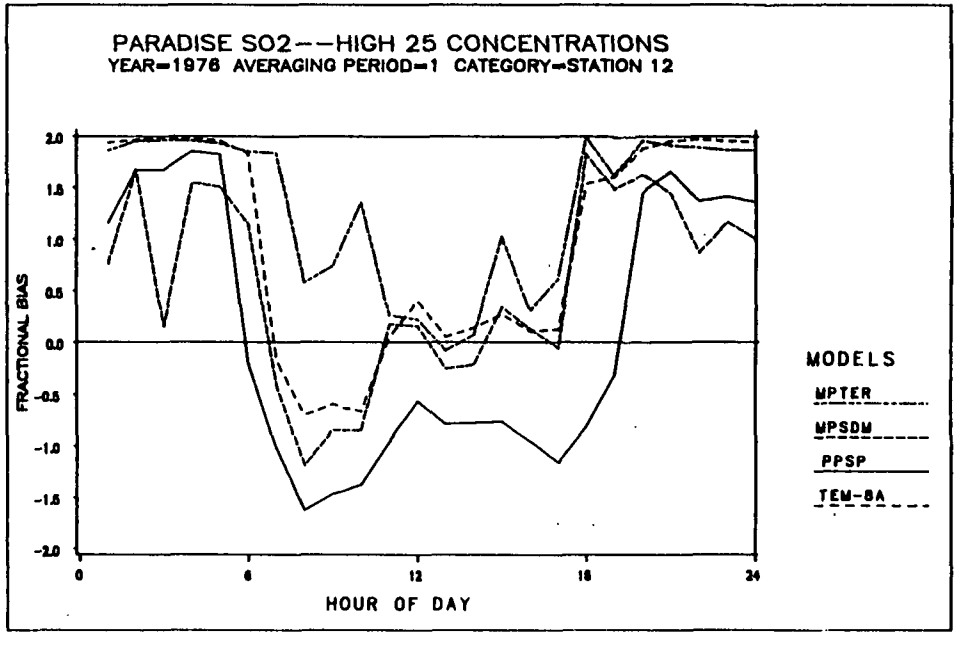


Figure B.18 (Continued)

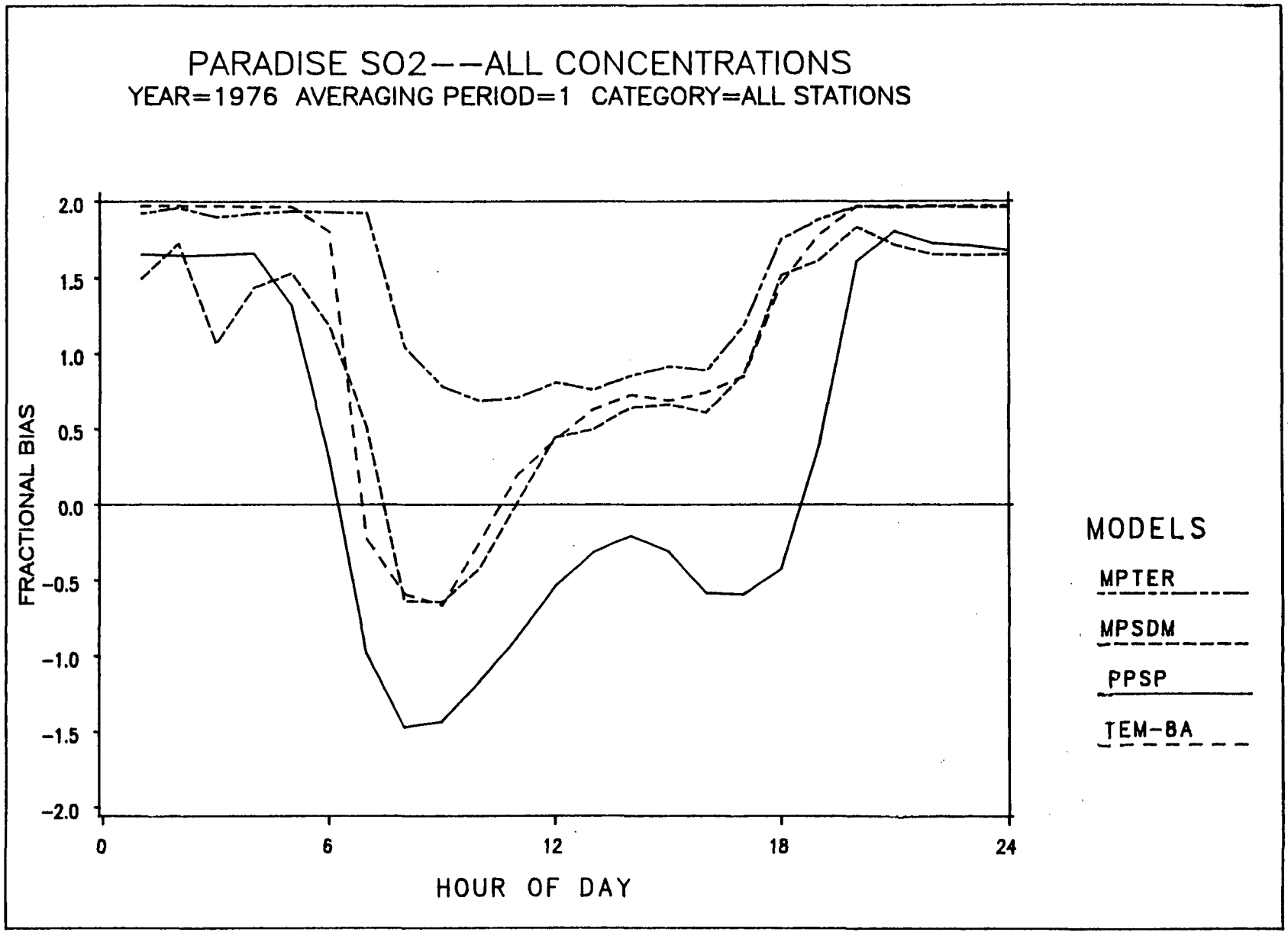


Figure B.19 Fractional Bias Of The Average Vs Hour Of The Day Using All Paired Values

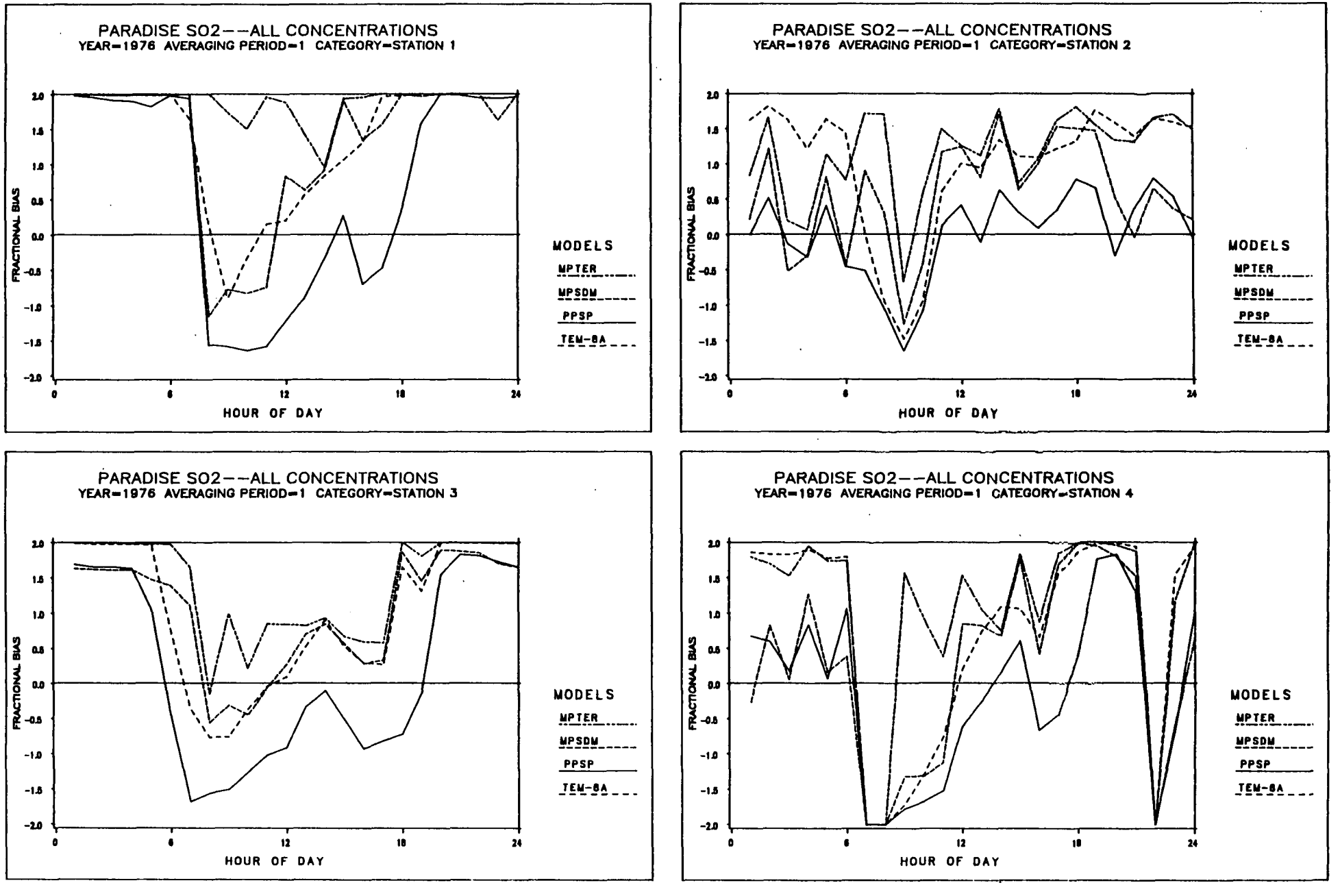
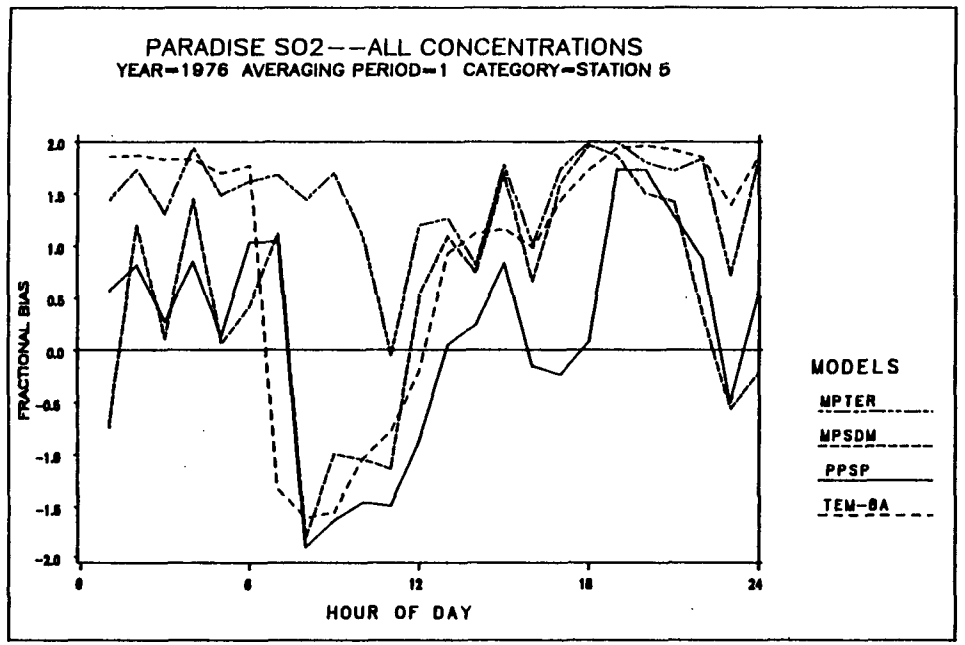
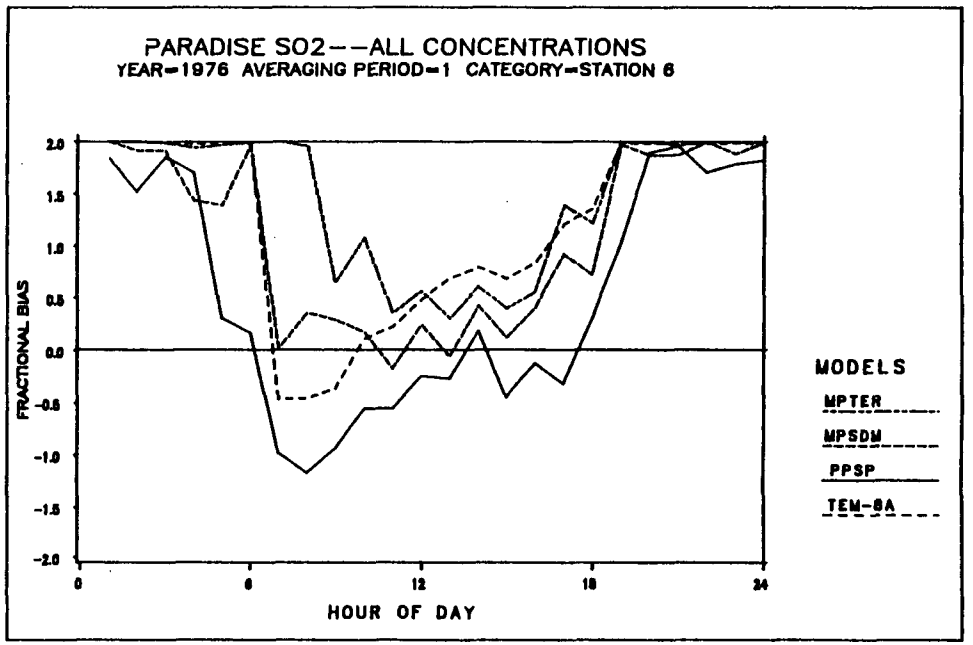
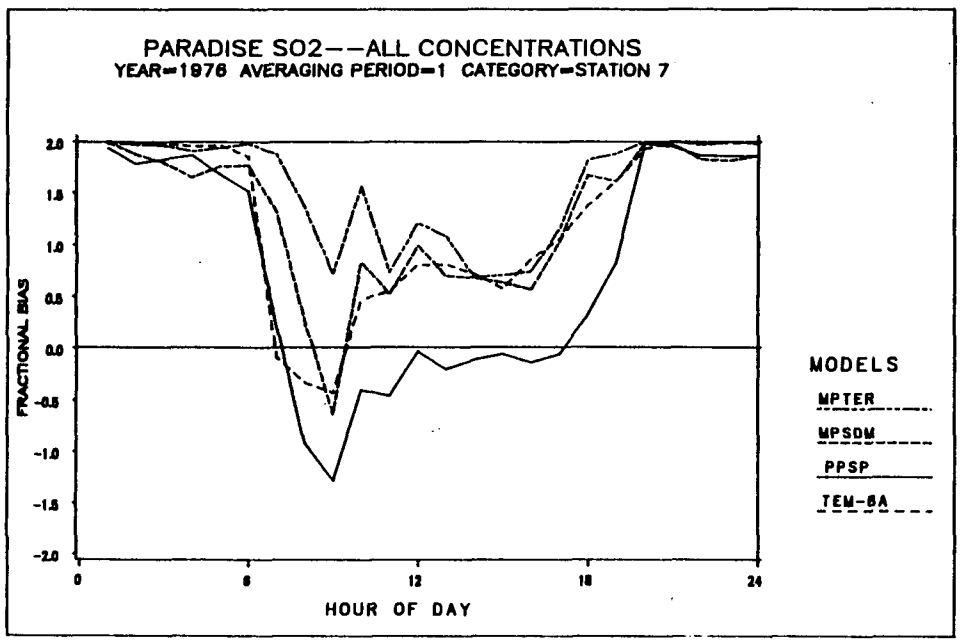


Figure B.20 Fractional Bias Of The Average Vs Hour Of The Day By Station Using All Paired Values







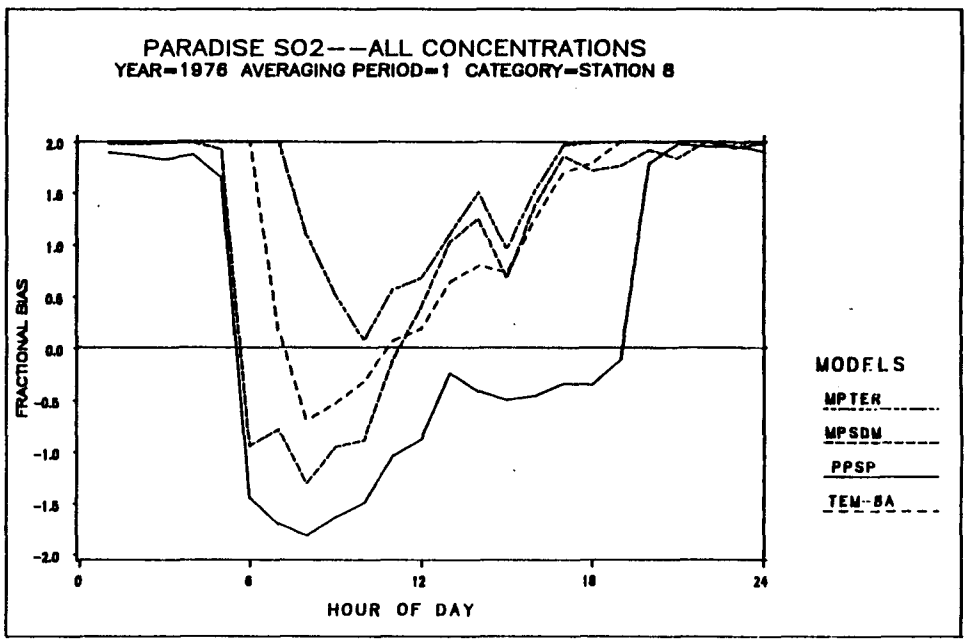
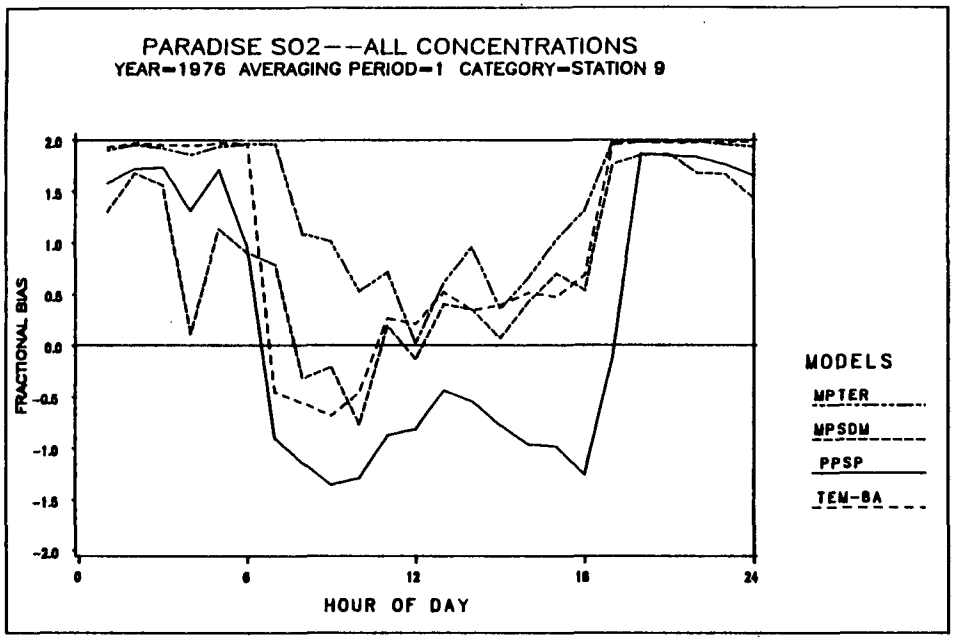
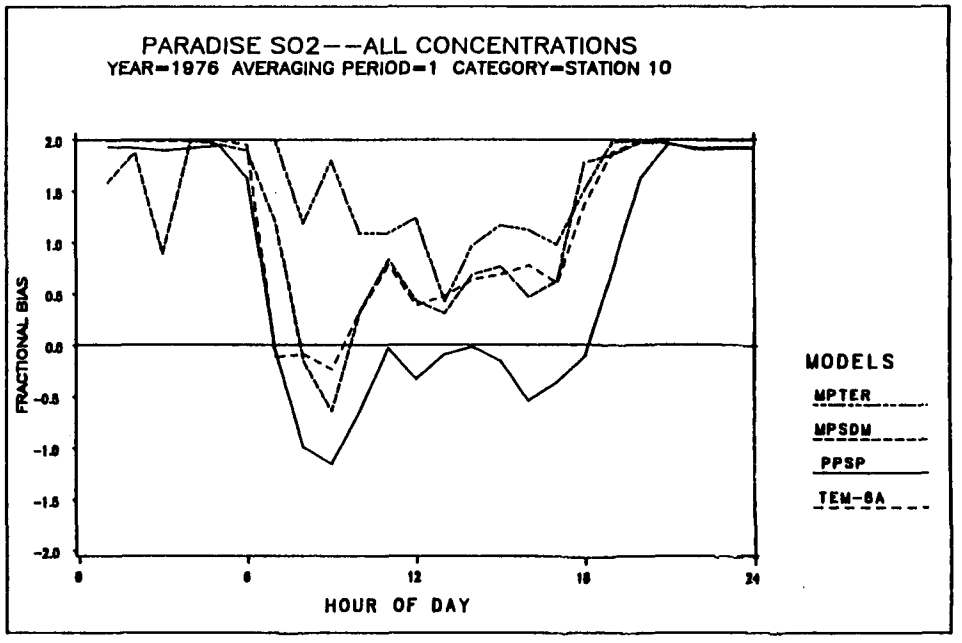
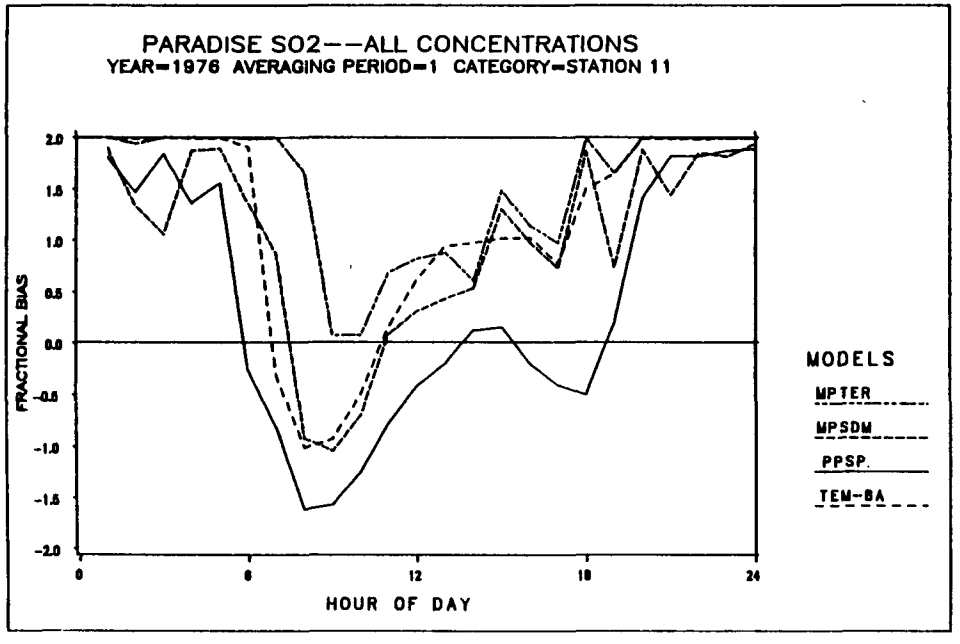
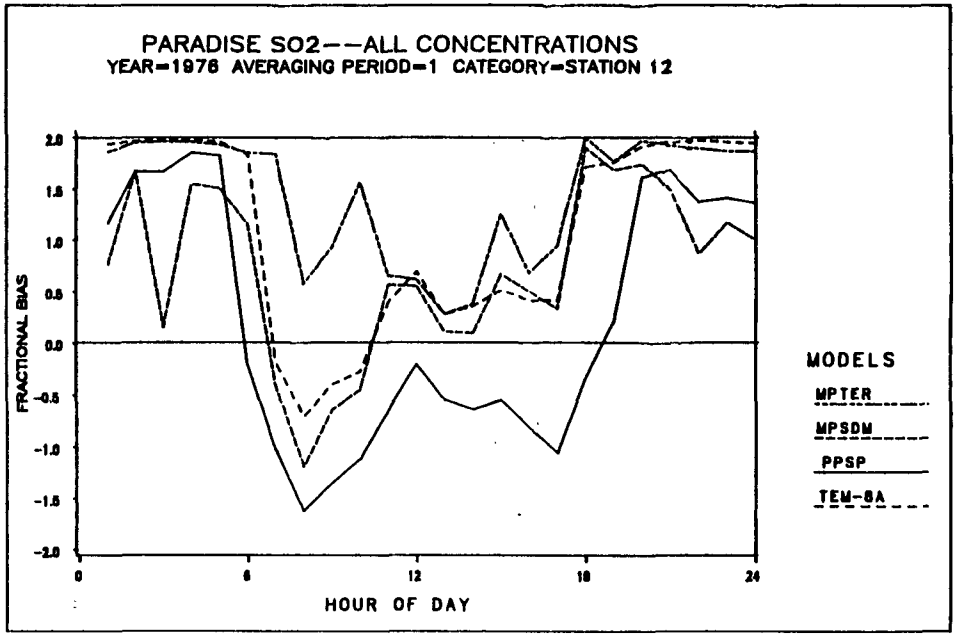


Figure B.20 (Continued)









(Figure B.20 (Continued)

APPENDIX C--MUSKINGUM RIVER PLANT

A sketch of the Muskingum River monitoring network is shown in Figure C.1. The four monitoring stations are located at distances ranging from approximately 4 km to 20 km from the plant. Elevations of the four receptors, relative to the base of the stack, range from approximately 75 to 130 meters.

The results for both 1975 and 1976 are presented for Muskingum River, primarily to provide a sense of the variation in performance that can be expected from year to year. The figures that follow are presented in the same sequence as for Kincaid and Paradise except results for both years are presented within each figure. Results for each year will be discussed only when differences appear to be worth noting.

From Figure C.2, it appears that the models as a group tend toward larger underpredictions and smaller overprediction in 1976 compared with 1975. Generally, the points plotted in 1976 are shifted to the right relative to their position in 1975. While this shift toward underprediction occurs consistently for the bias of the average, the bias in the scatter occasionally shifts in the direction of greater overprediction in 1976. For example, MPSDM is within a factor of two in 1975 for the scatter of both 1-hour and 3-hour averages of the highest 25 values. In 1976 this is not the case, as the fractional bias of the scatter for MPSDM falls outside and below the value (less than -0.67) corresponding to overprediction exceeding a factor of two.

In terms of the operational performance of the four models (Figures C.2-C.4), it appears that PPSP is the most seriously biased, overpredicting for most of the data categories. MPTER and MPSDM appear to be relatively

unbiased while TEM-8A tends toward underpredictions. Exceptions to this rule are evident for some averaging periods, data set types and years. The general trends noted previously still hold, i.e., the longer averaging periods and all data categories are associated with the largest underpredictions and smallest overpredictions.

Figures C.5-C.10 delve deeper into the performance across the four stations, the four stability categories and three wind speed categories. For the 25 highest values, performance among the stations varies somewhat (Figure C.5) but the pattern is fairly consistent among the four models. Station 4 is associated with the largest underpredictions (and smallest overpredictions). The spread in results among the four stations appears smaller in 1976 than in 1975. Again, 1976 and 1975 show differences in performance as exemplified by TEM-8A in which reasonably accurate predictions in 1975 contrast with noticeable underpredictions in 1976. Differences in performance among stability categories are also obvious (Figure C.6) Generally, the stable categories are associated with underpredictions while unstable categories are associated with relatively low bias. Differences in performance among wind speed categories (Figure C.7) are relatively small for MPTER but noticeably larger for the other three models. For the all data category (Figures C.8-C.10), there is no particular feature that warrants discussion other than to repeat that there is a general tendency toward greater underprediction compared with the data set comprised of the 25 highest values.

The next series of plots (Figures C.11-C.16) display the fractional bias of the average as a function of station to source distance. Since only

four stations are involved, the trend lines should be viewed with some caution. Examination of these plots broken down by stability category (Figure C.12--high 25 and Figure C.15--all data) reveals differences in performance among the models as a function of station to source distance for neutral and stable conditions. For very unstable and unstable categories, the trend line is relatively flat for each of the models. Examination of these plots broken down by wind speed (Figures C.13 and C.16), reveals no significant patterns, probably due to the relatively few number of stations that are available.

Finally, Figures C.17-C.20 display the familiar diurnal pattern seen for the other data bases. The models as a group show a tendency for underpredictions in the early morning and late evening hours and a tendency for relatively low bias in the mid-day hours. The tendency for underprediction by TEM-8A and overprediction by PPSP during the mid-day hour is particularly evident for the 25 highest stations where all stations are combined (Figure C.17).

