RITA *Intelligent Transportation Systems*
Joint Program Office

An Analysis of Temperature and Pressure Data from Connected Vehicles in the Developmental Testbed Environment

www.its.dot.gov/index.htm

Final Report — July 9, 2009

FHWA-JPO-10-062/NTL # 14968

U.S. Department of Transportation

Research and Innovative Technology Administration

Notice

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD MM YYYY)	2. REPORT TYPE	3. DATES COVERED
09 07 2009	Final	February 2008 – January 2013

4. TITLE AND SUBTITLE
AN ANALYSIS OF TEMPERATURE AND PRESSURE DATA FROM CONNECTED VEHICLES IN THE DEVELOPMENTAL TESTBED ENVIRONMENT

5a. CONTRACT NUMBER
DTFH61-08-D-00012

5b. GRANT NUMBER

6. AUTHOR(S)
Sheldon Drobot, Michael Chapman, Tara Jensen, Gerry Wiener, Elena Schuler, Christian Johansen

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER
5

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University Corporation for Atmospheric Research, 3090 Center Green Dr. Boulder, CO 80301

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
United States Department of Transportation
Research and Innovative Technology Administration
Federal Highway Administration, Office of Operations
1200 New Jersey Ave., SE
Washington, DC 20590

10. SPONSORING/MONITOR'S ACRONYM(S)
FHWA, HOTO

11. SPONSORING/MONITOR'S REPORT NUMBER(S)

012a. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

12b. DISTRIBUTION CODE

13. SUPPLEMENTARY NOTES

Mr. Paul Pisano (COTM)

14. ABSTRACT (Maximum 200 words)
With funding and support from the USDOT RITA and direction from the FHWA Road Weather Management Program, NCAR is developing a Vehicle Data Translator (VDT) that incorporates vehicle-based measurements of the road and surrounding atmosphere with other weather data sources and creates road and atmospheric hazard products. In support of VDT development, this report (1) analyzed archived probe message data from the PoC, DUAP, and DTE09 experiments, and (2) provided hardware recommendations for processing data. Major conclusions include: (a) The SRT, CRT, NST, and CAT provide a robust QCh set; (b) For DTE09, the Jeep Cherokees proved superior to the Ford Edges and Nissan Altima; (c) For all three data sets, temperature measurements are superior to pressure measurements; (d) Environmental conditions (precipitation, temperature) might affect the QCh pass rates, but vehicle characteristics (speed) and time of day do not. The effects of environmental conditions on these datasets were not statistically significant for the most part, and physically the differences were small; (e) For the temperature observations that passed QCh in DTE09, the resulting statistics indicate that the vehicle data is very similar to KDTW; and (f) Storage of vehicle and ancillary data will require considerable disk space.

15. SUBJECT TERMS
ITS, connected vehicle, road weather, probe data

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			None	80	Paul Pisano
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER
Unclassified	Unclassified	Unclassified			(202) 366-1301

Preface/ Acknowledgements

This work is sponsored by the U.S. Department of Transportation Road Weather Management Program and the Intelligent Transportation Systems Joint Program Office under contract DTH61-08-D-00012.

Disclaimer Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. This report does not constitute a standard, specification, or regulation.

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

The United States Government does not endorse products or manufactures. Trade and manufactures' names appear in the report only because they are considered essential to the objective of the document.

The opinions, findings, and conclusions expressed in this document are those of the authors and not necessarily those of the U.S. Department of Transportation, Federal Highway Administration, or any other sponsor.

Table of Contents

List of Figures

Chapter 1 Introduction and Task Overview

The use of vehicle sensor data to improve weather and road condition products, as envisioned as part of the Research and Innovative Technology Administration (RITA), could revolutionize the provision of road weather information to transportation system decision-makers, including travelers. For example, vehicle-based probe data will significantly increase the density of weather observations near the surface and also provide unique datasets for deriving and inferring road-condition information. However, the amount of data flowing through a fully functional connected vehicle network could be immense, and many prospective users likely will not be capable of handling this vast quantity of data in its native form.

Complementing results from the Task 3 Report (NCAR 2009a), the goals of this report are to (1) analyze archived probe message data from the 2008 Detroit Proof-of-Concept (PoC) experiment, the Michigan Department of Transportation VII Data Use Analysis and Processing (DUAP) data set, and the Development Testbed Environment 2009 (DTE09), and (2) provide hardware recommendations for processing vehicle and ancillary data. Specifically, this report focuses first on assessing how many of the air temperature and pressure observations, the two observations collected by the vehicles with a ground truth comparison available, pass Quality Checking (QCh) tests, and whether the pass rates vary by day, time of day, vehicle speed, ambient temperature, or precipitation class. Secondly, this report compares the observations passing QCh to a known standard and computes the correlation coefficient, mean absolute error, bias, and t-test for significance of the difference between the vehicle data and known standard. For this report, the known standard used was the Detroit Automated Surface Observing System (KDTW ASOS). The KDTW ASOS is a weather station that is part of the nation's weather observing network and is located at the Detroit Metropolitan Wayne County Airport in Romulus, Michigan. Lastly, this report outlines the hardware requirements for processing vehicle and ancillary data.

This report is organized into 6 sections. Section 2 discusses the data used in this analysis. Section 3 presents the results of the analyses. Section 4 provides the hardware requirements, and section 5 summarizes the key findings and remaining uncertainties. Section 6 provides the cited references.

Chapter 2 Data

Three data sets were used in the analysis for Task 5. The first is the DTE09, which contained over 270,000 air temperature and 260,000 barometric pressure observations collected over the DTE09 on 11 days in April 2009 (Figure 1). Vehicles were driven on predetermined routes in the DTE09 area for each testing day with data being collected by the vehicles and logged via the On-board Equipment (OBE). Temperature was measured by a sensor mounted in the vehicle's grill, and pressure was derived using the mass air flow and manifold absolute pressure measurements (MAF and MAP). Additional details of the DTE09 testing can be found in NCAR (2009b).

The second data source is the Proof of Concept (PoC) experiment conducted in the Michigan Testbed in the fall of 2008. Overall, there are just over 35,000 temperature and 32,000 pressure observations (Figure 2). Additional details regarding the PoC experiment can be found in NCAR (2008) and Dion and Robinson (2009).

Lastly, the third data set is the Michigan Department of Transportation VII Data Use Analysis and Processing (DUAP) data set. This data set contains nearly 4 million observations collected within the Detroit Metropolitan area from April 2008 through January 2009 (Figure 3). Additional information about DUAP can be found in Dion and Robinson (2009).

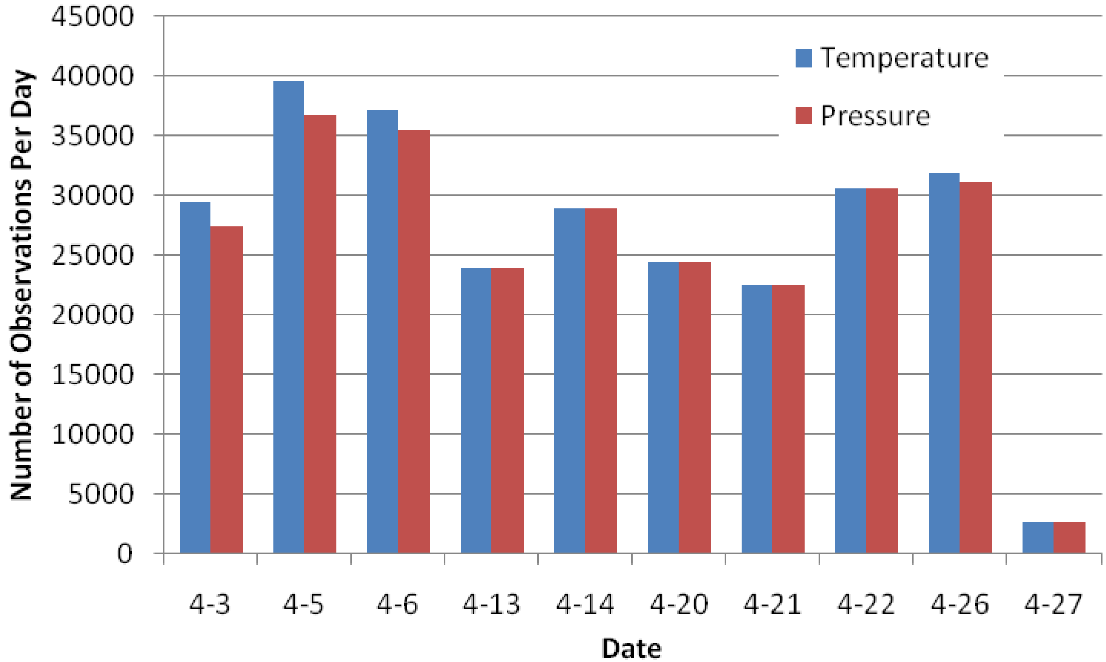


Figure 1. Number of observations collected during DTE09.

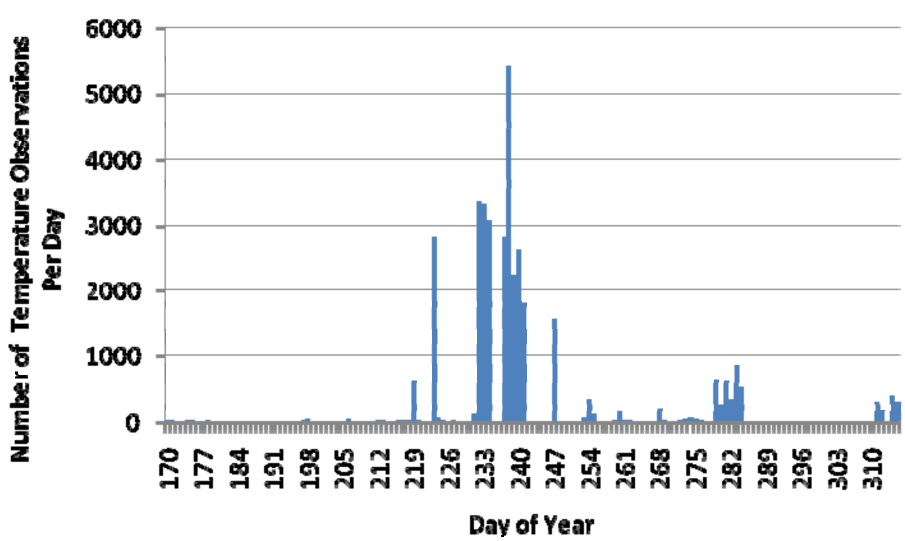


Figure 2. Number of temperature (top) and pressure (bottom) observations collected during PoC.

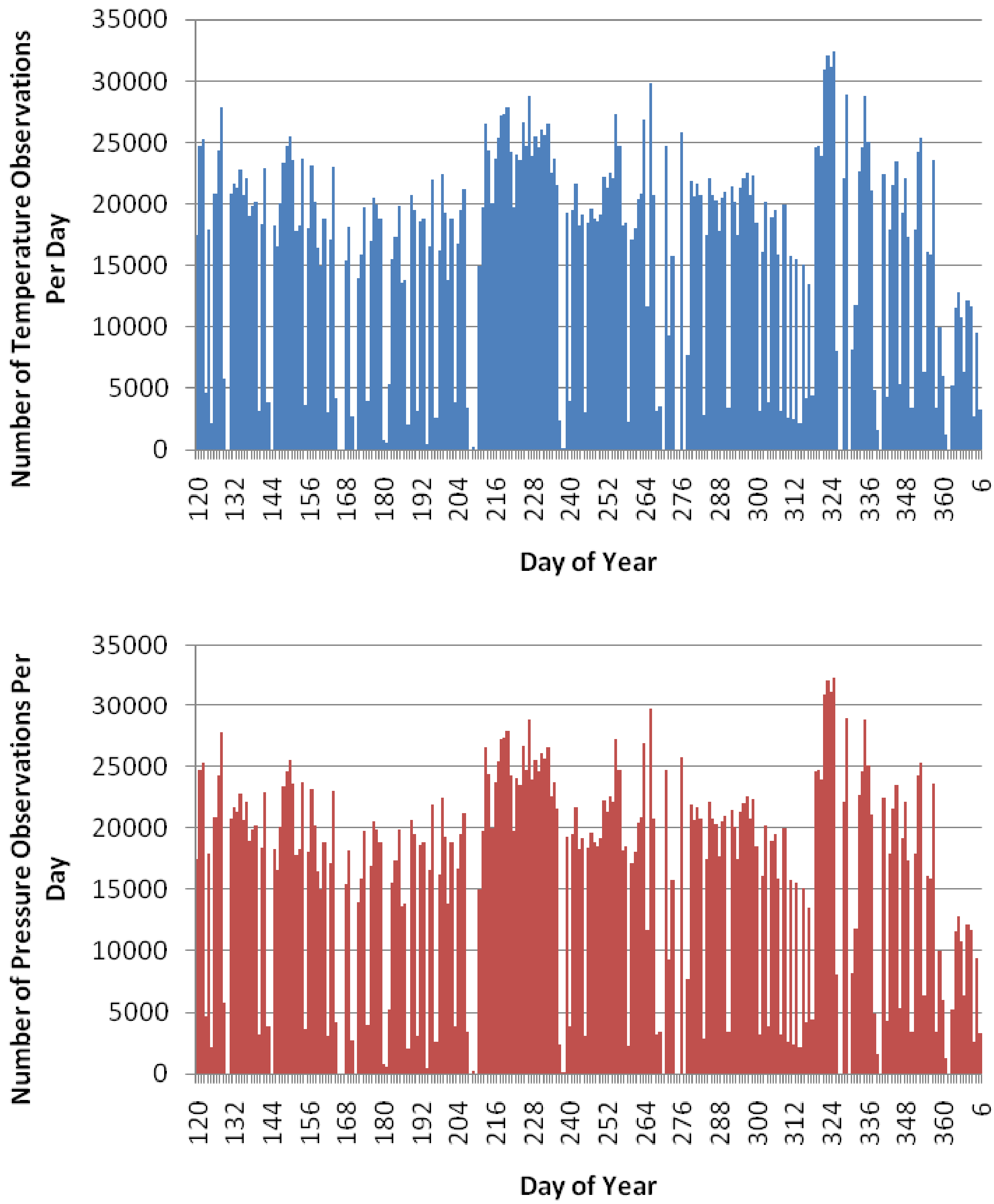


Figure 3. Number of temperature (top) and pressure (bottom) observations collected during DUAP.

Chapter 3 Results

3.1. Quality Check (QCh) Analysis

This section begins with an examination of how many individual temperature and pressure observations pass the relevant QCh tests. For each of the three data sets (DTE09, PoC, and DUAP), we also determine if the pass rate is affected by day, time of day, vehicle speed, precipitation condition (yes/no), and ambient temperature. This section also provides some comments on the failed data points to understand what led to the QCh failures.

Each temperature and pressure observation is assessed based on five of the six QCh tests discussed in NCAR (2008) and Drobot et al. (2009). The Remote Observation Test (ROT) was not run because we currently do not have high-resolution satellite-based temperature and pressure data. The five tests performed include:

- Sensor Range Test (SRT)
- Climatological Range Test (CRT)
- Neighboring Vehicle Test (NVT)
- Neighboring Surface Station Test (NST)
- Model Analysis Test (MAT)

The SRT identifies observations that fall outside the range of the known sensor hardware specifications. For temperature, the SRT specifications are set as [-40, 151°C], and these limits are based on metadata from the vehicle logs. For pressure, the SRT specifications are [580, 1090 mb], again based on metadata.

The CRT identifies observations that fall outside of location-specific climatological ranges. This is a more complex task than the SRT because the climatological range varies over times, dates, locations, and seasons. As a default, the CRT specifications are set as static historical minimum and maximum values. For temperature, we used the historical minimum and maximum values from Detroit (KDTW) compiled by the National Climatic Data Center (NCDC; http://lwf.ncdc.noaa.gov/oa/ncdc.html). For pressure, NCDC does not compile statistics, so we computed minimum and maximum values recorded at the KDTW station during the duration of the field seasons. The specifications are [-30, 40°C] for temperature and [960, 1020 mb] for pressure.

The NVT is a “nearest neighbor” test that compares the observation to neighboring vehicles in the road segment. If the observation value falls outside of a dynamic threshold, then the observation will fail the test. For a given temperature or pressure observation, the NVT uses all measurements on the road segment collected within the last five minutes. For this report specifically, all measurements in the testing area were used because overall there were a limited number of observations in the Detroit area, and including all measurements allowed a more discriminating test. Additionally, the Detroit area was deemed small enough spatially that all observations ought to be representing similar meteorological conditions. The algorithm then passes the temperature observation if it is within 2.5°C of the mean of its neighbors. A pressure observation passes if it falls within 5 mb of the mean of its neighbors. These initial thresholds were chosen using an educated guess and are subject to adjustment as more data sets are collected and analysis done.

The NST compares vehicle data with the KDTW measurement, which is located approximately 30 miles from the Testbed. The KDTW data are compiled on one-minute intervals, and the NST selects the closest-in-time observation. The algorithm then passes the temperature observation if it is within 2.5°C of the KDTW temperature, and it passes a pressure observation if it is within 5 mb of the KDTW pressure. These initial thresholds were chosen using an educated guess and are subject to change as work continues.

The MAT compares temperature and pressure data from the vehicles with the Rapid Update Cycle (RUC) Surface Assimilation System (RSAS) data for the pixel closest in space to the observation. The RSAS data assimilates surface observations into the RUC weather model to provide detailed model surface analyses, and a pixel of RSAS data is defined by its 15 km grid spacing. The RSAS is currently ingested at NCAR every hour, so the data are not as current as KDTW in most cases. The algorithm passes the temperature observation if it is within 2.5°C of the RSAS temperature, and it passes a pressure observation if it is within 5 mb of the RSAS pressure.

After running each observation through the five tests, the observation is given one final QCh flag. This is termed the Combined Algorithm Test (CAT), and a value passes if it passes the SRT, CRT, and NST. Currently, the MAT is not included because we believe that the NST is a superior test, and the NVT is not used because if there is a preponderance of poor observations, the NVT could give a misleading result.

3.1.1. DTE09

Looking first at the QCh test results for all days emphasizes the superiority of the vehicle-based temperature measurements compared with pressure measurements. At least 75% of the temperature observations pass each QCh test, and 78% of the observations pass the CAT (Figure 4). In contrast, only 29% of the pressure observations pass the CAT, largely related to the low pass rate for the pressure NST (31%).

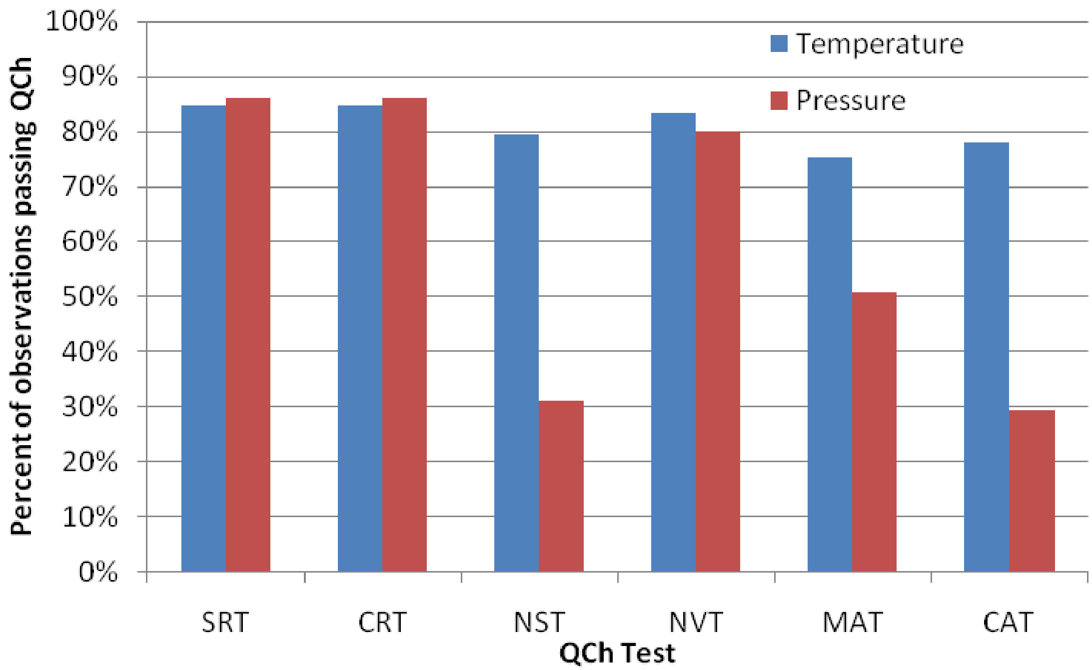


Figure 4. Percentage of DTE09 observations passing QCh tests.

For the SRT and CRT, all of the failed observations are null values ("-9999"). For the NST, NVT, MAT, and CAT, not all of the failed QCh values are null values. Focusing on the CAT results, the failed vehicle temperatures tend to be lower than KDTW, and the pressure readings tend to be lower as well. There is no clear bias for temperature, except for the slight negative trend, or a typical offset distance. Pressure is more clearly negatively biased, although some of the observations have a positive bias.

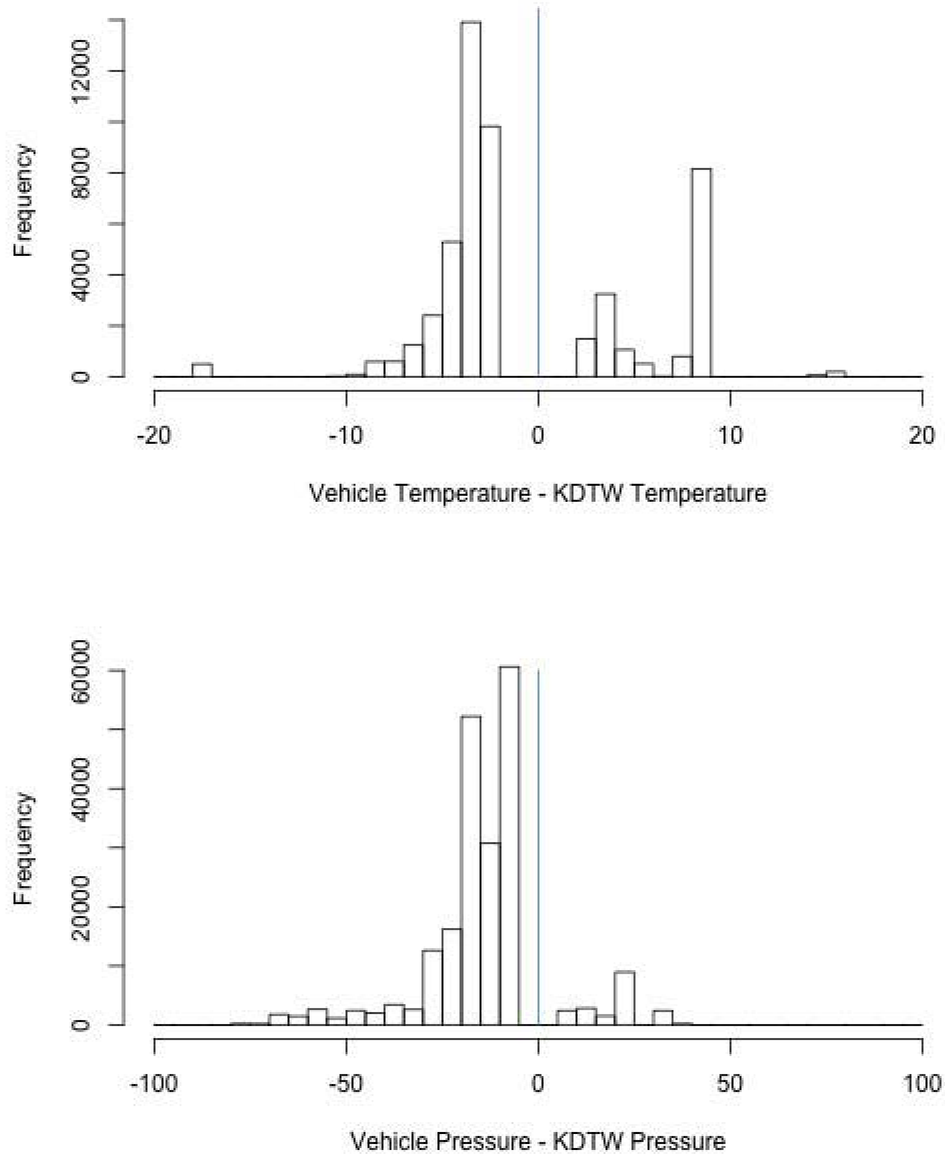


Figure 5. Histograms of CAT-failed DTE09 temperature and pressure observations. The blue line signifies the location of 0 difference, so that skewing to the left of this line shows CAT-failed observations tended negative, skewing to the right shows they tended positive.