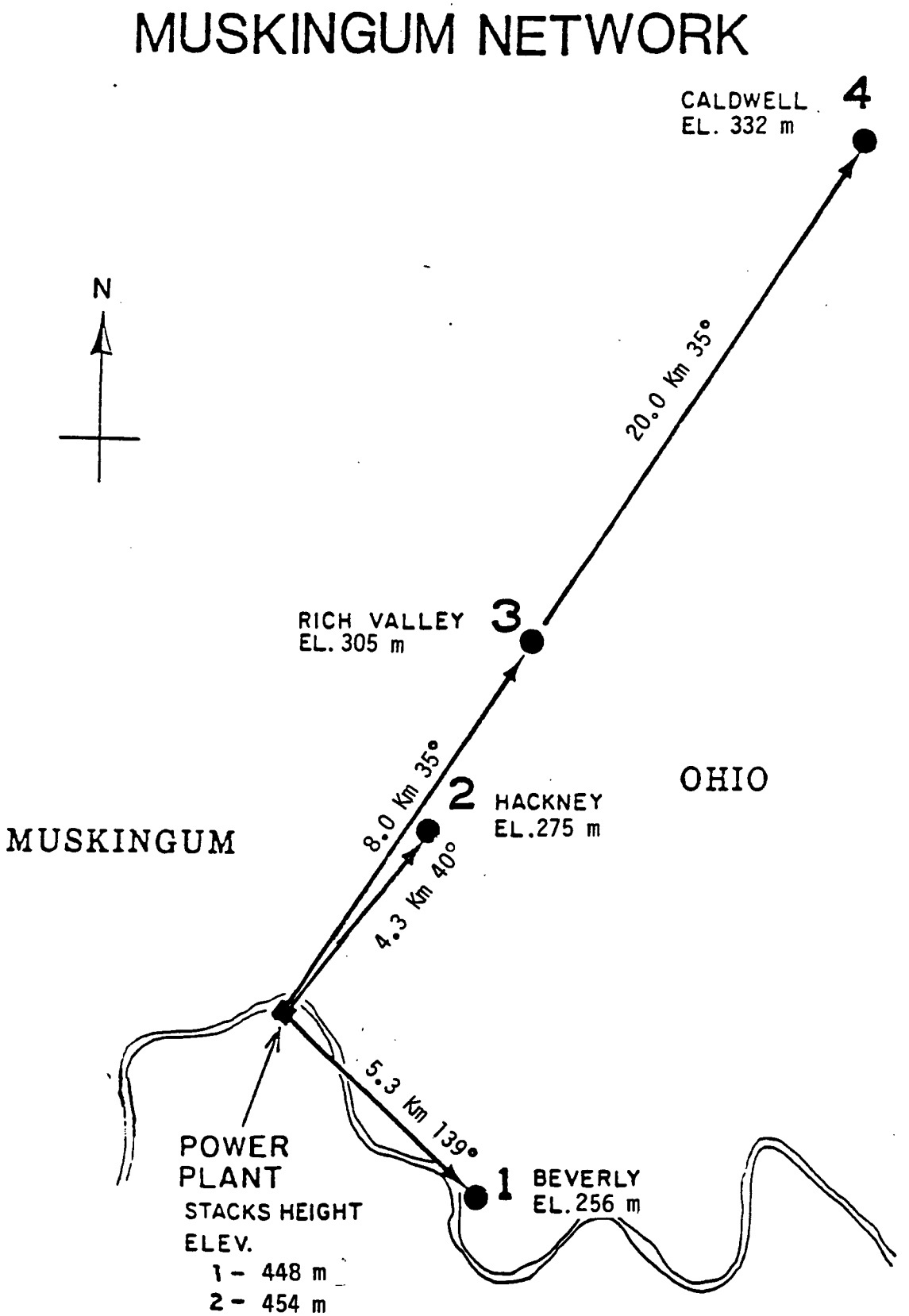


Figure C.1 Field Monitoring Network Near The Muskingum River Plant

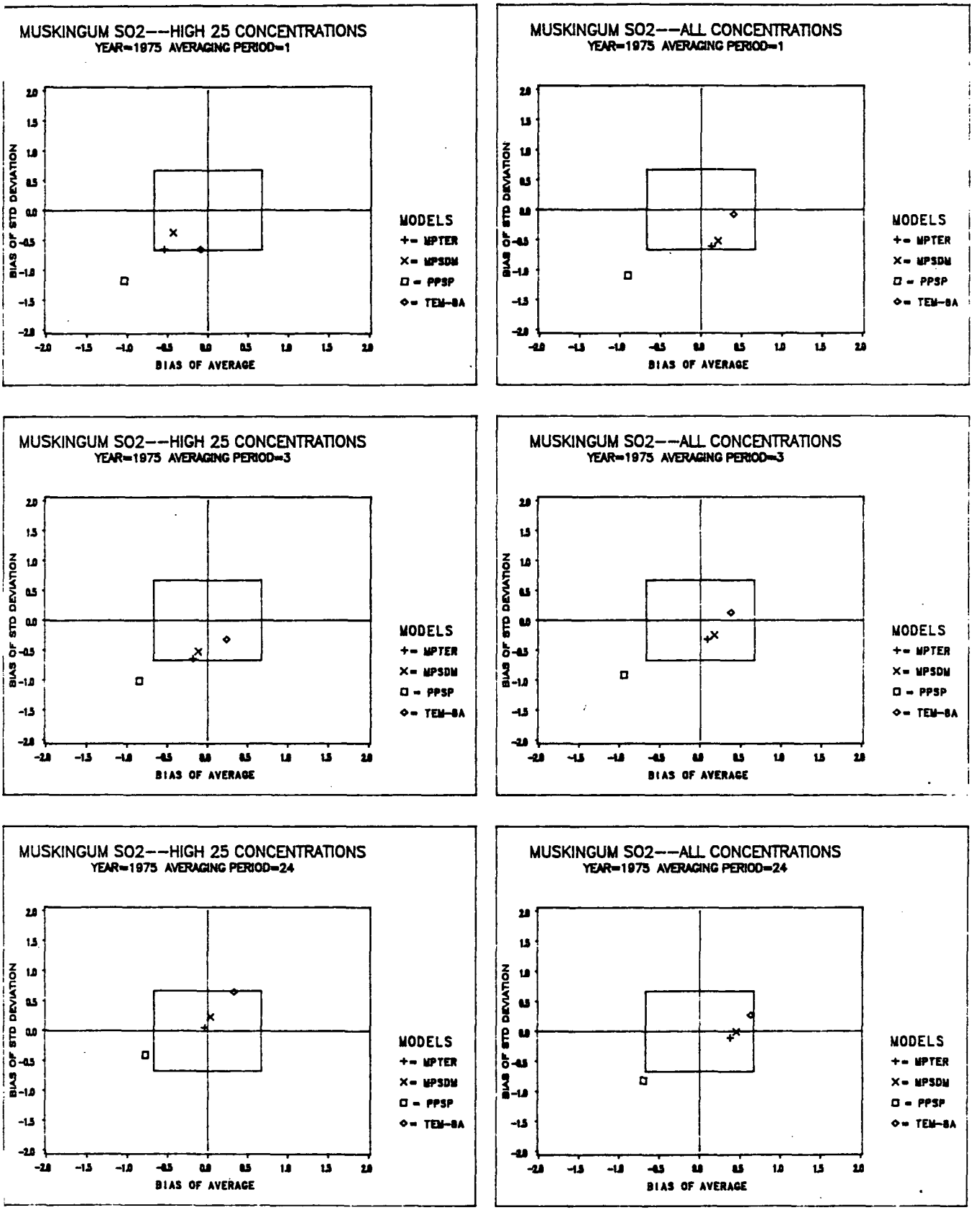
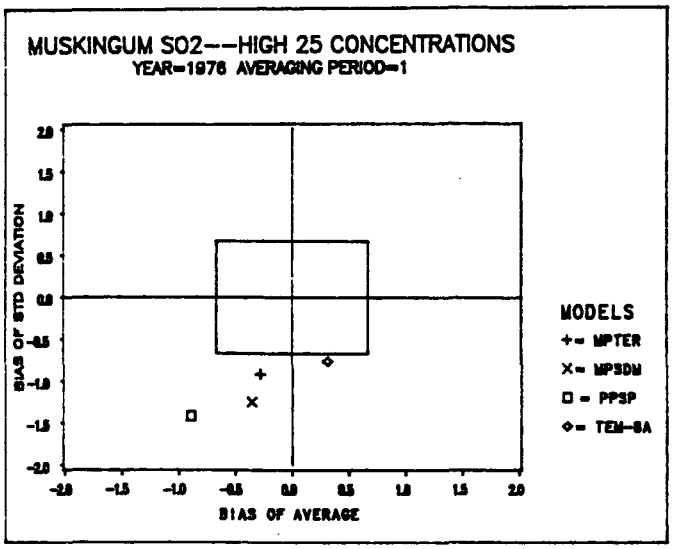
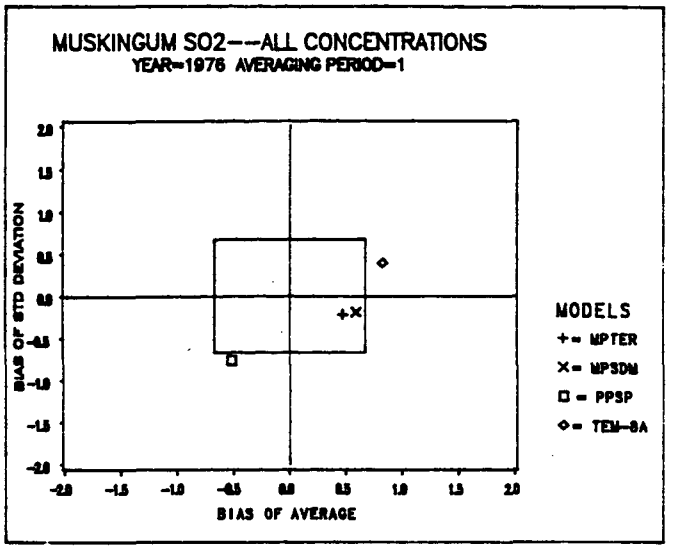
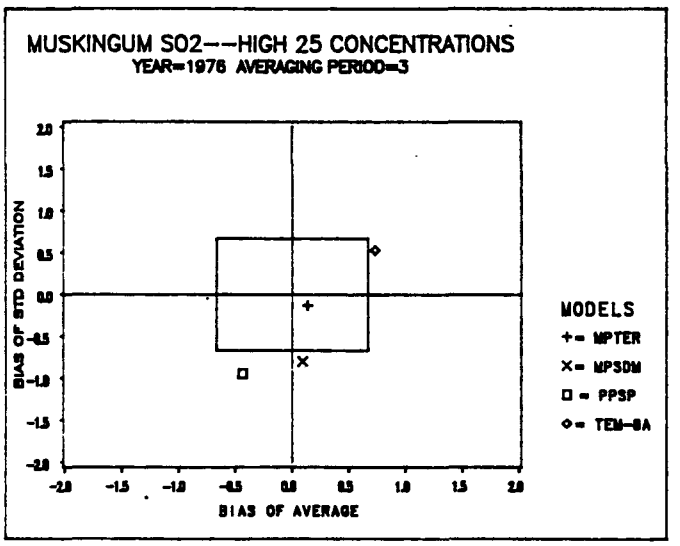
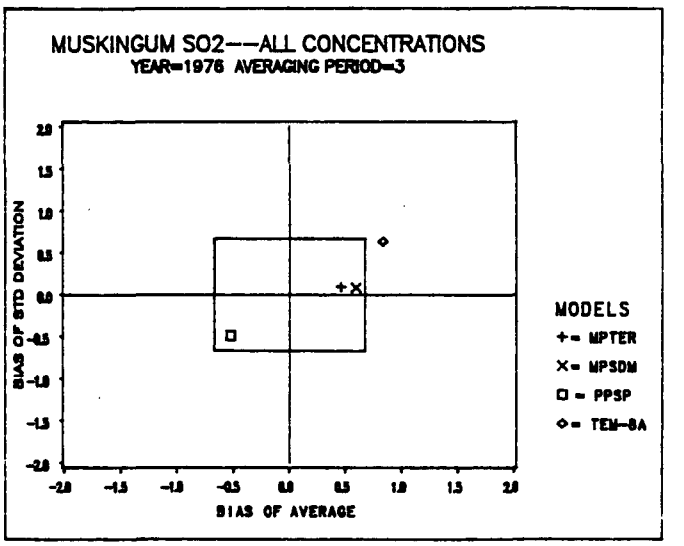


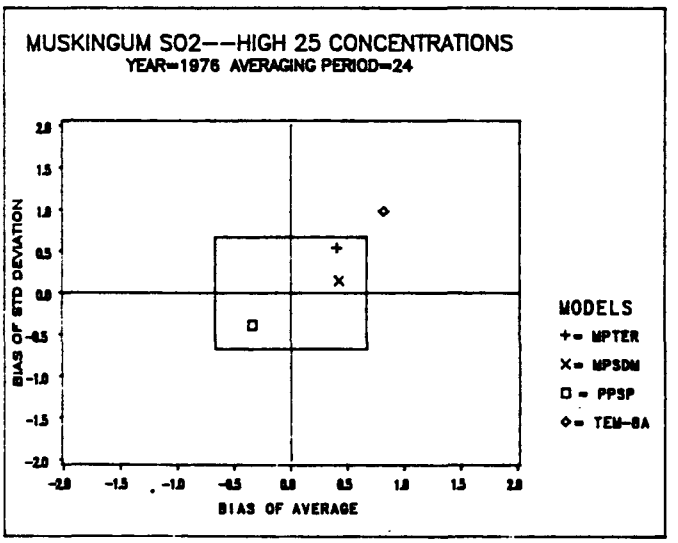
Figure C.2 Fractional Bias By Averaging Period Using High 25 Values And All Paired Values











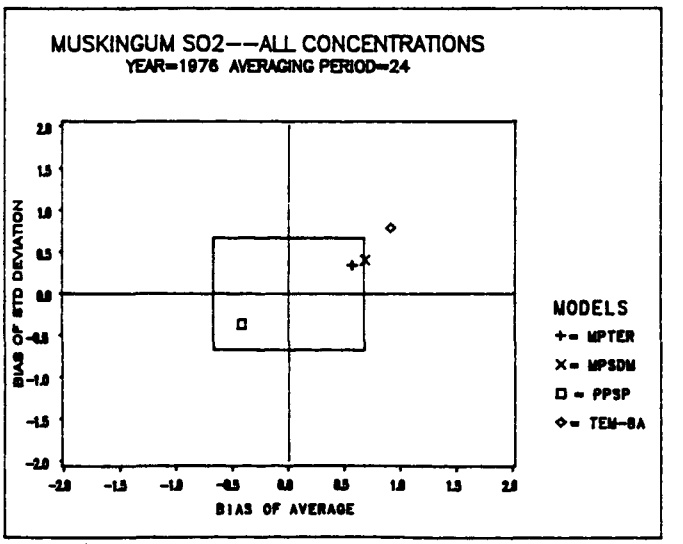


Figure C.2 (Continued)

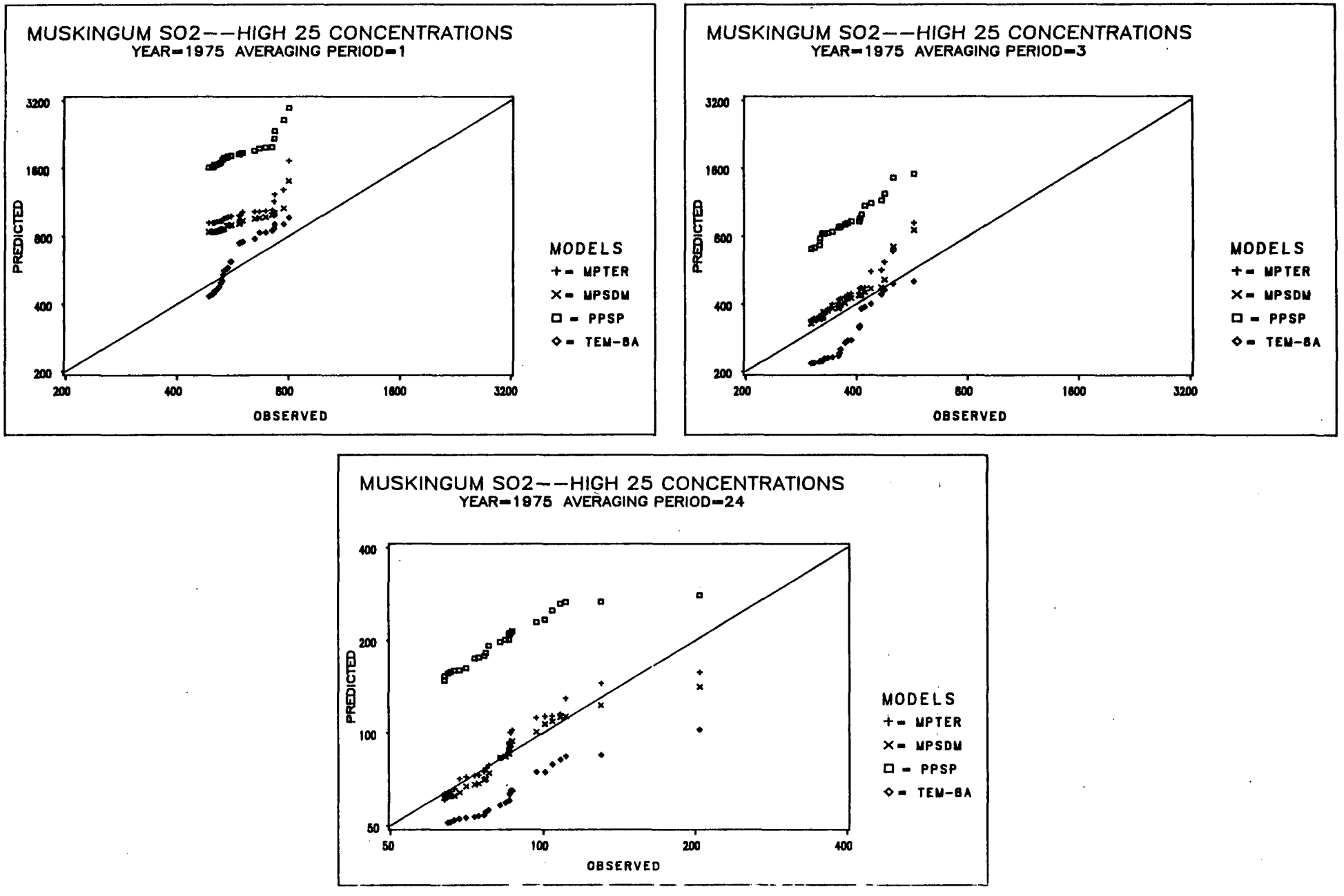
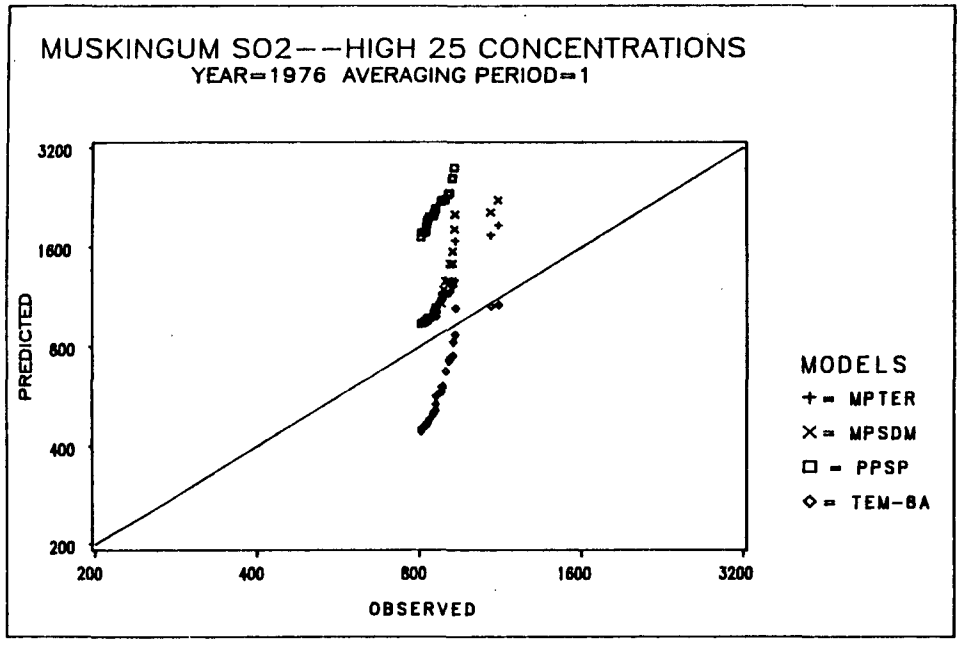
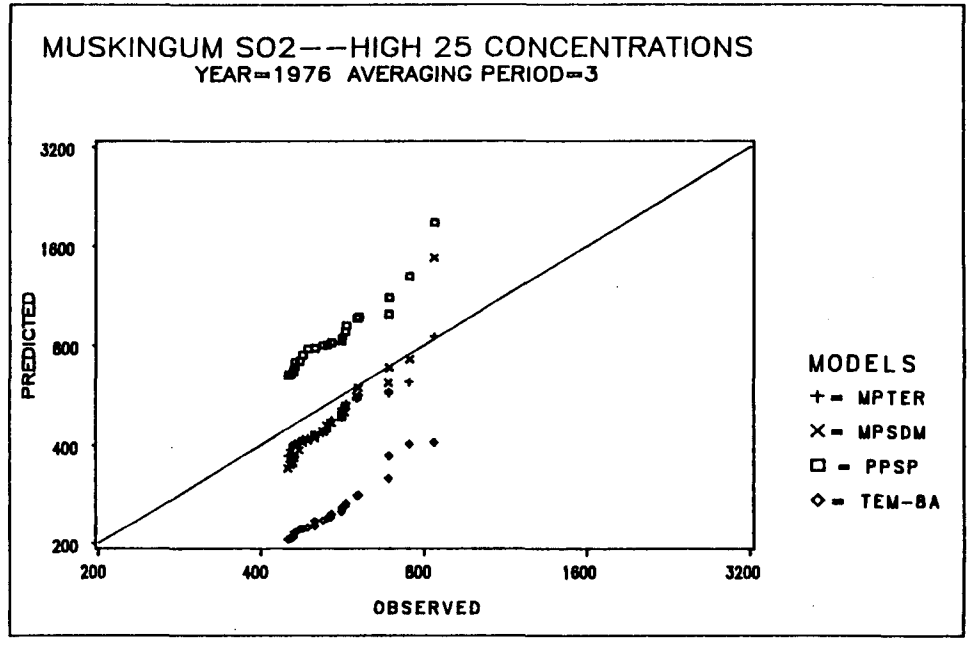


Figure C.3 Predicted Quantiles Vs Observed Quantiles By Averaging Period Using High 25 Values





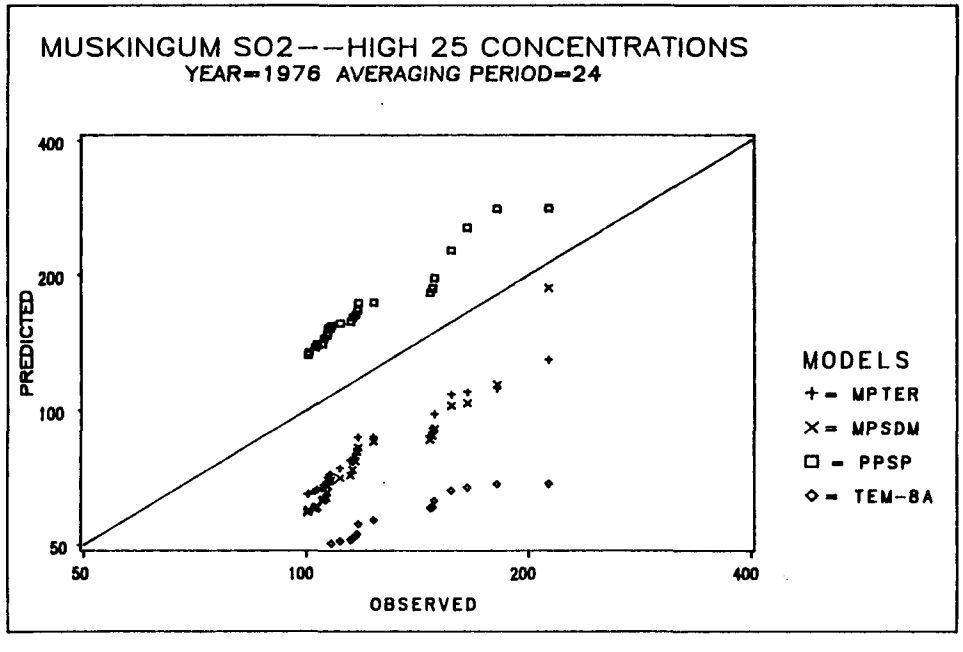
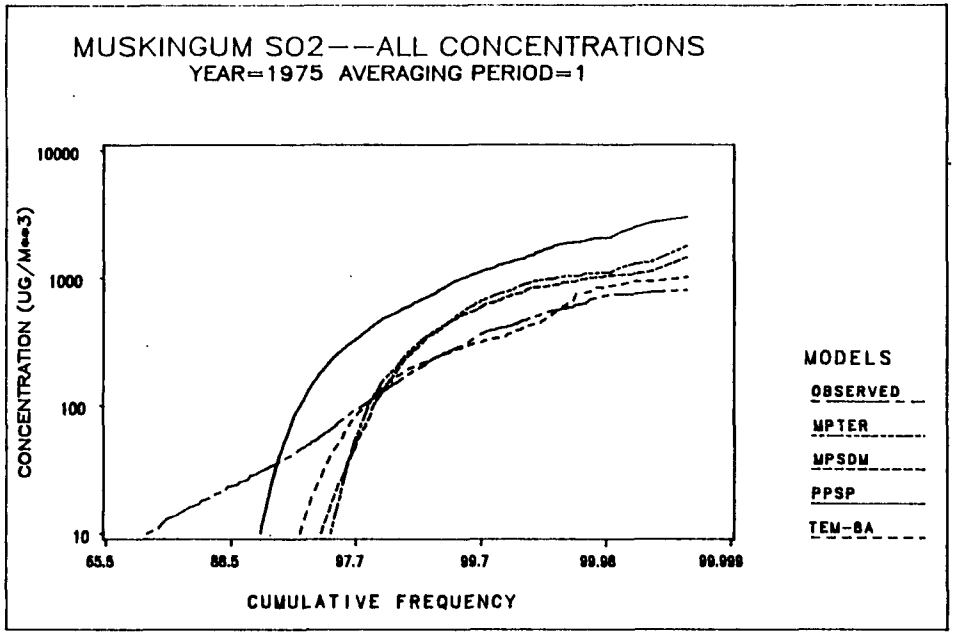
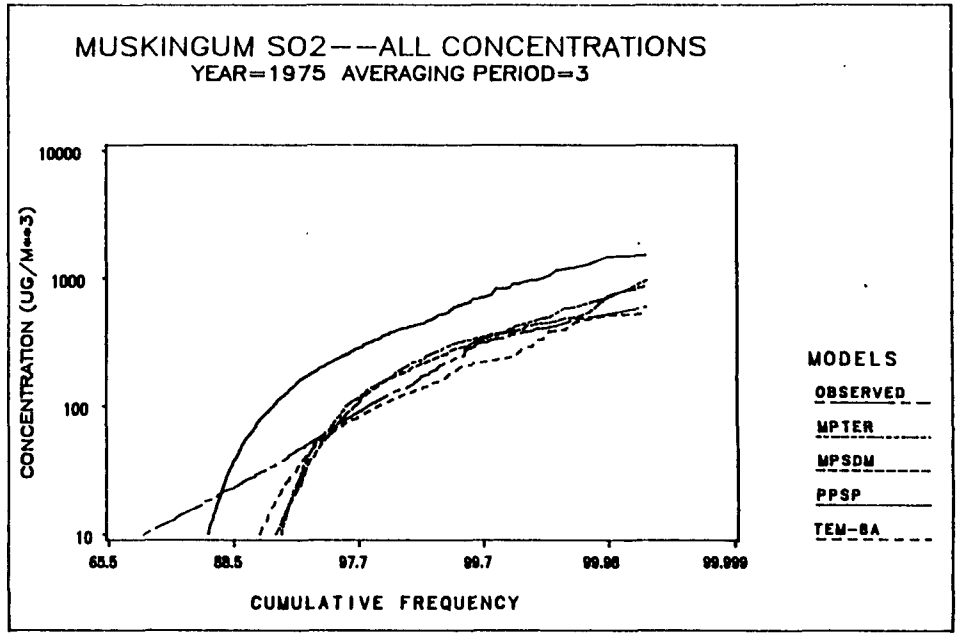


Figure C.3 (Continued)





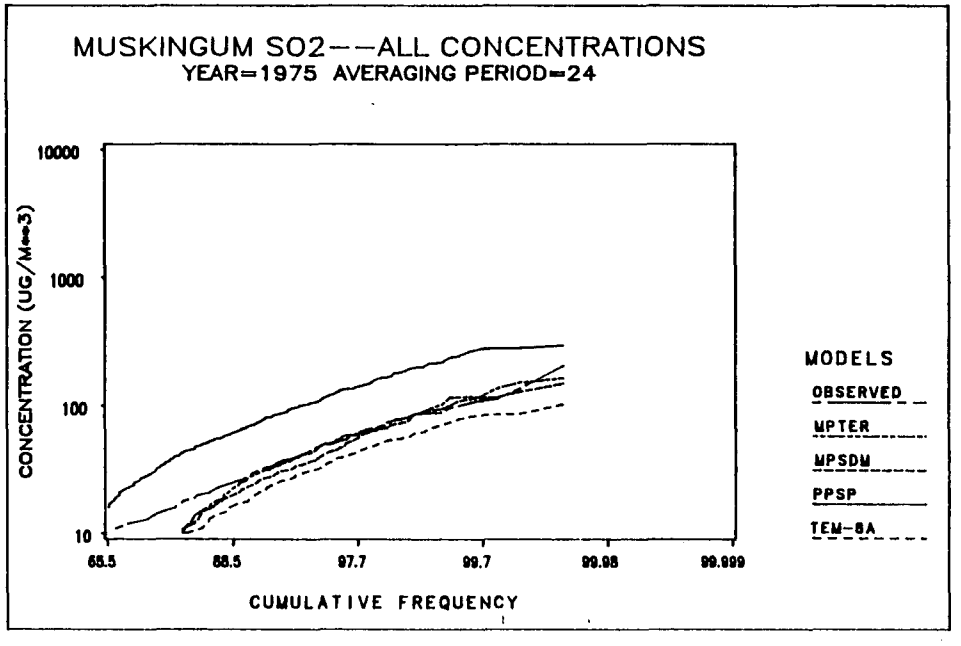
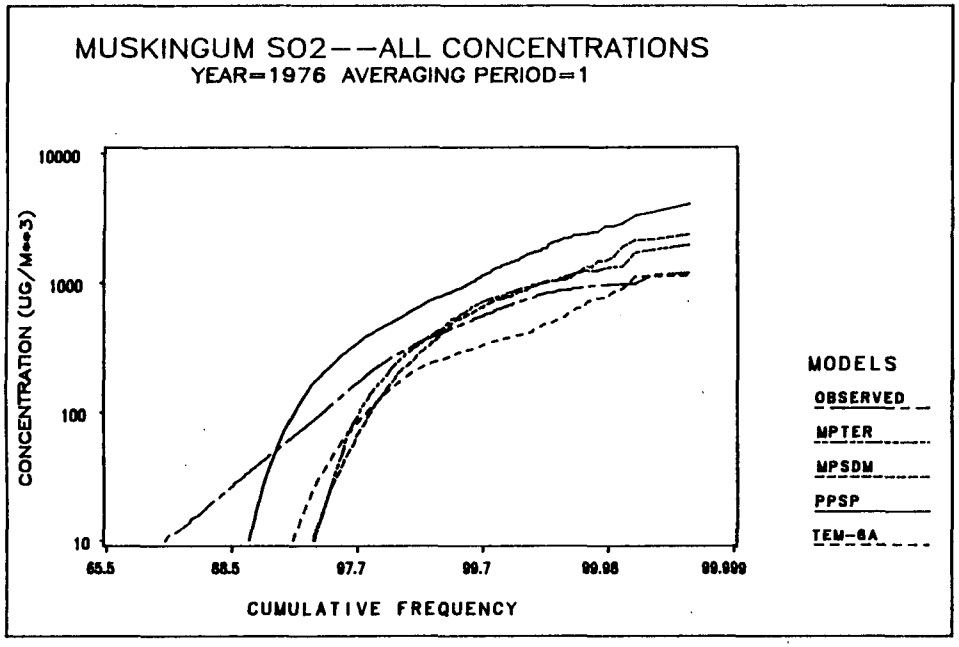
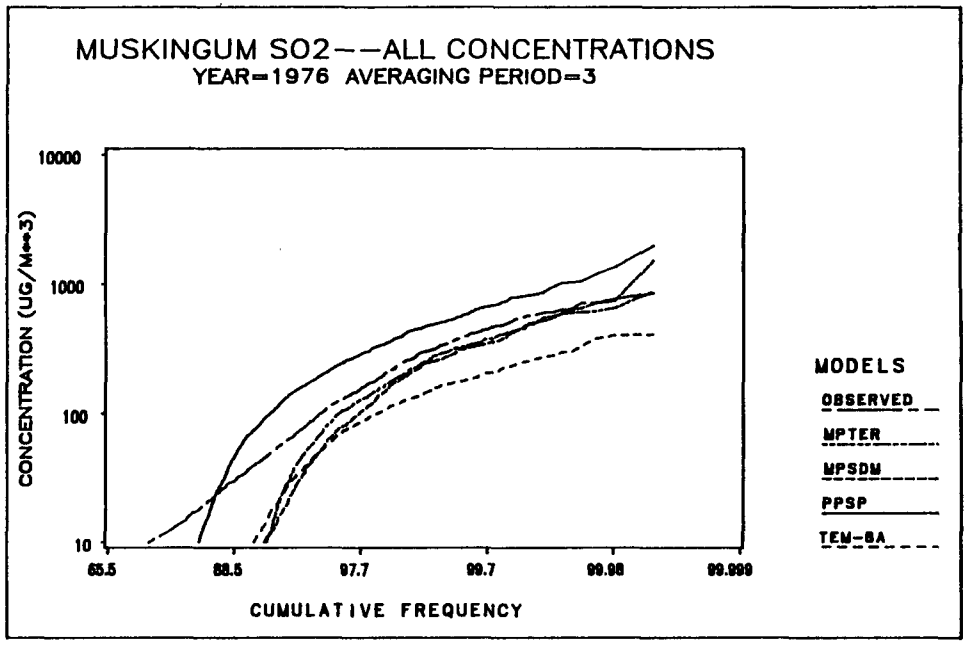


Figure C.4 Cumulative Frequency Distributions By Averaging Period Using All Paired Values





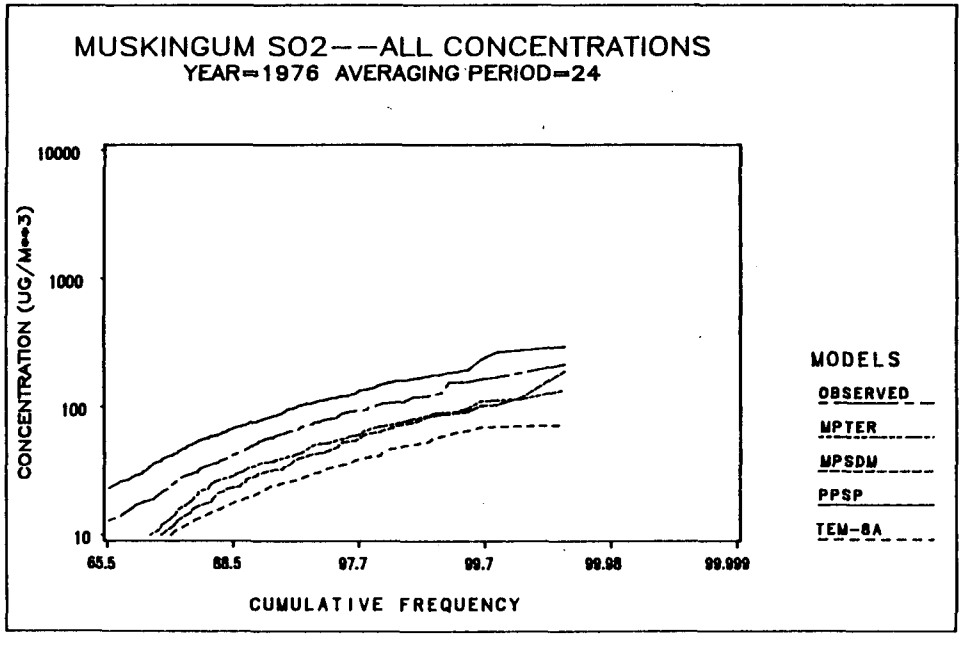


Figure C.4 (Continued)

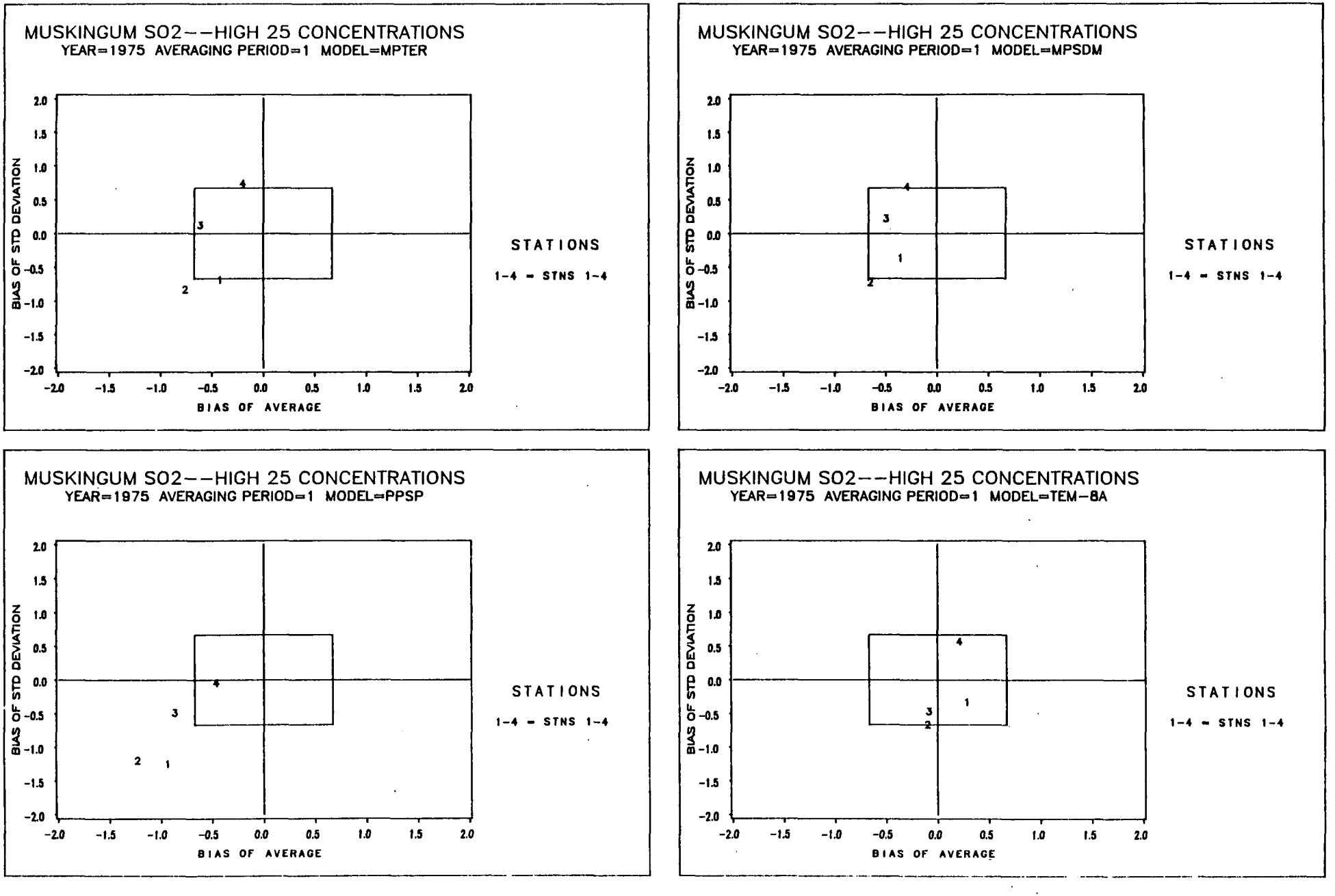
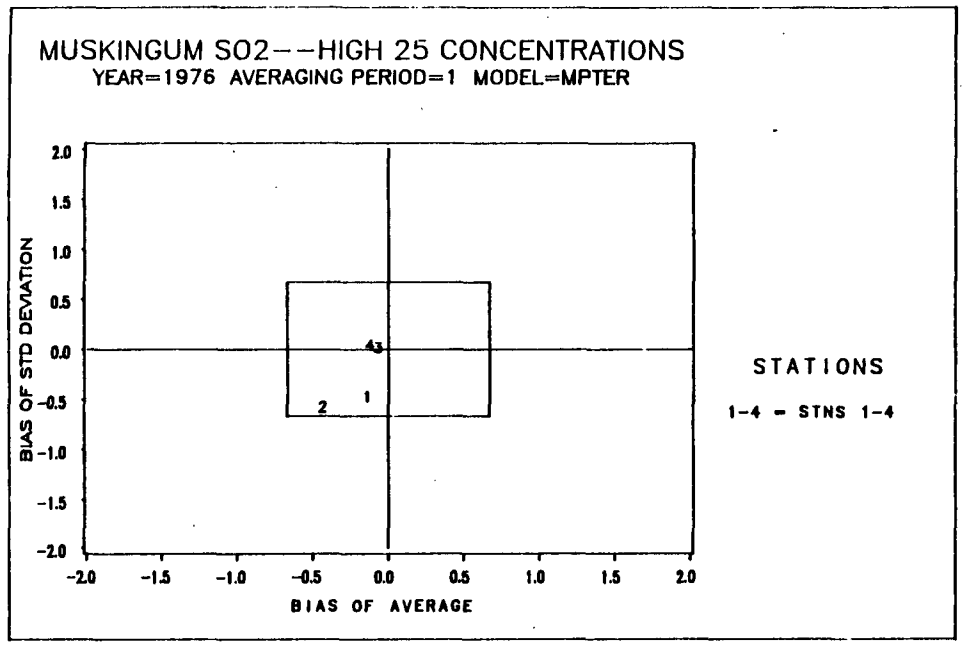
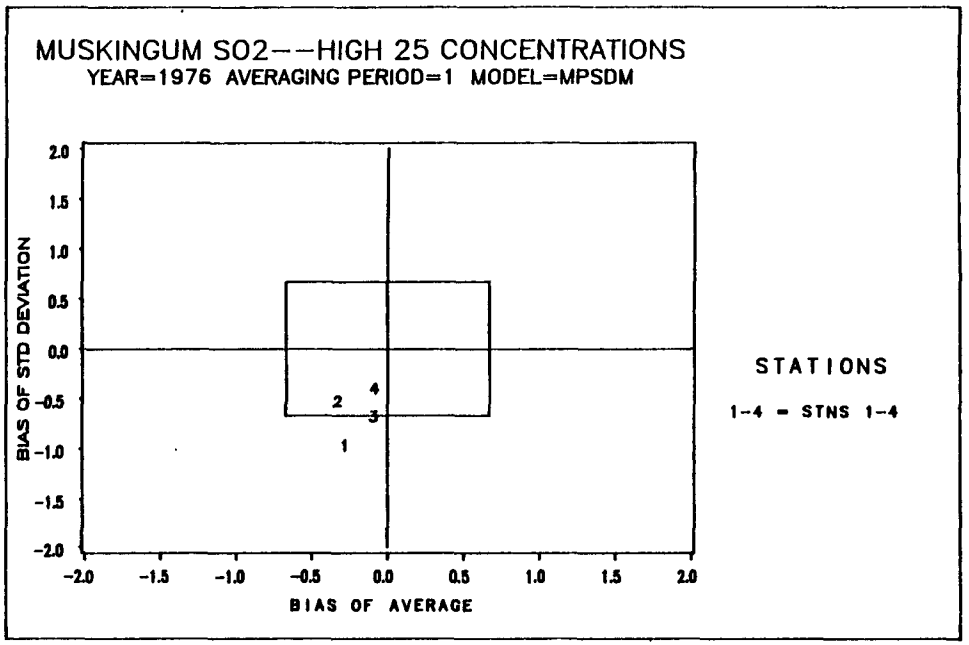
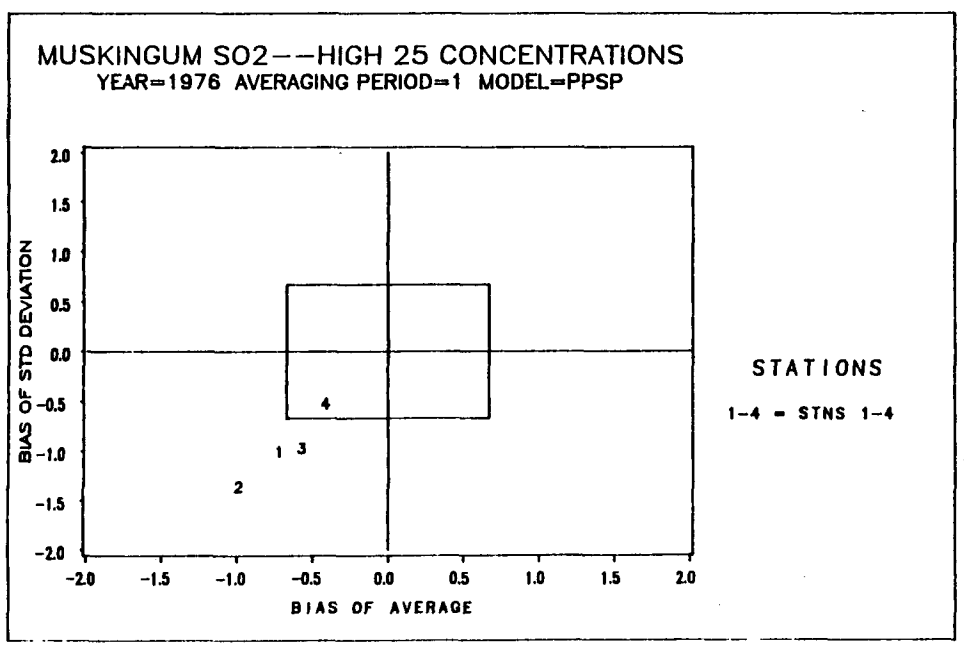


Figure C.5 Fractional Bias By Model Using 25 High Values For Each Station







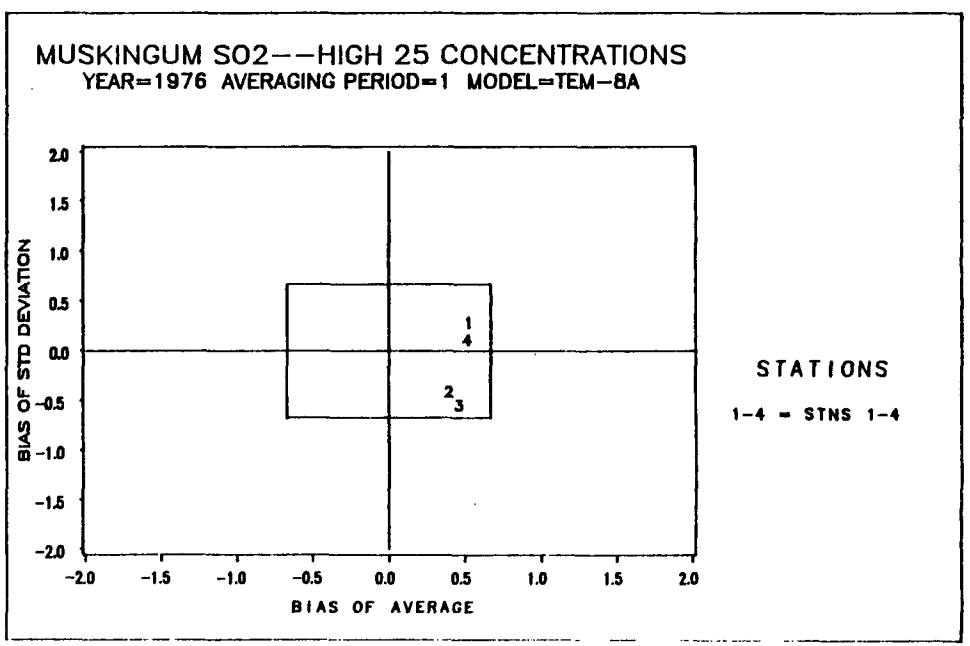


Figure C.5 (Continued)

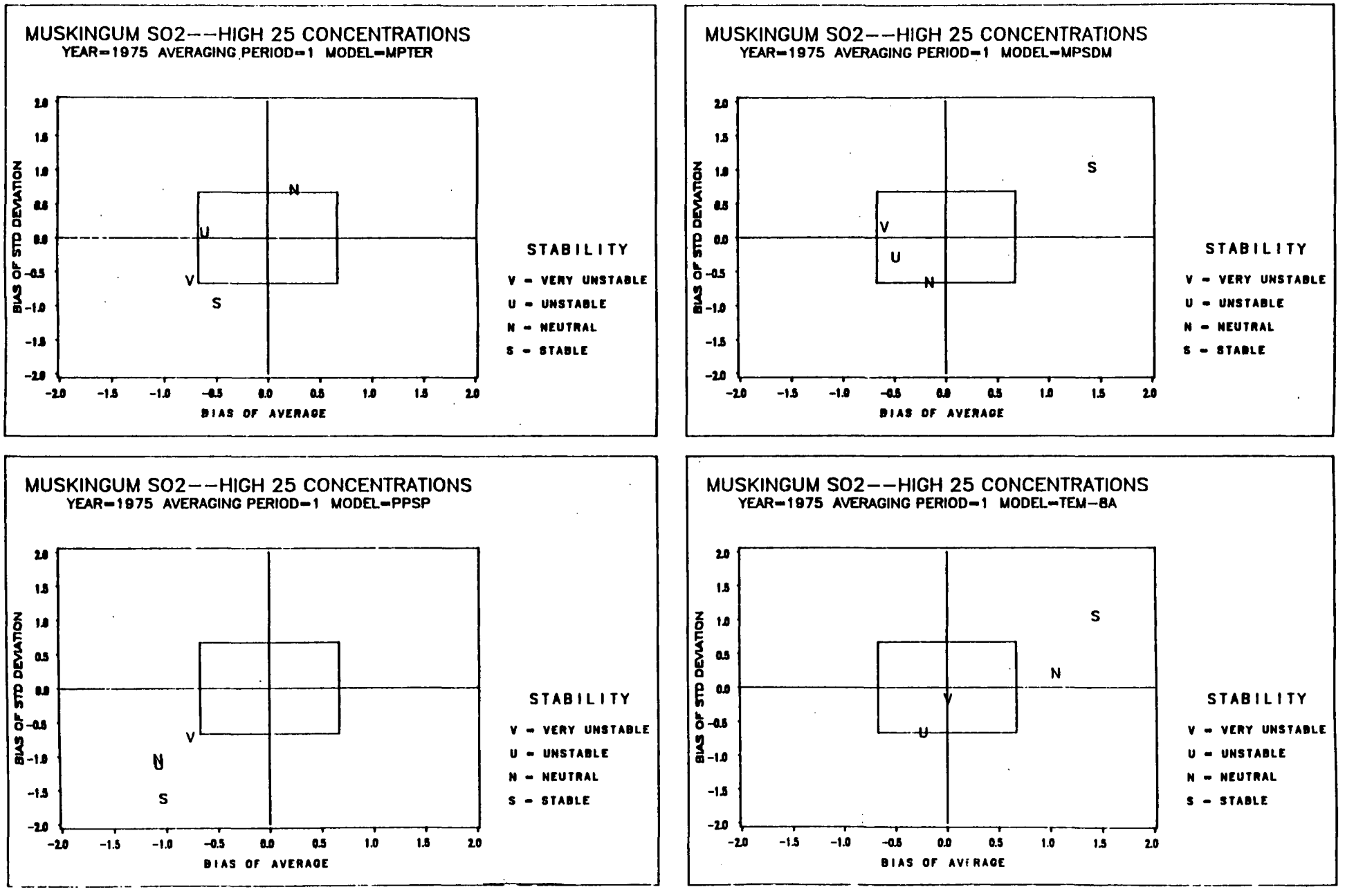
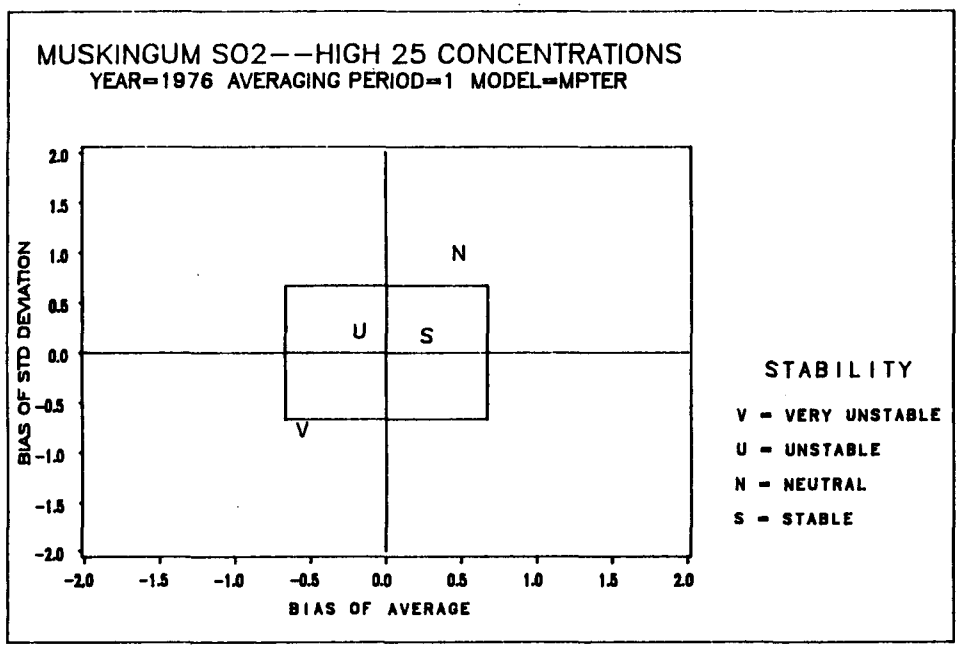
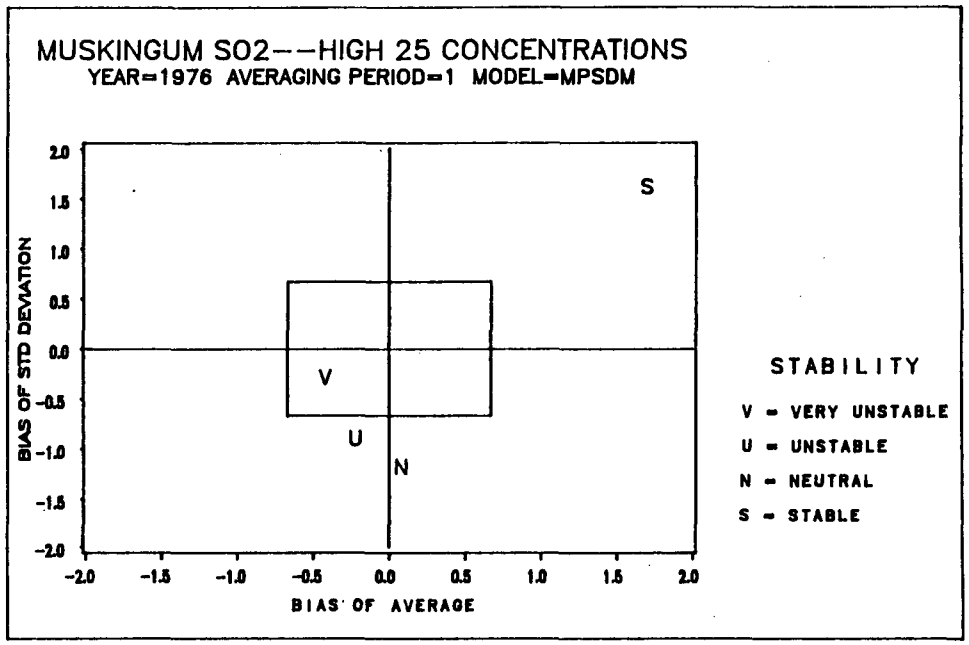
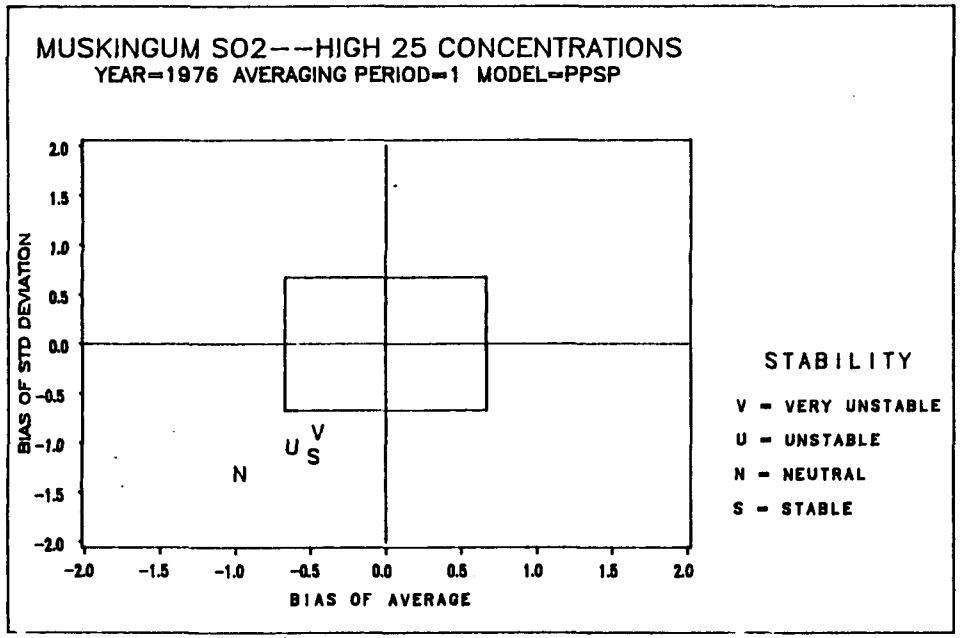


Figure C.6 Fractional Bias By Model Using 25 High Values For Each Stability Class