The pass rates for temperature and pressure vary when stratified by day (Figure 6). On April 27, 100% of the temperature observations pass the SRT, CRT, and NVT, and 88% of the observations pass the CAT. However, the highest NST and MAT values occur on April 13, when 97% of the temperature observations pass. Largely due to the high NST percentage, the CAT is also highest on April 13 (96% pass rate). For pressure, 100% of the observations also pass the SRT and CRT on April 27, as well as 98% of the values for the NVT. However, the NST pass rates are generally low across all days. On only one day (April 27) does the NST pass rate exceed 50%, with a low of only 17% on April 22. April 27th was one of the two days when the vehicles were driven closest to the KDTW station, which may indicate that the distance between the testbed and KDTW partly explains some of the failed observations. There also is no clear daily trend in the pass rates; high pass rate days and low pass rate days are intermingled.

The pass rates for temperature and pressure show considerable variability when stratified by vehicle (Figure 7). For the SRT and CRT, the Nissan observations have considerable problems, with a temperature pass rate less than 50% (there are no Nissan pressure observations). The NVT results are similar in pattern to SRT and CRT, but with slightly lower values, especially for Ford 2. Some clear problems for the Ford vehicle pressure measurements are apparent in the NST and MAT results. For NST, Ford 2 passes 0% of the time, while Ford 1 has a pass rate of only 5%. Ford 3 is better, with a pass rate of just over 40%, which is the second highest of all the vehicles. The best pass rate is 61% for Jeep 2; the pass rates for the remaining Jeeps are less than 40%, suggesting that the pressure problems are not related solely to the Fords.

Looking at each vehicle-day combination individually provides some additional insight into pass-fail rates and errors (Figure 81). Ford 1 returned observations each day, but the results were wildly inconsistent. On April 27, each of the Ford 1 temperature observations passed, but on April 6, all of the Ford 1 temperature observations failed. These temperature failures on April 6 are related to some kind of sensor malfunction; all Ford 1 temperature observations on April 6 registered 10°C, even though the KDTW temperature was much colder. Ford 2 only returned observations on seven of the test days, but on the days it did return data, at least 50% of the temperature observations passed the CAT, including two days (April 14 and April 27) when 100% of the temperature observations passed. However, these results are slightly misleading. On these two days, only 354 and 240 snapshots were acquired from Ford 2, whereas the other vehicles generally transmitted over 3000 observations on those days. Ford 2 pressure observations were roughly as bad as Ford 1, and on the best day (April 14), only 12% passed the CAT. Ford 3 collected data on all days and had the best pass rate for pressure of all the Fords, with a similar pass rate for temperature. On April 27, all 305 pressure observations passed the CAT. The pressure pass rate exceeded 50% on three days, and the lowest was a 3% pass rate on the first day of testing. For temperature, Ford 3 had a 50% or greater pass rate on six days. The 0% pass rate on April 27 was a similar problem to Ford 1, where the temperature was stuck at 10°C even though KDTW registered much higher temperatures.

The Jeep Cherokee vehicles tended to have higher CAT pass rates for both temperature and pressure. Jeep 1 passed 100% of its temperature observations on five days (April 3, 6, 13, 20, and 27) and had only one poor day (April 14), where only 36% of the temperature observations passed. This was largely a case where the failed observations were null values. Jeep 1 pressure observations are better than the Fords, and on April 20, 3551 of the 3936 pressure observations passed the CAT. High failure rates on several days (April 6, 14, and 22) were largely due to the presence of null values. Jeep 2 was the best performing vehicle, passing 100% of the temperature observations on eight of the ten test days. Pressure pass rates are lower, but not because of null values. Rather, the pressure observations are simply too far from the KDTW measurement to register as a pass. Jeep 3 also had a high pass rates for temperature (six perfect days) and the low pass rate on April 14 was due to the null values. Jeep 3 pressure readings were similar in pass rate to Jeep 1, exceeding 50% on only two days (April 6 and 27). The 0% pass rate on April 3 was related to the Jeep 3 pressure values registering a near-uniform 950 mb even though KDTW values were 15–20 mb higher. Jeep 4 had a reasonably good pass rate for temperature, with five perfect days, but also one particularly bad day (April 26), when less than 1 in 4 observations passed. As with the other Jeep vehicles, this related to the presence of many null values. Jeep 6 was nearly as good as Jeep 2, with six perfect days, but also a poor pass rate on April 22 (36%). The final Jeep was among the top performers for temperature (seven perfect days) but the worst Jeep for pressure. On only one day (April 5) did the pressure even reach a 25% pass rate, with the errors again being largely due to differences compared to KDTW, not null values. The Nissan only ran on four days, and the pass rates were generally poor (4%, 22%, 69%, 88%).

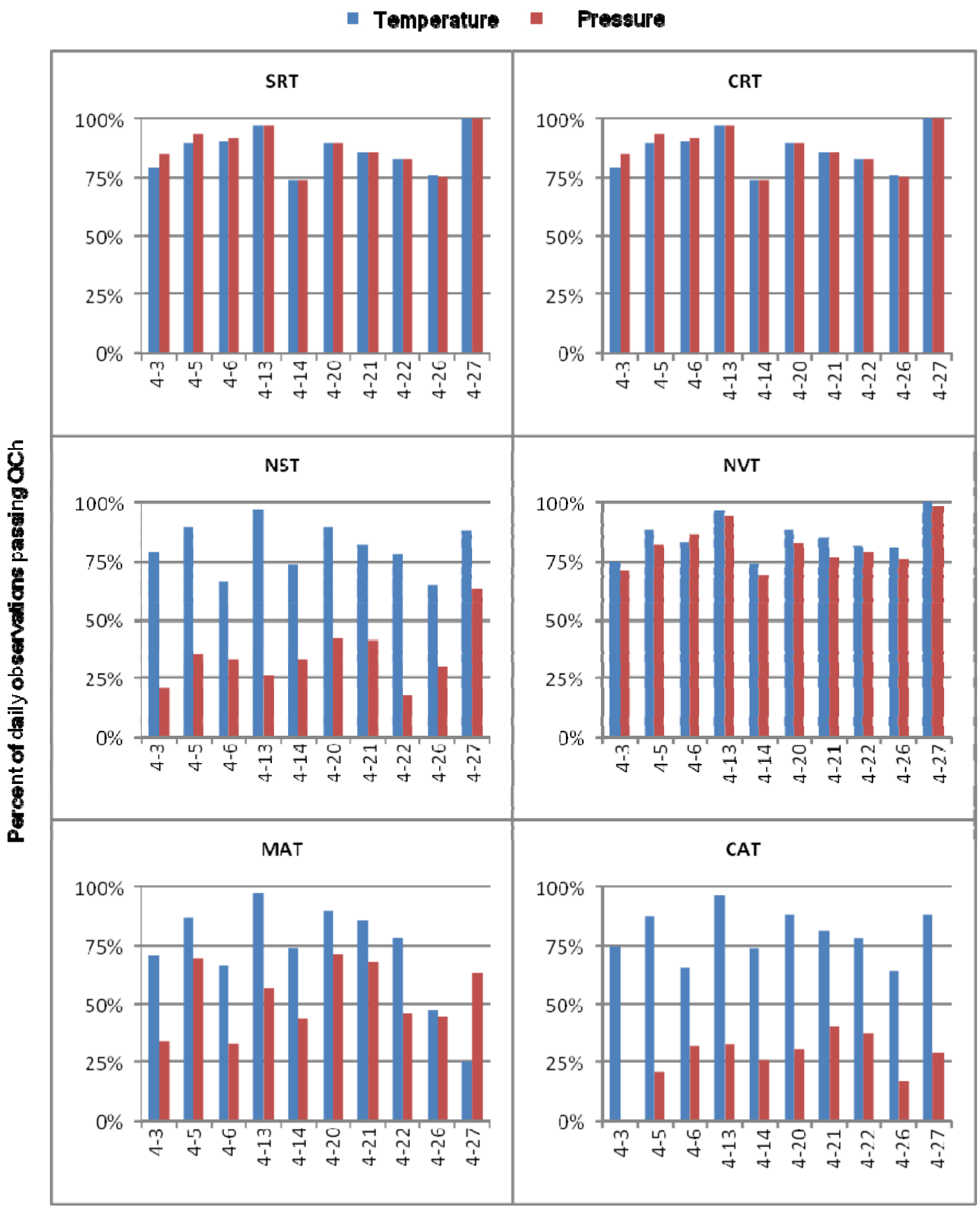


Figure 6. Percentage of DTE09 observations passing QCh tests stratified by date.

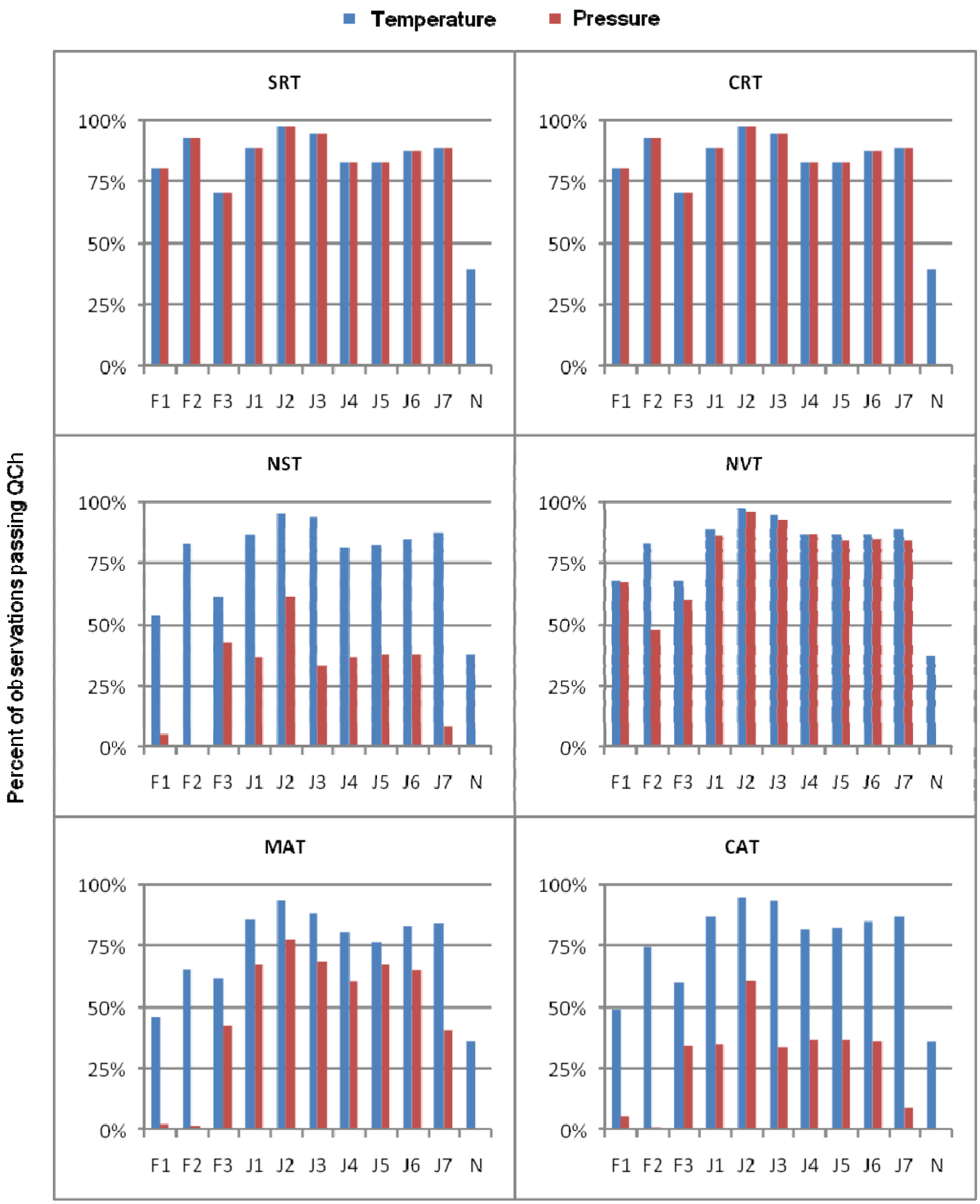


Figure 7. Percentage of DTE09 observations passing QCh tests stratified by vehicle.

Looking at each vehicle-day combination individually provides some additional insight into pass-fail rates and errors (Figure 8[1]). Ford 1 returned observations each day, but the results were wildly inconsistent. On April 27, each of the Ford 1 temperature observations passed, but on April 6, all of the Ford 1 temperature observations failed. These temperature failures on April 6 are related to some kind of sensor malfunction; all Ford 1 temperature observations on April 6 registered 10°C, even though the KDTW temperature was much colder. Ford 2 only returned observations on seven of the test days, but on the days it did return data, at least 50% of the temperature observations passed the CAT, including two days (April 14 and April 27) when 100% of the temperature observations passed. However, these results are slightly misleading. On these two days, only 354 and 240 snapshots were acquired from Ford 2, whereas the other vehicles generally transmitted over 3000 observations on those days. Ford 2 pressure observations were roughly as bad as Ford 1, and on the best day (April 14), only 12% passed the CAT. Ford 3 collected data on all days and had the best pass rate for pressure of all the Fords, with a similar pass rate for temperature. On April 27, all 305 pressure observations passed the CAT. The pressure pass rate exceeded 50% on three days, and the lowest was a 3% pass rate on the first day of testing. For temperature, Ford 3 had a 50% or greater pass rate on six days. The 0% pass rate on April 27 was a similar problem to Ford 1, where the temperature was stuck at 10°C even though KDTW registered much higher temperatures.

The Jeep Cherokee vehicles tended to have higher CAT pass rates for both temperature and pressure. Jeep 1 passed 100% of its temperature observations on five days (April 3, 6, 13, 20, and 27) and had only one poor day (April 14), where only 36% of the temperature observations passed. This was largely a case where the failed observations were null values. Jeep 1 pressure observations are better than the Fords, and on April 20, 3551 of the 3936 pressure observations passed the CAT. High failure rates on several days (April 6, 14, and 22) were largely due to the presence of null values. Jeep 2 was the best performing vehicle, passing 100% of the temperature observations on eight of the ten test days. Pressure pass rates are lower, but not because of null values. Rather, the pressure observations are simply too far from the KDTW measurement to register as a pass. Jeep 3 also had a high pass rates for temperature (six perfect days) and the low pass rate on April 14 was due to the null values. Jeep 3 pressure readings were similar in pass rate to Jeep 1, exceeding 50% on only two days (April 6 and 27). The 0% pass rate on April 3 was related to the Jeep 3 pressure values registering a near-uniform 950 mb even though KDTW values were 15–20 mb higher. Jeep 4 had a reasonably good pass rate for temperature, with five perfect days, but also one particularly bad day (April 26), when less than 1 in 4 observations passed. As with the other Jeep vehicles, this related to the presence of many null values. Jeep 6 was nearly as good as Jeep 2, with six perfect days, but also a poor pass rate on April 22 (36%). The final Jeep was among the top performers for temperature (seven perfect days) but the worst Jeep for pressure. On only one day (April 5) did the pressure even reach a 25% pass rate, with the errors again being largely due to differences compared to KDTW, not null values. The Nissan only ran on four days, and the pass rates were generally poor (4%, 22%, 69%, 88%).

[1] Only CAT results are shown in Figure 8 due to the large volume of possible figures.

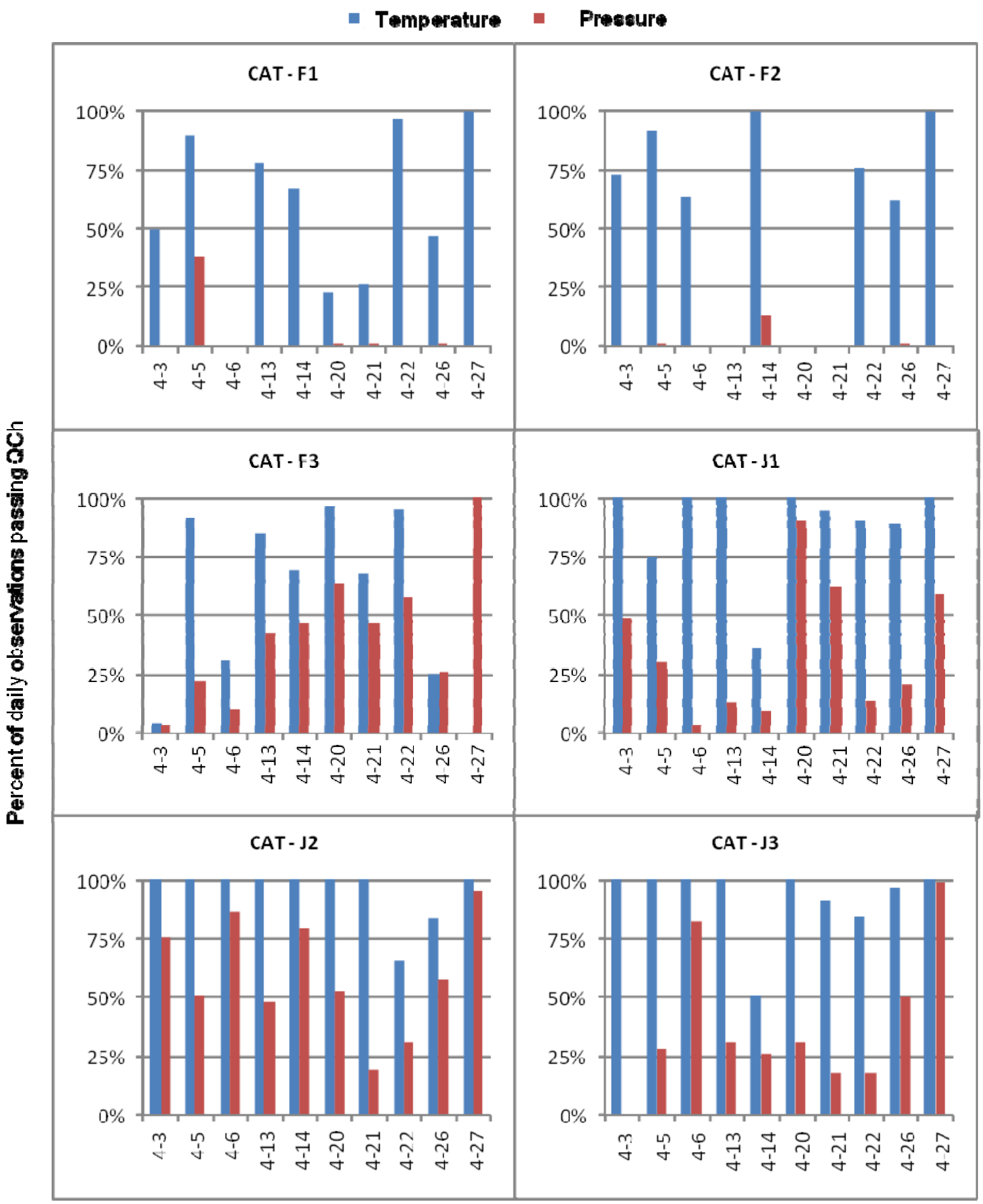


Figure 8. Percentage of DTE09 observations passing CAT QCh test stratified by vehicle and date.

Stratified by time of day (Figure 9), there are slight differences in pass rates, particularly for the NST, MAT, and CAT. For the NST and CAT, there are slightly higher pass rates in the mid morning hours, and again in the evening. However, these variations are generally on the order of only 10–15%. Testing the statistical significance of these differences reveals that the temperature sample has a p-value (probability of obtaining this result if the null hypothesis – nothing is happening – is true) of 0.018, and pressure has a p-value of 0.016. These lie right on the edge of statistical significance, with common maximum p-values being 0.05 for a 95% confidence interval and 0.01 for a 99% confidence interval. Given these p-values, the small sample of different hours available, and the relatively trivial differences physically, the pass rate variations by time for both temperature and pressure are basically negligible.

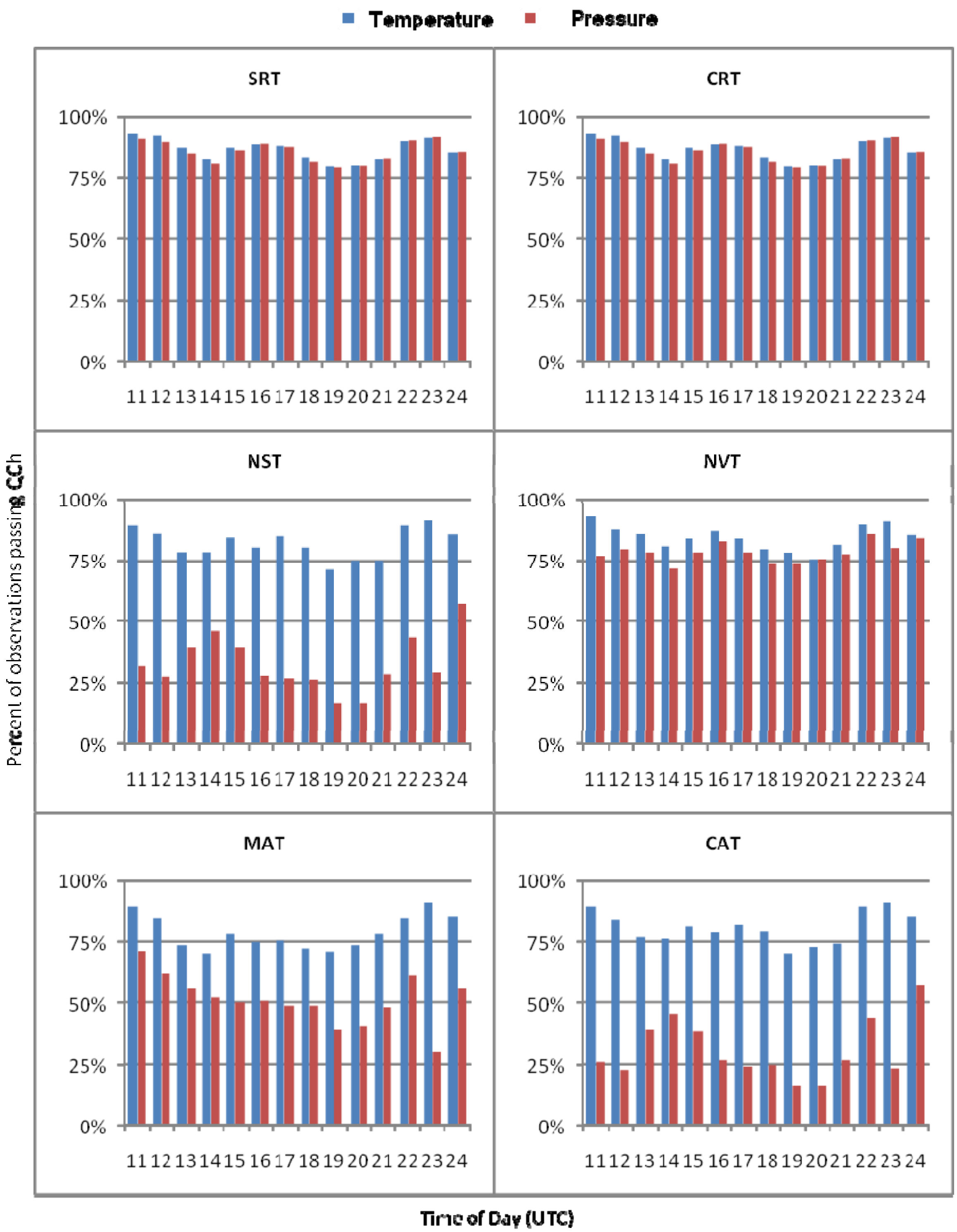


Figure 9. Percentage of DTE09 observations passing QCh test stratified by time of day.

On the surface, there is some evidence to suggest that vehicle speed plays a role in the QCh pass rates (Figure 10). For both temperature and pressure, the pass rate is typically less for vehicle speed equal to zero than the other speed classes, especially for the SRT, CRT, NVT, and for temperature, the NST, MAT, and CAT. On closer examination, the reduced pass rate for the speed equals zero grouping is related to the presence of null values. In no case does the failed SRT or CRT value equal anything else than a null value. It is not clear why there is a higher percentage of null values when speed equals zero.