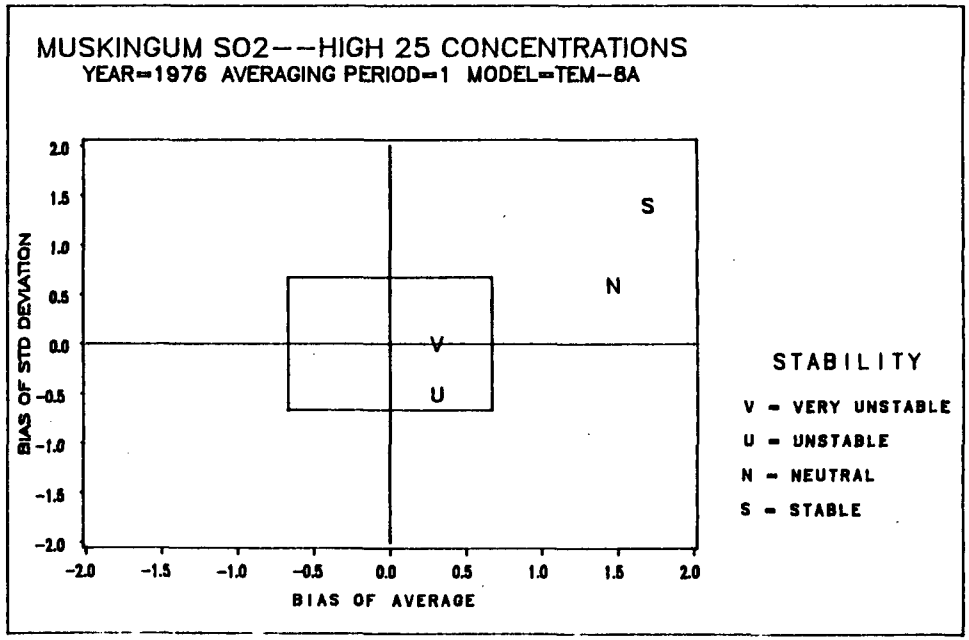


Figure C.6 (Continued)

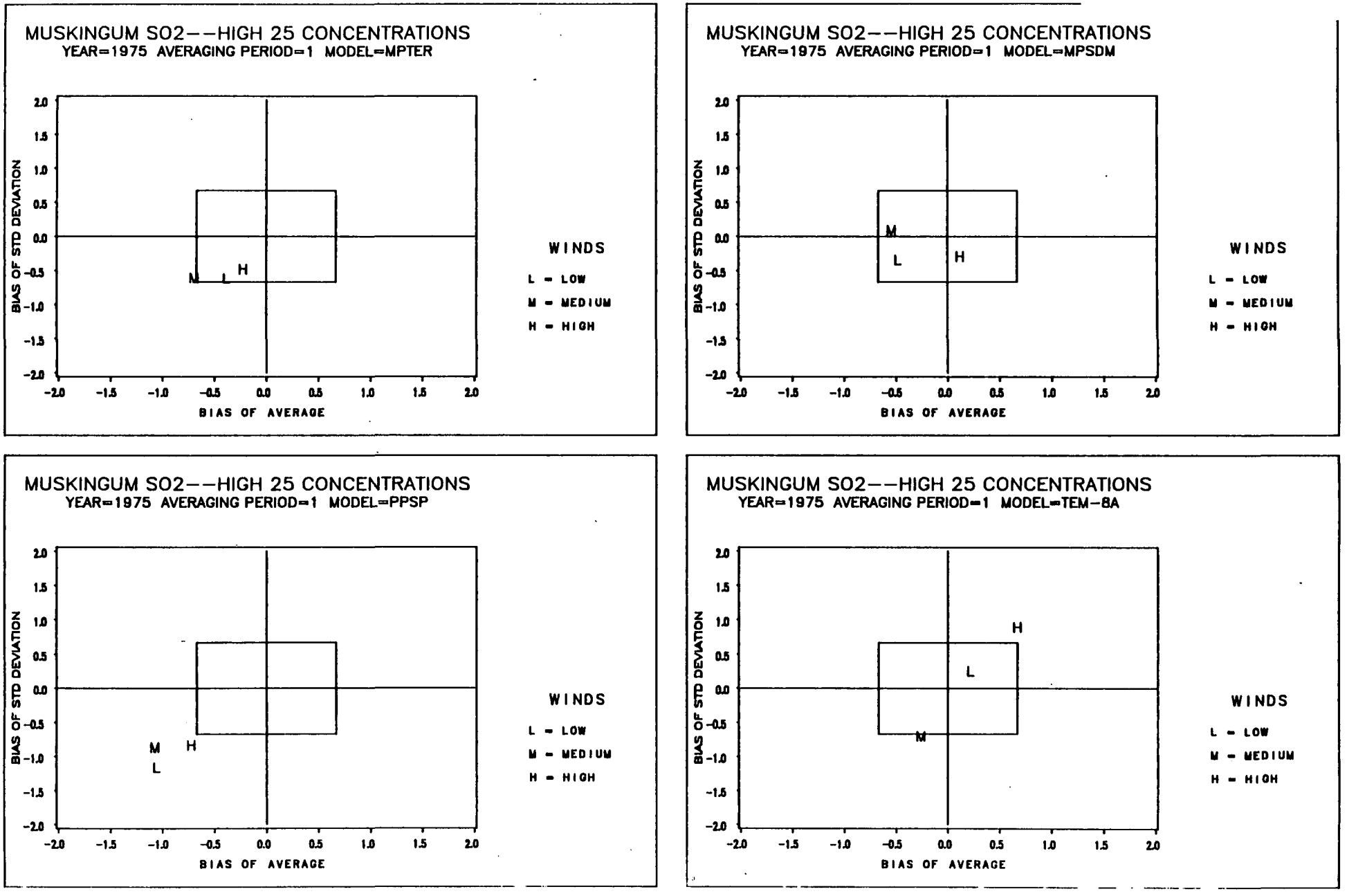
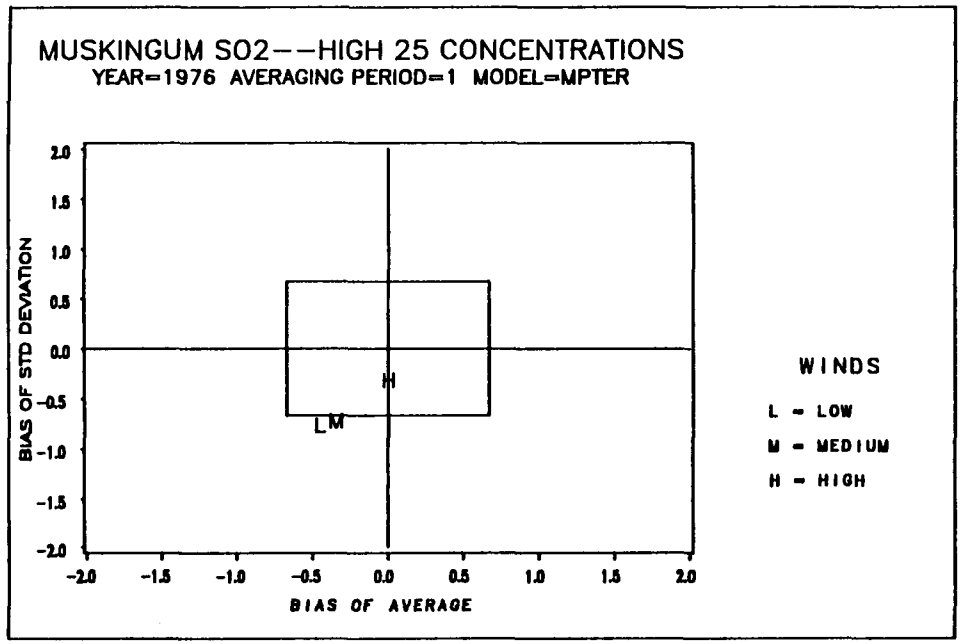
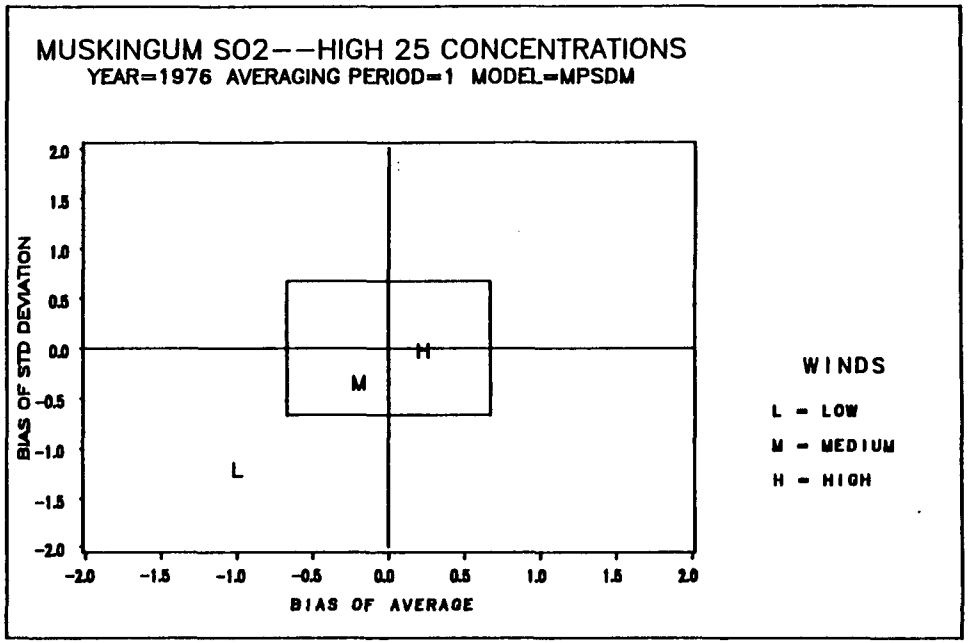
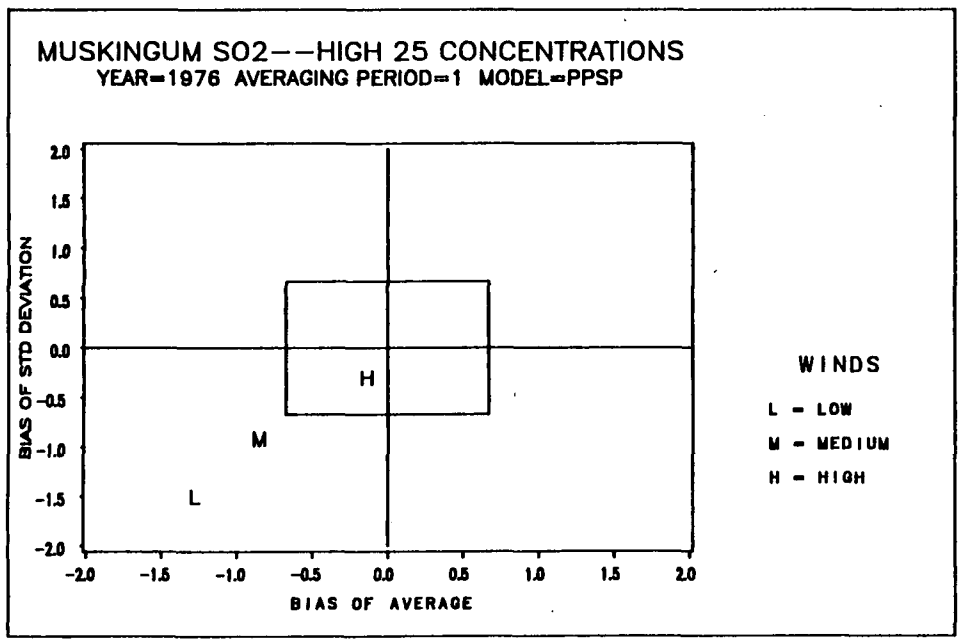


Figure C.7 Fractional Bias By Model Using 25 High Values For Each Wind Speed Class







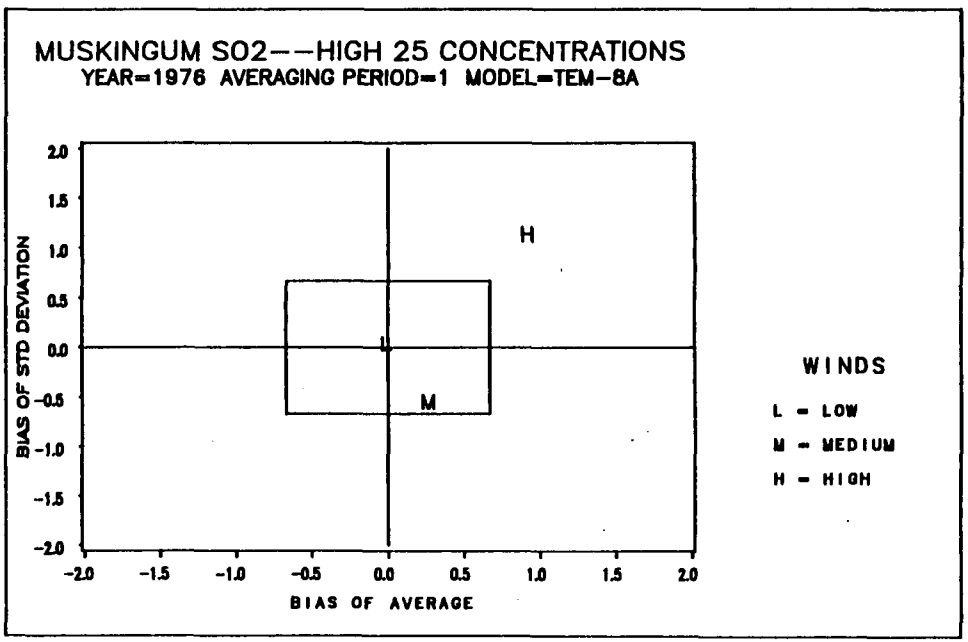


Figure C.7 (Continued)

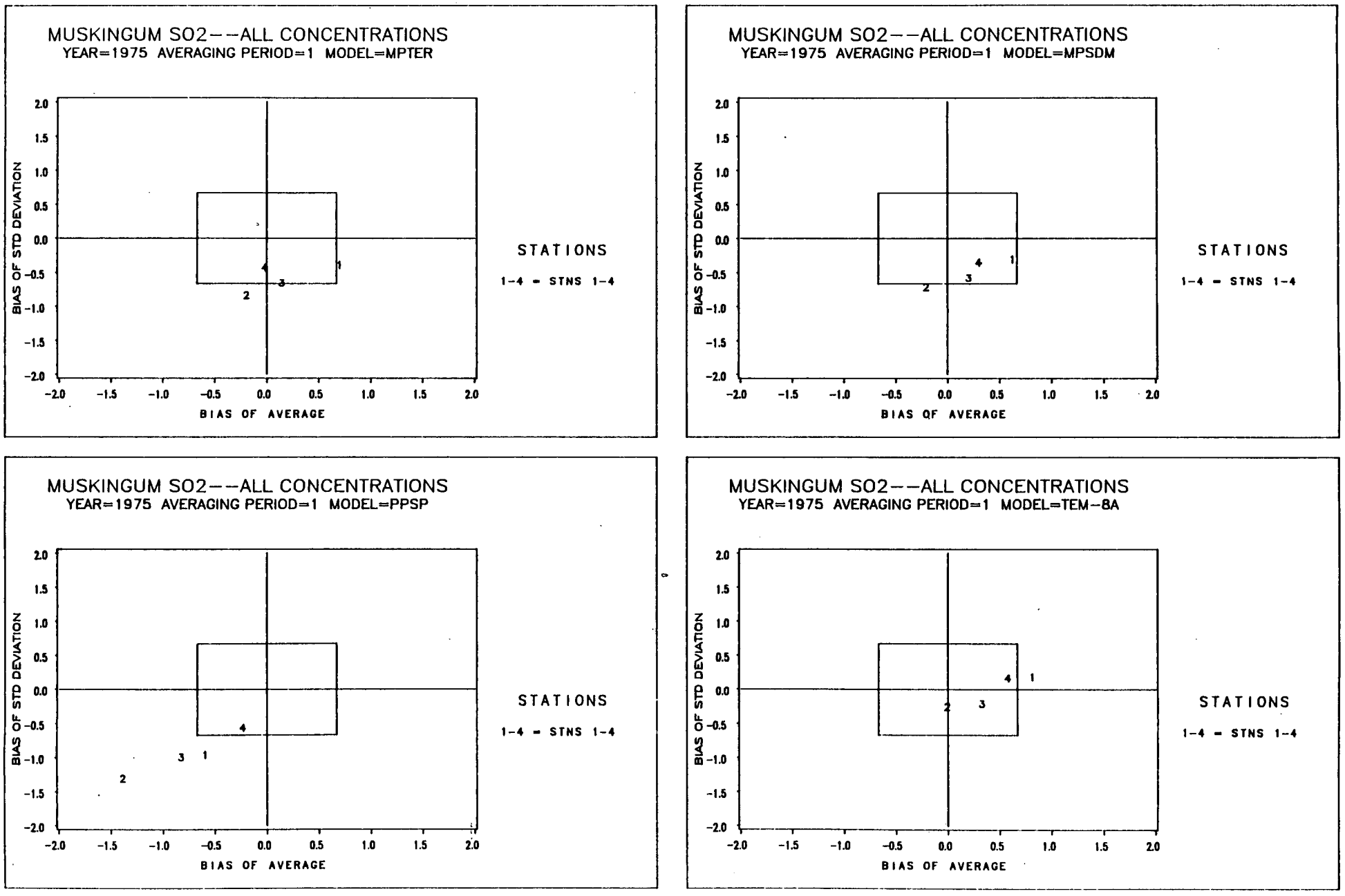
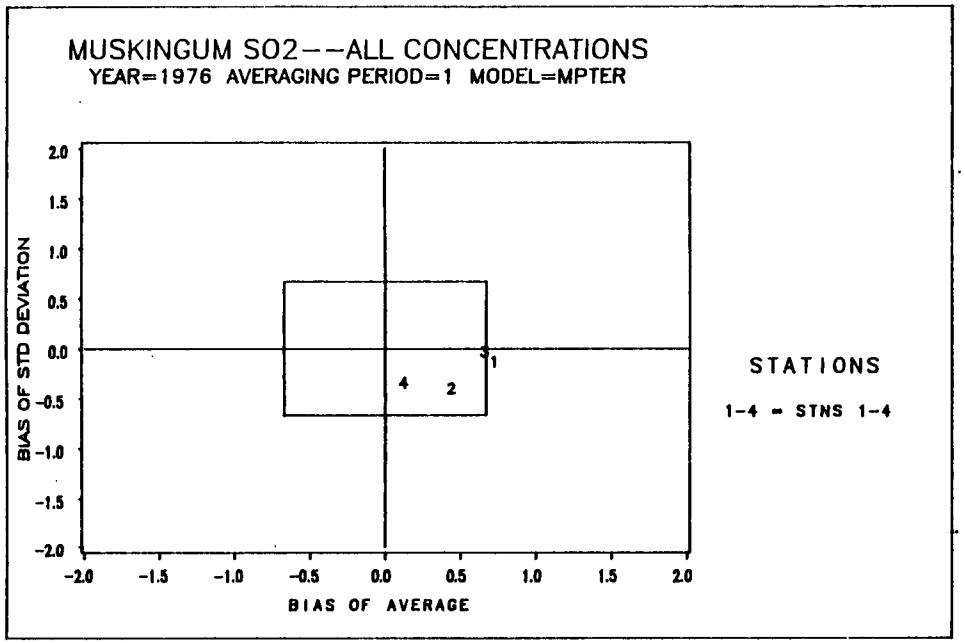
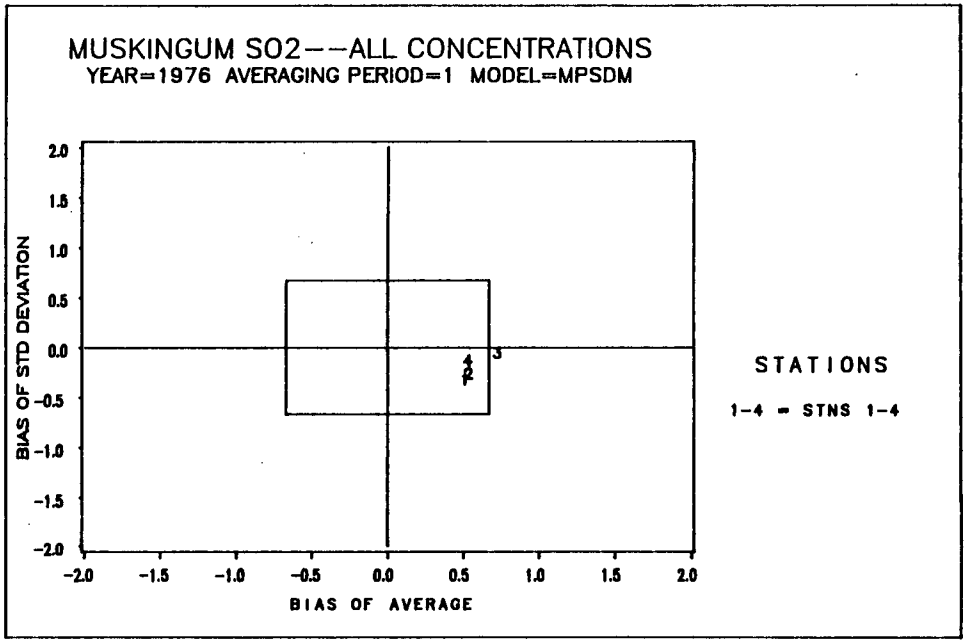
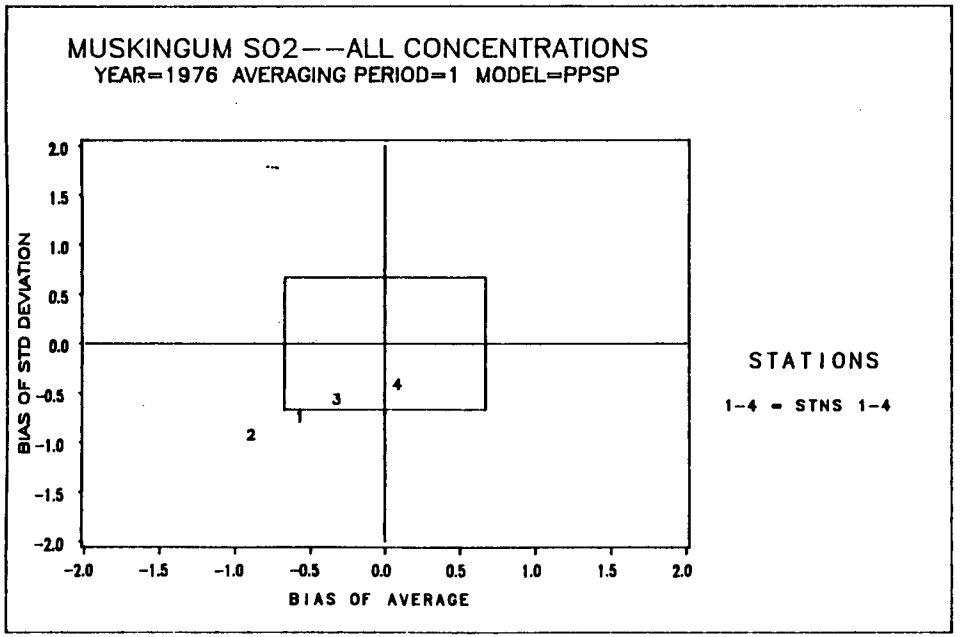


Figure C.8 Fractional Bias By Model Using All Paired Values For Each Station







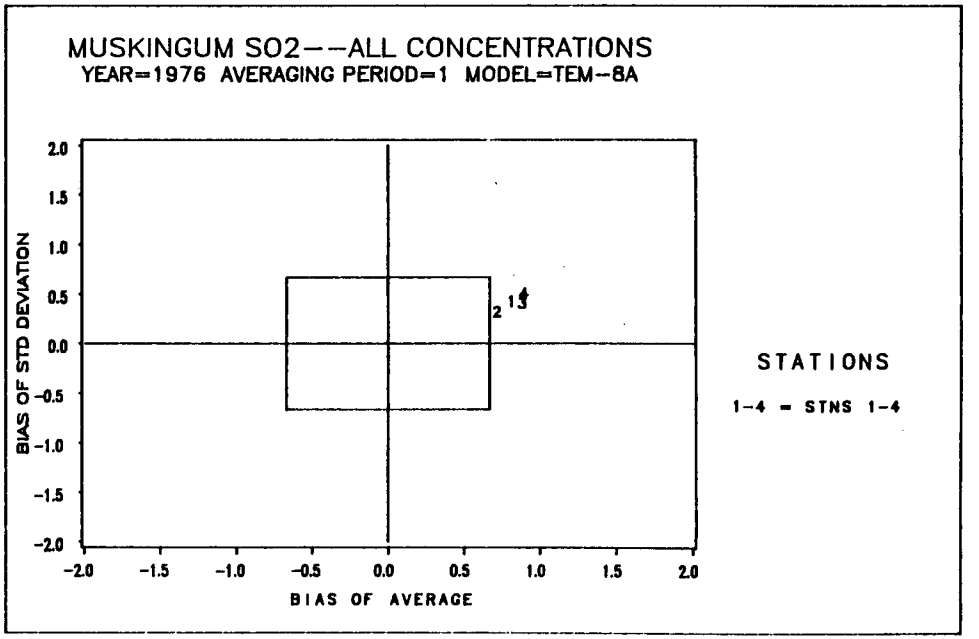


Figure C.8 (Continued)

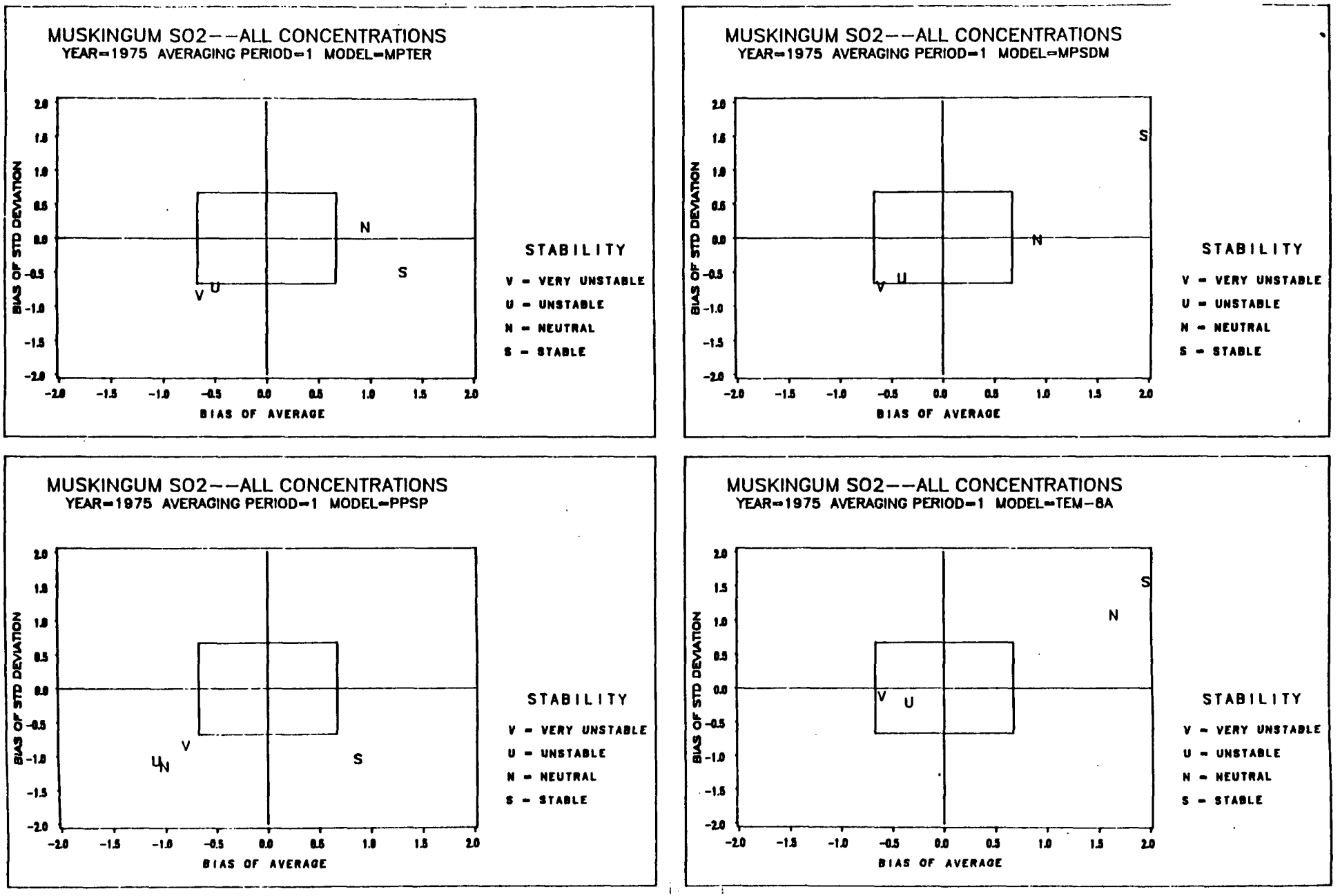
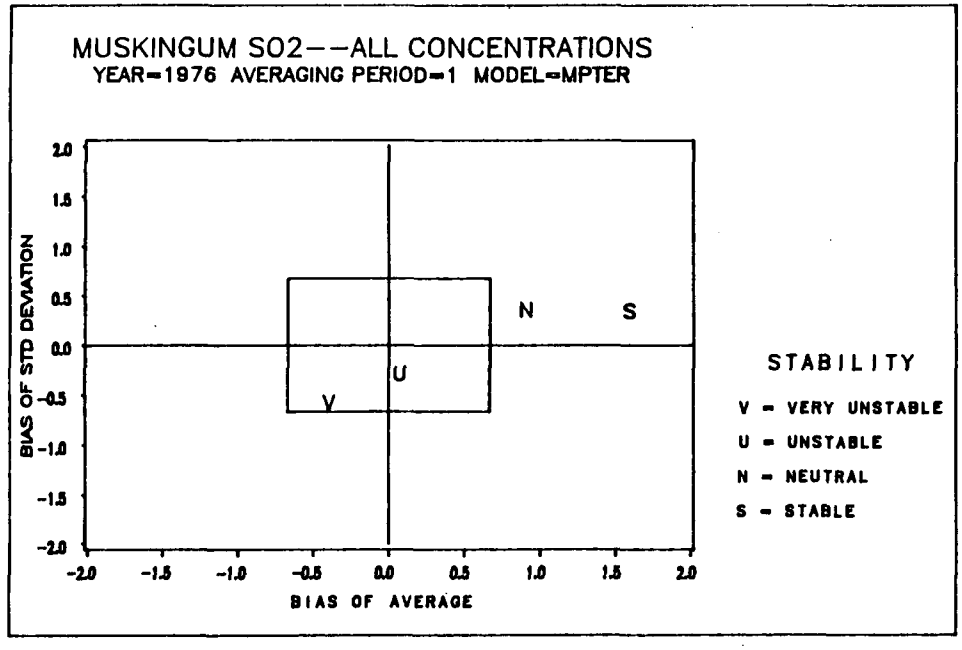
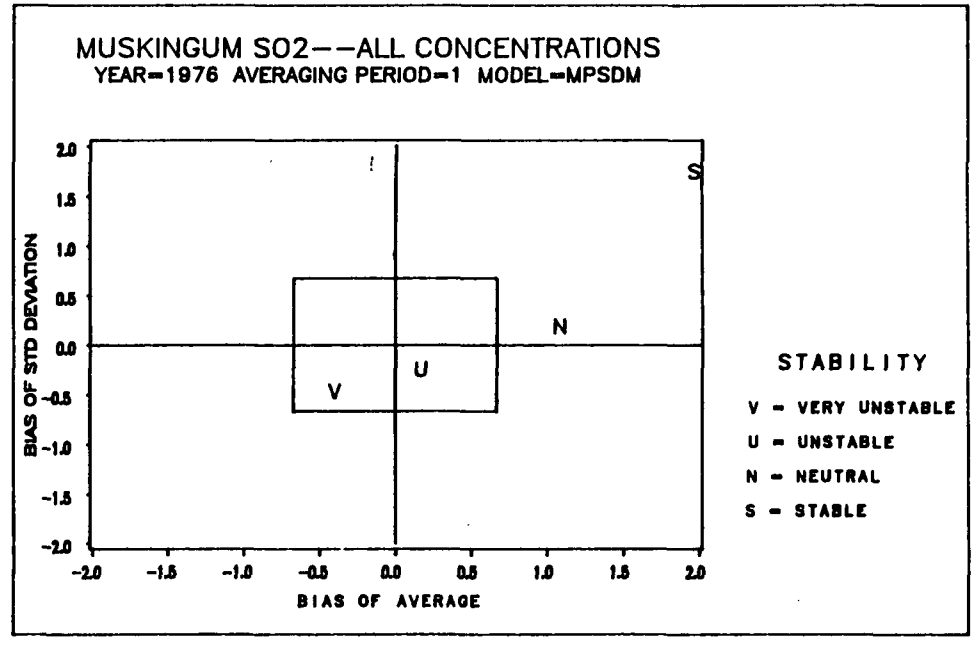
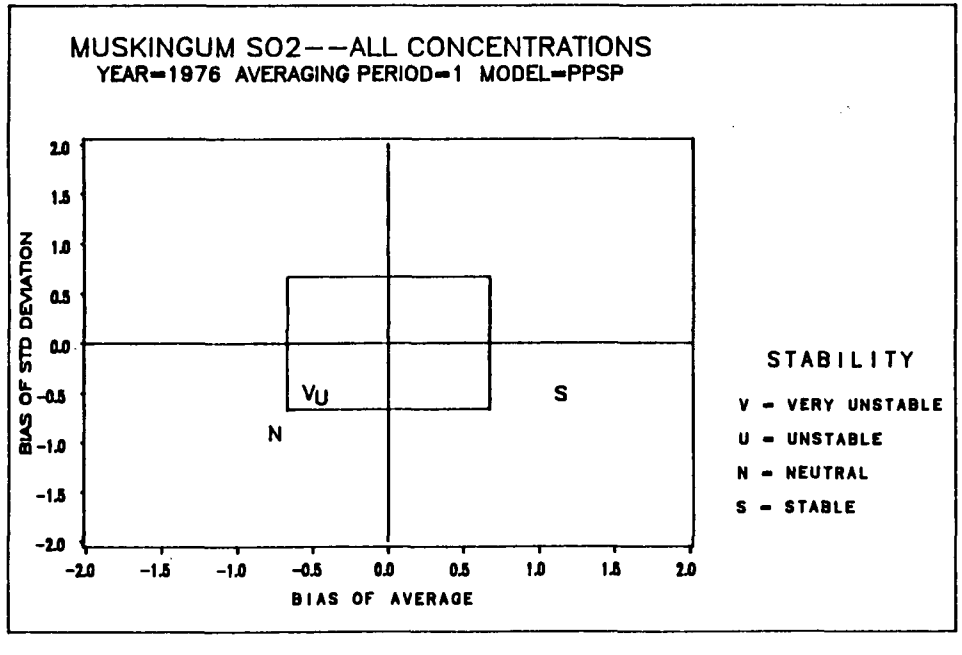


Figure C.9 Fractional Bias By Model Using All Paired Values For Each Stability Class







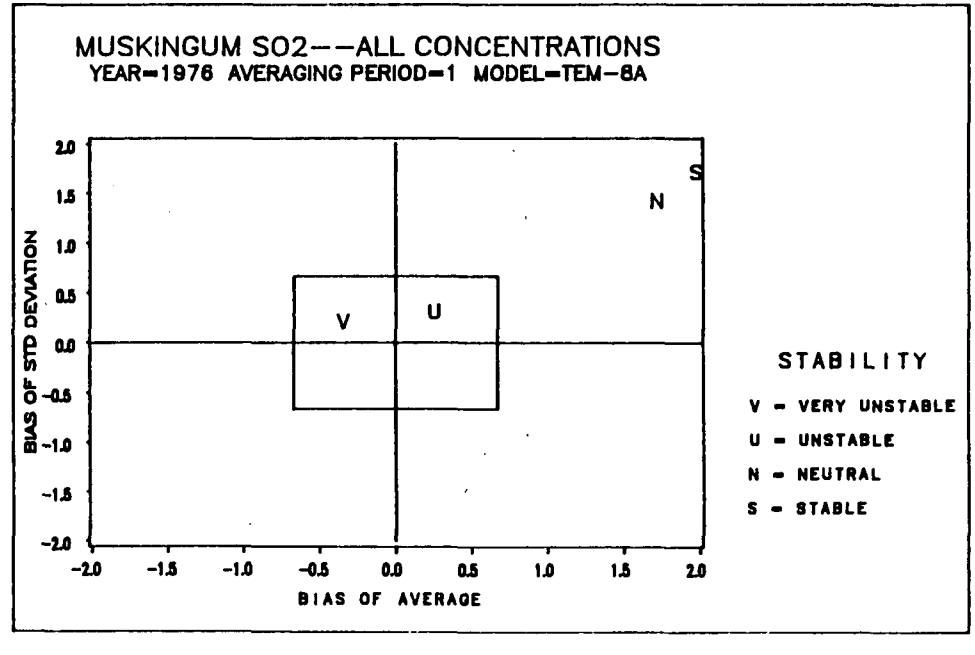


Figure C.9 (Continued)

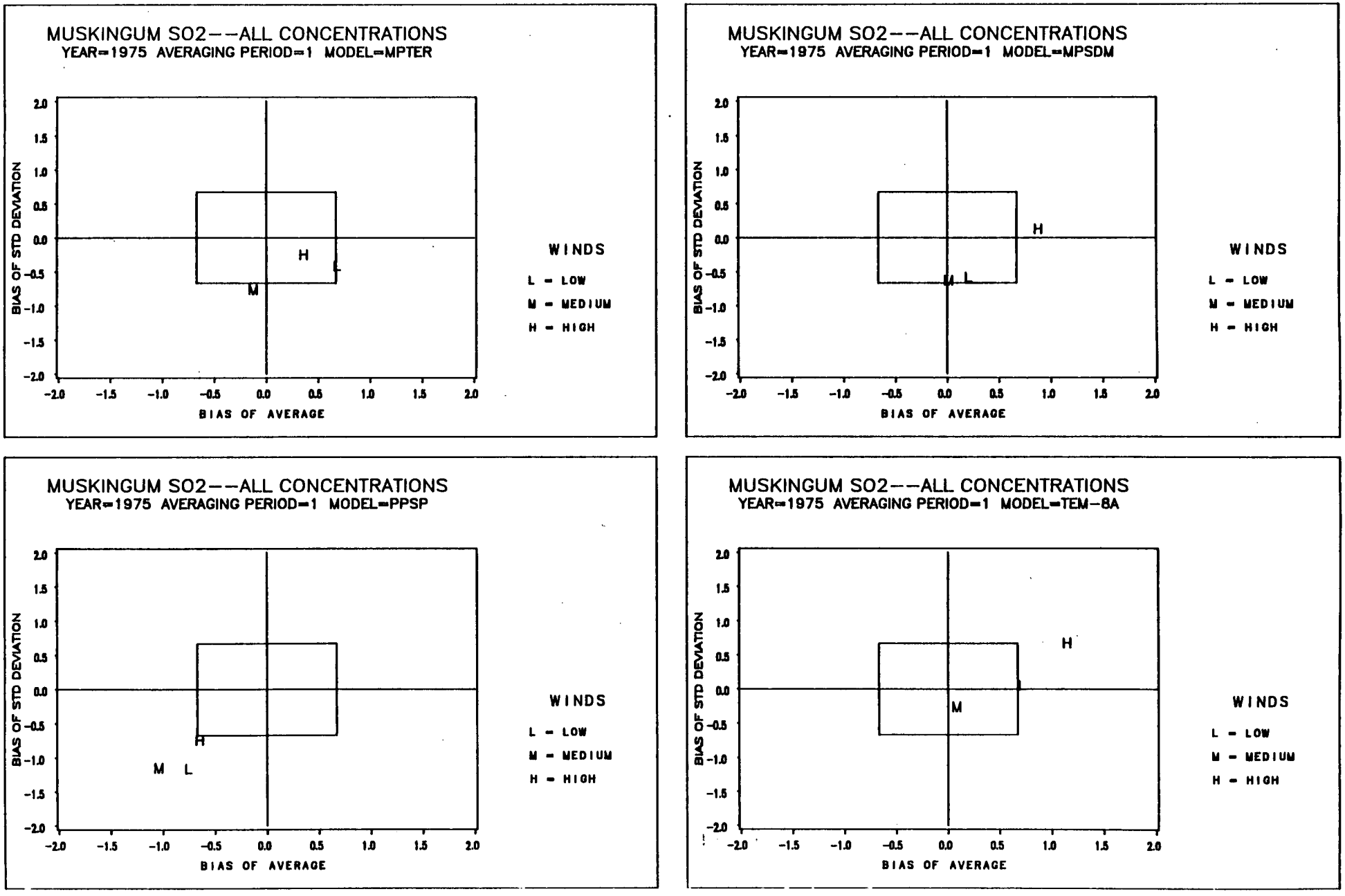
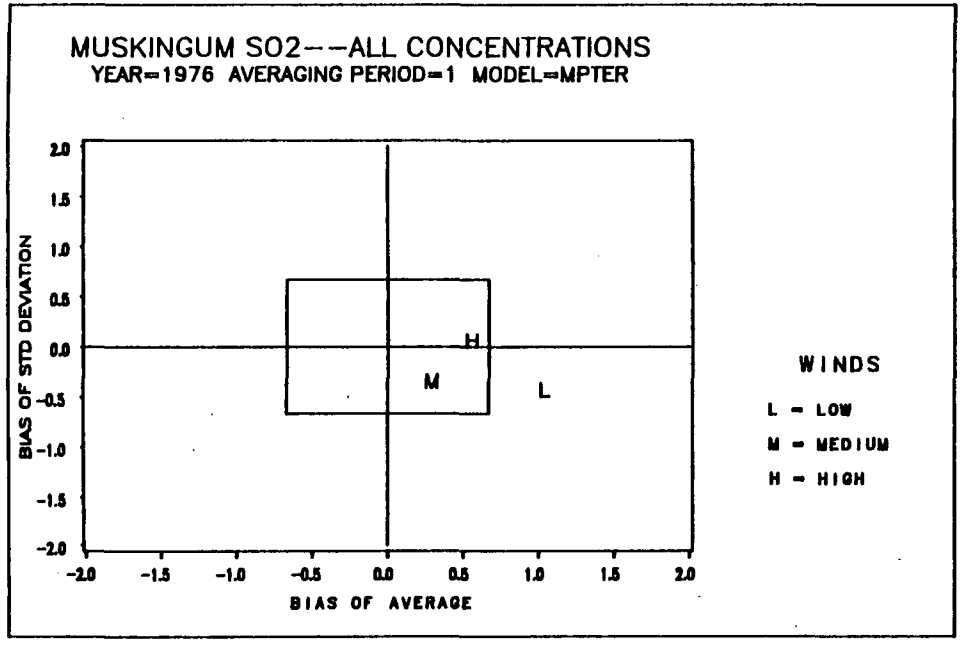
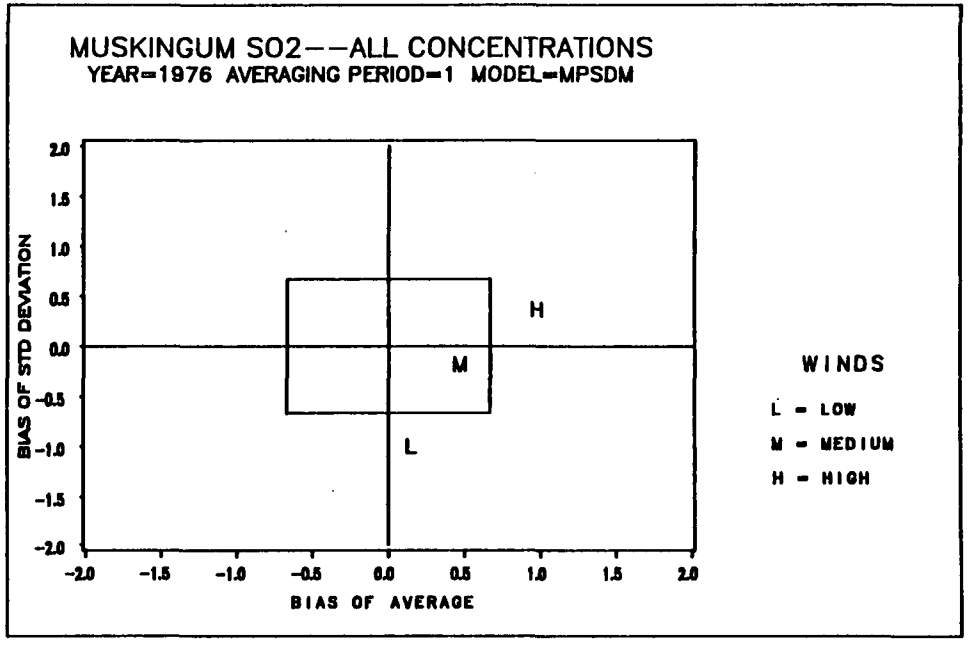
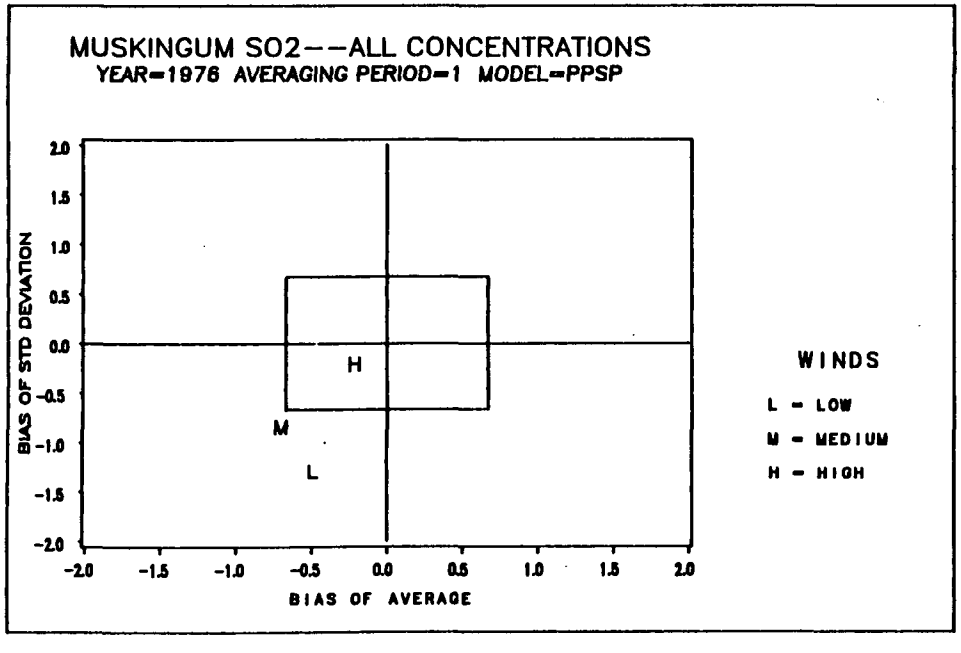


Figure C.10 Fractional Bias By Model Using All Paired Values For Each Wind Speed Class







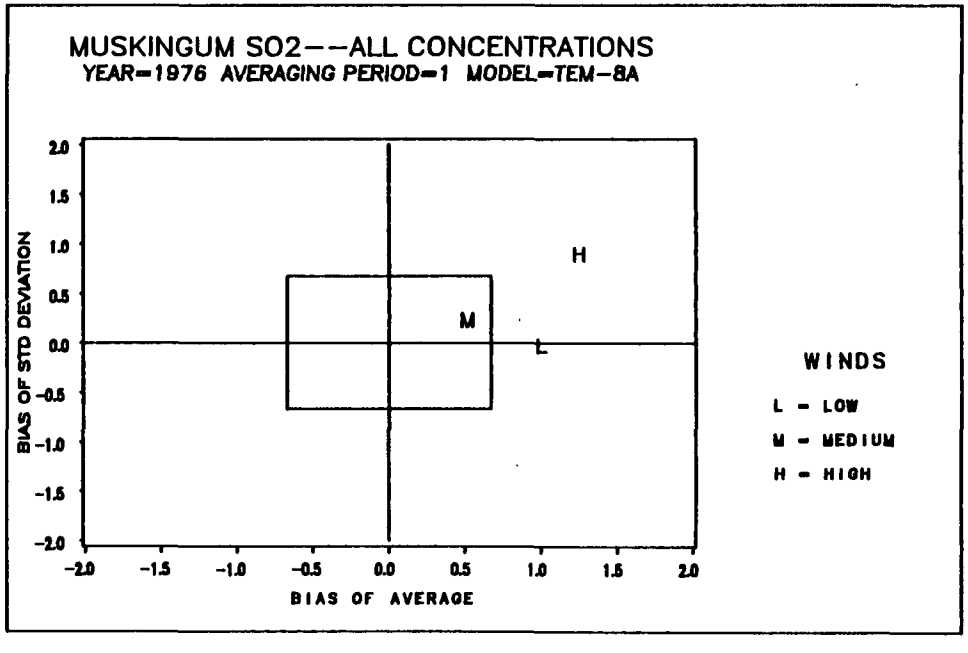


Figure C.10 (Continued)

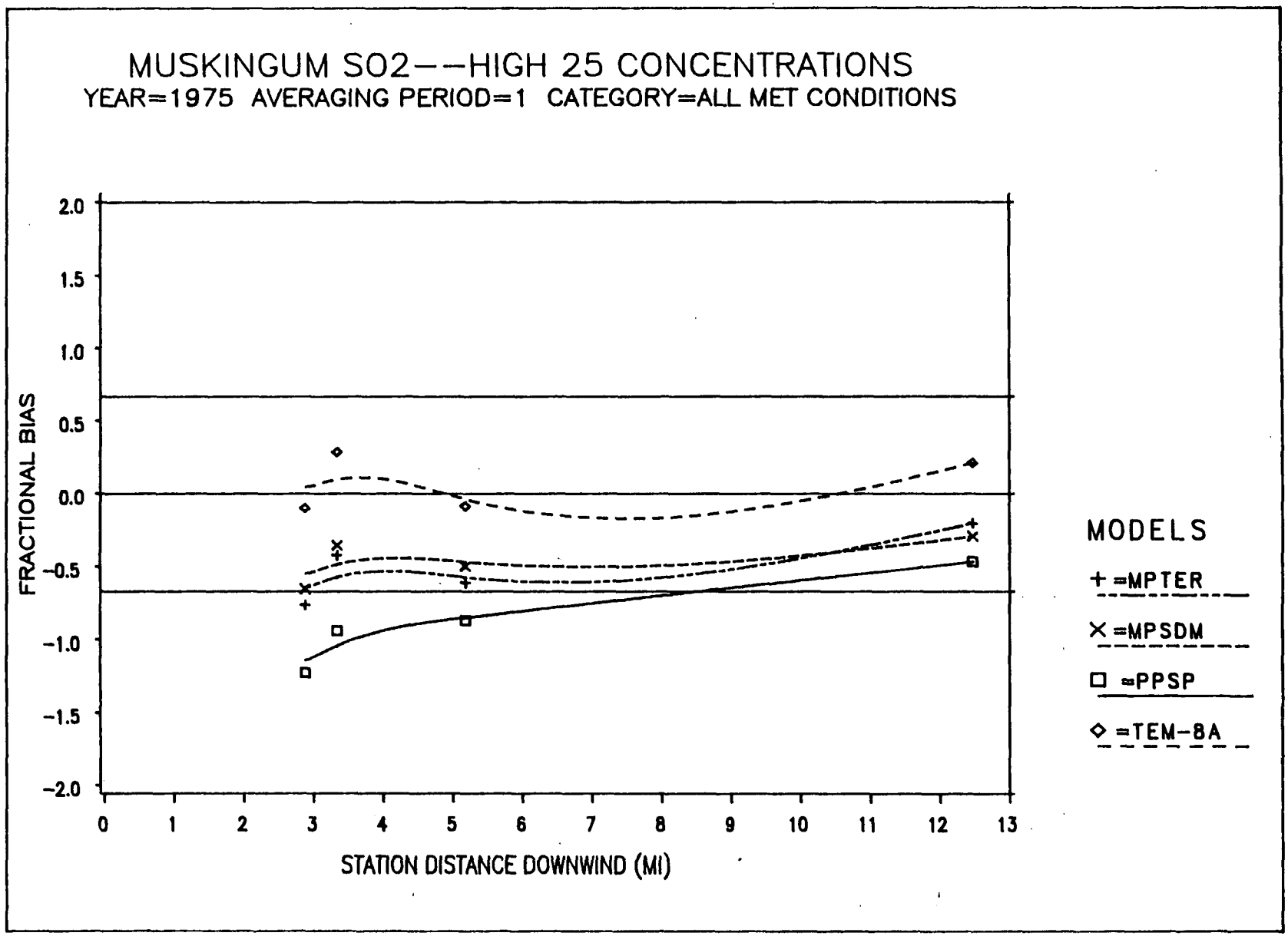


Figure C.11 Fractional Bias Of The Average Vs Station Distance Using High 25 Values

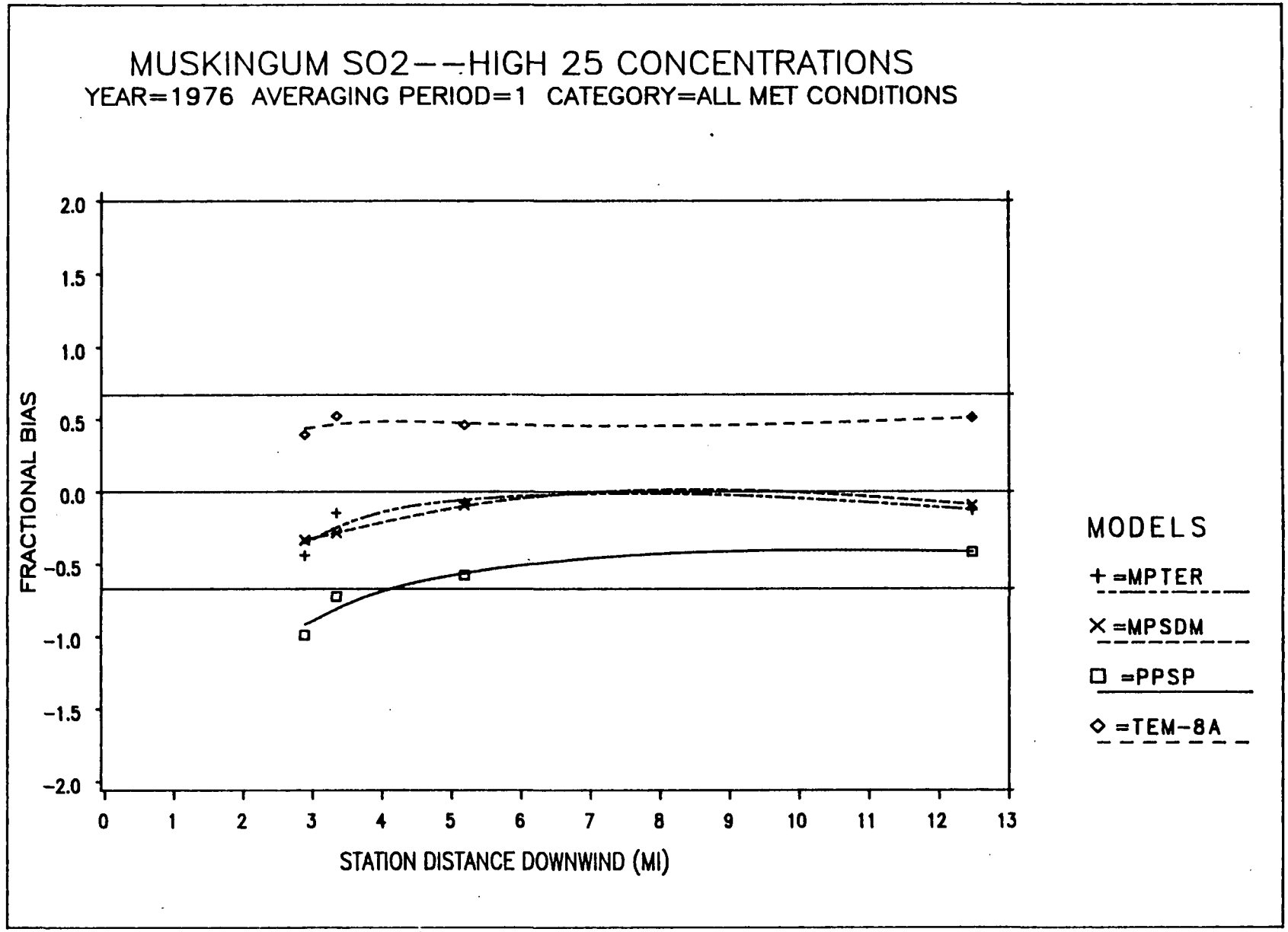


Figure C.11 (Continued)