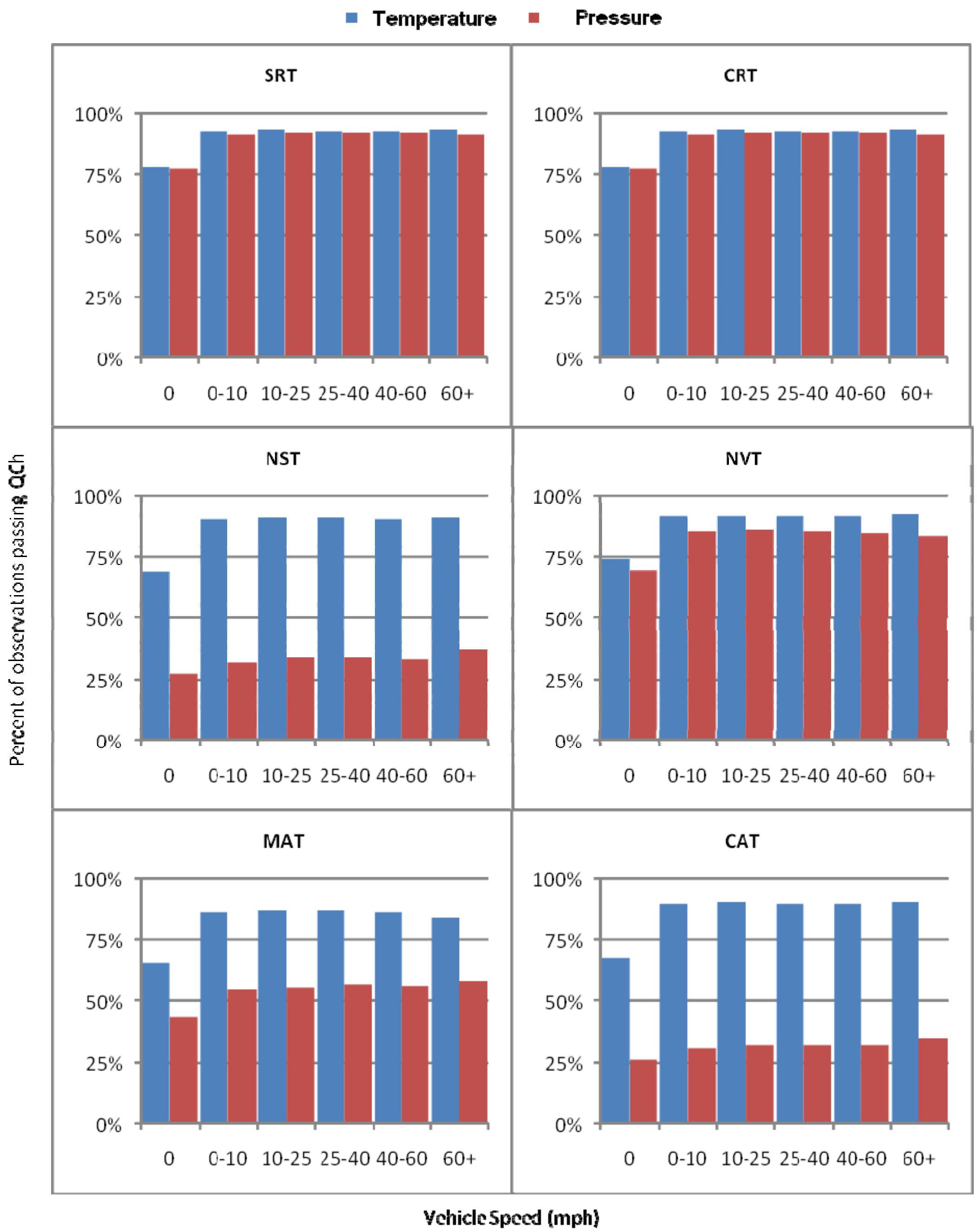


Figure 10. Percentage of DTE09 observations passing QCh test stratified by vehicle speed.

The presence or absence of precipitation, as defined by 1 minute KDTW ASOS data matched with vehicle observations, also influences the QCh pass rates for temperature and pressure (Figure 11). Excluding the MAT, pass rates for the other tests are lower during precipitation than non-precipitation conditions. The p-values for these differences between precipitation and absence of precipitation are

0.012 for temperature and 0.745 for pressure, indicating that the difference in the pass rates is not statistically significant. In these graphs, there are no snow days, so all precipitation is liquid.

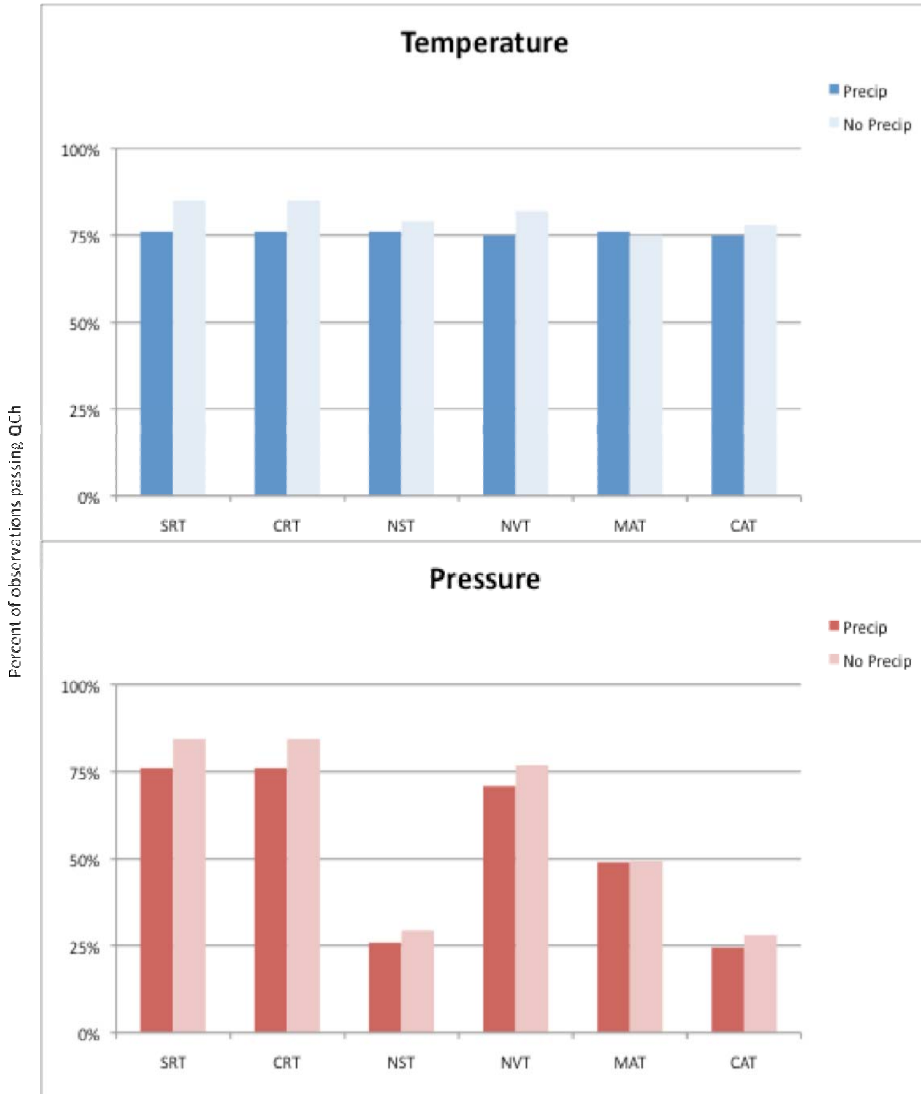


Figure 11. Percentage of DTE09 observations passing QCh test stratified by precipitation class (yes/no).

Lastly, results based on stratification of the KDTW temperatures are inconclusive (Figure 12). There is some indication that with higher temperatures, the NST, MAT, and CAT pass rates for temperature are lower, but this is at least partly related to the presence of null values and is not statistically significant.

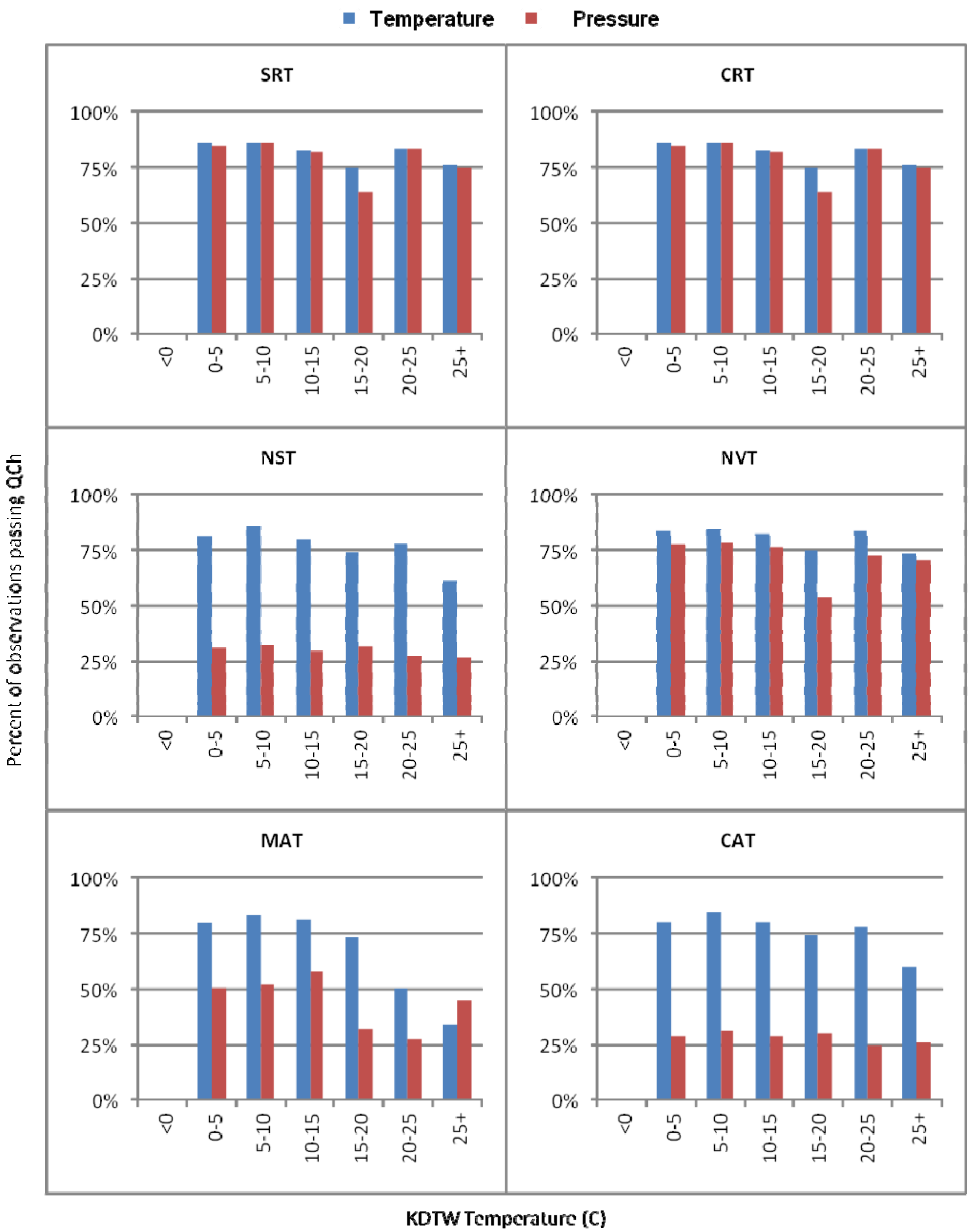


Figure 12. Percentage of DTE09 observations passing QCh test stratified by KDTW temperature.

3.1.2. PoC

The set of PoC vehicles is a superset of the set of DTE09 vehicles, and it is therefore not surprising that the QCh patterns for the two data sets mirror one another. Overall, the PoC QCh tests also reveal that vehicle-based temperature measurements are superior to pressure (Figure 13). Over 80% of the temperature observations passed the QCh tests, but less than 50% of the pressure observations passed the NST, MAT, or CAT.

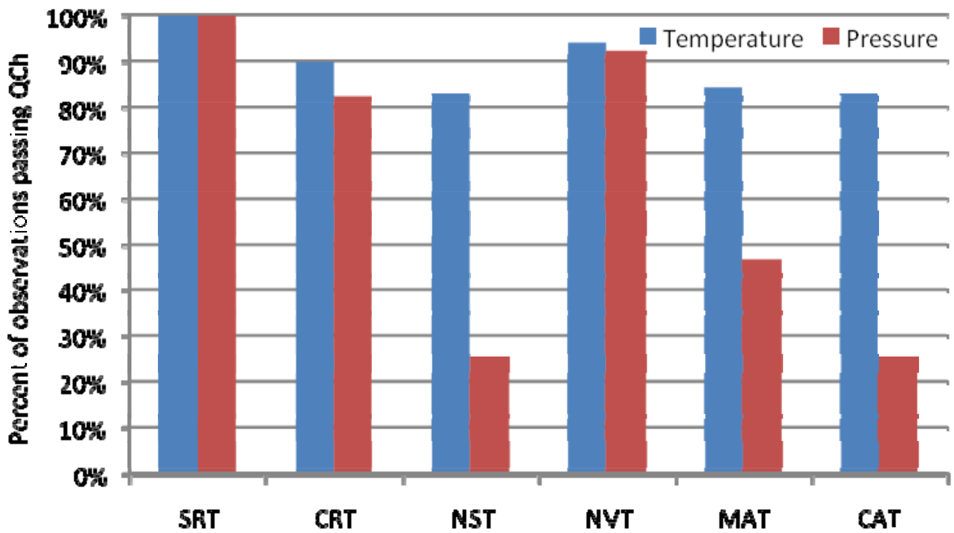
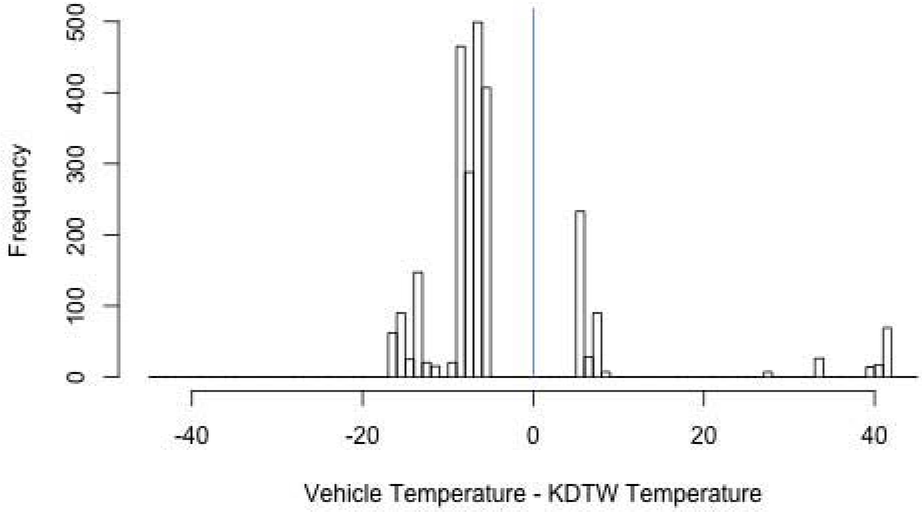


Figure 13. Percentage of PoC observations passing QCh tests.

Following the error analyses from DTE09, the CAT QCh failures for PoC temperature and pressure are further investigated in Figure 14. The pressure failures mirror the DTE09 measurements, with the vehicle recording much lower values. The temperature pattern is also similar to DTE09, with vehicle measurements generally being colder than KDTW temperatures.

The variation in pass rates by day observed in DTE09 also appears in PoC (Figure 15). Throughout the month of August, there is day-to-day variability in the temperature CAT QCh pass rates, but for September the QCh pass rates increased noticeably, which may correspond to the increased QCh pass rates observed with cooler temperatures discussed later in this section (Figure 19). The pressure CAT QCh results are less obvious because the failure rate is high on most days.



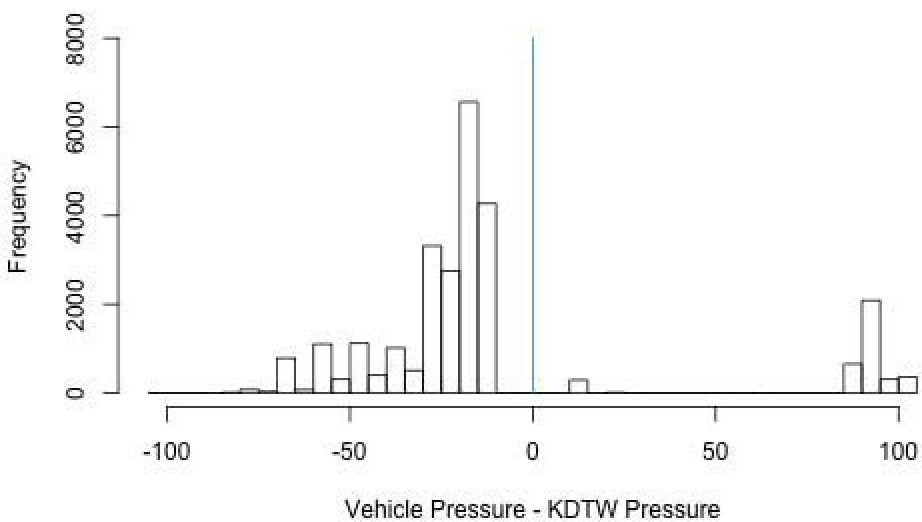


Figure 14. Scatter plots of CAT-failed PoC temperature and pressure observations.

Separating the data by time of day shows results roughly similar to DTE09 for temperature, with a slight decrease in the pass rate after 12 UTC (Figure 16, Universal Time Coordinate (UTC) is Local Standard Time +05 hours), a local maximum around 18 UTC, and then another slight decline. The pressure data are not conclusive. Excluding a 100% pass rate for 09–10 UTC, the pass rates are uniformly low. The 100% pass rate may be related to there only being 46 observations collected during that time period.

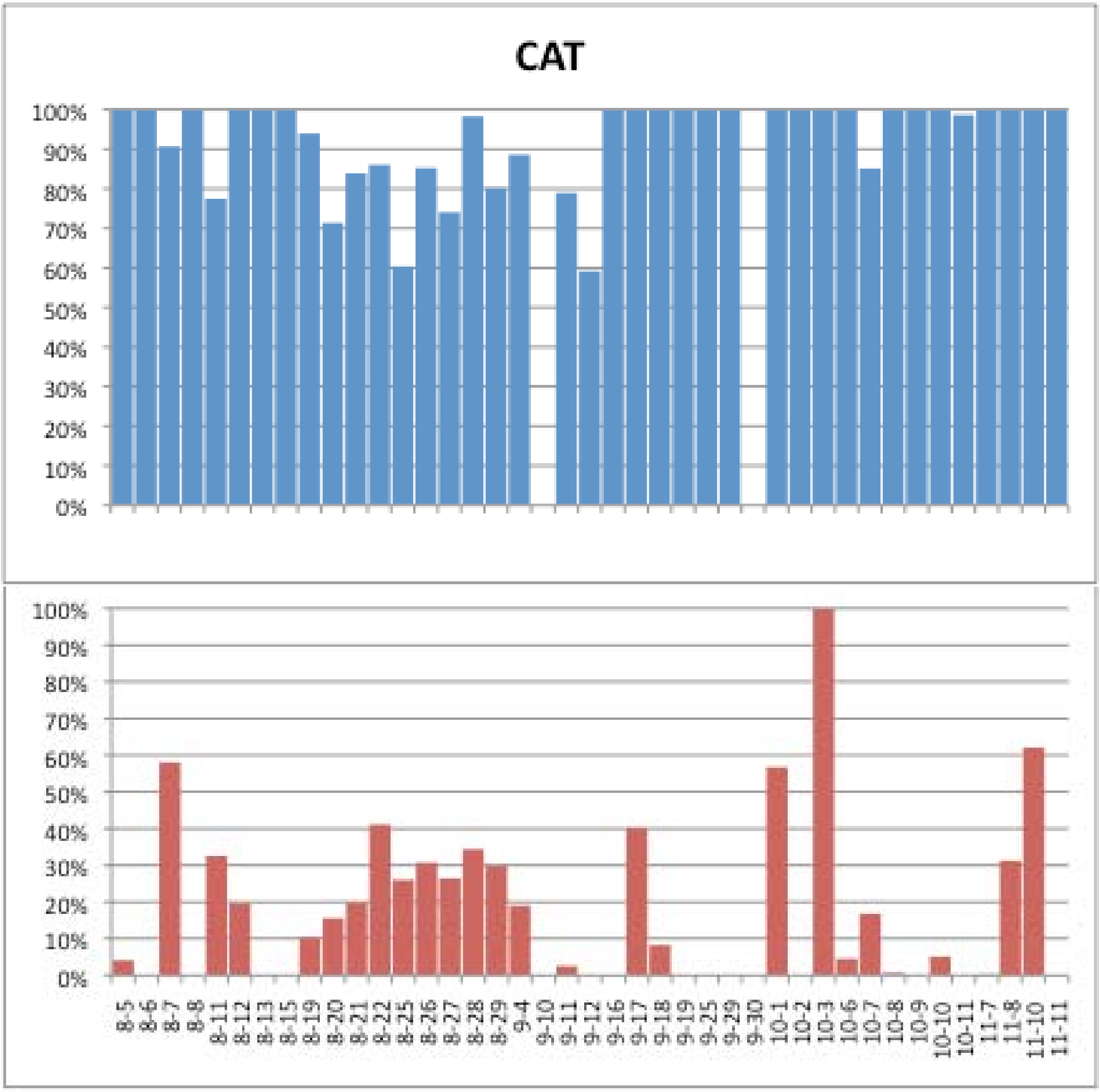


Figure 15. Percentage of PoC observations passing CAT QCh test for temperature (top) and pressure (bottom) stratified by day.

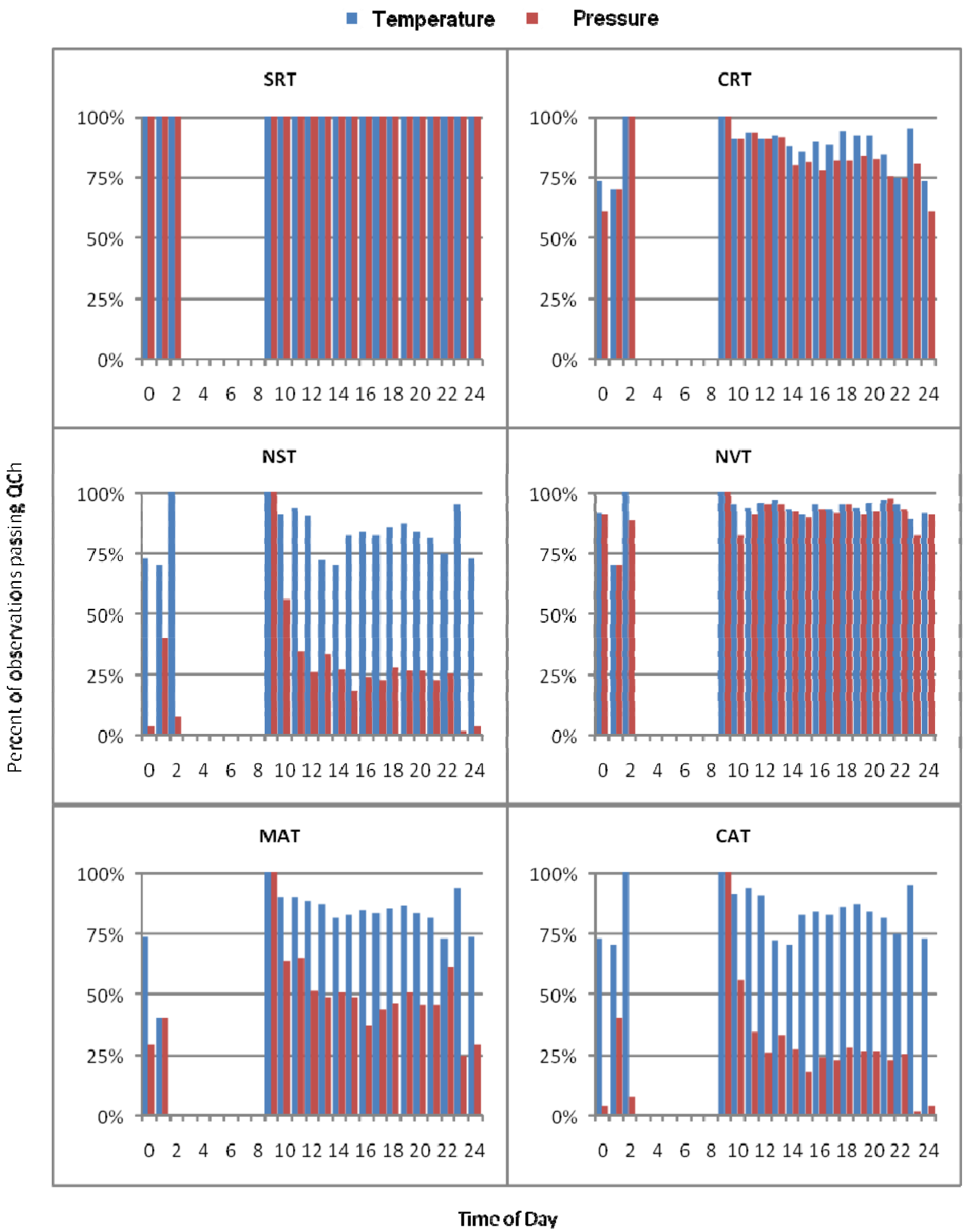


Figure 16. Percentage of PoC observations passing CAT QCh test stratified by time of day.

For PoC, vehicle speed appears to have no effect on the QCh pass rates (Figure 17). For every test, the pass rates hardly vary no matter what vehicle speed is used. In comparison, there is evidence to suggest that precipitation increases the QCh pass rates, opposite to the DTE09 findings (Figure 18). The exception to this exists in the pressure data, where the NST, MAT, and CAT all have lower pass rates in precipitation. As with DTE09, there is some indication that with higher temperatures, the NST, MAT, and CAT pass rates for temperature are lower (Figure 19).