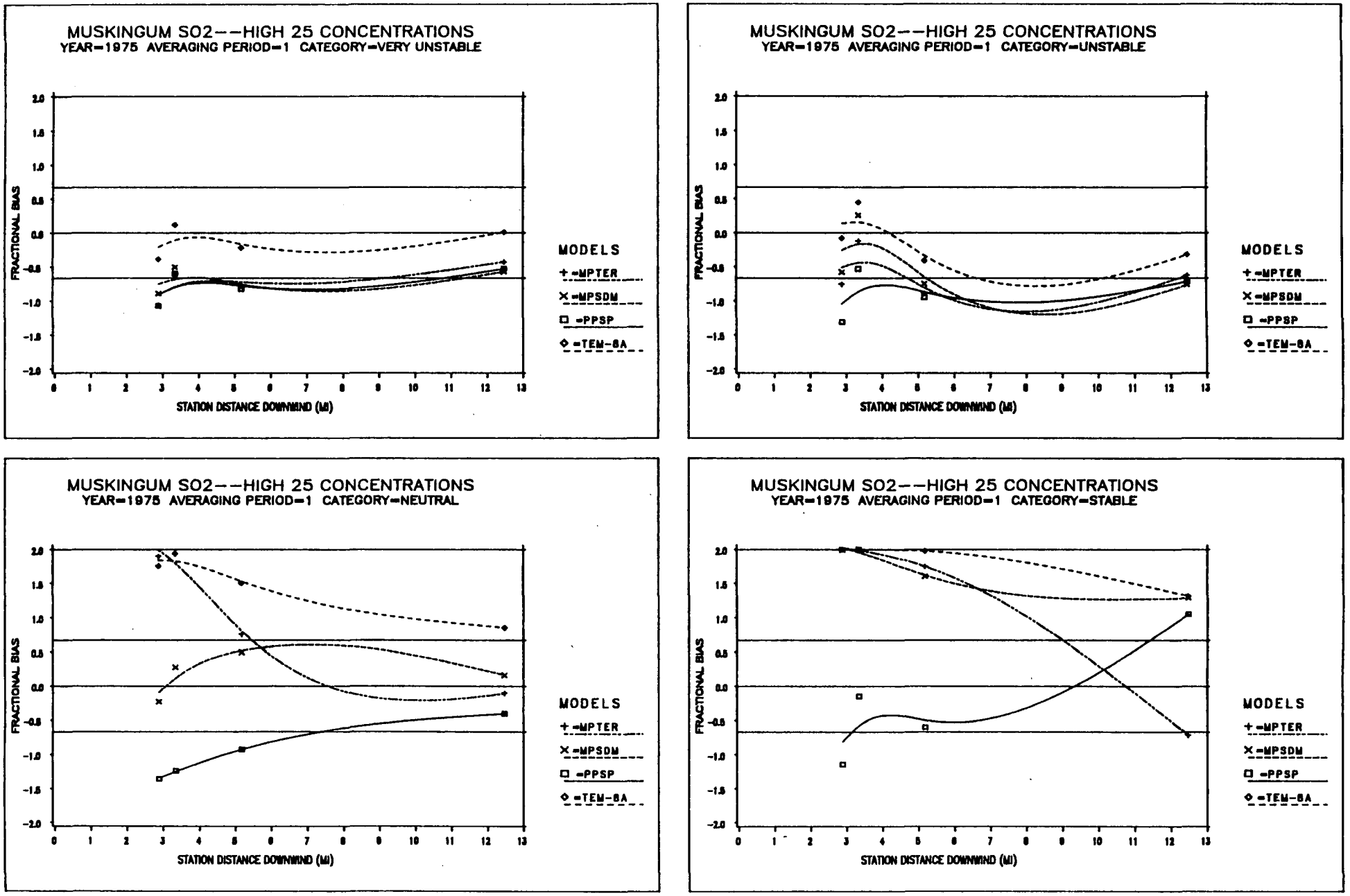
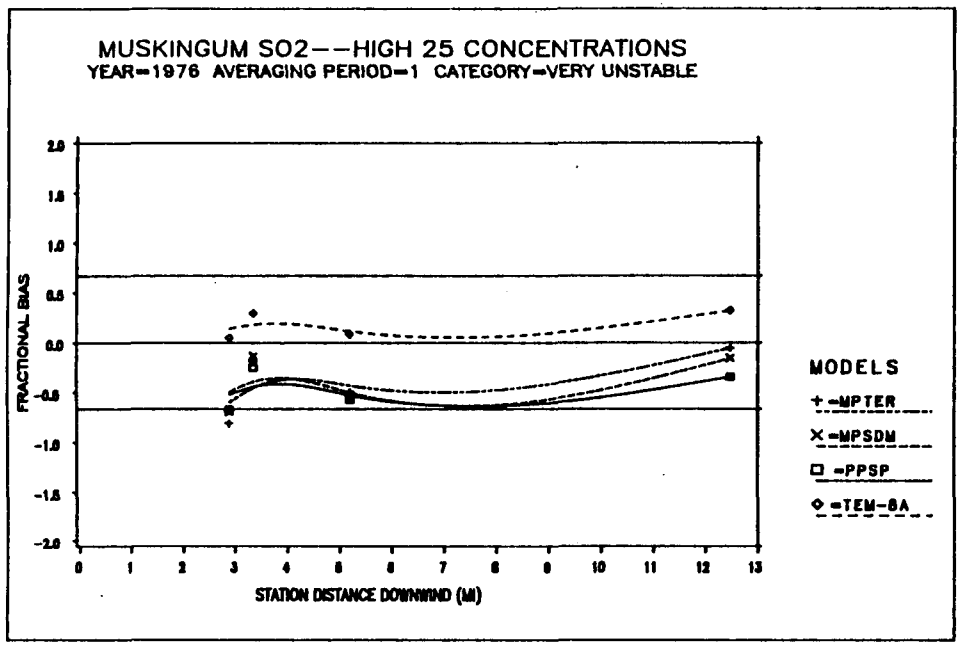
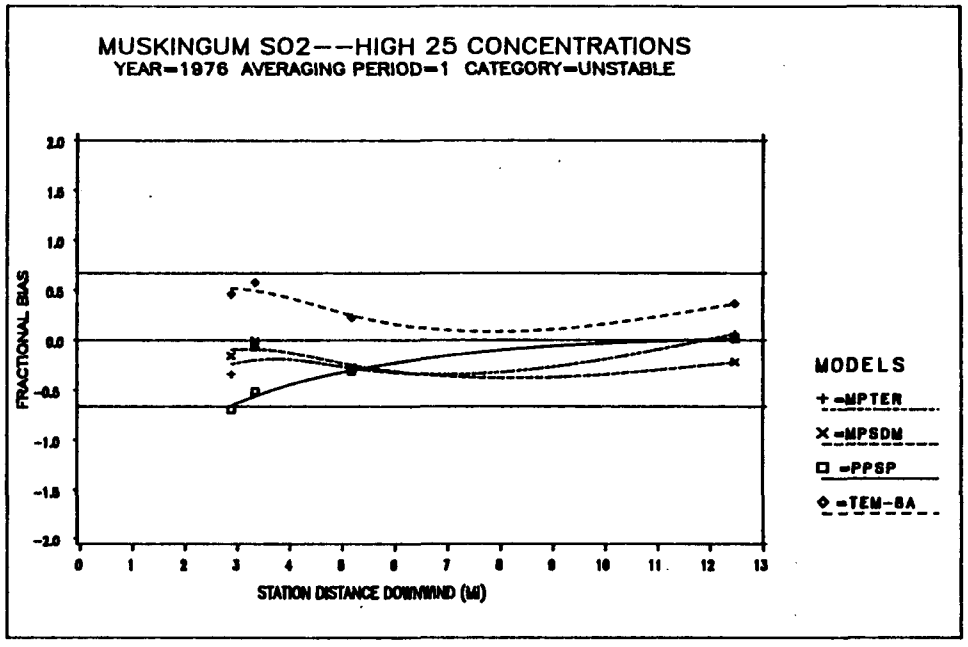
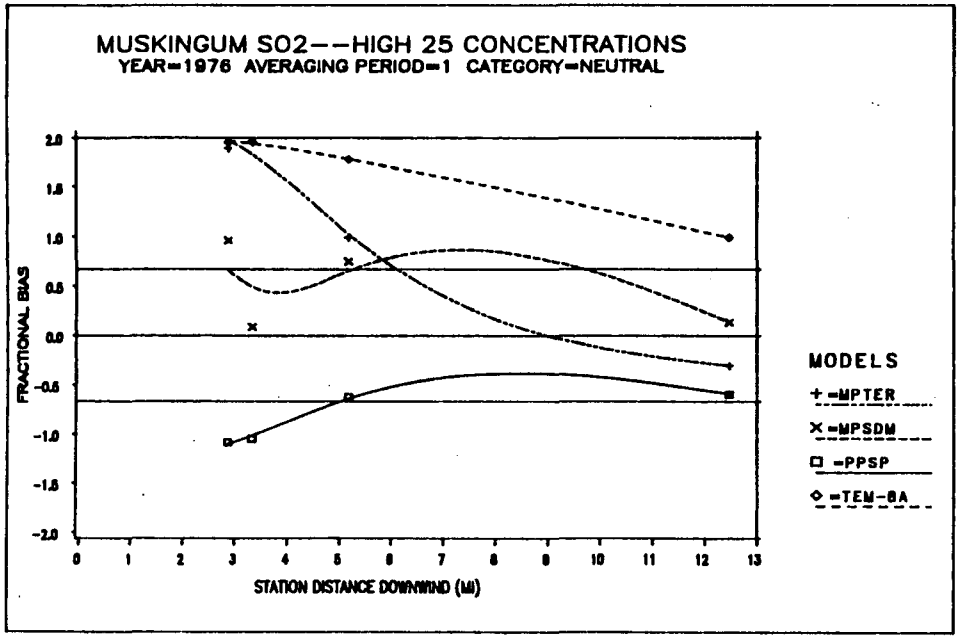


Figure C.12 Fractional Bias Of The Average Vs Station Distance By Stability Class Using High 25 Values







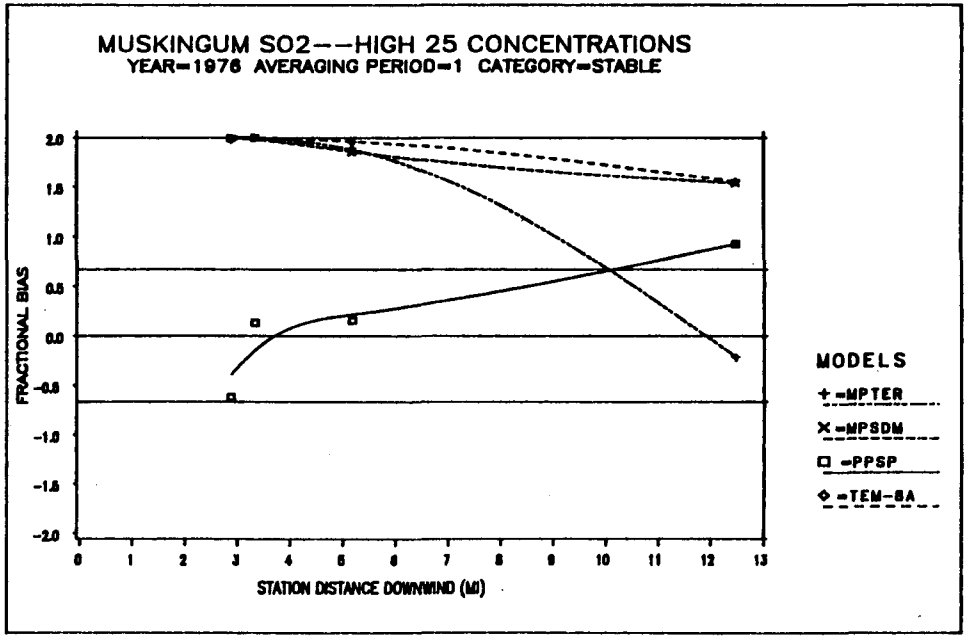
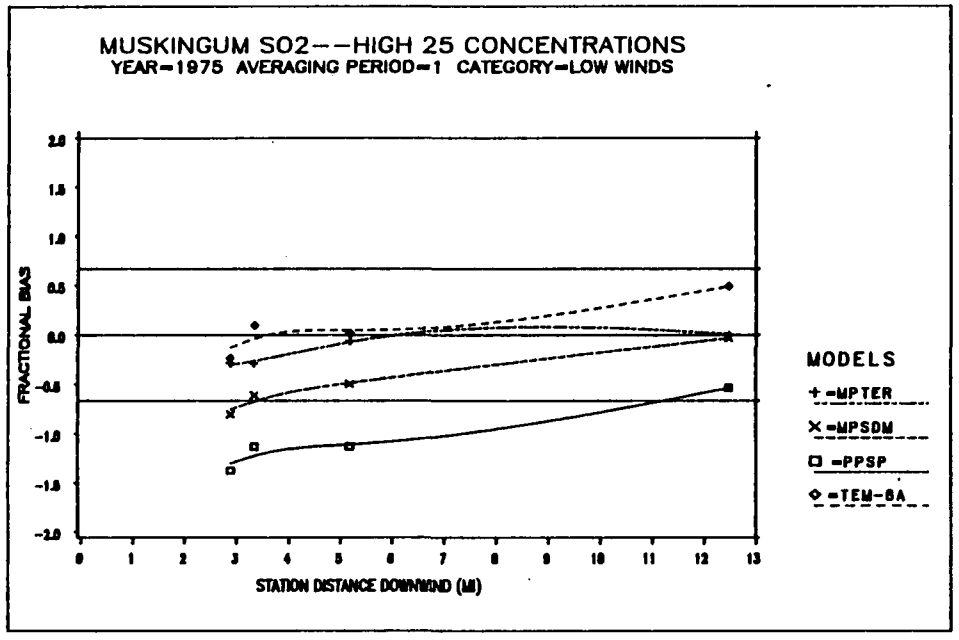
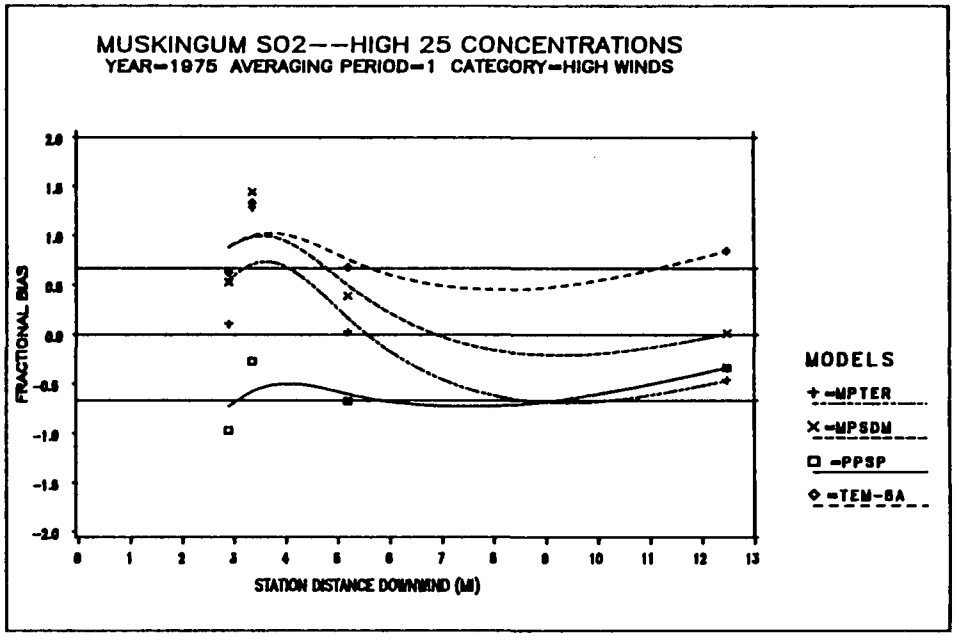


Figure C.12 (Continued)





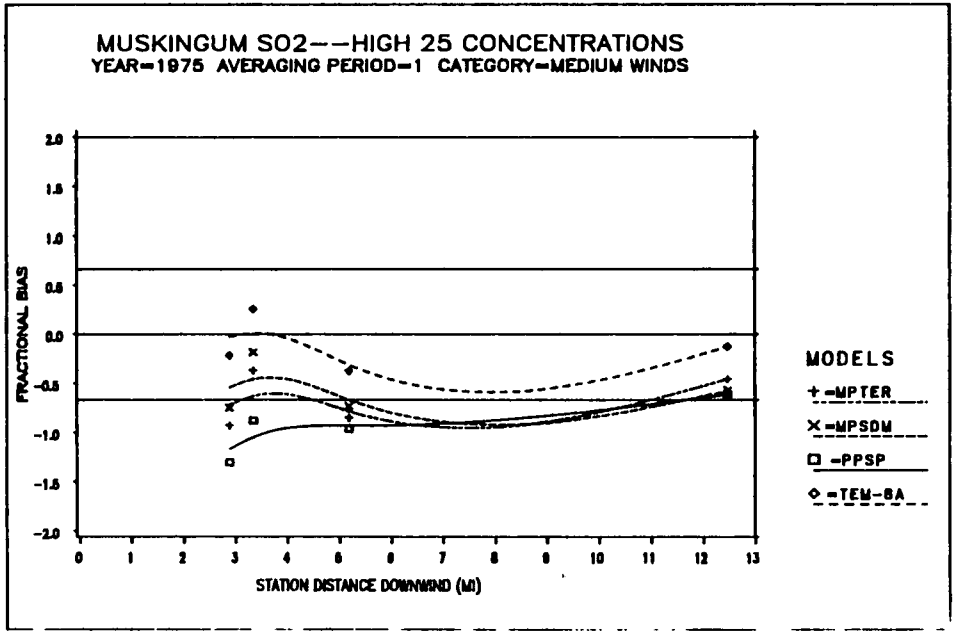
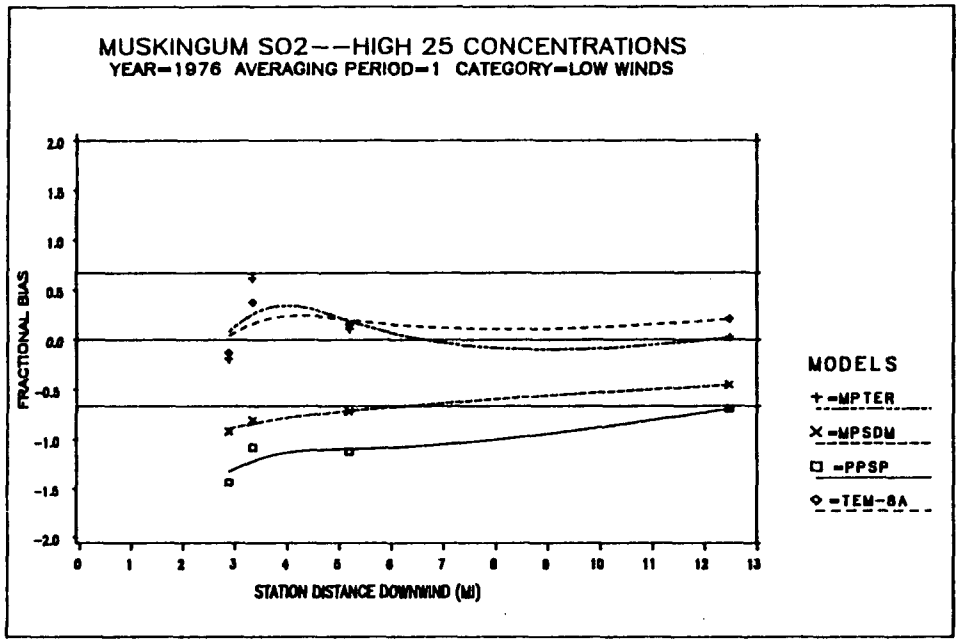
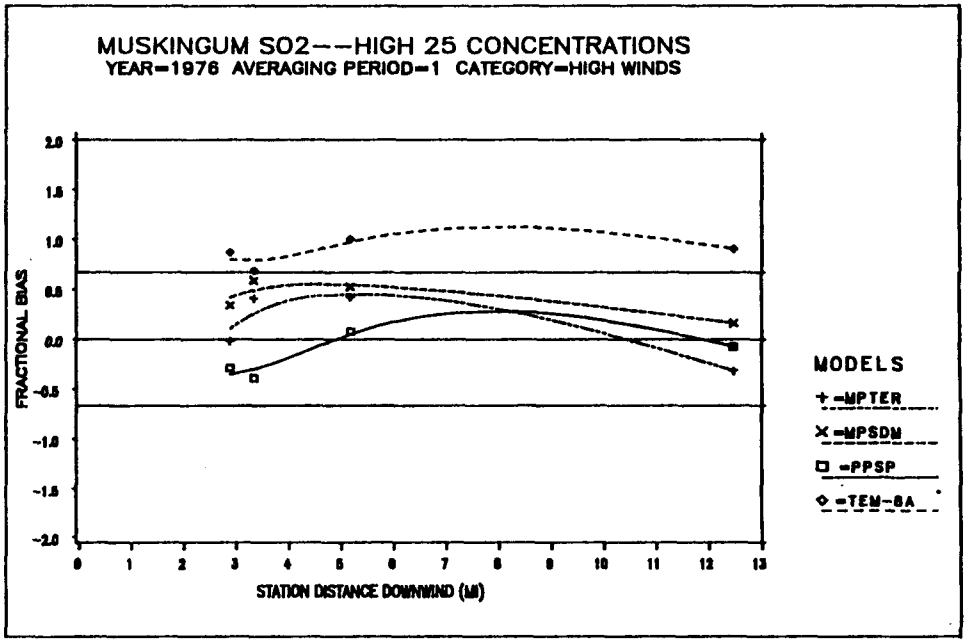


Figure C.13 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using High 25 Values





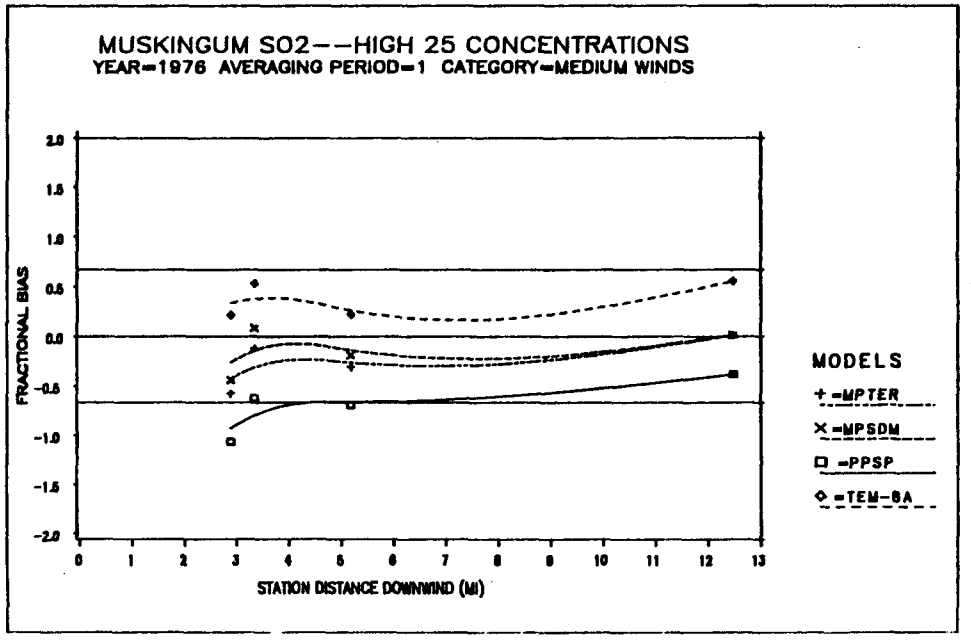


Figure C.13 (Continued)

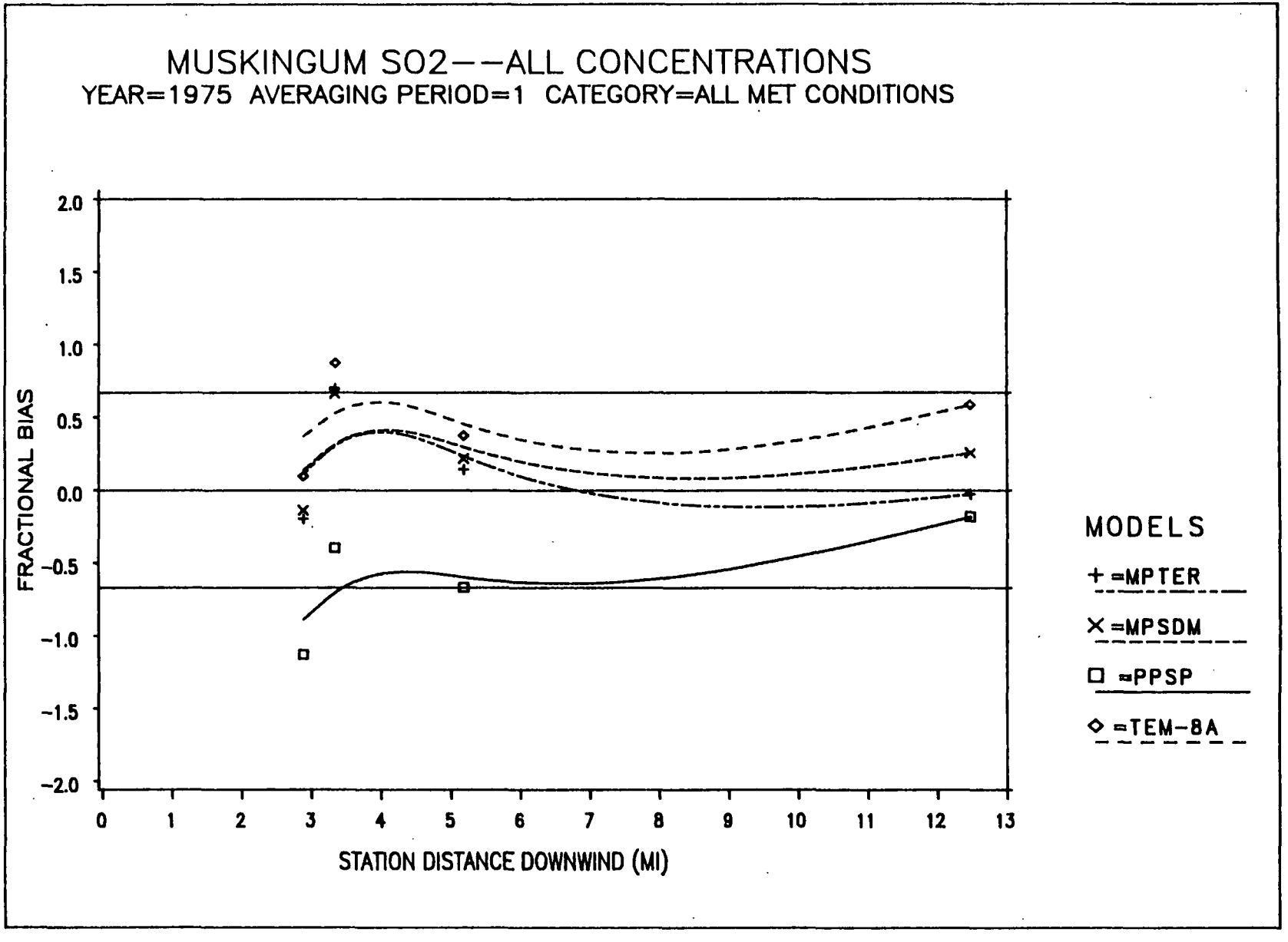


Figure C.14 Fractional Bias Of The Average Vs Station Distance Using All Paired Values

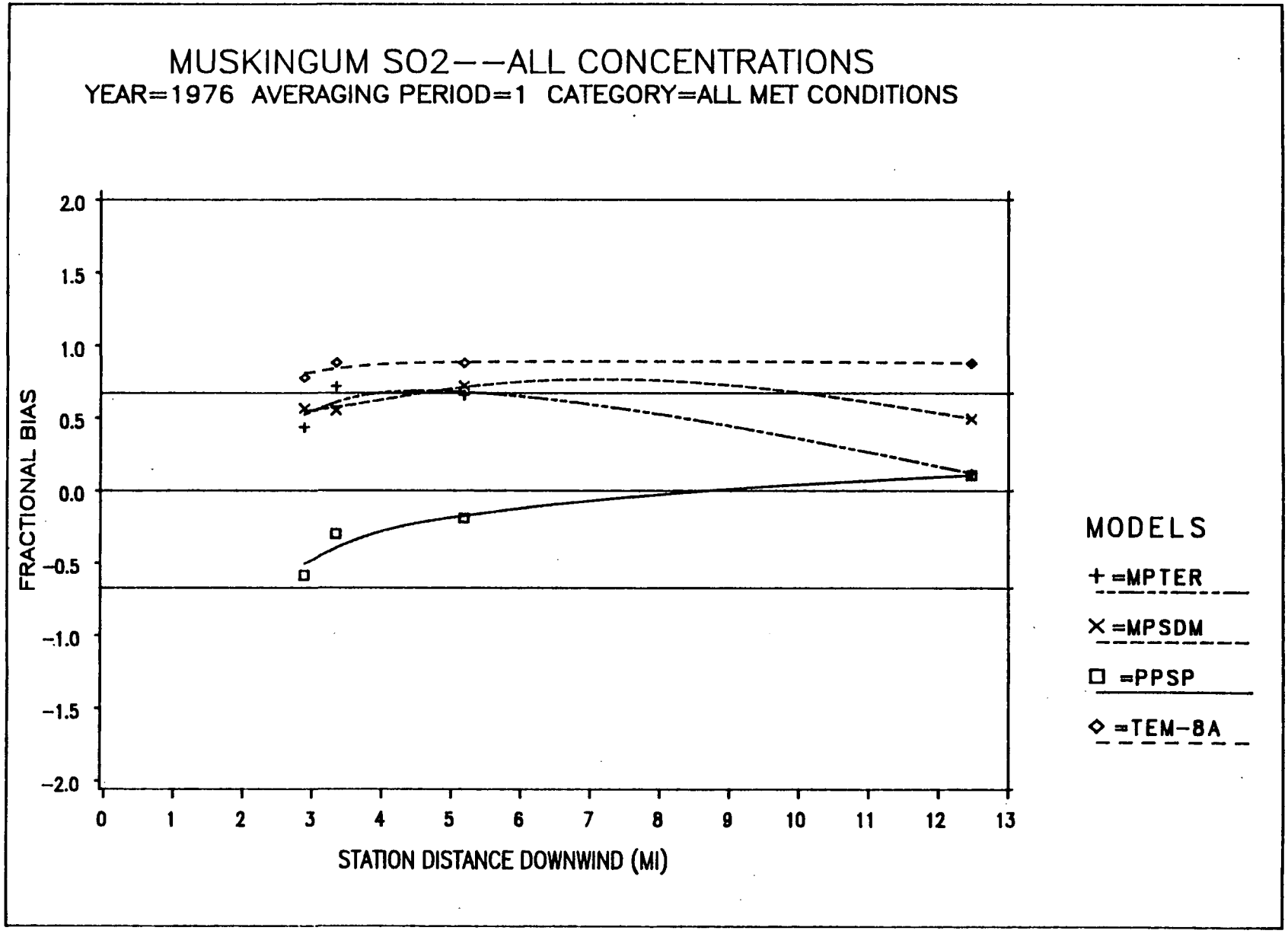


Figure C.14 (Continued)

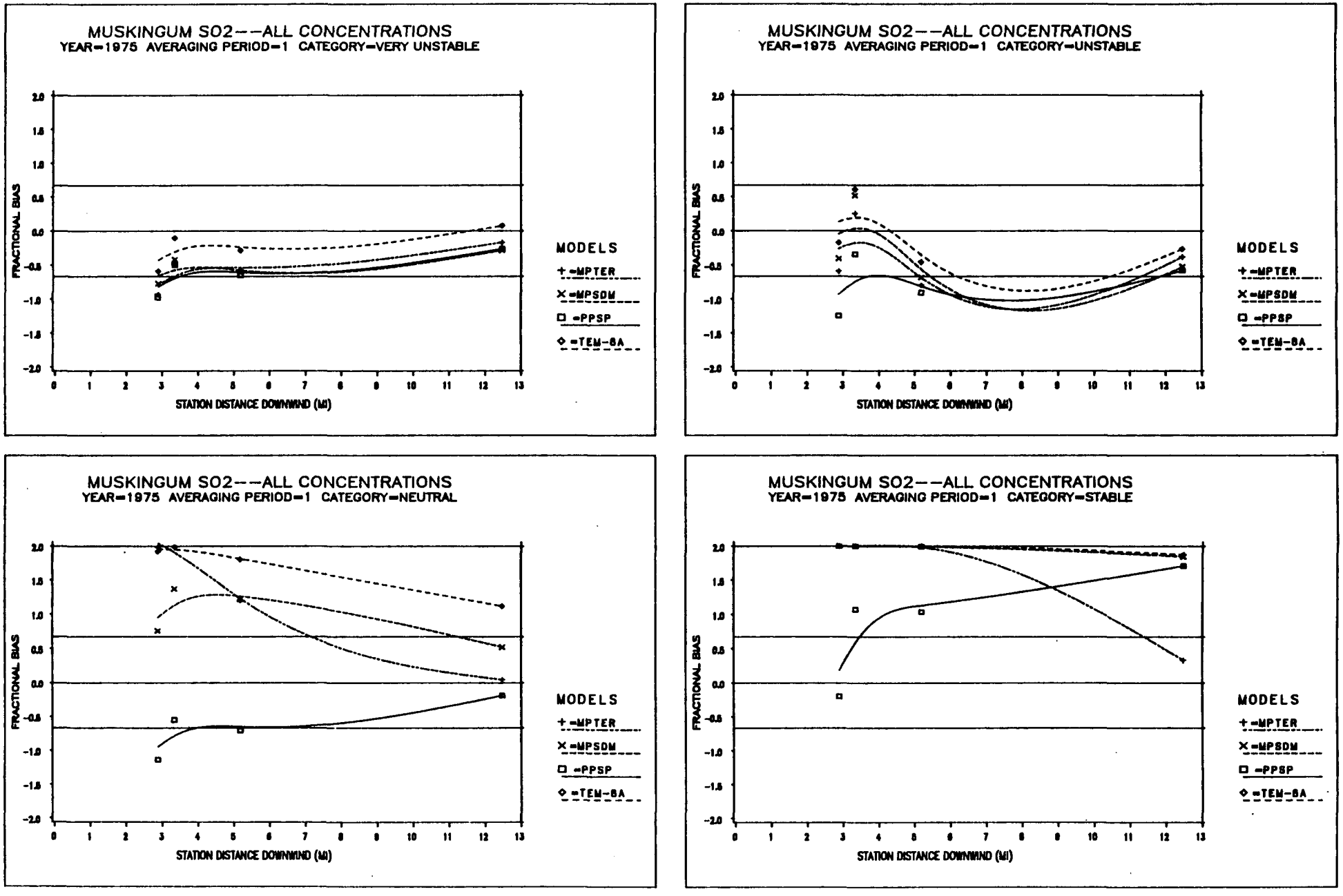
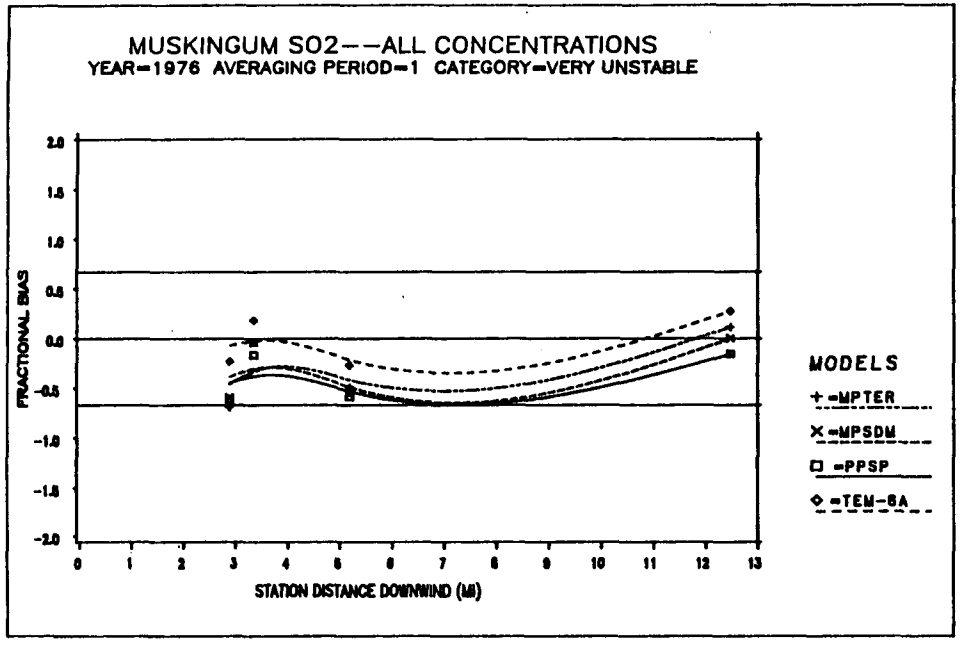
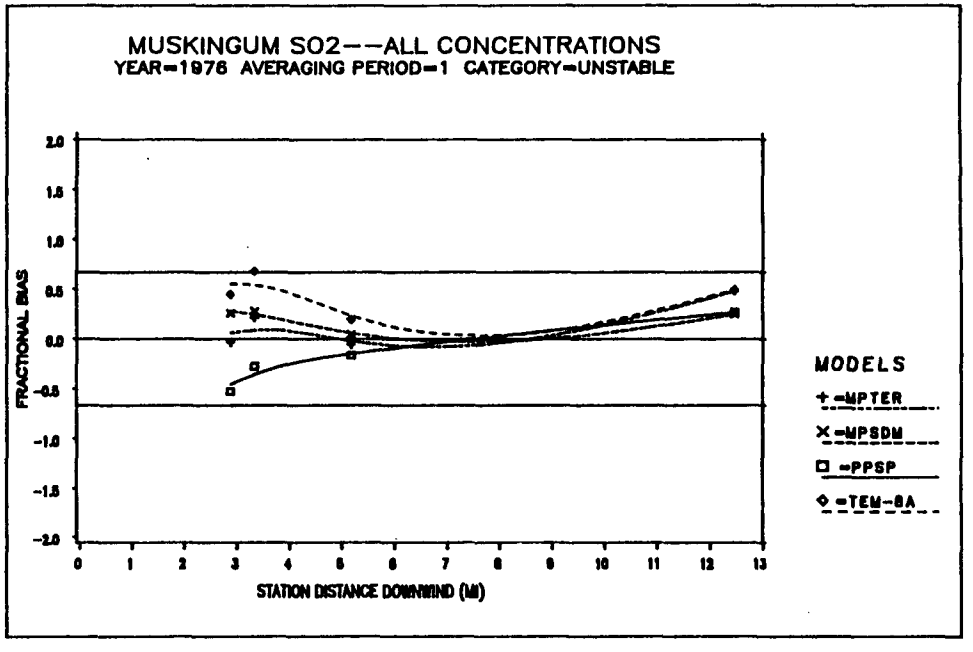
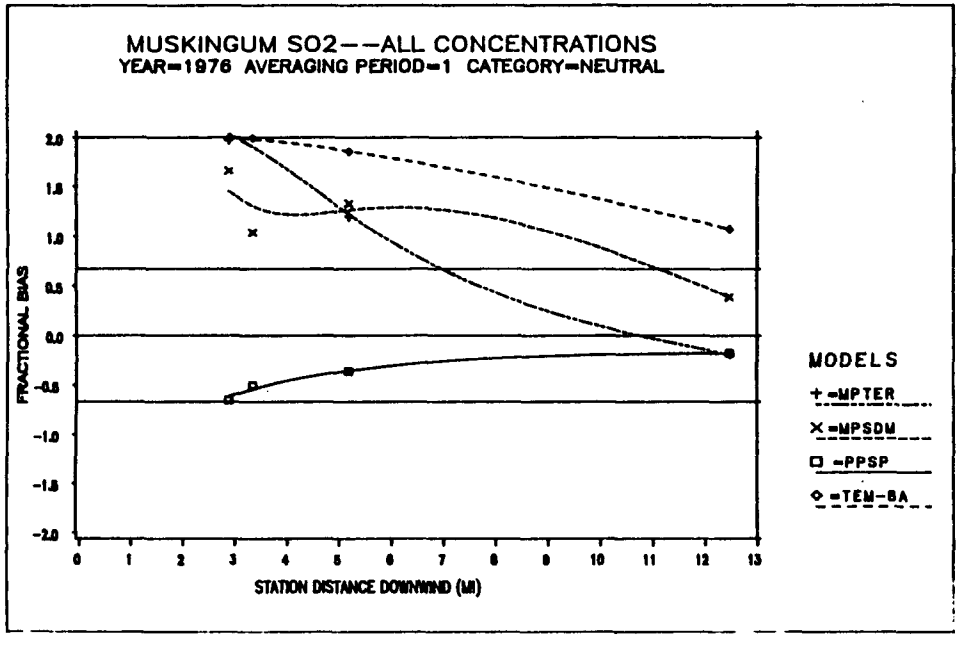


Figure C.15 Fractional Bias Of The Average Vs Station Distance By Stability Class Using All Paired Values







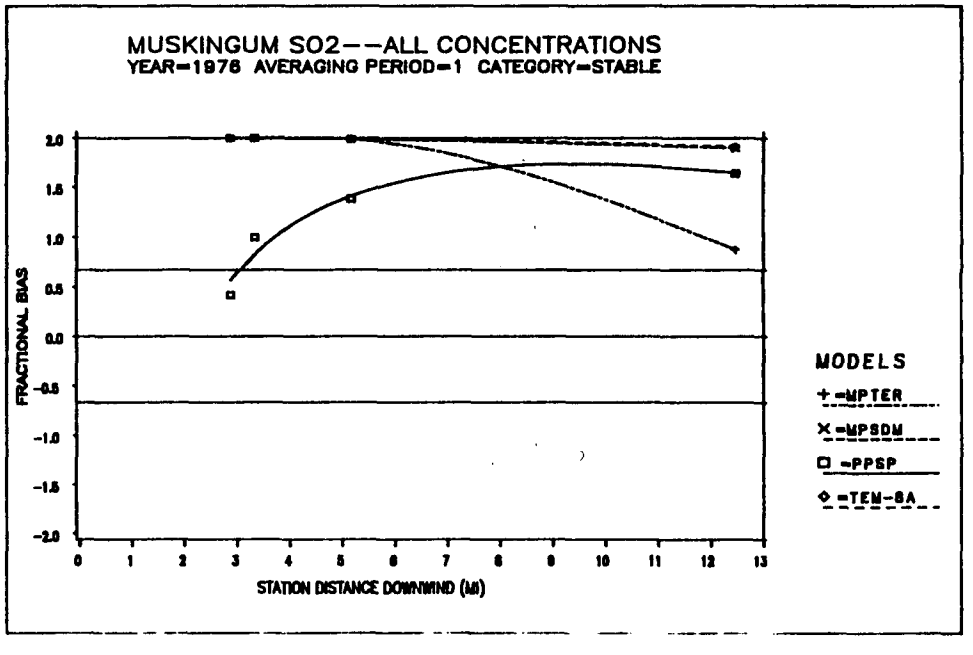
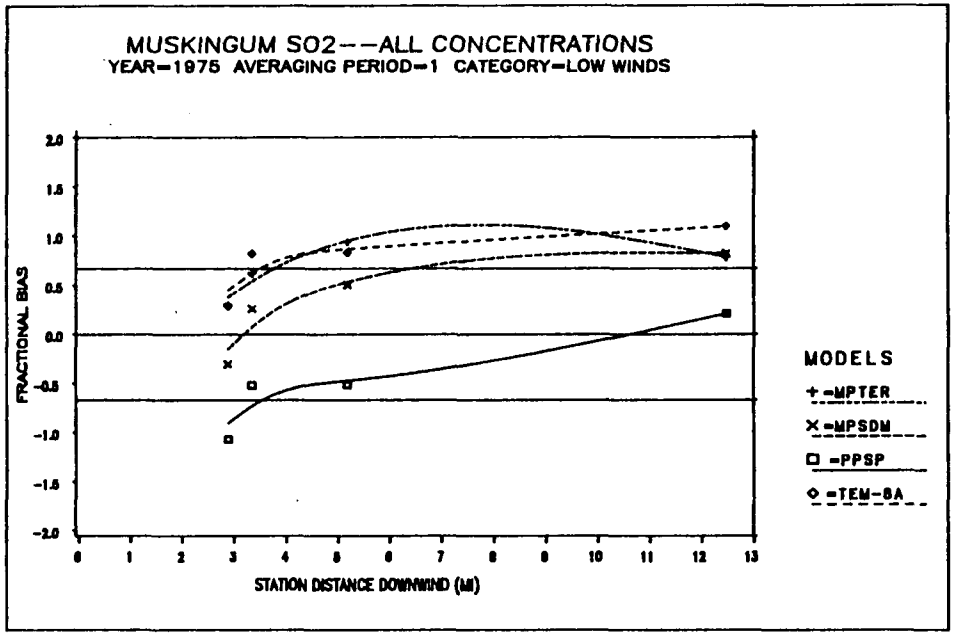
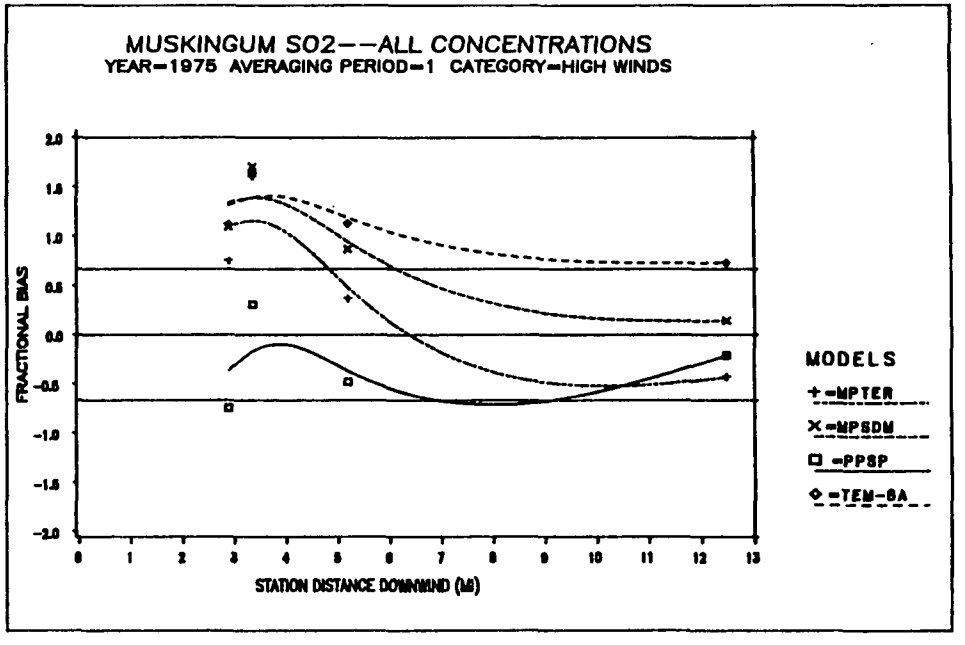


Figure C.15 (Continued)





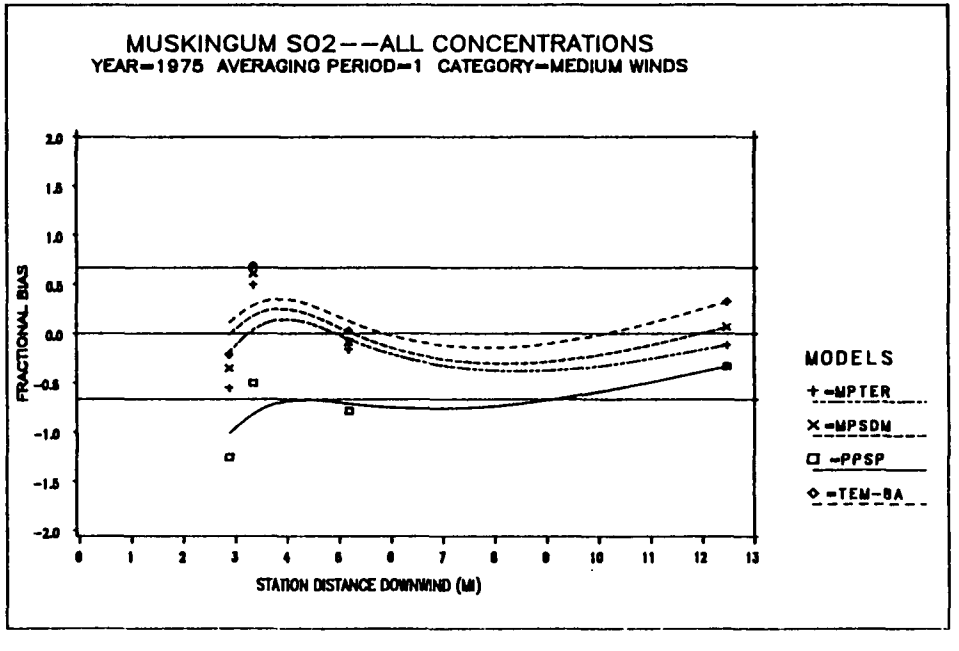
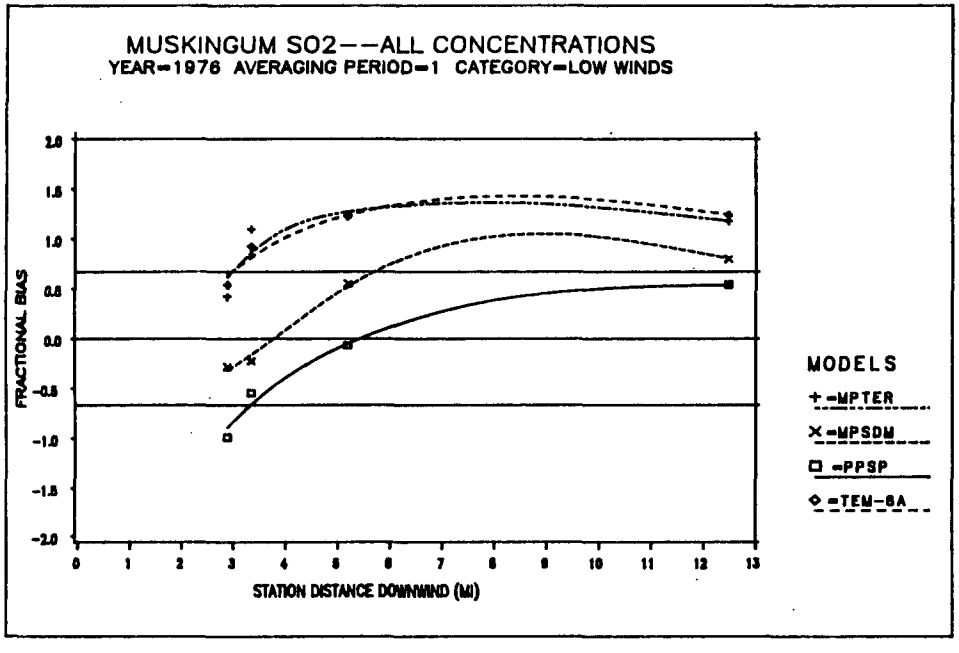
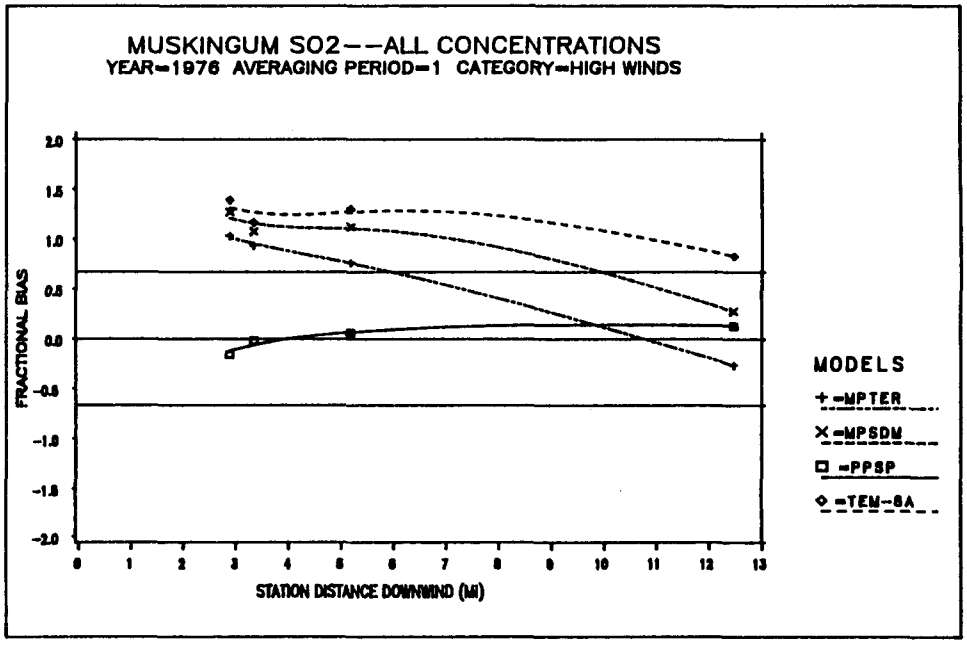


Figure C.16 Fractional Bias Of The Average Vs Station Distance By Wind Speed Class Using All Paired Values





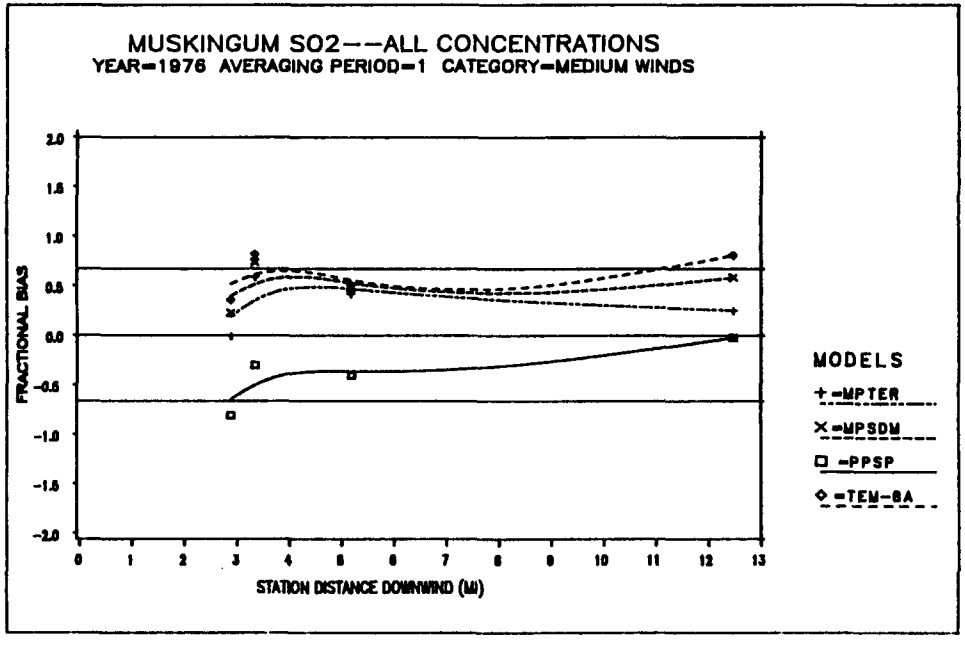


Figure C.16 (Continued)

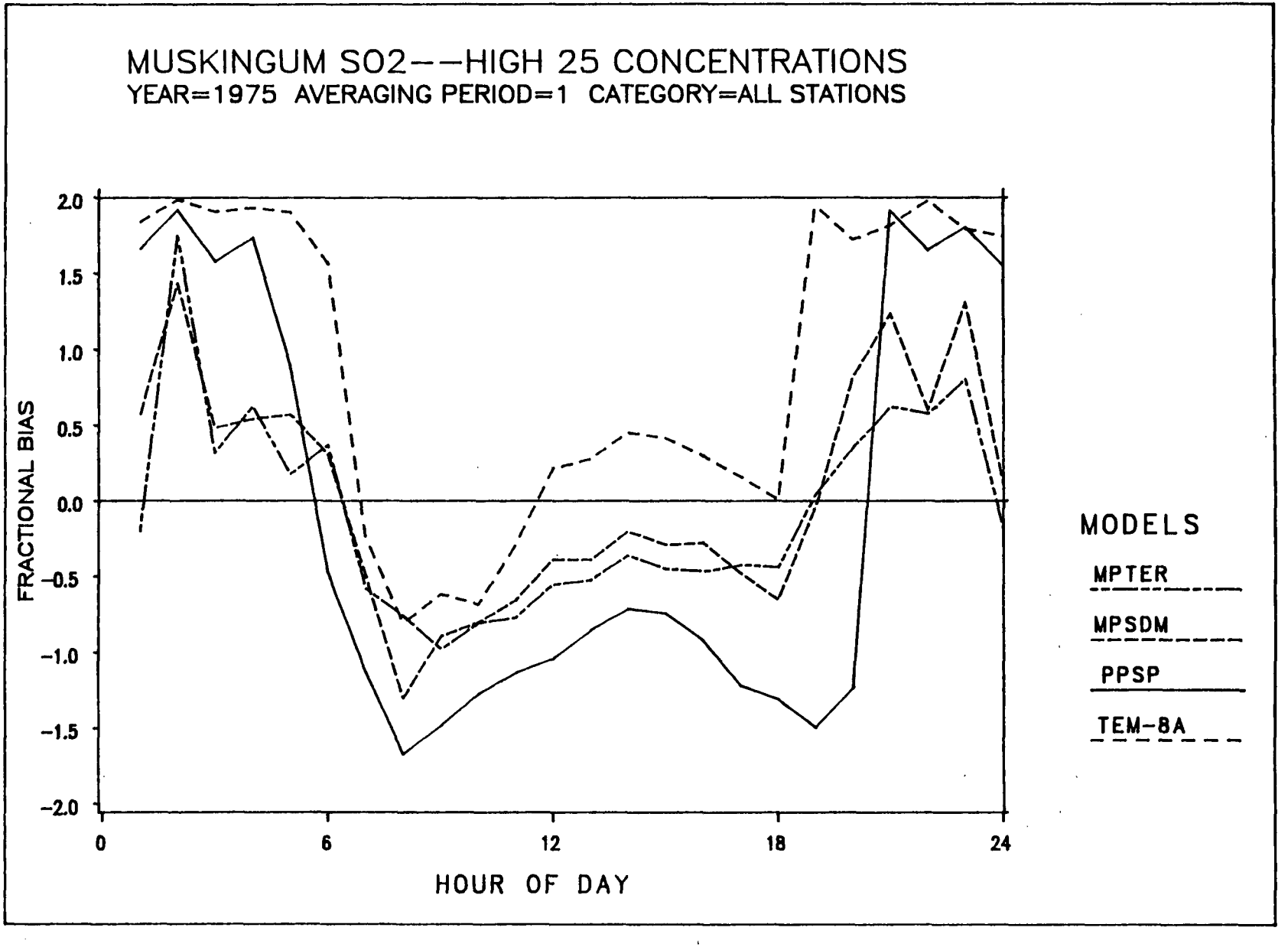


Figure C.17 Fractional Bias Of The Average Vs Hour Of The Day Using High 25 Values

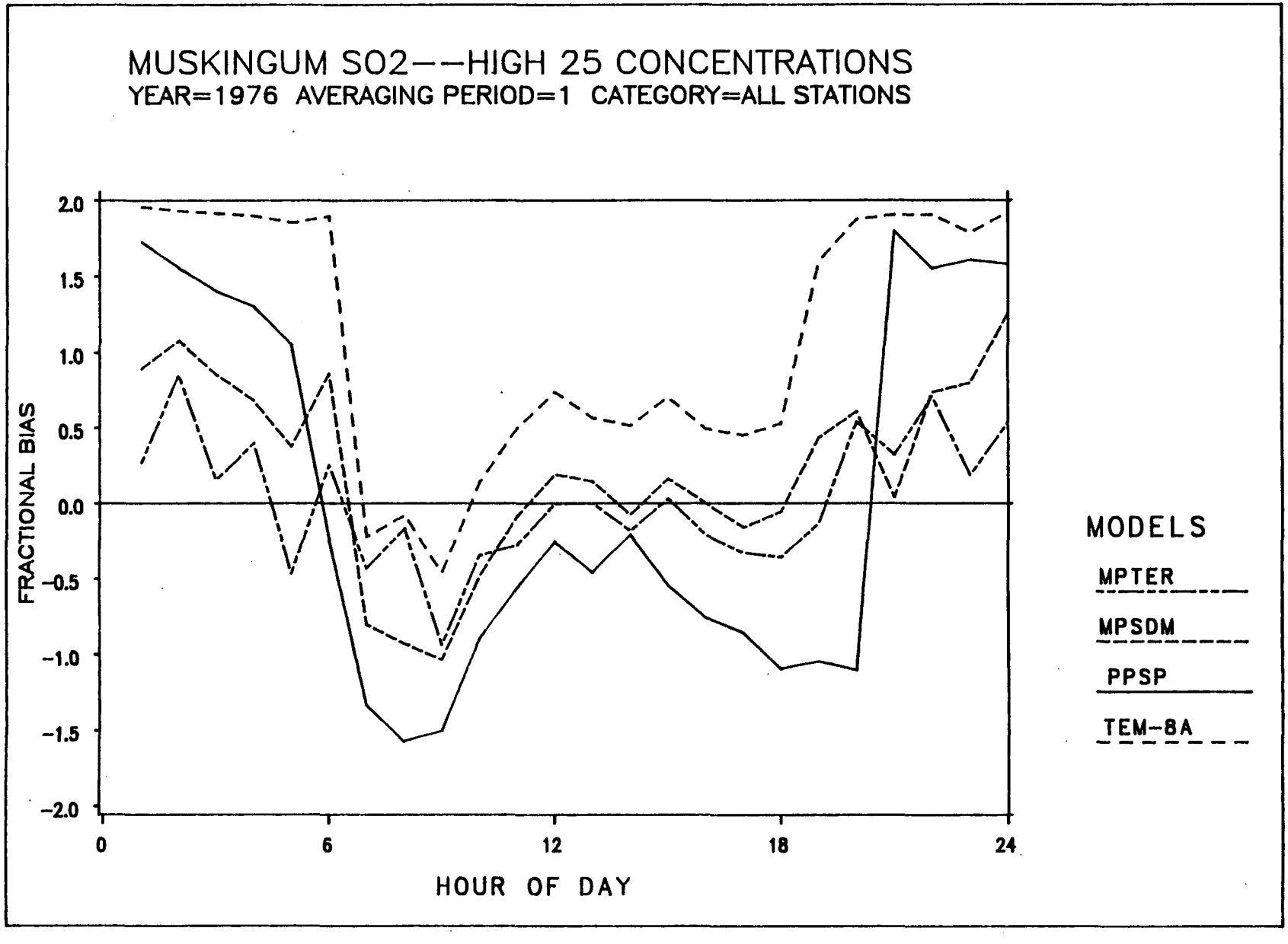


Figure C.17 (Continued)