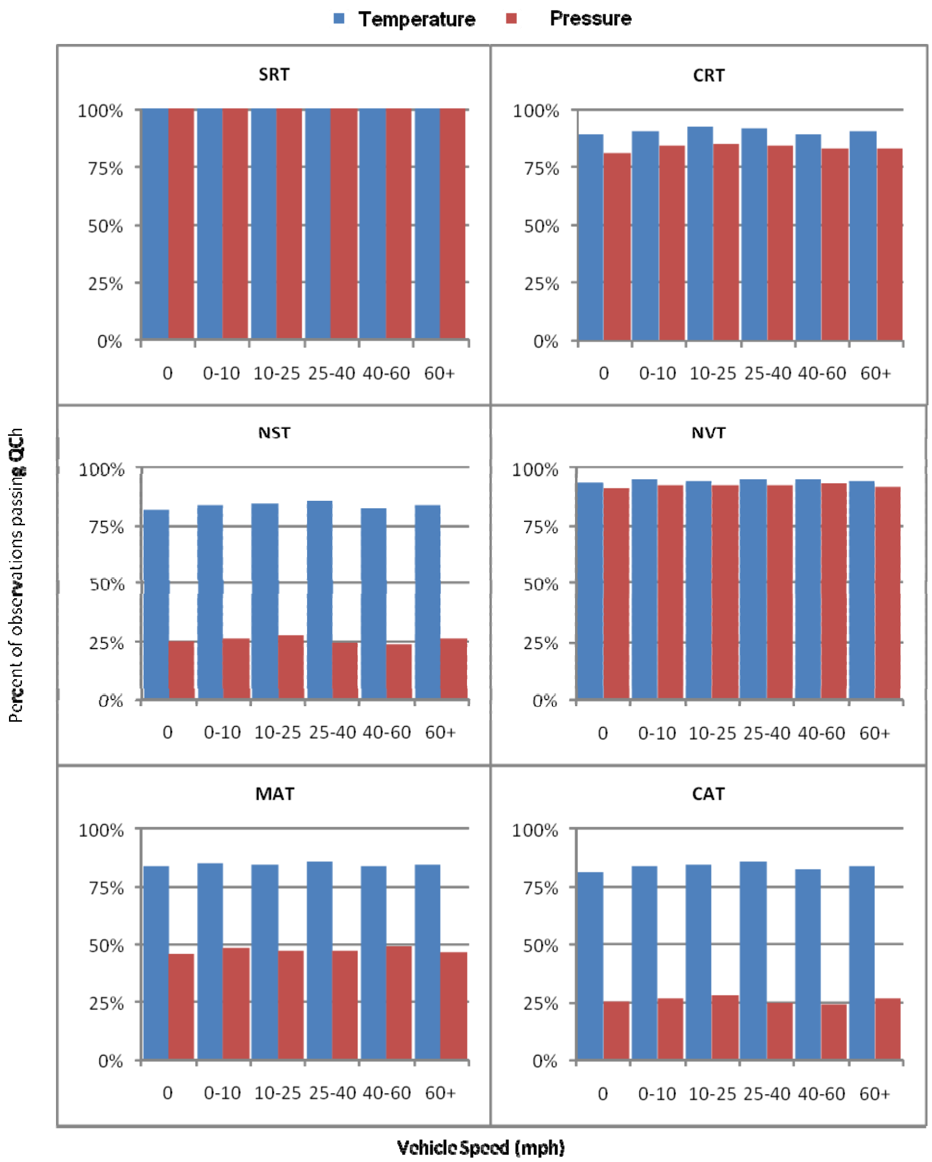


Figure 17. Percentage of PoC observations passing CAT QCh test stratified by vehicle speed.

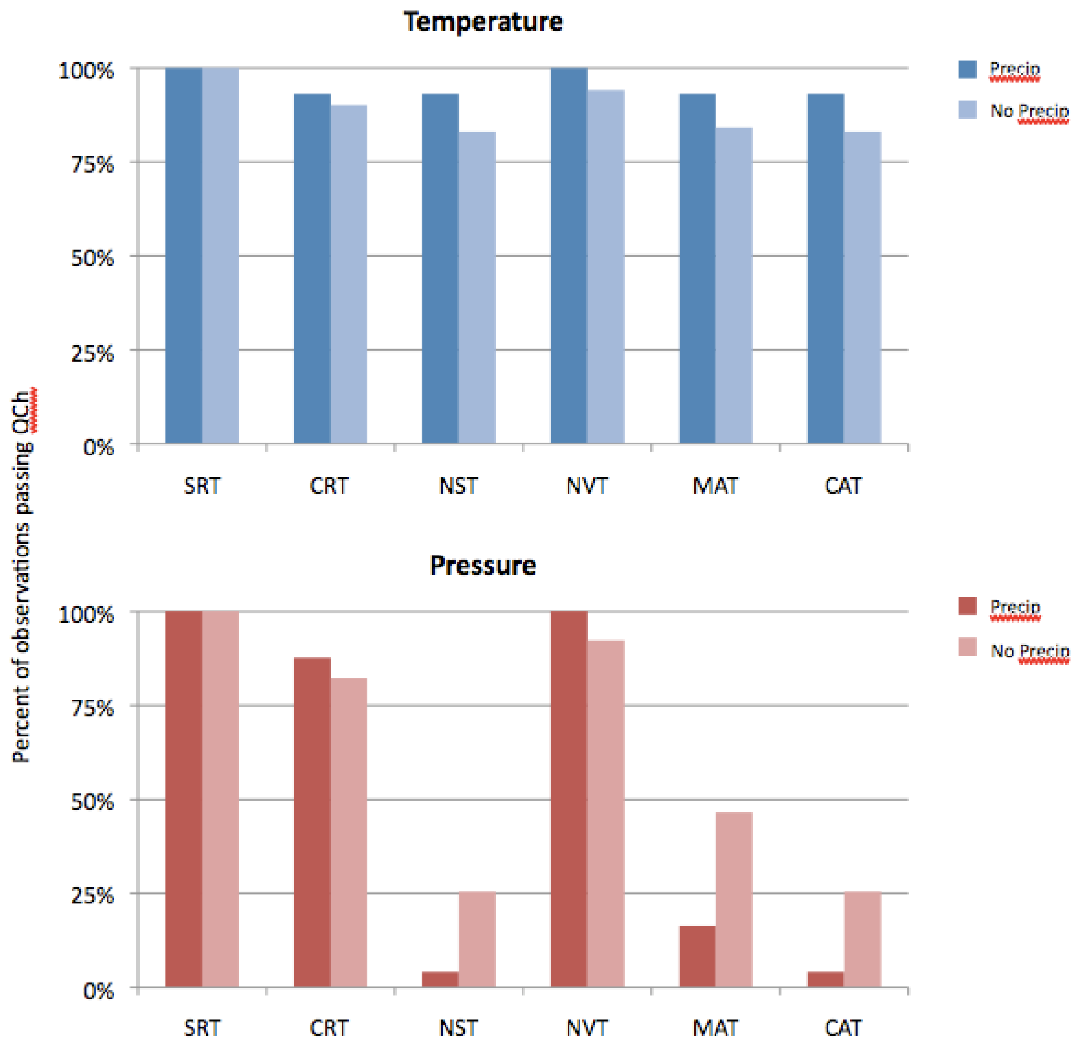


Figure 18. Percentage of PoC observations passing CAT QCh test stratified by precipitation class.

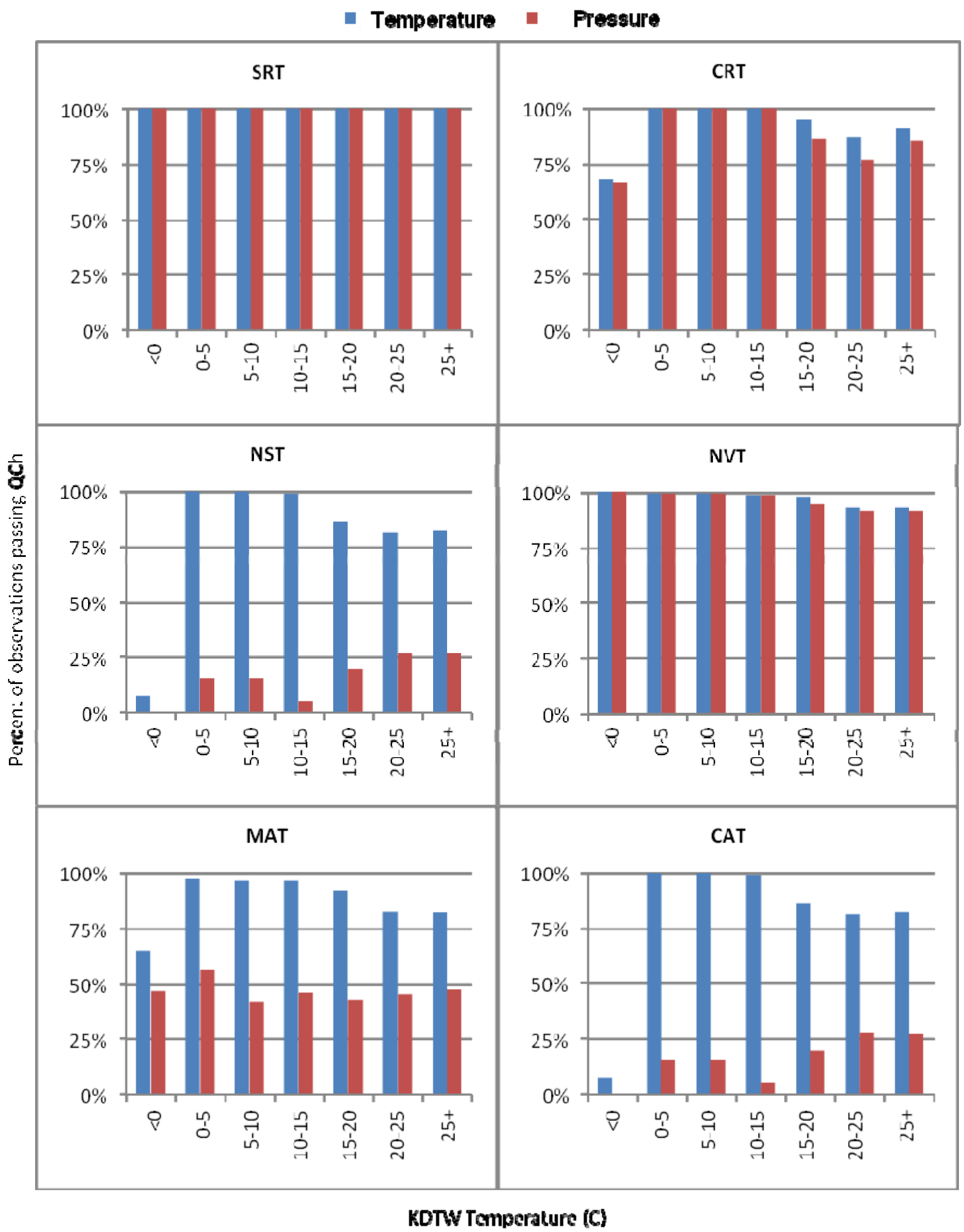


Figure 19. Percentage of PoC observations passing CAT QCh test stratified by temperature.

3.1.3. DUAP

Quality checking results for the DUAP data share some similarities to the DTE09 and PoC data. The overall pass rates for temperature are generally higher than pressure, especially for the SRT, CRT, and NVT (Figure 20). Unlike the DTE09 and PoC data, the pressure data pass rate is higher for DUAP for the NST, MAT, and CAT tests. The CAT pass rate for DUAP temperature (76%) is similar to DTE09 (78%) and PoC (83%), but the DUAP pressure pass rate for CAT is much higher in DUAP (77%) than DTE09 (29%) or PoC (25%).

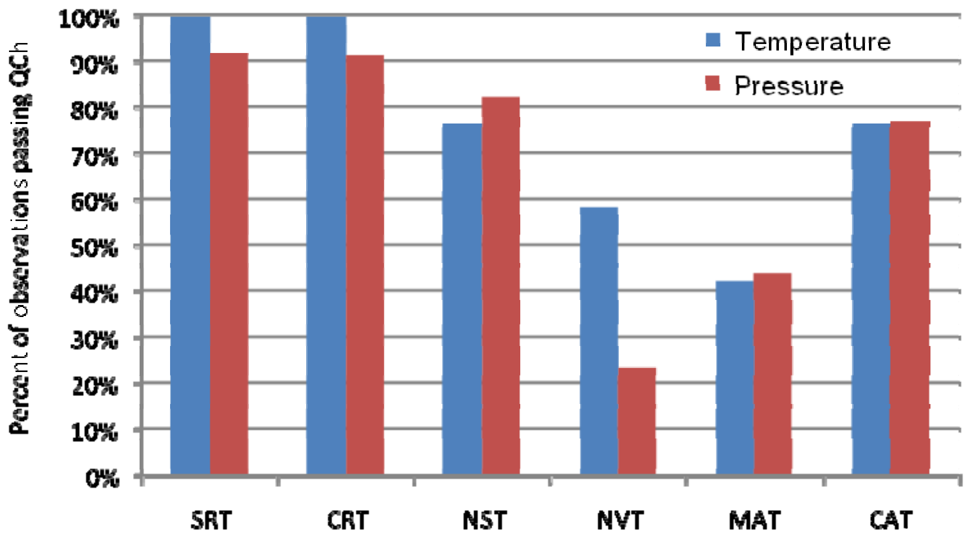


Figure 20. Percentage of DUAP observations passing QCh tests.

Looking at the failed CAT QCh observations for temperature and pressure (Figure 21), there is a tendency for the pressure failures to be lower than the KDTW pressure, similar to DTE09 and PoC. The temperature results suggest the DUAP failures tend to occur where the DUAP vehicle records a too-warm reading, which is opposite the trend for DTE09 and PoC temperature failures.

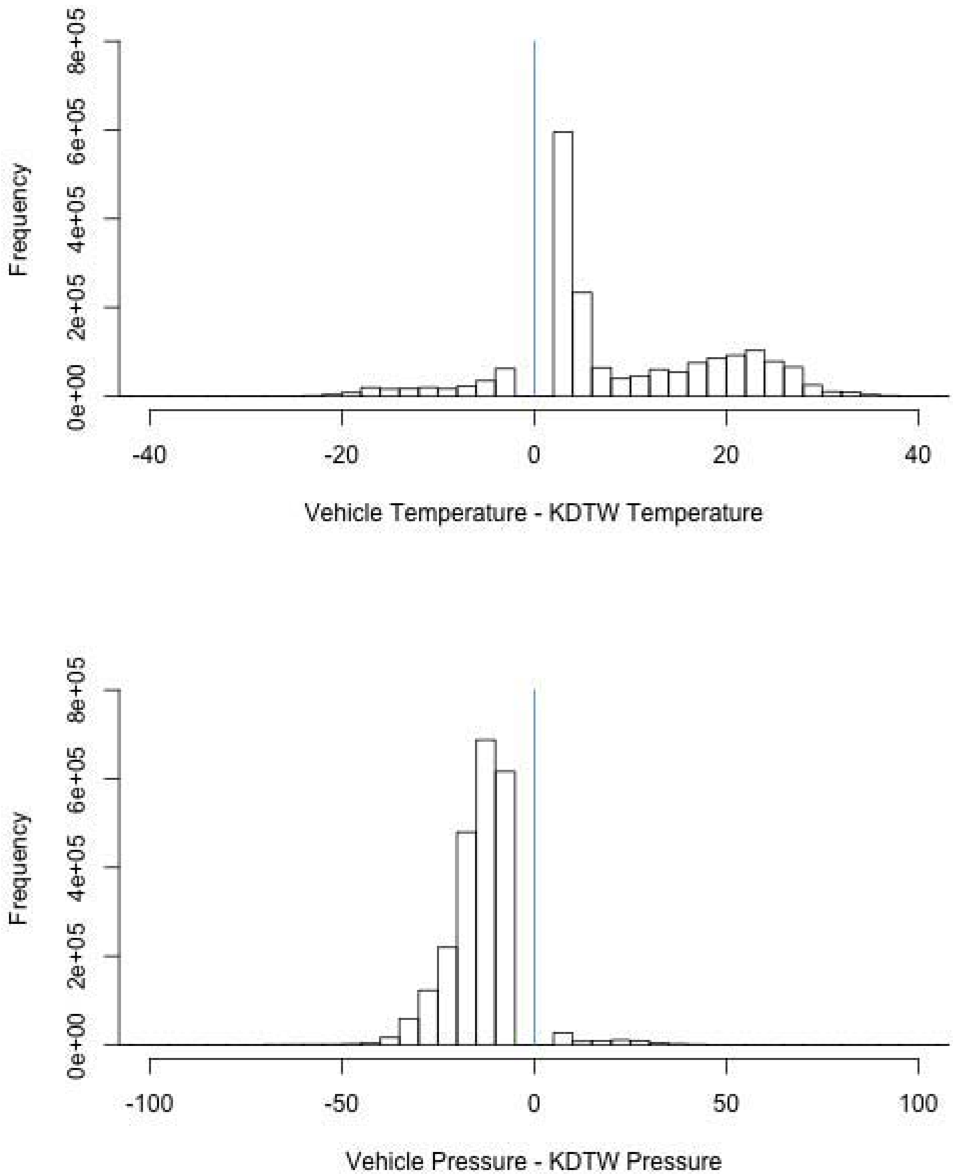


Figure 21. Scatter plots of CAT-failed DUAP temperature and pressure observations.

Stratified by day, there is little difference in the CAT pass rates for temperature or pressure. Excluding a low percentage day once every few months, the pass rate holds fairly constant at or above 90% for temperature and pressure (Figure 22). In comparison, there are slightly larger differences in the pass rates when subdivided by time of day (Figure 23). For the temperature CAT, pass rates are above

80% for a few hours around 12 UTC, slowly declining to around 70% by 24 UTC, before generally increasing again. The pressure CAT pass rate pattern mirrors temperature, but it is slightly less pronounced. Both of these trends are statistically significant with p-values less than 0.01.

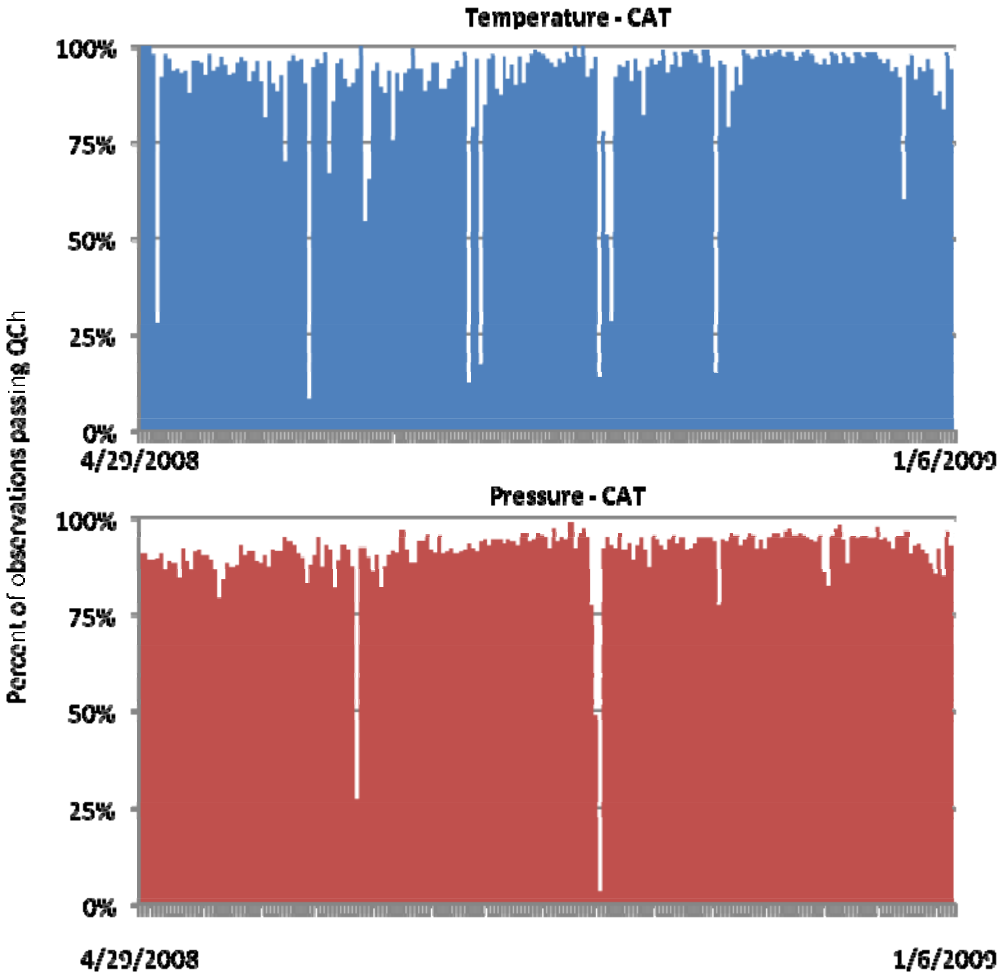


Figure 22. Percentage of DUAP observations passing CAT QCh tests for temperature (top) and pressure (bottom) stratified by day.

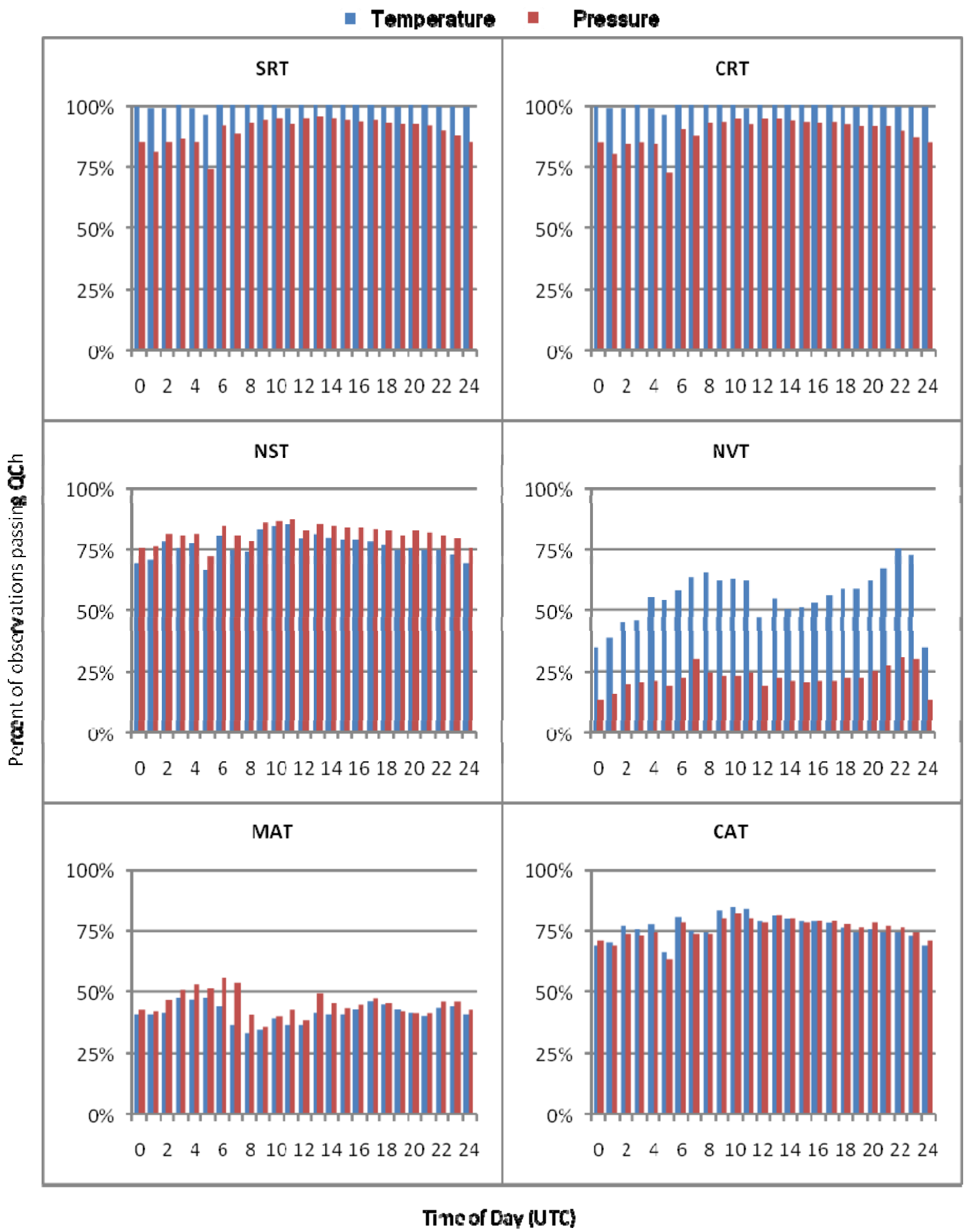


Figure 23. Percentage of DUAP observations passing QCh tests by time of day.

Vehicle speed appears to have no discernable influence on the pass rates for any test for either temperature or pressure (Figure 24). The largest differential in pass rate based on speed is only 3% (for temperature NST and CAT).

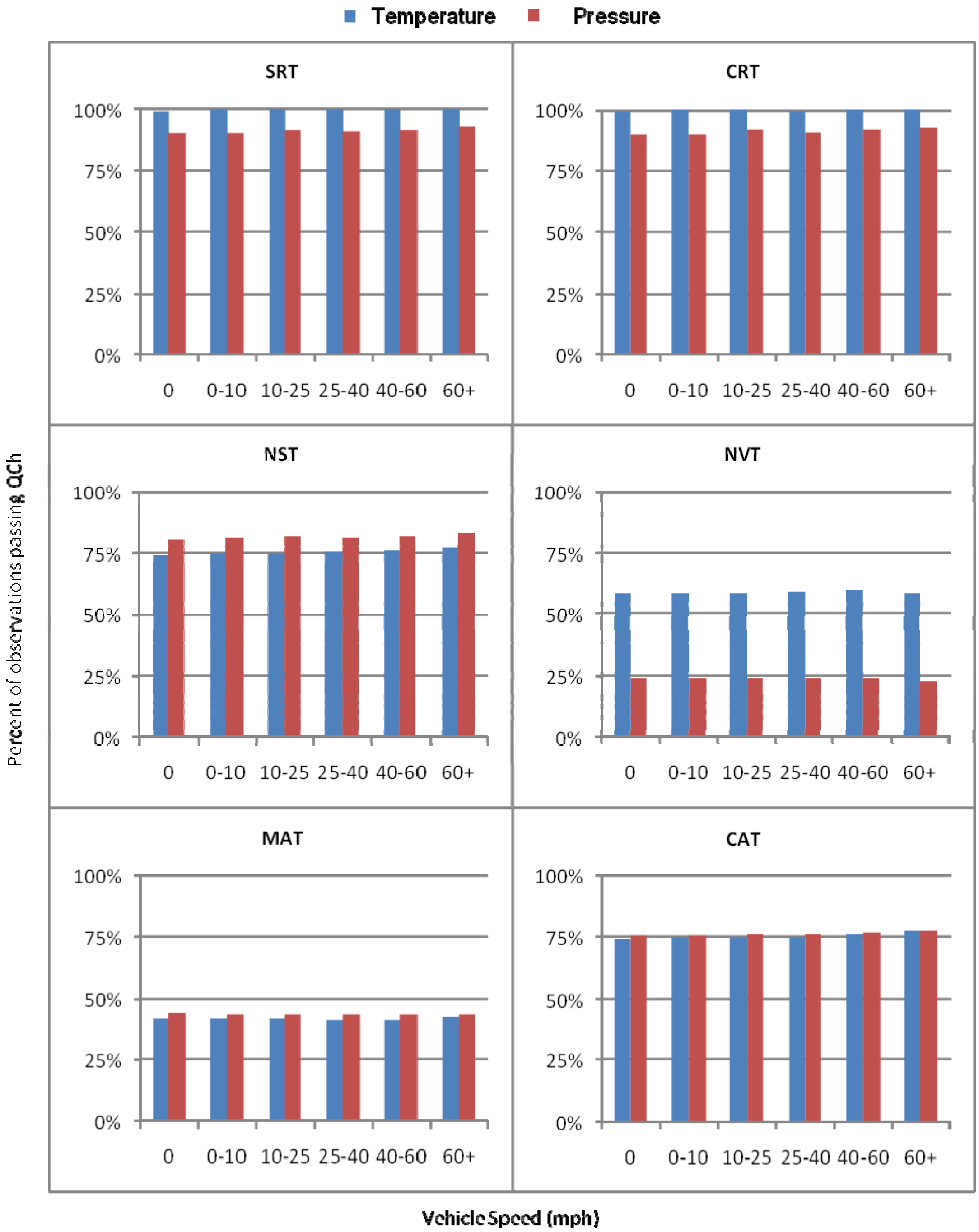


Figure 24. Percentage of DUAP observations passing QCh tests by vehicle speed.

In contrast to day, time of day, and vehicle speed, environmental conditions do affect the pass rate. For example, when the temperature is below 0°C, the pass rate for temperature and pressure is significantly less than when the temperature is above 0°C for the NST and CAT (Figure 25). There are also some indications that as temperatures increase past 15°C, the pass rate for temperature and pressure NVT and MAT begin to decline. The influence of precipitation is opposite for DUAP as compared to DTE09 (Figure 26). In DUAP, pass rates are higher for NST, NVT, MAT, and CAT for both vehicle temperature and pressure measurements in the presence of precipitation.

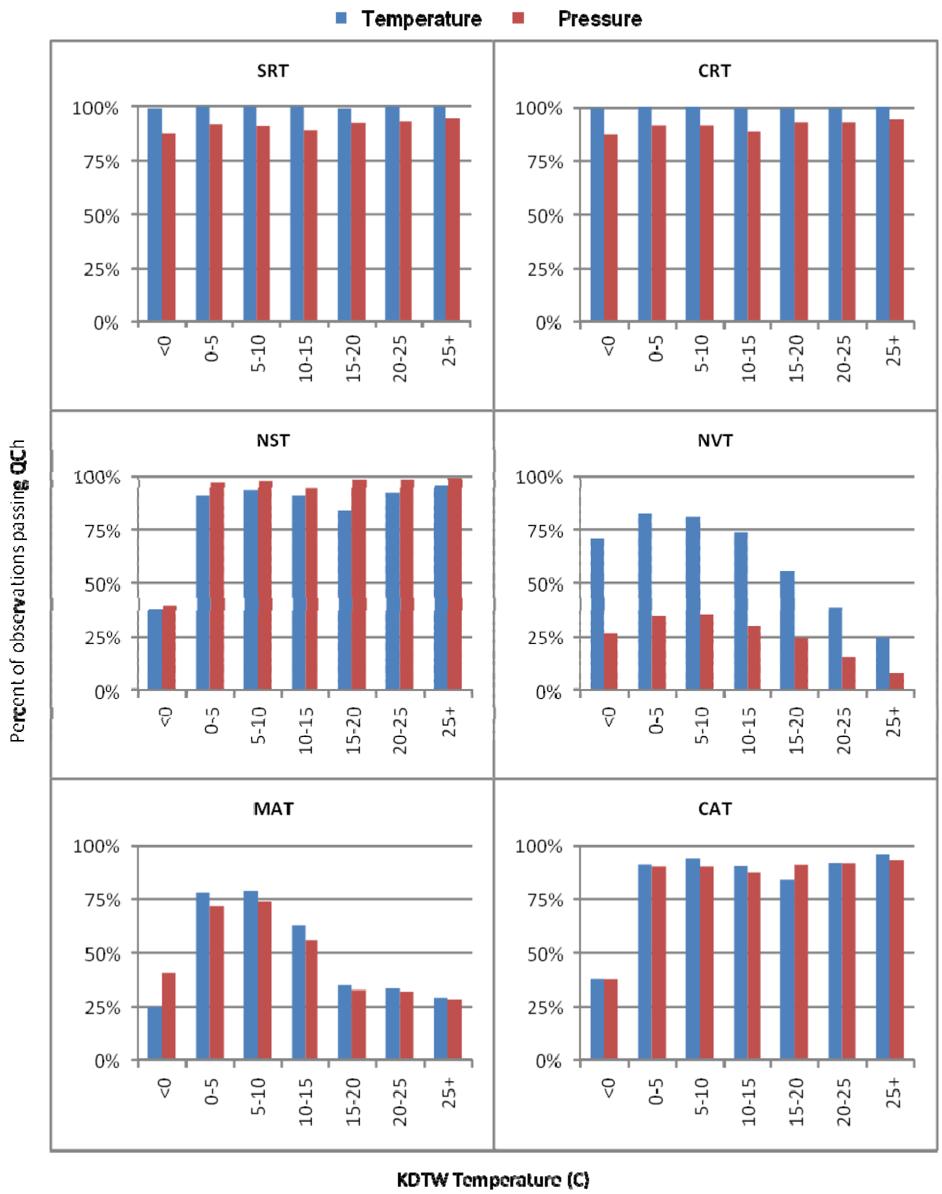


Figure 25. Percentage of DUAP observations passing QCh tests by KDTW temperature.

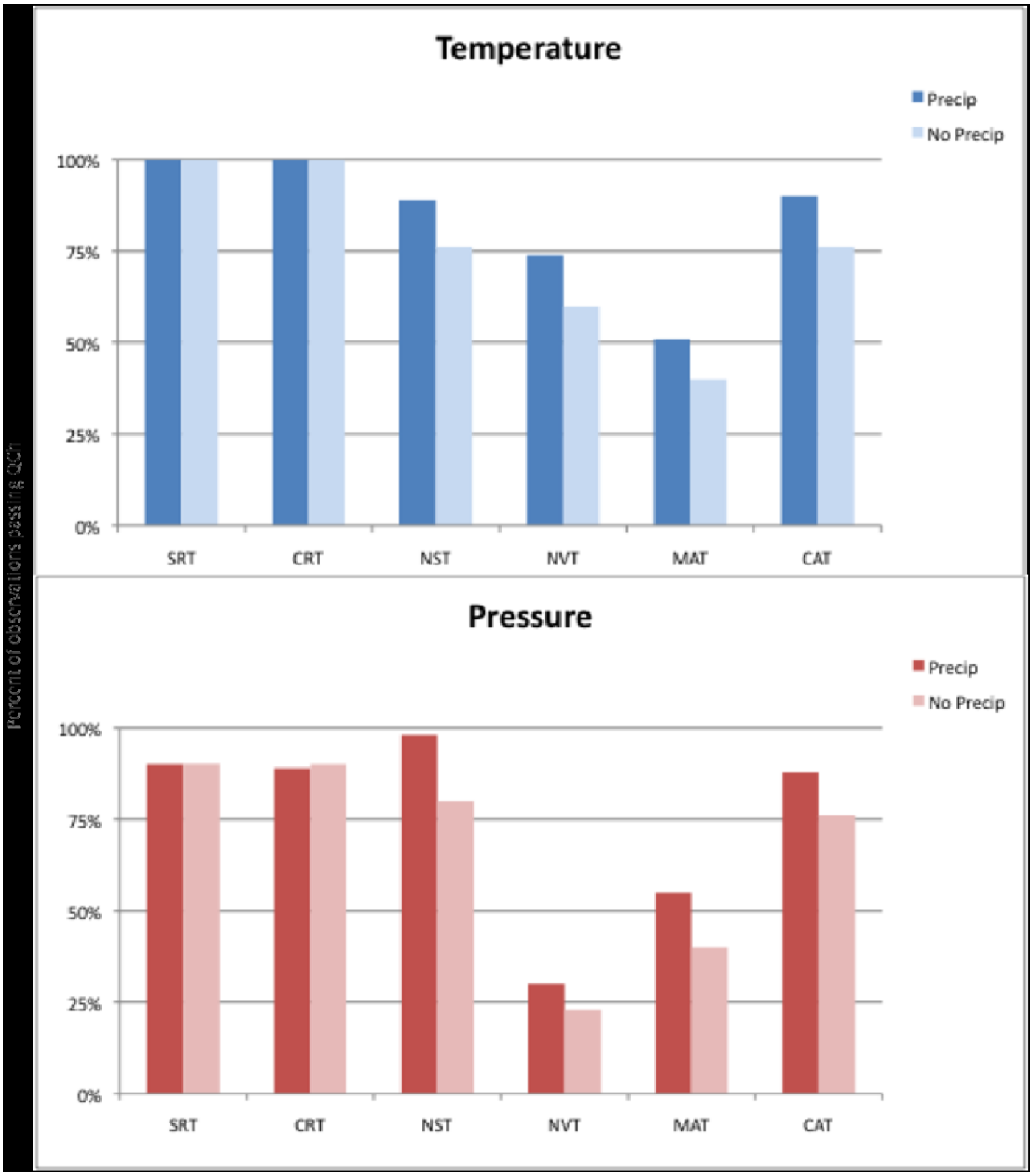


Figure 26. Percentage of DUAP observations passing QCh tests by precipitation class (yes/no).

3.2. Post-Quality Check (QCh) Analysis

In this section, we examine the accuracy and bias of the temperature and pressure data that have passed the SRT, CRT, and NST QCh tests in the last section (i.e., the CAT test). The analysis for accuracy uses the correlation coefficient, the mean absolute error (MAE), bias, and the Student's T-Test for paired observations. The correlation coefficient measures the strength of the linear relationship between two independent variables. The MAE is used to measure how close the variable being observed is to a comparison, ground truth observation. Bias is a measurement of how far the observation over- or underestimates the observation it is being compared with. The T-Test is used to test if there is a significant difference between the mean of the samples being compared. With one exception (April 27, Ford 2 temperature vs. KDTW temperature), all results of the T-Test were the same, and indicated that the vehicle data and the KDTW data were statistically different from each other.

When comparing the temperature statistics for the three (e.g. DTE09, PoC, and DUAP) field study periods (Figure 27), the results show a small negative bias for the PoC data (Bias=0.12) and larger bias results for the DTE09 (Bias=-0.95) and DUAP (Bias=1.26). The results show good positive correlation for the DTE09 and DUAP (0.96 and 0.99, respectively), but little correlation in the PoC (0.20) data. MAE results are similar for all three data sets with values ranging from 1.37 to 1.78°C.

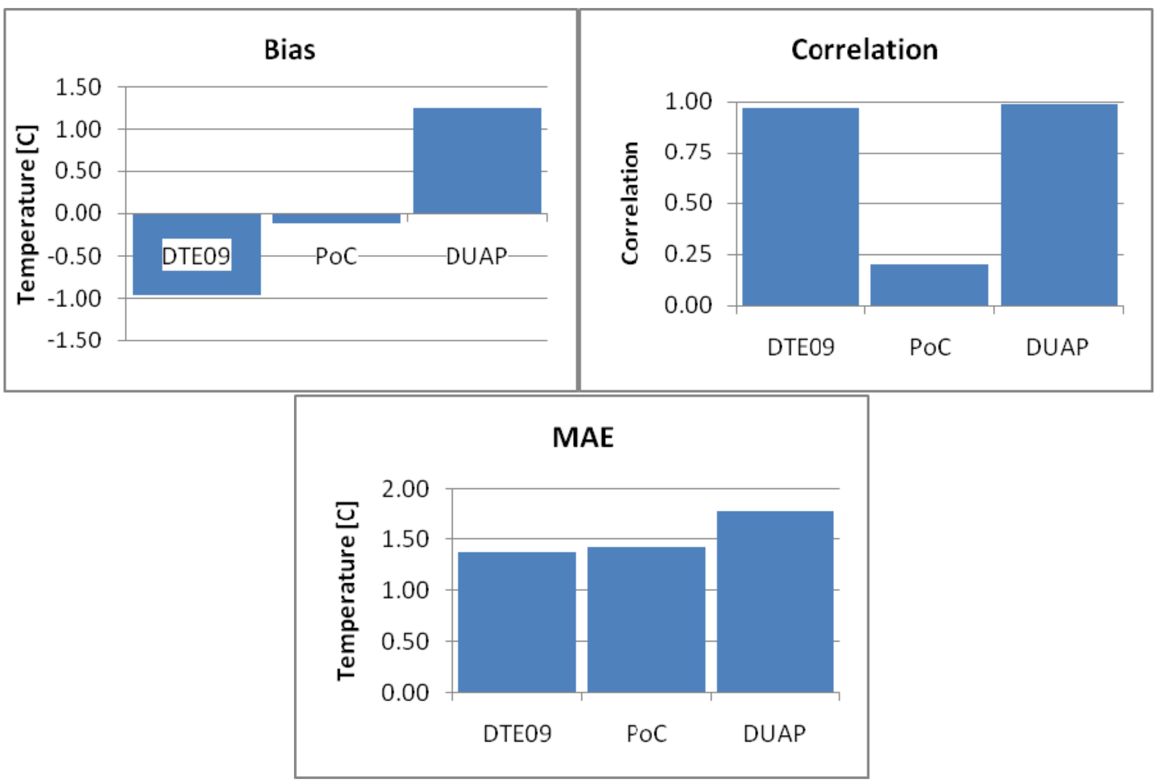


Figure 27. Statistics (Bias, Correlation, and MAE) for the DTE09, PoC, and DUAP field studies for the temperature variable.

The results for the pressure variables (Figure 28) are much less favorable than the results discussed for temperature. All three data sets have a large negative bias with the DUAP data showing the best results (bias of -3.74 mb). All three data sets show positive correlation with values ranging from 0.63 (PoC) to 0.93 (DTE09). Error results are also poor for all three datasets and once again the DUAP (MAE=4.42 mb) stands out as being the most accurate. These results are likely due to the lack of precision (10-mb resolution) in the pressure measurements for the DTE09 and PoC datasets, with the DUAP dataset's better accuracy likely due, in part, to its 2 mb resolution.

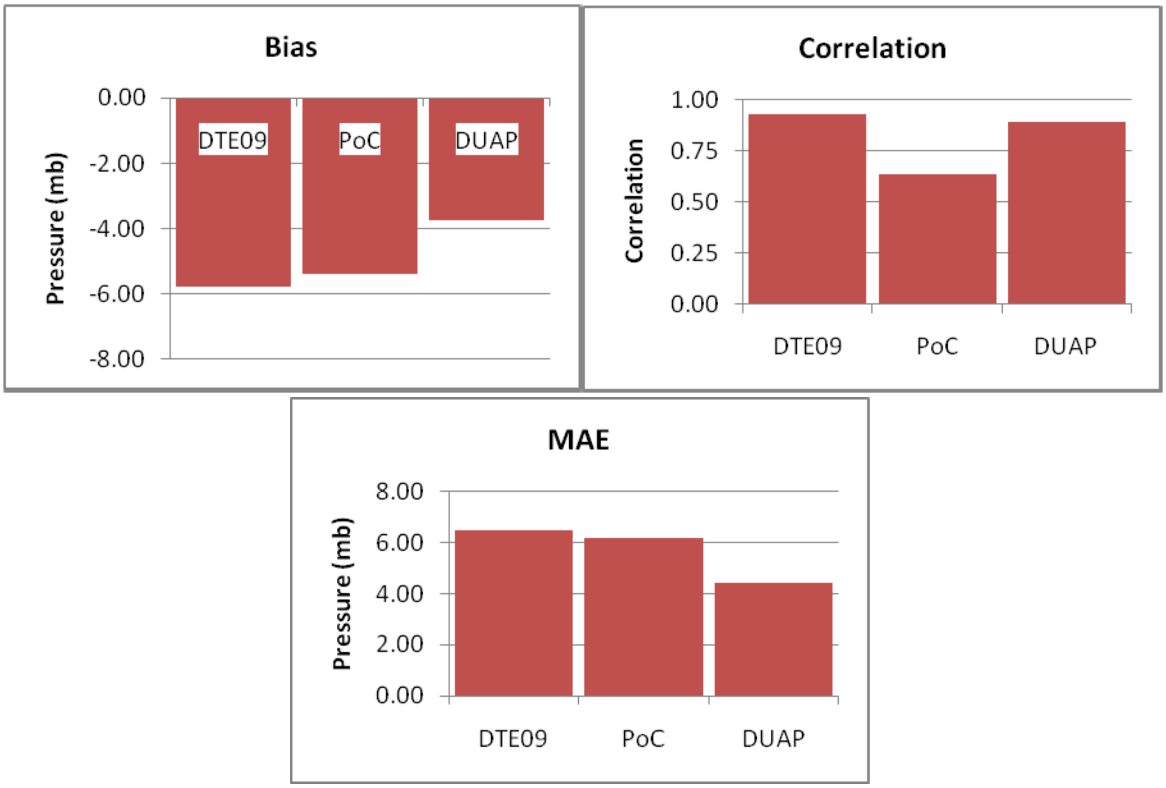


Figure 28. Statistics (Bias, Correlation, and MAE) for the DTE09, PoC, and DUAP field studies for the pressure variable.

The results are stratified (in the following subsections) as follows: hour, day, vehicle identifier (for DTE09 only), precipitation conditions, speed category and temperature category.

3.2.1. DTE09

Results for the DTE09 experiment are discussed in this subsection.

The first stratification for DTE09 was by hour of day (Figure 29 and Figure 30). The results for temperature (Figure 29) are consistent (bias less than 0; correlation = 0.88 to 0.99; MAE = 1.07 to 1.92) over a majority of the hours with the exception of the late afternoon (2200 to 00 UTC; bias = -0.08 to 0.48, correlation = -0.08 to 0.09, MAE = 0.61 to 1.03). This is likely because the sample sizes are much smaller and less variable towards the end of the testing day and the vehicles were idle but still recording data for several minutes upon return to the Roush facility in order for data to be downloaded from the OBE.

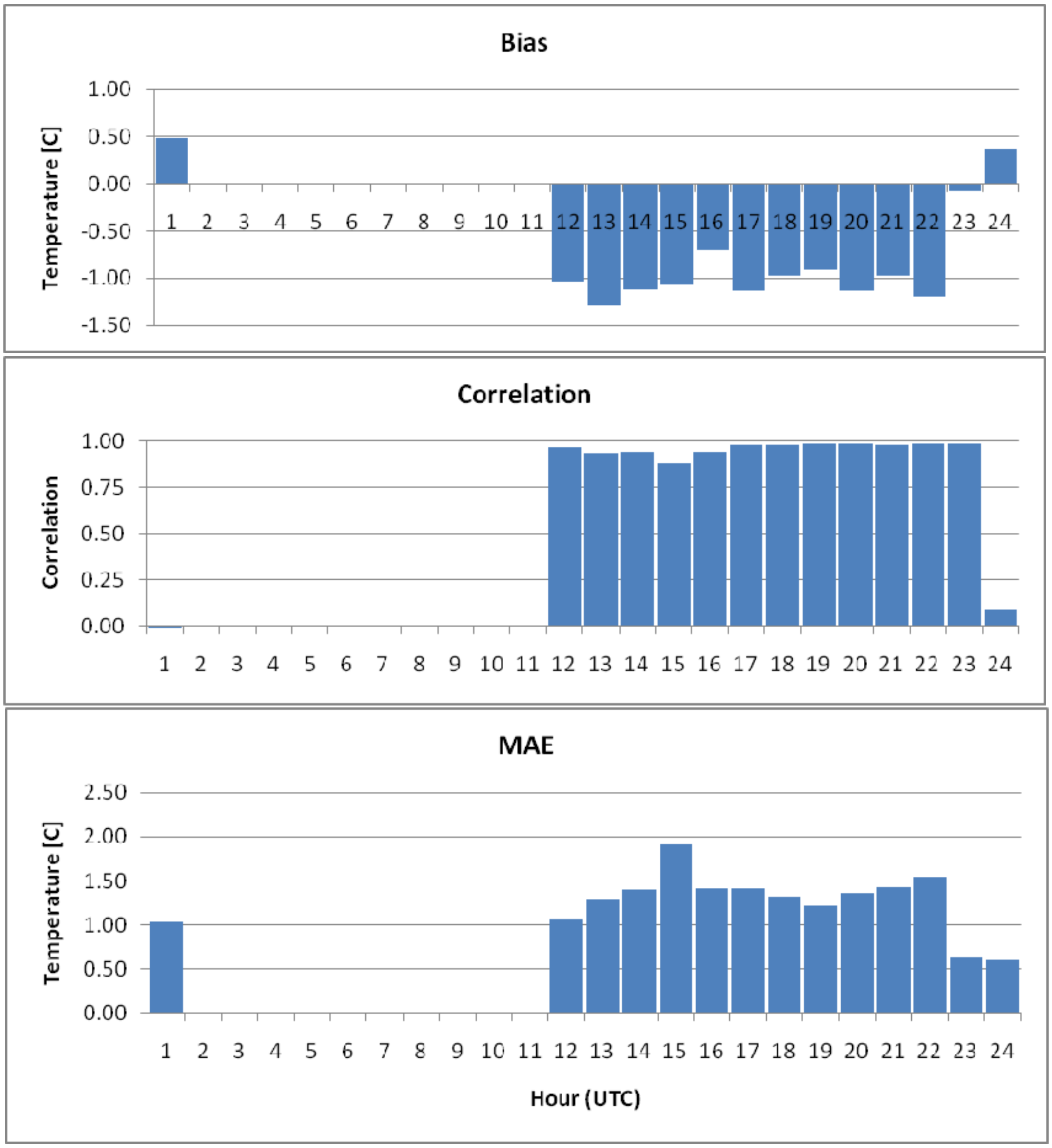


Figure 29. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the temperature variable stratified by hour of day (UTC).

The results for pressure are also consistent with good correlation over most of the time period (with the exception of 00 UTC), but poor results for both bias (bias ~ -4 to -9) and MAE (MAE ~ 6 to 9). These results are also likely due to the lack of precision for the pressure variable.

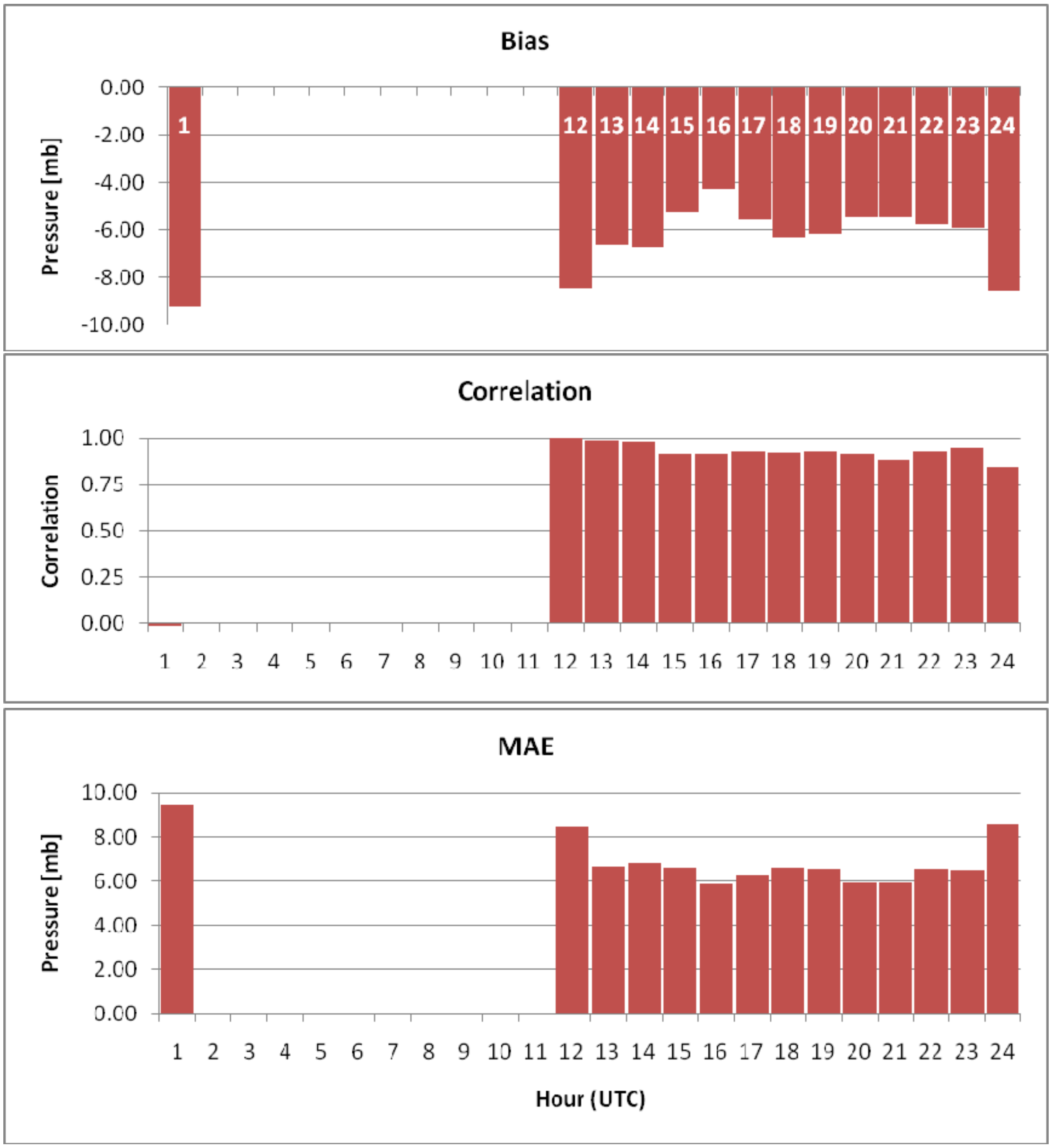


Figure 30. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the pressure variable stratified by hour of day (UTC).

When stratifying the DTE09 data over different speed categories (e.g. 0, >0 – 10, >10 – 25, >25 – 40, >40 – 60, and >60 mph; Figure 31) the results are favorable and remarkably consistent for the three statistics (bias = -.80 to -1.10, correlation=0.96 to 0.98, and MAE=1.32 to 1.41)

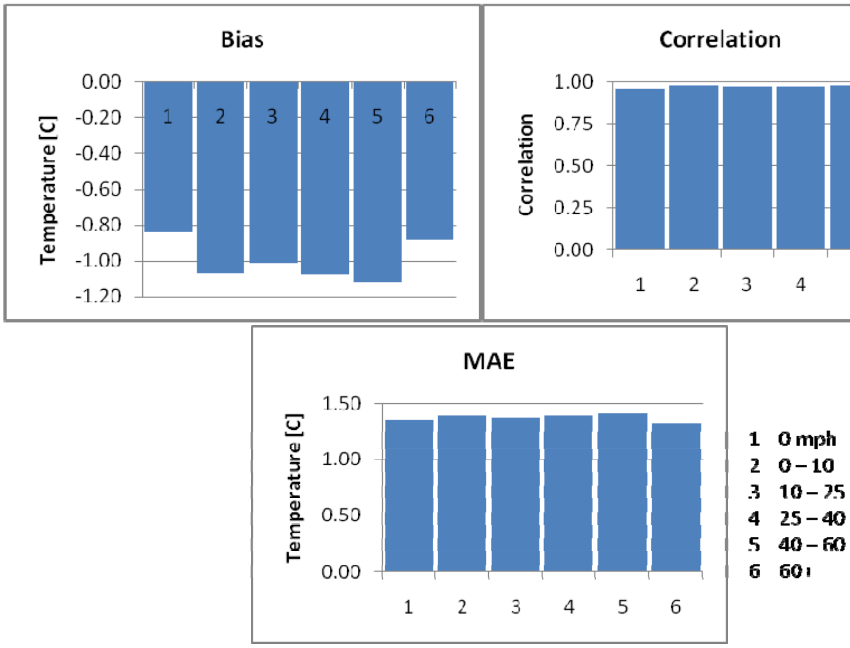


Figure 31. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the temperature variable stratified by speed.

For the same stratification by speed (as discussed previously), the pressure variable again shows consistency but overall the results are much like the hourly statistics. The data has high correlation but the bias (~-5.5 to -6) and MAE (~6 to 6.75) show that the pressure data is less accurate.

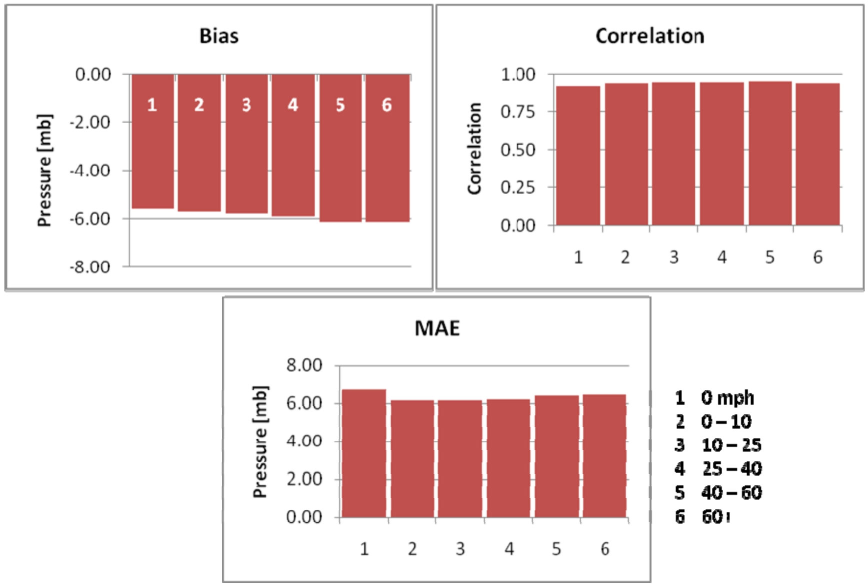


Figure 32. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the pressure variable stratified by speed.

When the data are stratified over temperature categories (e.g. <=0, >0–5, >5–10, >10–15, >15–20, >20–25, and >25°C; Figure 33), the results show some variability. The temperature shows minimal bias (<-1) for the lower temperature ranges (from 0 – 15°C), but a spike in the bias (-2.75 and -3.25) and MAE (2.75 and 3.25) for the 15 – 25°C is evident. It is not apparent why this occurred at this time.

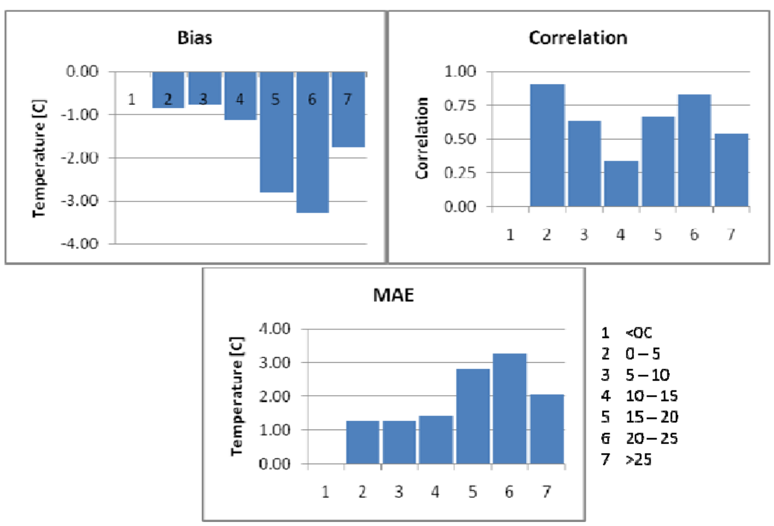


Figure 33. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the temperature variable stratified by temperature categories.

Pressure (Figure 34) is also less consistent for this analysis than for the previous two stratifications (hourly and speed). Correlation is good for the lower temperature ranges (e.g. 0 – 15°C) but drops off at the higher temperatures. Pressure again shows poor result for both the bias (~-5 ti -9.5) and MAE (~6 to 0), which is consistent with the previous two analyses.

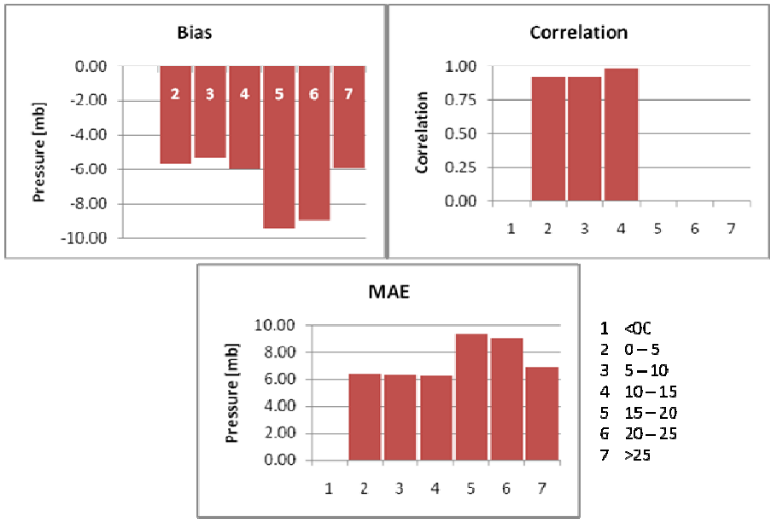


Figure 34. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the pressure variable stratified by temperature categories.

The DTE09 temperature data were also stratified by precipitation occurrence (Yes/No; Figure 35). There is little change in the temperature statistics due to the occurrence (or lack thereof) of precipitation. Both scatterplots show remarkable correlation and accuracy between the DTE09 vehicle data and the KDTW weather station.

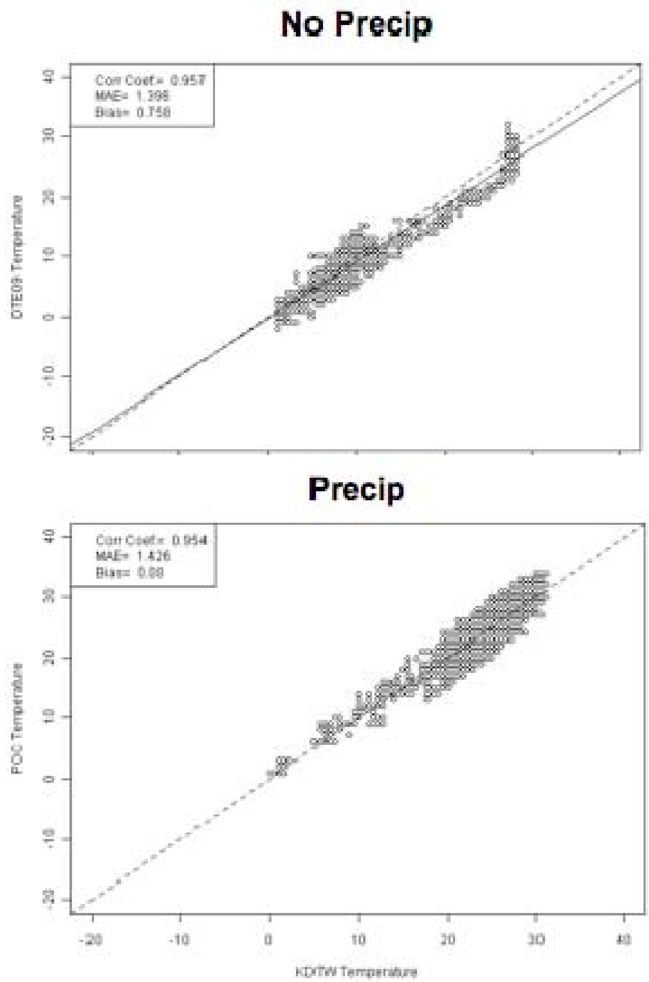


Figure 35. Scatterplot and statistics (Bias, Correlation, and MAE) for the DTE09 field study for the temperature variable stratified by precipitation occurrence. The dotted line represents a correlation of 1.0, the solid line shows the fit for the data.

The pressure data for DTE09 were also stratified by precipitation occurrence (Yes/No; Figure 36) and the results not only highlight that the pressure from the vehicles is consistently below that of the KDTW weather station, but also highlights the coarse resolution of the pressure data.

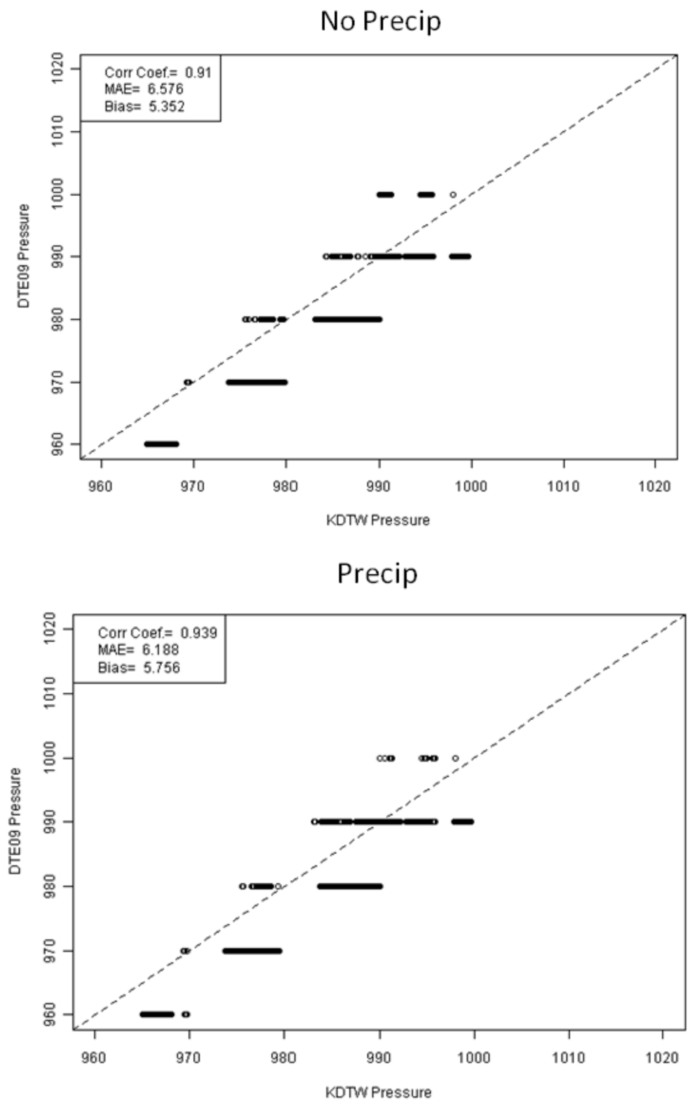


Figure 36. Scatterplot and statistics (Bias, Correlation, and MAE) for the DTE09 field study for the pressure variable stratified by precipitation occurrence.

The DTE09 dataset is unique in that (unlike PoC and DUAP) the individual vehicles could be parsed and the data stratified by individual or by groups of vehicles. In this analysis the vehicles are analyzed individually. The results show no discernable trend amongst the different makes/models of vehicles for temperature (Figure 37). Something to keep in mind though is that these data are quality controlled and in the QCh section the results showed that a significantly larger number of the temperature observations from the Ford Edges were eliminated when QCh tests were applied.

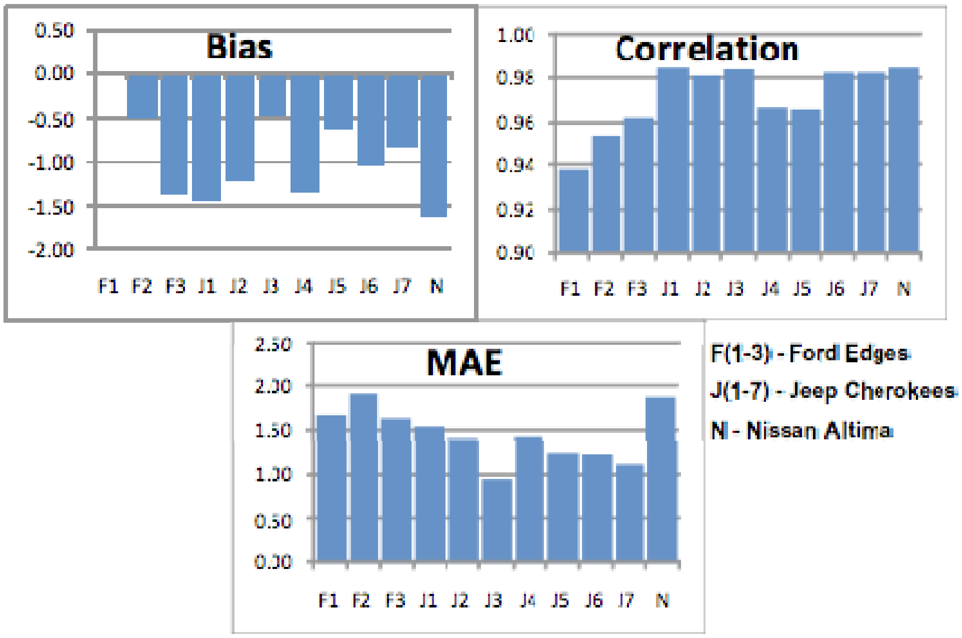


Figure 37. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the temperature variable stratified by individual vehicle.

The pressure results by individual vehicle (Figure 38) are consistent with all of the other stratifications, which showed high positive correlation but poor accuracy.

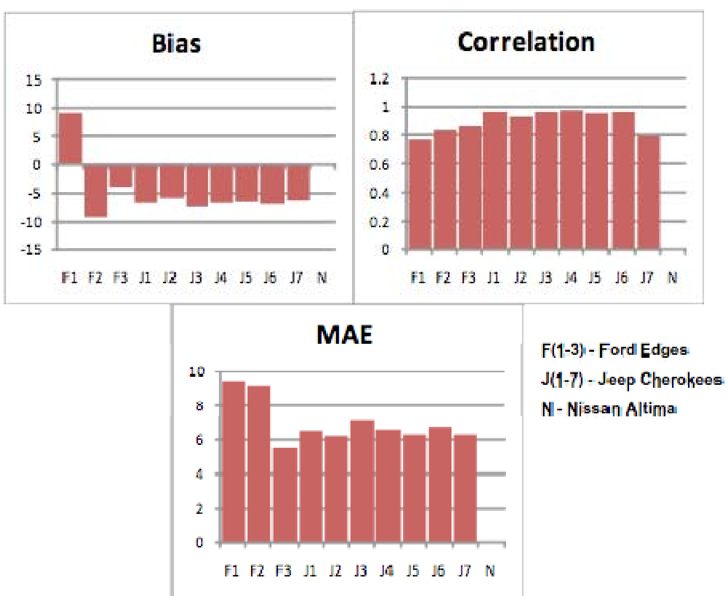


Figure 38. Statistics (Bias, Correlation, and MAE) for the DTE09 field study for the pressure variable stratified by individual vehicle.

In order to ascertain how well the data compares with the KDTW weather station, boxplots (as opposed to scatterplots) were generated in order to easily analyze the distribution of the temperature data per individual vehicle versus the KDTW weather station. Boxplots are a convenient way of depicting graphically the distribution of the data about the mean. On each of the boxplots in Figure 39 there are five number summaries (bottom horizontal line – sample minimum, lower end of box – lower quartile, middle line in box – median, upper end of box – upper quartile, and upper horizontal line – sample maximum).

Remarkable consistency is evident amongst the individual vehicles (Figure 39) with the slight negative Bias (previously discussed) apparent for most vehicles. Again, the results show no obvious difference amongst the different make/models of the vehicles for temperature. The one exception lies with the Nissan (N), which was driven by the Demonstration Test Director and experienced OBE problems throughout testing that most likely contributed to its wide interquartile range.

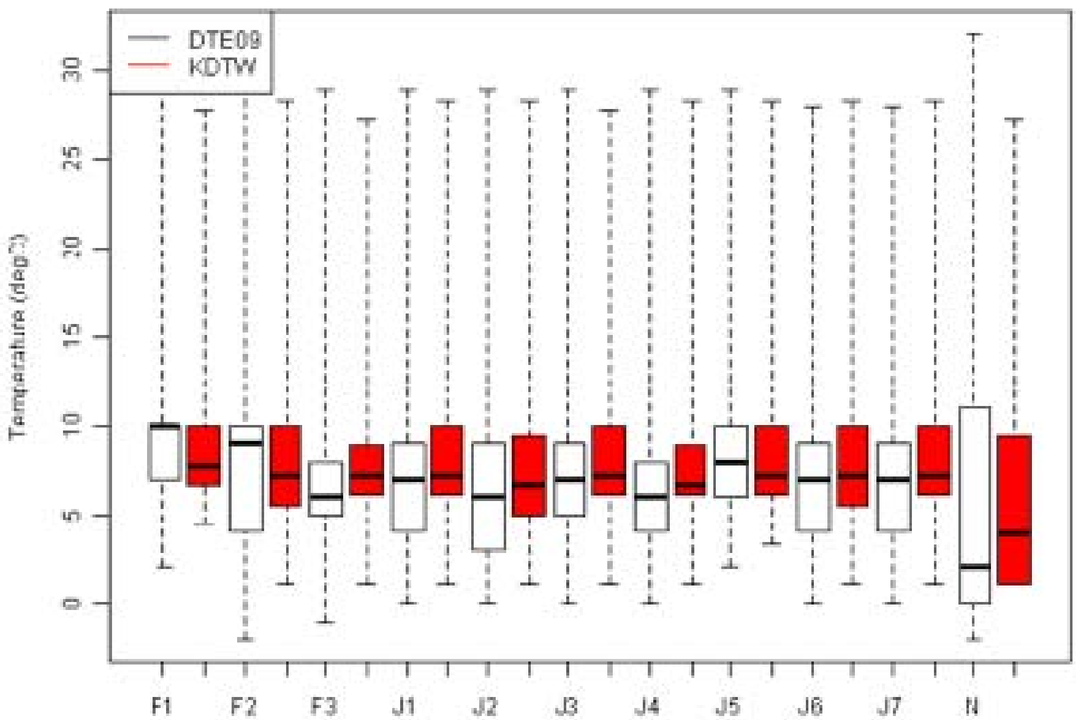


Figure 39. Boxplot for the DTE09 field study for the temperature variable stratified by individual vehicle.

The pressure variable results (Figure 40) also show concurrence with the previous analysis, which showed a large negative bias between the vehicles and the KDTW weather station. The results show a larger amount of variability for the Jeep Cherokees when compared with the Ford

Edges. Note that the Nissan did not report barometric pressure.

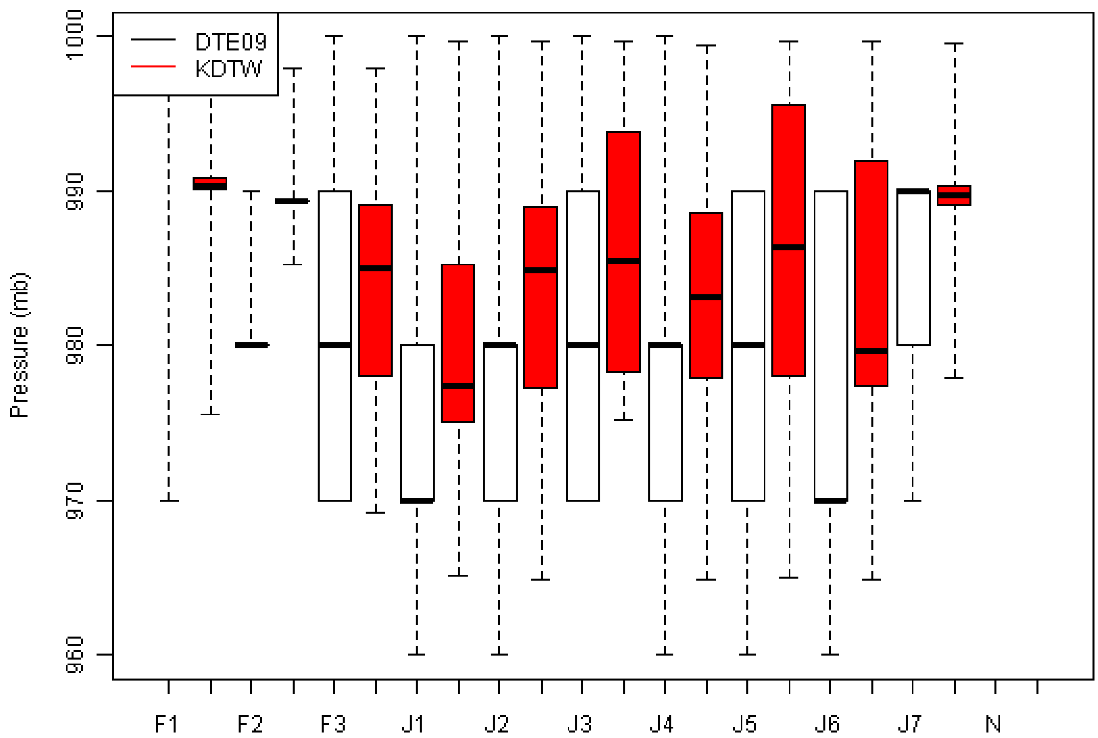


Figure 40. Boxplot for the DTE09 field study for the pressure variable stratified by individual vehicle.

Overall, the DTE09 vehicle air temperature measurements appear to provide a reliable representation of the environment they were driving in compared to the ground truth KDTW, with at least some of the differences attributable to distance from KDTW. Barometric pressure derived from the vehicles does not show as good of agreement, with a large negative bias compared to KDTW. This is likely, in part, due to the poor reporting resolution (10 mb) of the vehicles, which needs to be addressed in the future by working with vehicle manufacturers. Also, when looking at individual vehicles, the Jeep Cherokees appear to provide superior observations compared to the Ford Edges or the Nissan Altima, and more research would need to be done to determine why this is the case, with a first possibility being the type of sensors used and their placement on the vehicle.

3.2.2. PoC

This section highlights the results of an analysis of the PoC field study. When stratifying the data by hour (Figure 41), there is high positive correlation as well as favorable bias and MAE results. The data from this time period are not as consistent as the DTE09, which implies that variability is introduced at least because the testing procedures are different between the two time periods.

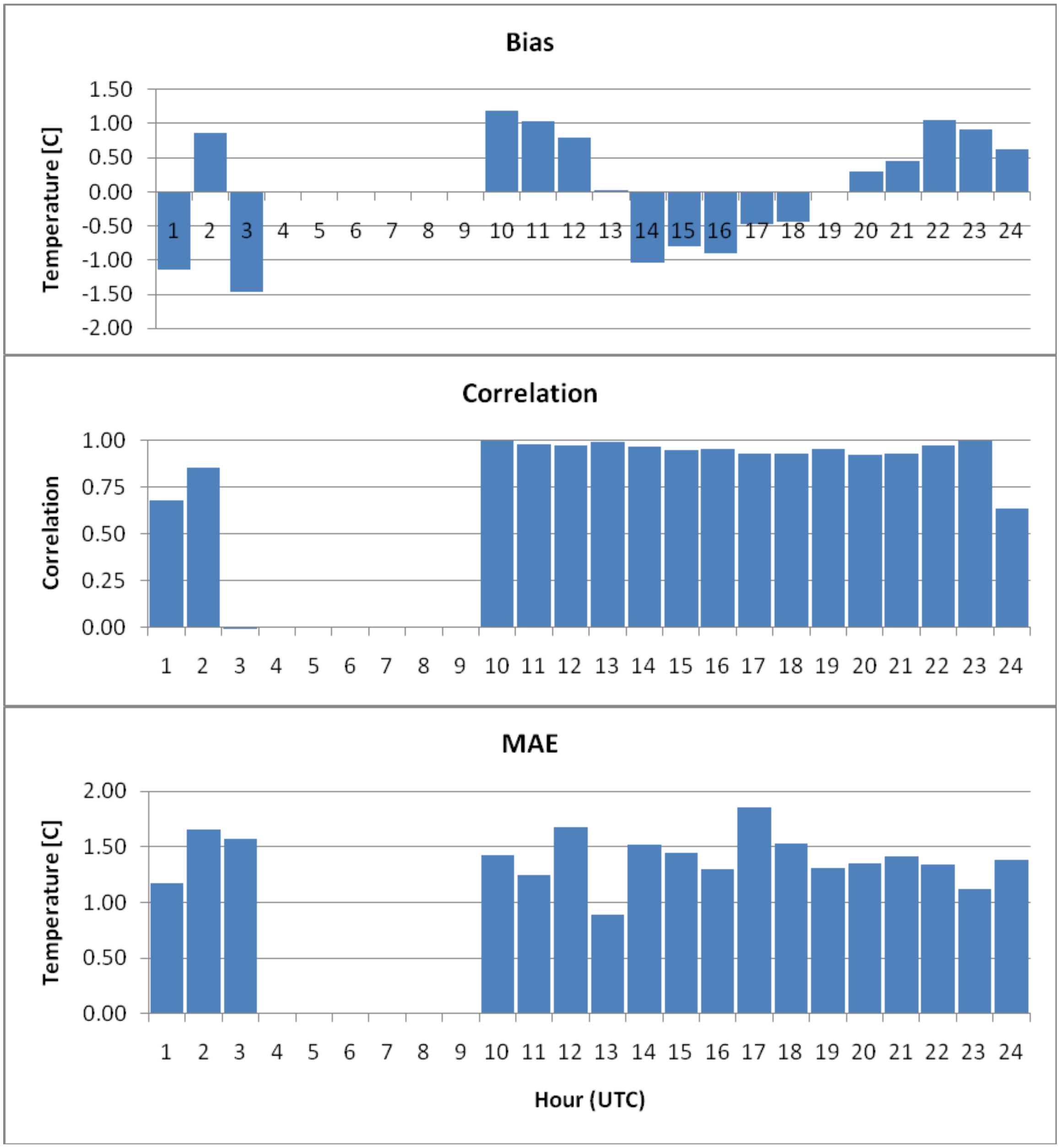


Figure 41. Statistics (Bias, Correlation, and MAE) for the PoC field study for the temperature variable stratified by hour of day (UTC).

The pressure variable accuracy (bias and MAE) results (Figure 42) are consistent with the DTE09 results, which show poor accuracy for the vehicle-based pressure observations. The correlation results are positive and reasonable but are not as consistent as the results from the DTE09 analysis.

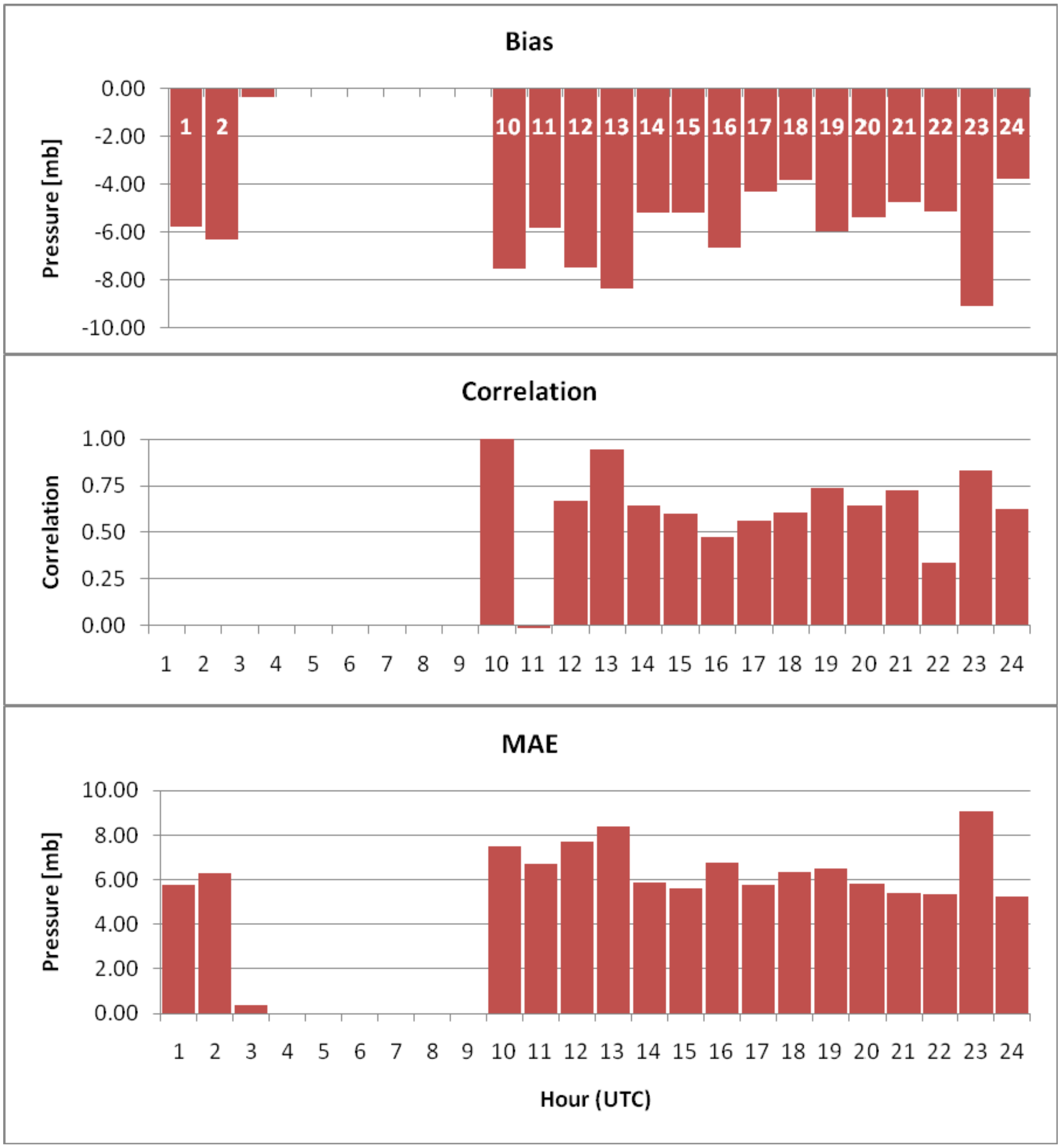


Figure 42. Statistics (Bias, Correlation, and MAE) for the PoC field study for the pressure variable stratified by hour of day (UTC).

Stratified by speed, the PoC data have a slight negative bias (~0.05 to 0.19), which decreases slightly as speed increases. The correlation and MAE statistics are very consistent over the different speed categories.

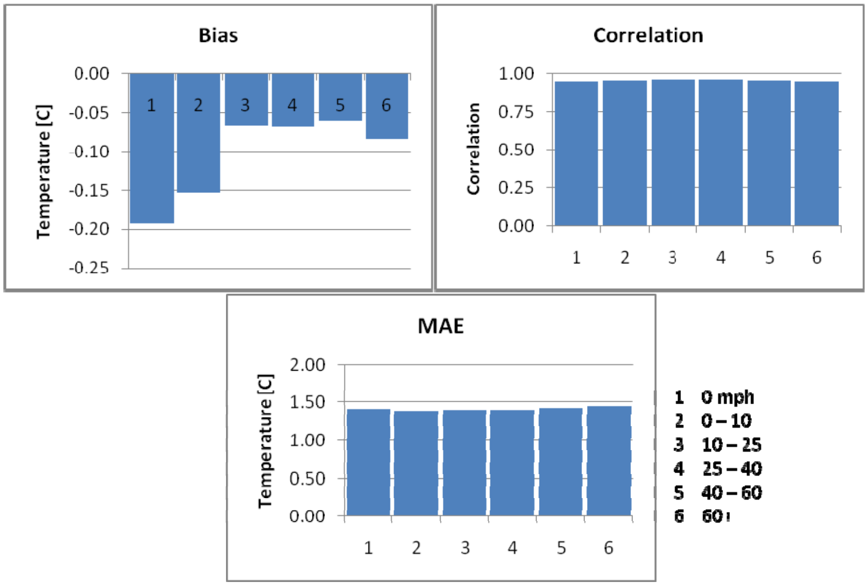


Figure 43. Statistics (Bias, Correlation, and MAE) for the PoC field study for the temperature variable stratified by speed category.

While the results are again not favorable for the pressure variable, all three of the resulting statistics are consistent, which when coupled with the results from DTE09, leads to a conclusion that pressure is not necessarily sensitive to the speed that the vehicle is moving.

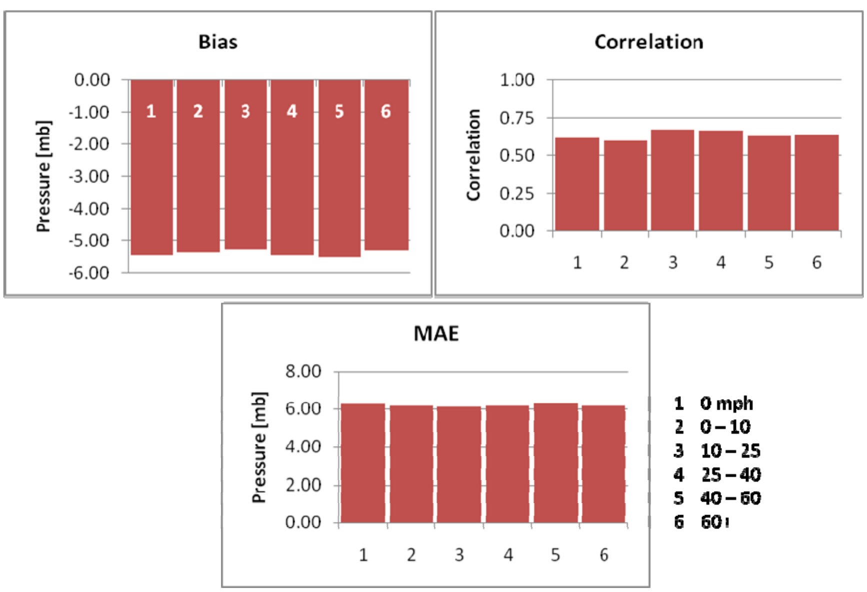


Figure 44. Statistics (Bias, Correlation, and MAE) for the PoC field study for the pressure variable stratified by speed category.

Temperature from the vehicles (Figure 45) over different temperature ranges shows a slight warm Bias (~1°C) up to 10°C and a very slight cold bias from 10 to 25+°C. This may be an artifact of having a small sample size at the colder temperatures. The temperature data shows good positive correlation and possibly a slight decrease in accuracy (MAE = 1 to 1.5) as temperature increases.

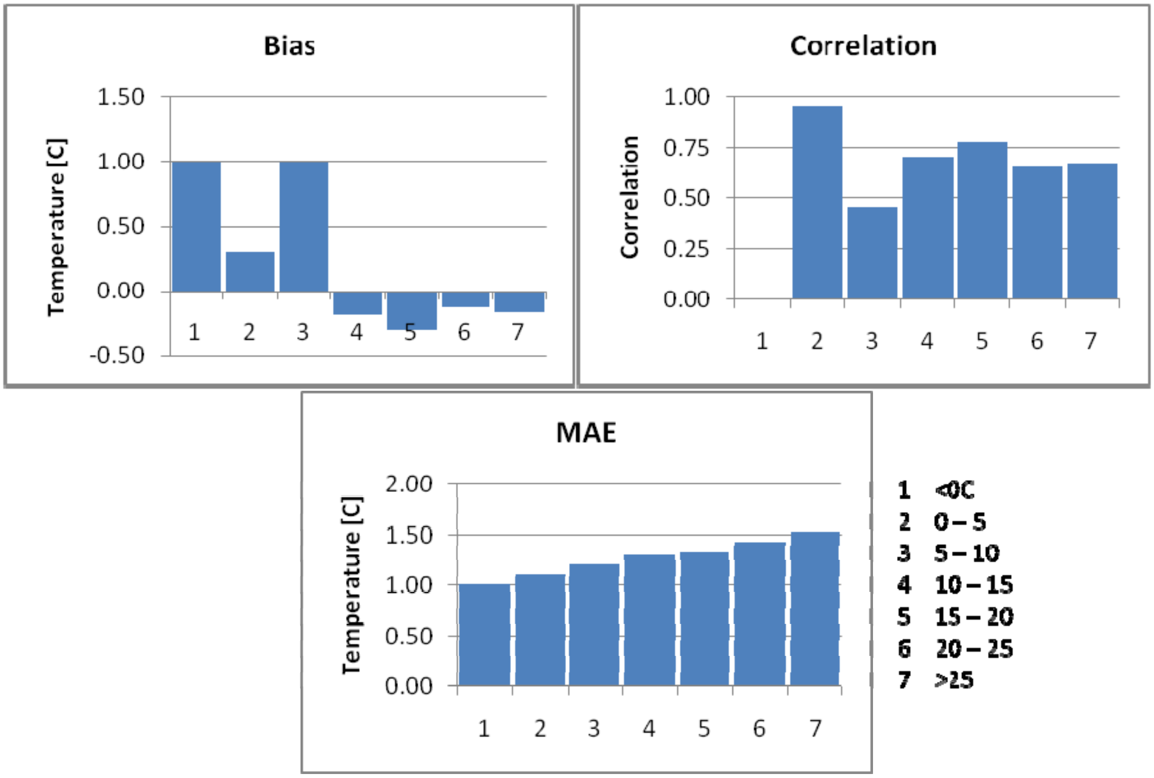


Figure 45. Statistics (Bias, Correlation, and MAE) for the PoC field study for the temperature variable stratified by temperature category.

The pressure variable results (Figure 46) show an increase in bias (from -8 to -4 mb) as the temperature increases but a decrease in correlation. MAE results are consistent for the different temperature categories. These results may also be an artifact of having a smaller sample size at the lower temperatures.

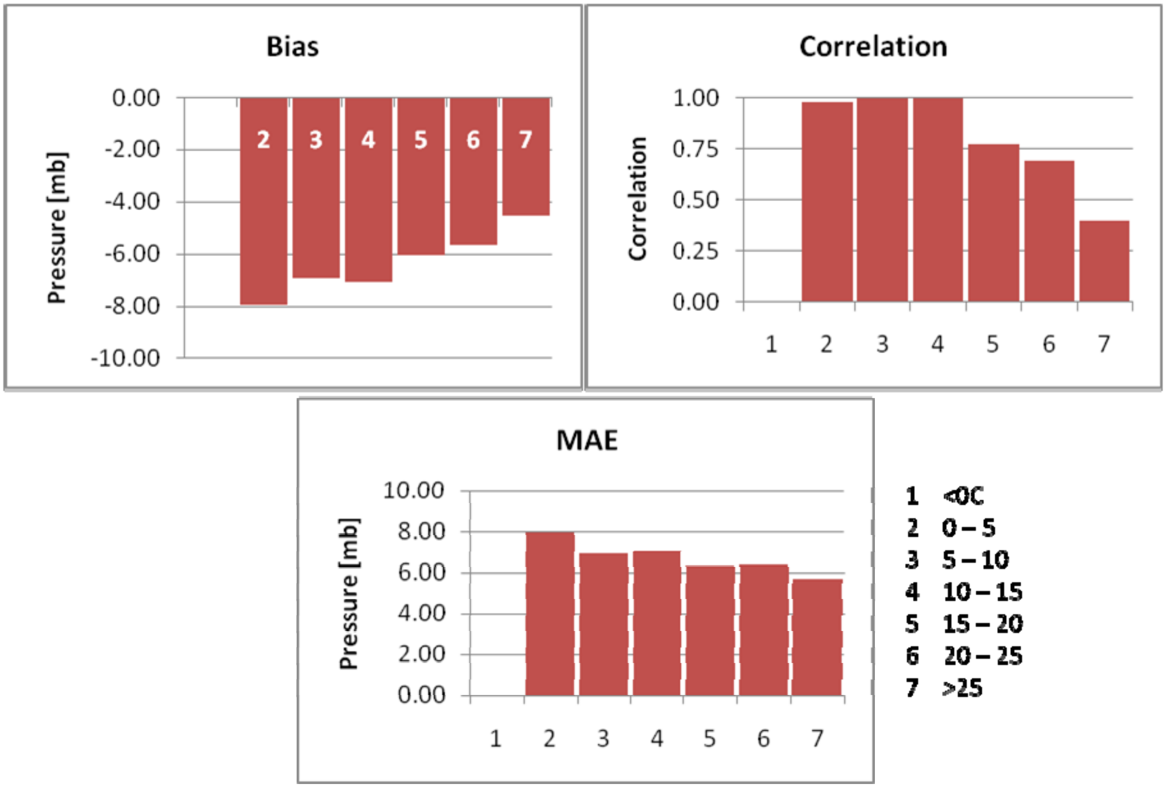


Figure 46. Statistics (Bias, Correlation, and MAE) for the PoC field study for the pressure variable stratified by temperature category.

The PoC temperature results (Figure 47) for different precipitation occurrences (Yes/No) are consistent with the results from DTE09 and show remarkable accuracy from the vehicles measurements but not discernable difference when precipitation is (or is not) occurring. However, these results need to be considered cautiously as there were only 3 samples for no precipitation.

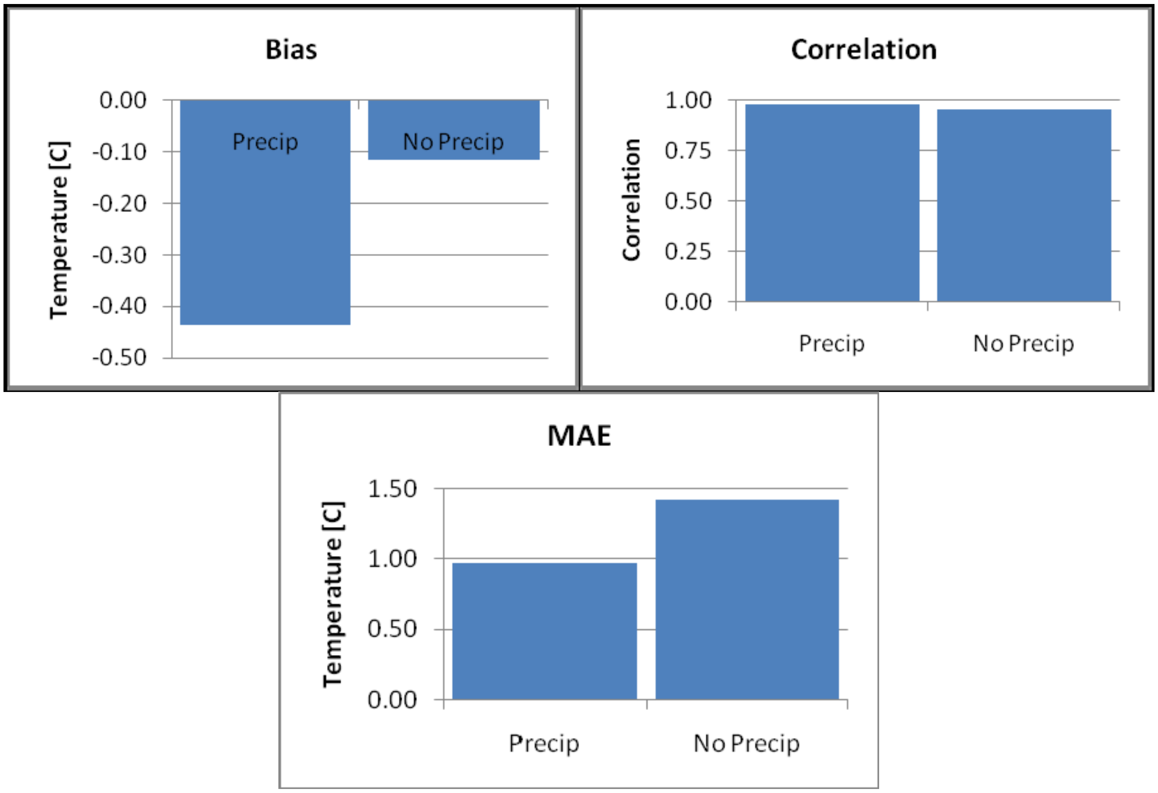


Figure 47. Statistics (Bias, Correlation, and MAE) for the PoC field study for the temperature variable stratified by precipitation occurrence.

The pressure results for the PoC stratified by precipitation (Figure 48) are also consistent with the results from DTE09, which illustrate the large negative bias in the data; again, with only 3 precipitation cases, we cannot draw conclusions from the data.

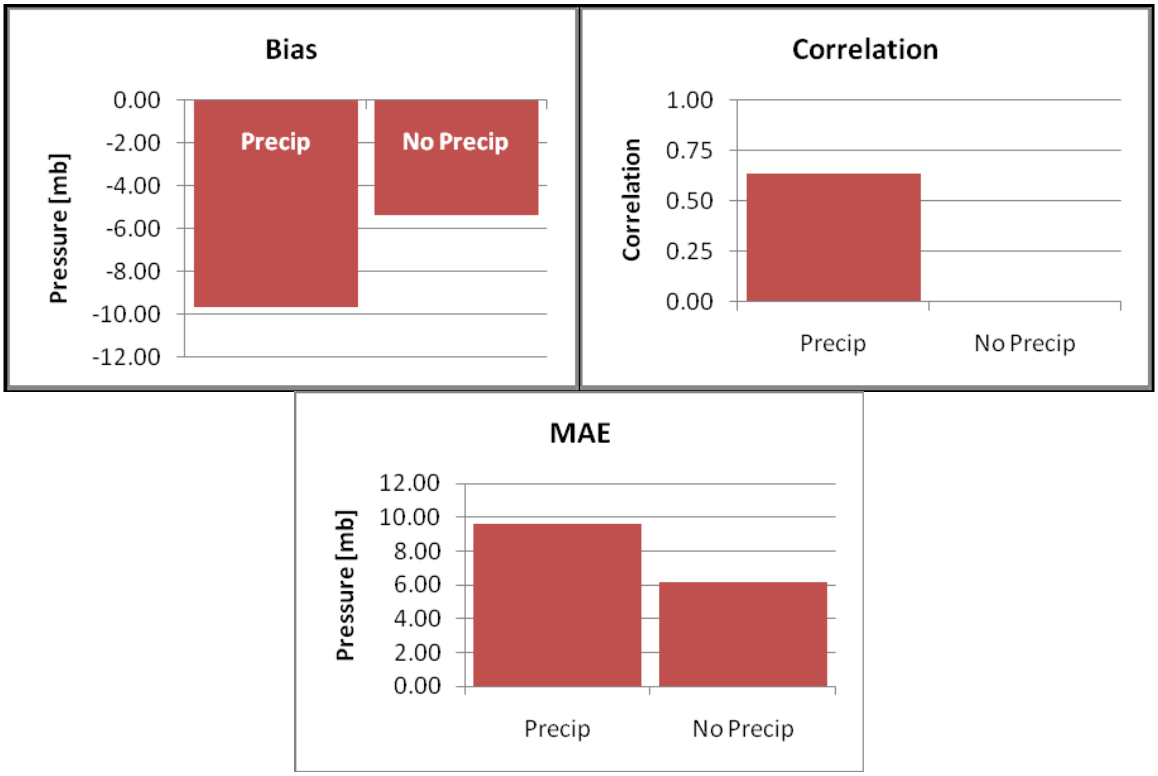


Figure 48. Scatterplot and statistics (Bias, Correlation, and MAE) for the PoC field study for the pressure variable stratified by precipitation occurrence.

The pressure results for the PoC stratified by precipitation (Figure 48) are also consistent with the results from DTE09, which illustrate the large negative bias in the data; again, with only 3 precipitation cases, we cannot draw conclusions from the data.

Overall, the PoC results are consistent with DTE09, with the most obvious result being the accurate air temperature observations but inadequate barometric pressure measurements. As with the DTE09 data, part of this problem may be solved by working with manufacturers to improve the reporting resolution.

3.2.3. DUAP

The results in this section are from an analysis of the DUAP dataset. When the temperature data is stratified by hour (Figure 49), the results show marked accuracy (bias = 0.85 to 1.55°C, and MAE=~1.75°C) and high positive correlation. The results also show a slight increase in the warm bias in the afternoon and evening hours but these values are within the resolution (1°C) of the sensor.

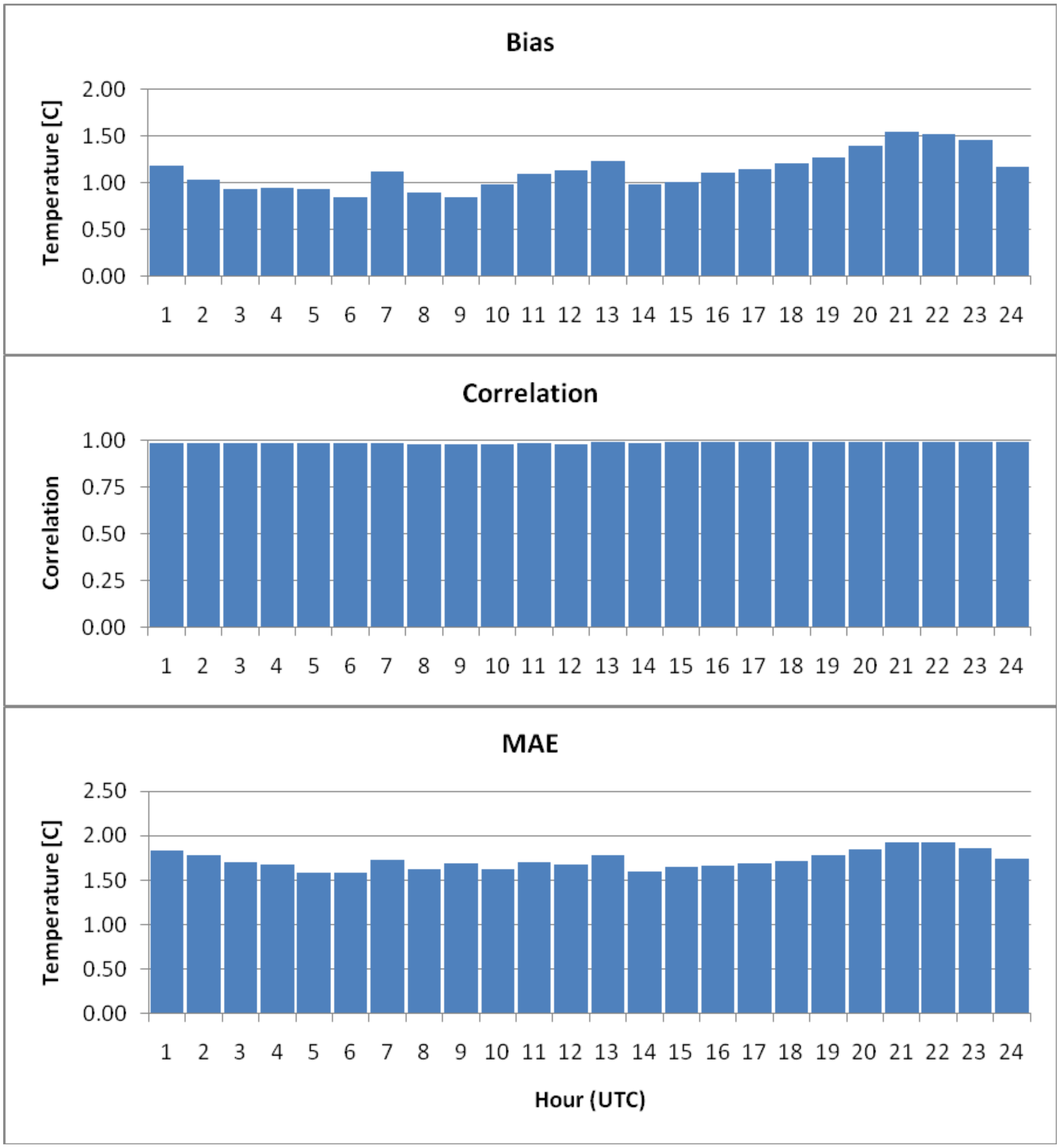


Figure 49. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the temperature variable stratified by hour of day (UTC).

The pressure results (Figure 50) also show a large negative bias, but it is less pronounced than for the PoC and DTE09 dataset. There does not appear to be a change in the bias or MAE during the nighttime hours, which is interesting as this is the only data set that has a reasonable number of night-time samples.

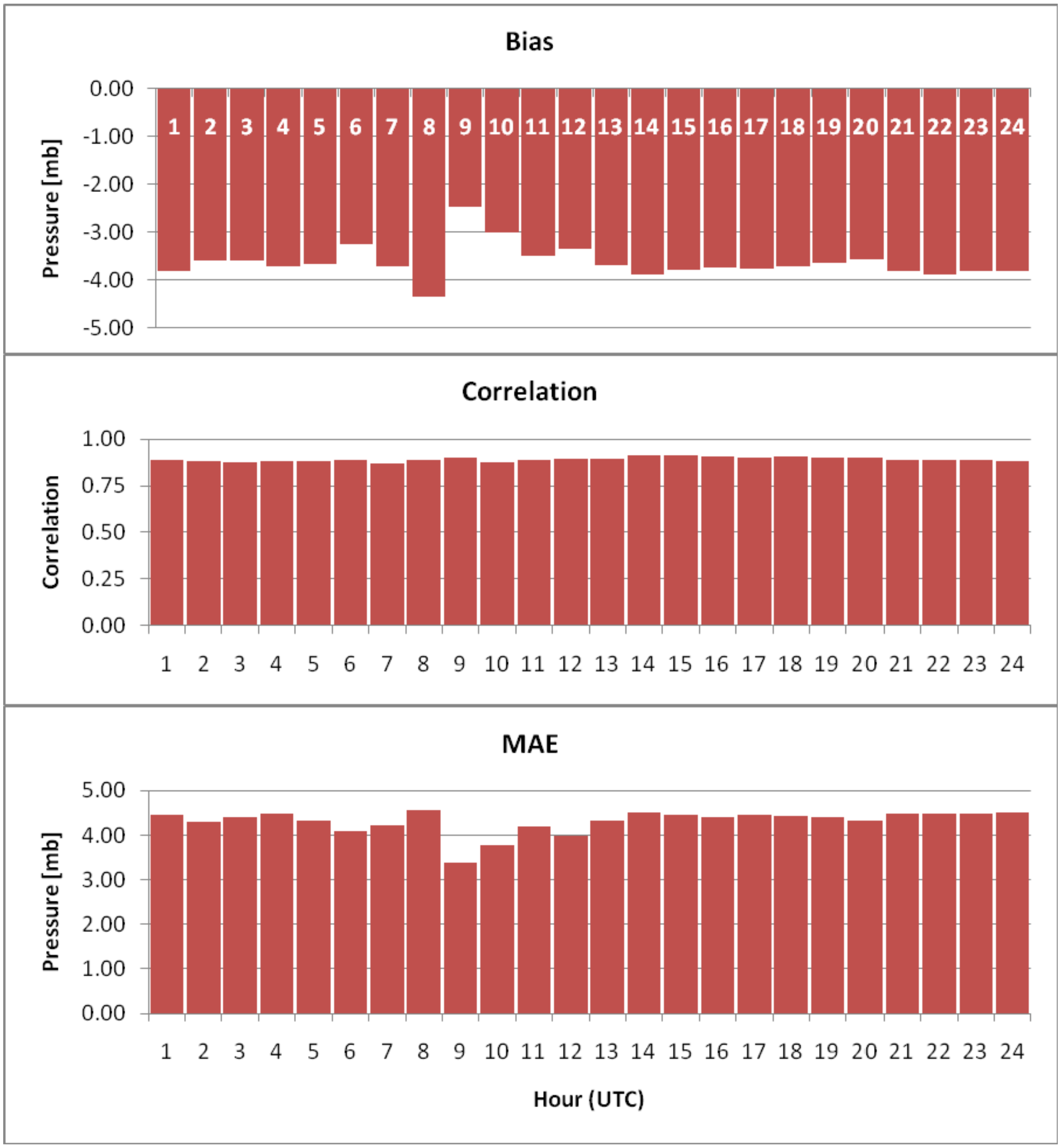


Figure 50. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the pressure variable stratified by hour of day (UTC).

Much like the previous analysis of both the DTE09 and PoC datasets, there appears to be no difference in the results (Figure 51) due to differences in speed of the vehicle and the temperature measurements are accurate and have a high positive correlation.

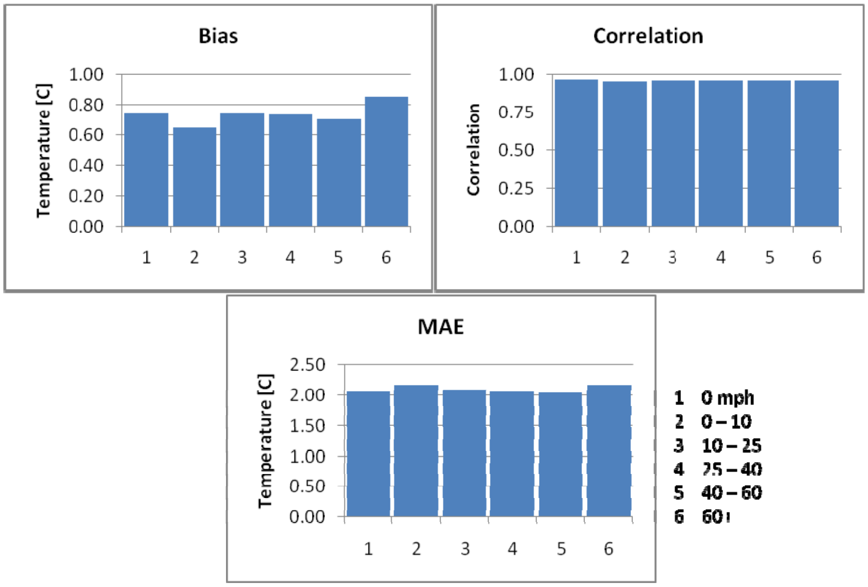


Figure 51. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the temperature variable stratified by speed category.

The DUAP pressure results (Figure 52) show poor accuracy but reasonably high correlation and consistency over the different speed categories.

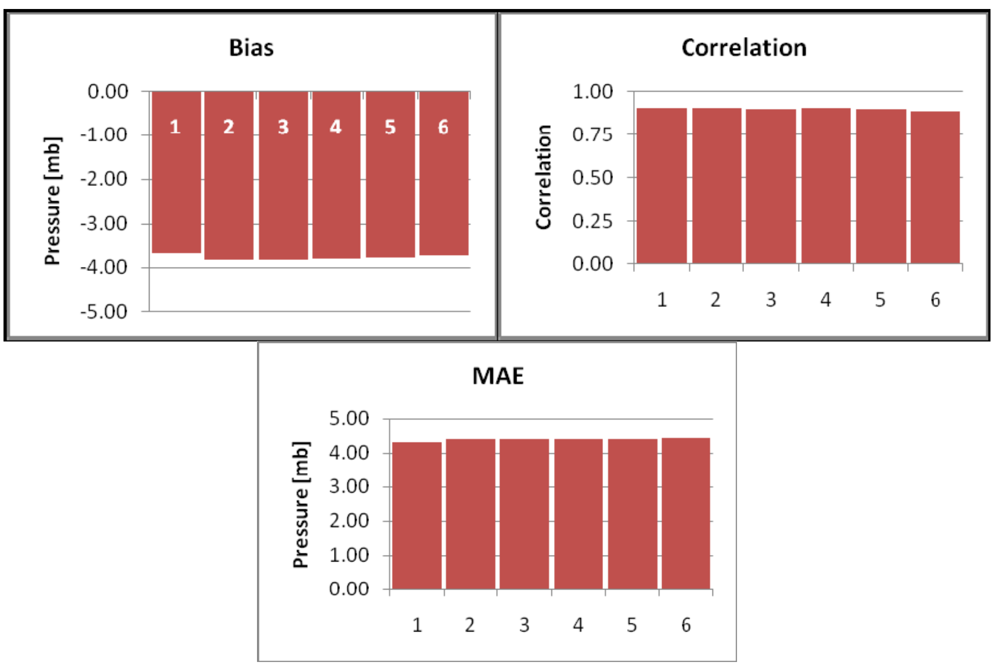


Figure 52. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the pressure variable stratified by speed category.

When stratifying the DUAP data by KDTW temperature occurrence, the results mirror findings from PoC and DTE09. As temperatures warm, the bias and MAE tend to increase (Figure 53). However, ambient temperature does not seem to influence pressure readings (Figure 54).

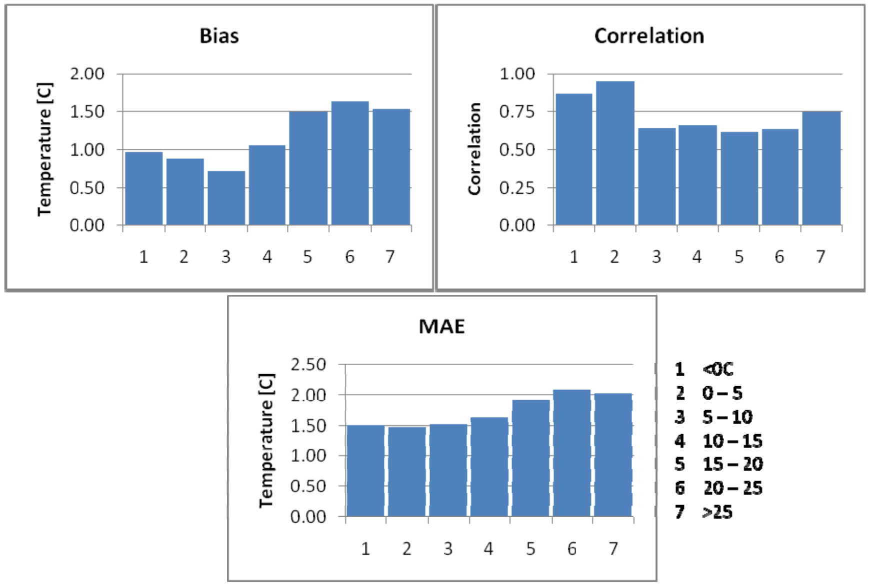


Figure 53. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the temperature variable stratified by KDTW temperature category.

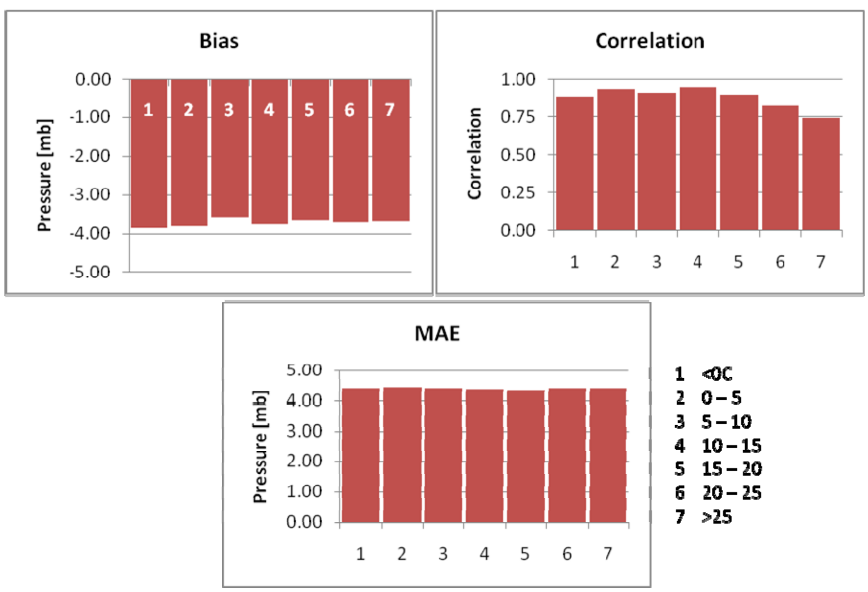


Figure 54. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the tem perature variable stratified by KDTW temperature category.

When stratifying the DUAP data by precipitation occurrence, the results are remarkable similar for both temperature (Figure 55) and pressure (Figure 56). The temperature data have a slight positive bias, unlike the PoC and DTE09 data. More analysis of these data will need to be performed to diagnose the cause for this difference. The pressure data (Figure 56) in the precipitation occurrence analysis again show a negative bias and high correlation. These results also show that there is no discernable difference between the vehicle measurements of pressure during precipitation (and non-precipitation) events, which is consistent with the analyses performed on the DTE09 and PoC datasets.

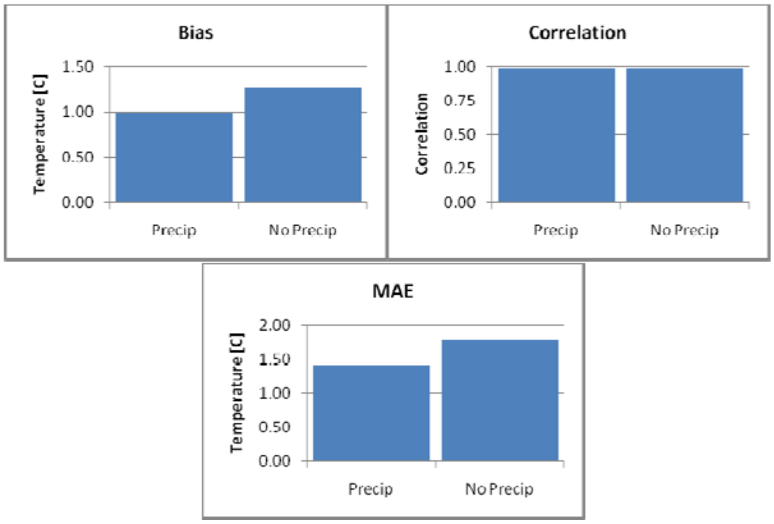


Figure 55. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the tem perature variable stratified by precipitation occurrence.

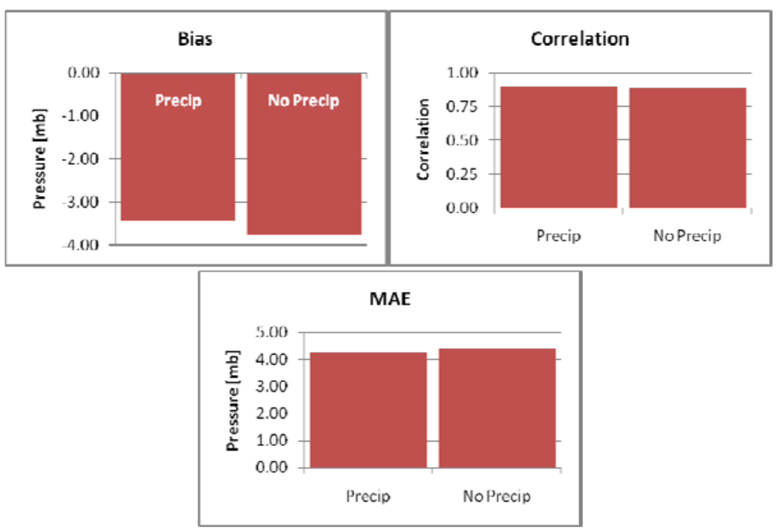


Figure 56. Statistics (Bias, Correlation, and MAE) for the DUAP field study for the pressure variable stratified by precipitation occurrence.

Overall, the DUAP analysis shows again that the air temperature observations from the vehicles are a reliable representation of their environment compared to the KDTW observations, and barometric pressure is less accurate but well correlated with KDTW observations. In this case, poor reporting resolution is not a cause of pressure inconsistencies as the DUAP dataset has a resolution of 4 mb. This indicates that more examination must be done regarding how barometric pressure is derived from the vehicles in order to determine if any improvements in measurements can be made or if any corrections must be made to the data after it is received.

Chapter 4 Weather Processing Requirements

This section contains a rough estimate of the hardware required for processing weather data from approximately 1 million cars in both rural and urban areas. We have also provided an estimate for the hardware required to process ancillary meteorological data such as radar, satellite and surface observation data sets. Additional hardware may be required in order to process data from >1 million cars. In coming up with this estimate, we have made a number of simplifying assumptions that are presented below. In order to refine this estimate, we would recommend that the current Vehicle Infrastructure Integration (VII) simulation system be extended to include 1-km gridded output. Currently, it only outputs road-segment weather information.

The general assumptions include:

- The weather processing machine will focus on the initial processing of satellite, radar, and other meteorological data sets from non-vehicle sensors.
- The vehicle data processing machines will integrate probe message observations with gridded weather information from the weather processing machine.
- The input and output data are not compressed.
- The data archive files on disk are not compressed.
- The Central Processing Units (CPU's) being recommended reflect hardware that is currently being purchased at the Research Applications Laboratory at NCAR.
- The Contiguous United States (CONUS) is divided into 8 tiles approximately 2000 x 2000 km in size each, though this varies between tiles with tile 3 being the smallest at approximately 1000 x 1500 km (Figure 57). This tiling breaks the CONUS into 8 regions each covered by Next-Generation Radar (NEXRAD) data. This tiling represents a potential regionalization of the system. For the 1 million car examples discussed below, all vehicles are situated only 1 of these tiles.
- Vehicles in the rural area will be moving at faster speeds, on average, than vehicles in the urban environment, and therefore will be generating fewer snapshots per unit time. This assumption is based off (Petty and Mahoney 2007), which gives the example that a vehicle traveling at 60 mph or greater will generate snapshots every 20 seconds as it travels between Road Side Equipments (RSEs) while a vehicle traveling at 20 mph or less will generate snapshots every 4 seconds.