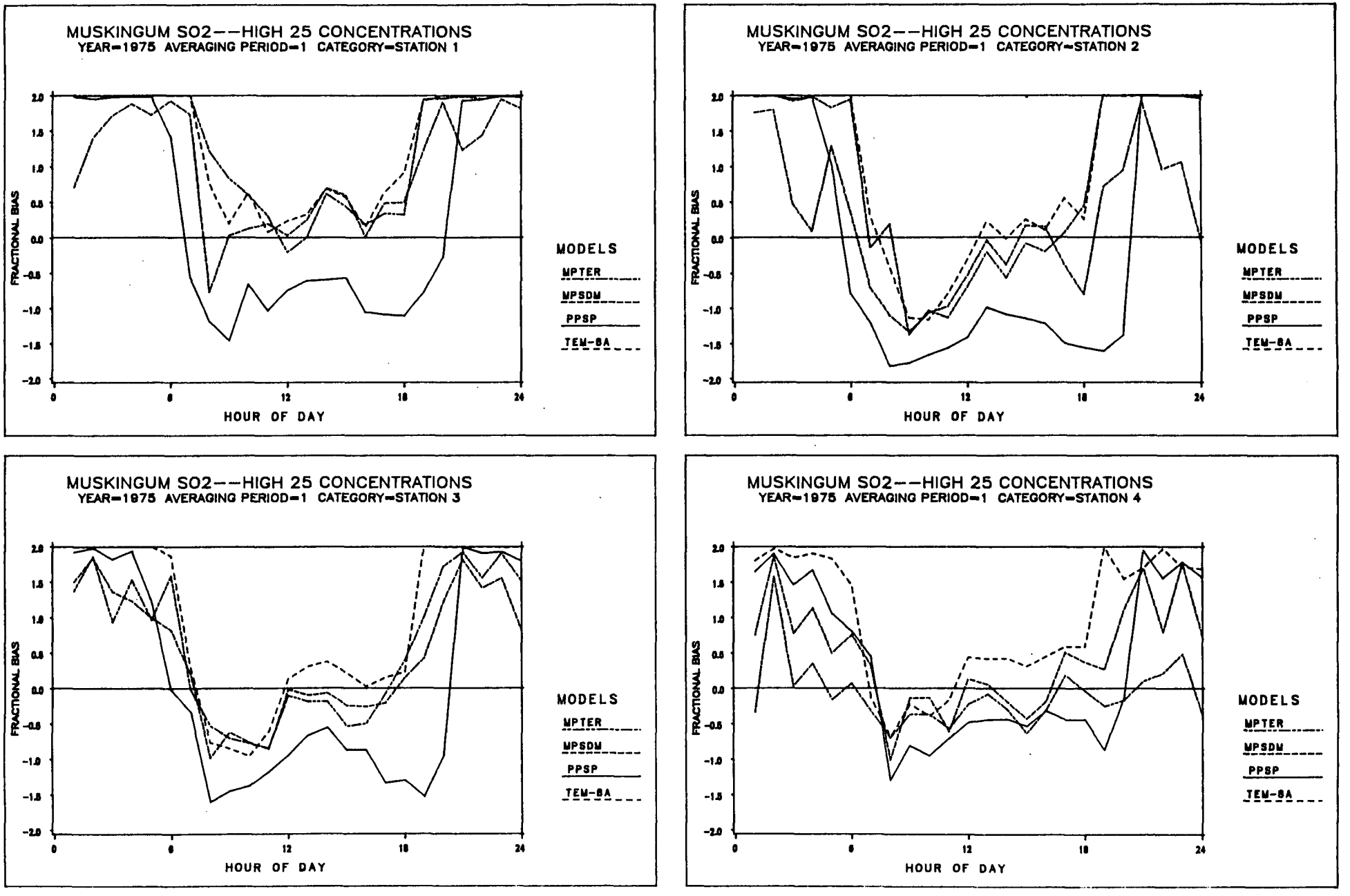
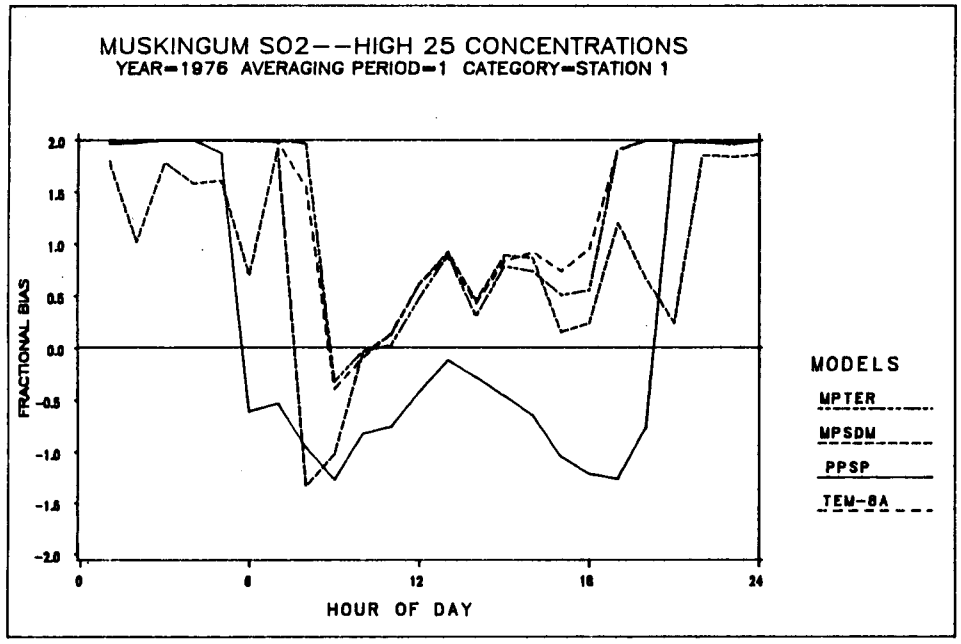
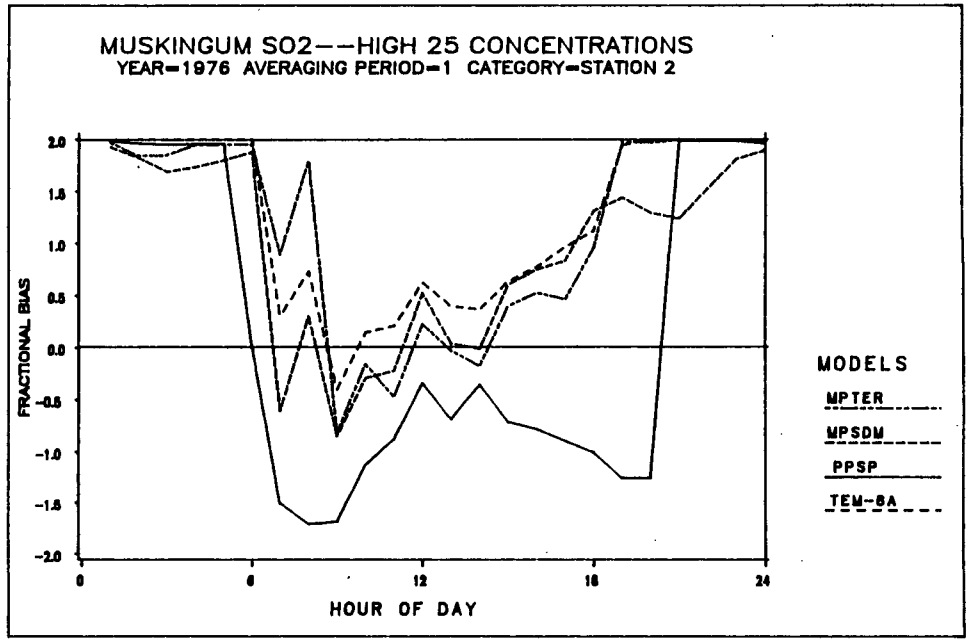
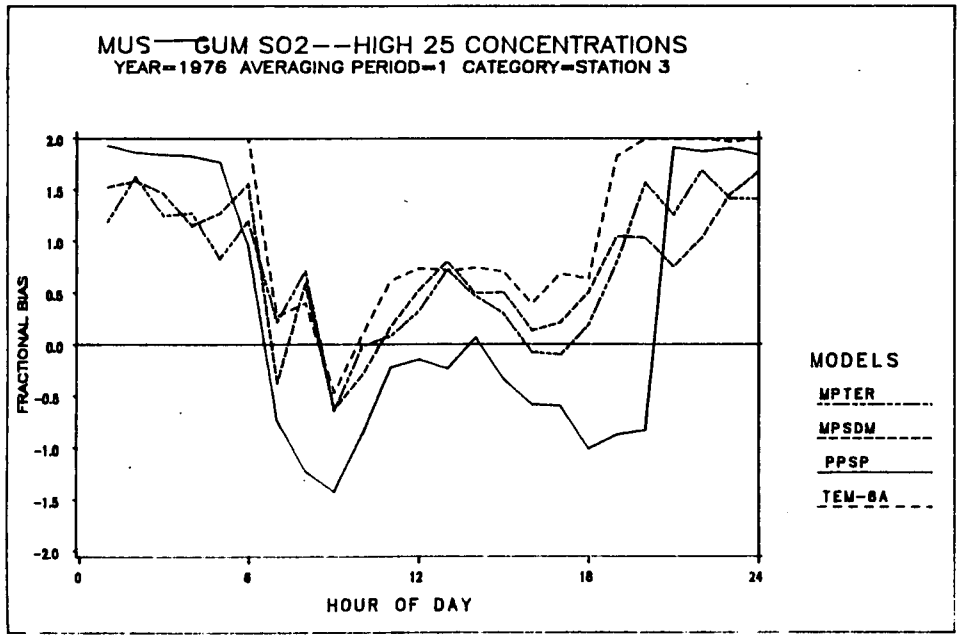


Figure C.18 Fractional Bias Of The Average Vs Hour Of The Day By Station Grouping Using High 25 Values







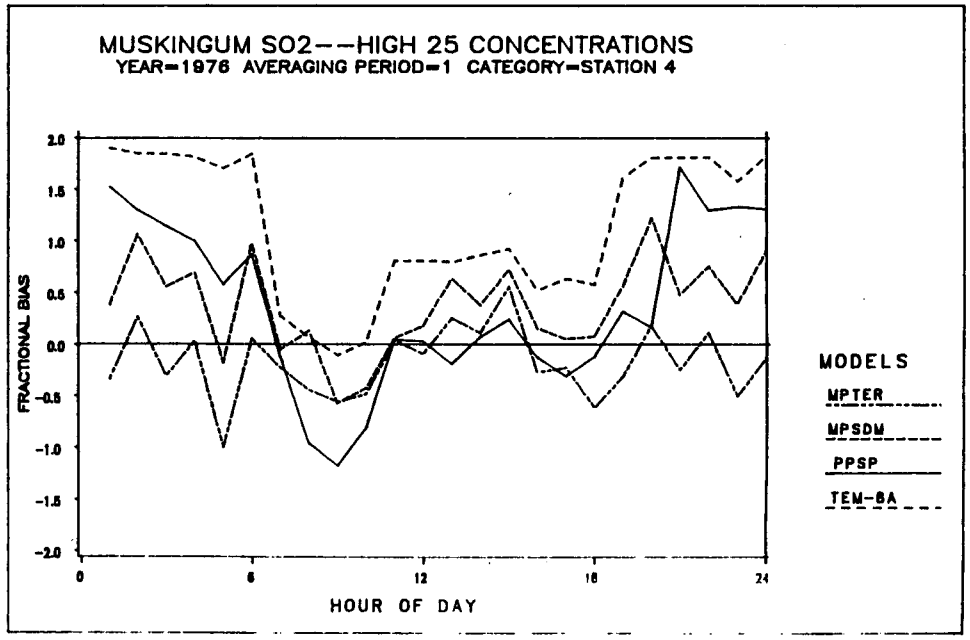


Figure C.18 (Continued)

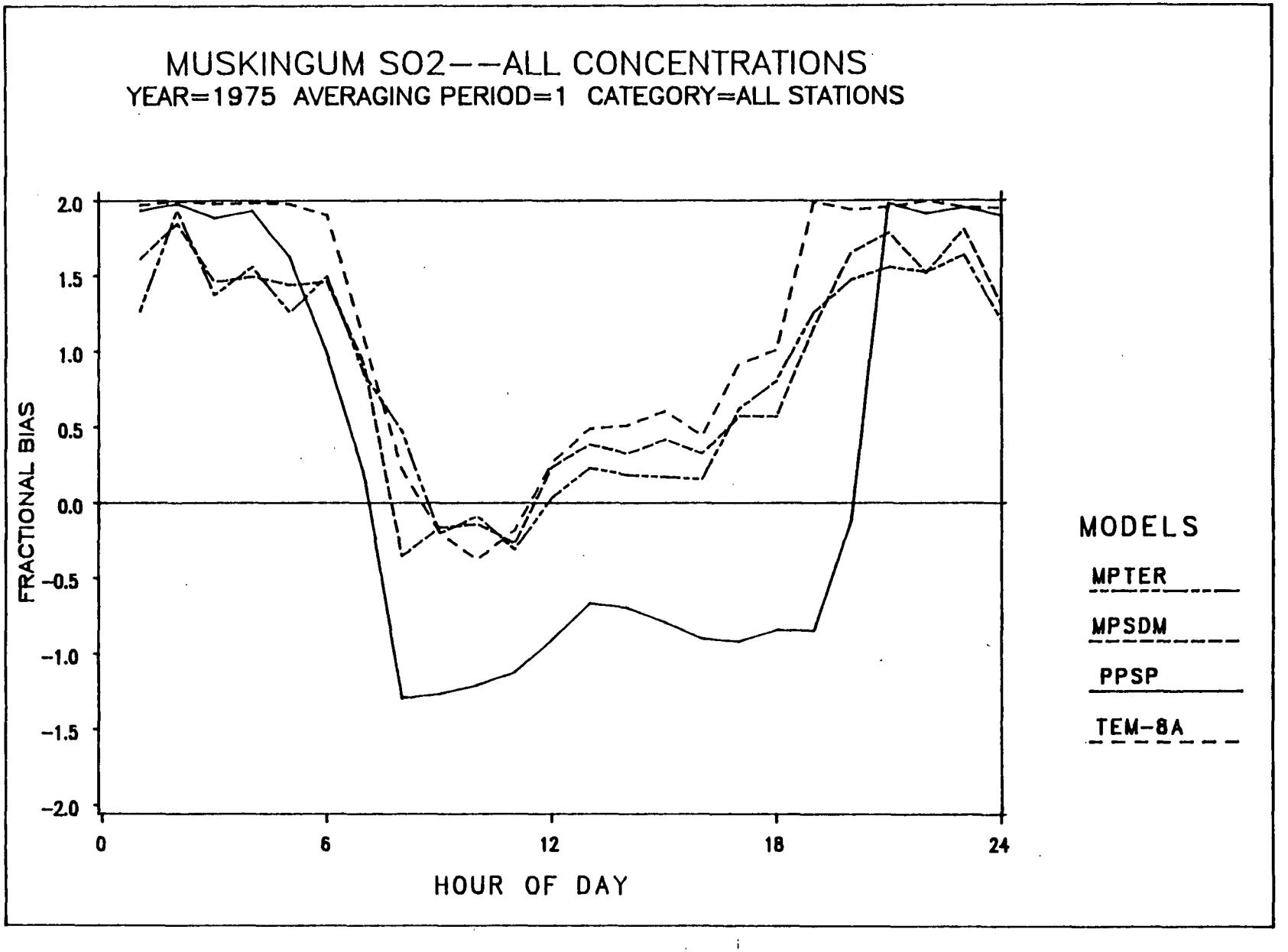


Figure C.19 Fractional Bias Of The Average Vs Hour Of The Day Using All Paired Values

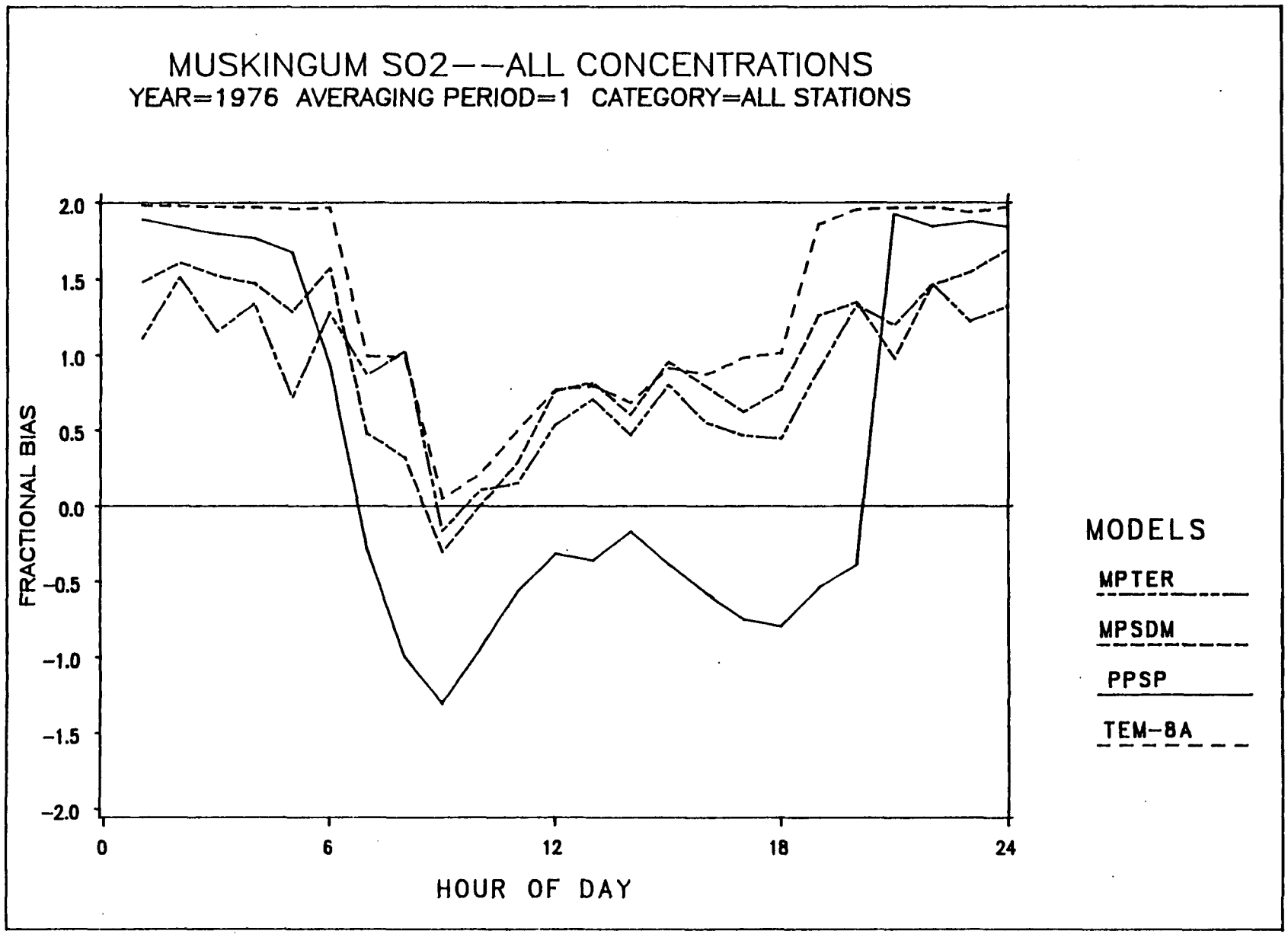


Figure C.19 (Continued)

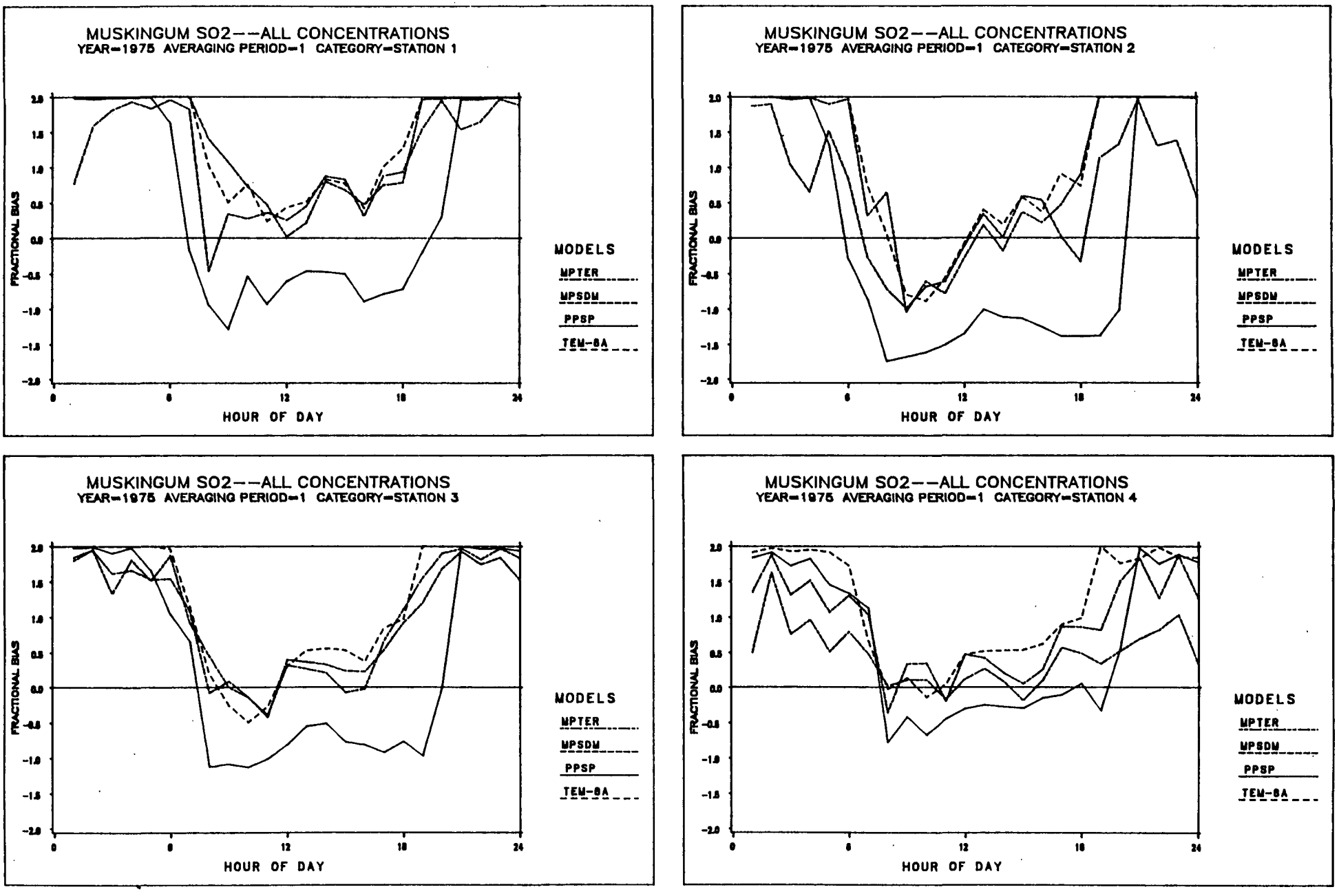


Figure C.20 Fractional Bias Of The Average Vs Hour Of The Day By Station Grouping Using All Paired Values

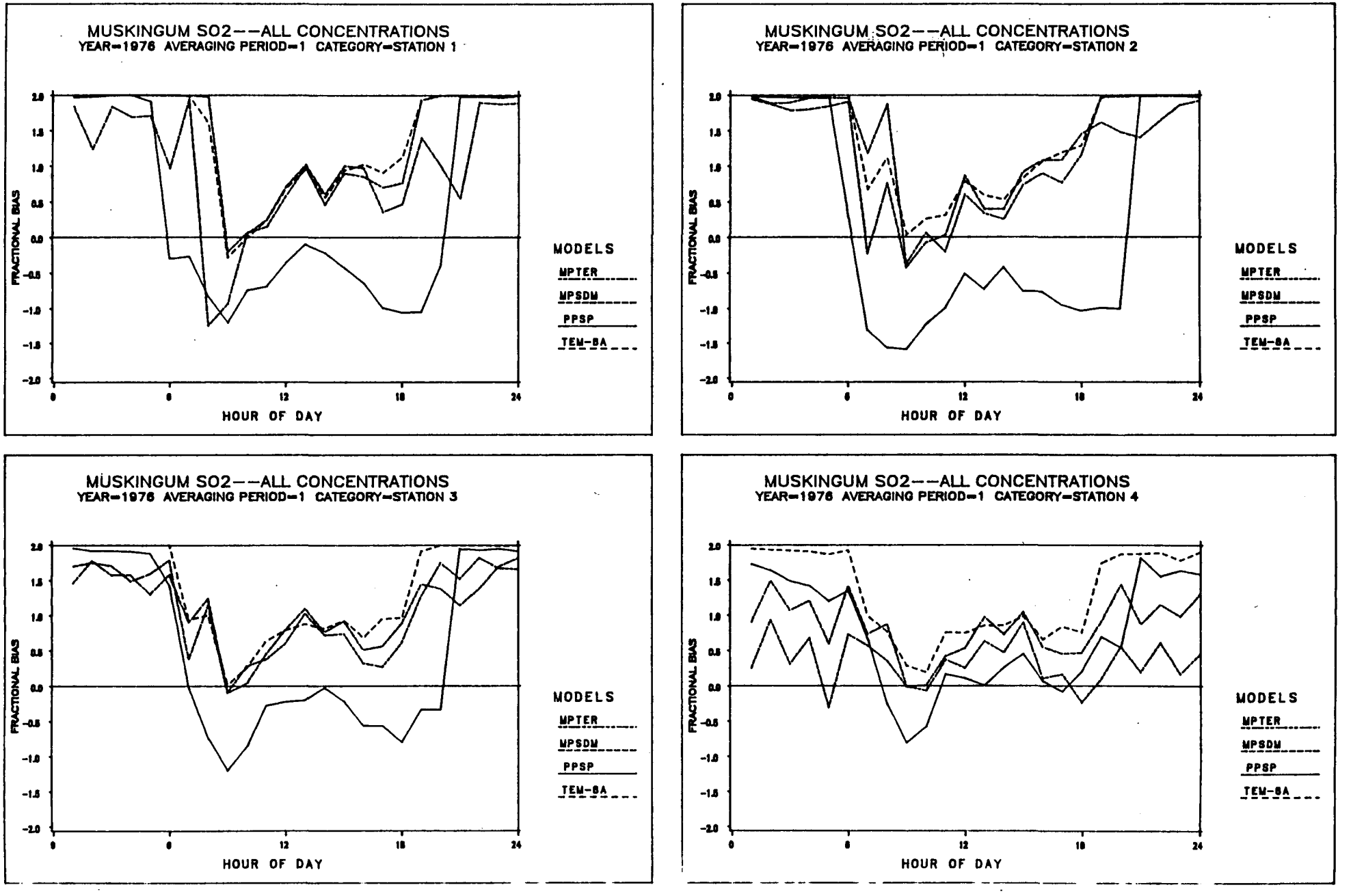


Figure C.20 (Continued)

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-450/4-83-003e	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE Evaluation of Rural Air Quality Simulation Models Addendum E: Graphic Summary of the Performance of Rural Air Quality Models		5. REPORT DATE July 1987
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) William M. Cox, Joseph A. Tikvart, Gerald K. Moss		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Source Receptor Analysis Branch, Monitoring and Data Analysis Division, U. S. Environmental Protection Agency		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE EPA-450/4-83-003e
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

This addendum uses a variety of graphic formats to display and compare the composite performance of four rural models using data bases assembled at Clifty Creek, Muskingum River, Paradise and Kincaid. The four models included MPTER (EPA), PPSP (Martin Marietta Corp.) MPSDM (ERT) and TEM-8A (Texas Air Control Board). Three Appendices to this addendum present the model performance separately for Kincaid (Appendix A), Paradise (Appendix B) and Muskingum River (Appendix C) using a graphic format similar to that used previously for Clifty Creek (Addendum B). Composite measures of performance are constructed using the information from each data base and presented as fractional bias plots where the composite fractional bias of the average vs fractional bias of the standard deviation is shown by averaging period, and selected meteorological categories. A net fractional bias is displayed for each model and averaging period which combines the fractional bias of the average and standard deviation, averaged over the data bases. The results indicate, inter alia, a consistent and significant bias (overprediction) for PPSP, a decreasing bias with increasing averaging period for MPSDM, and an increasing bias with averaging period for TEM-8A. Overall, MPTER indicated the most consistent performance and had the smallest composite bias among the four models.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution Mathematical Modeling Meteorology Sulfur Dioxide Statistical Measure Performance Evaluation	Air Quality Impact Assessment	

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE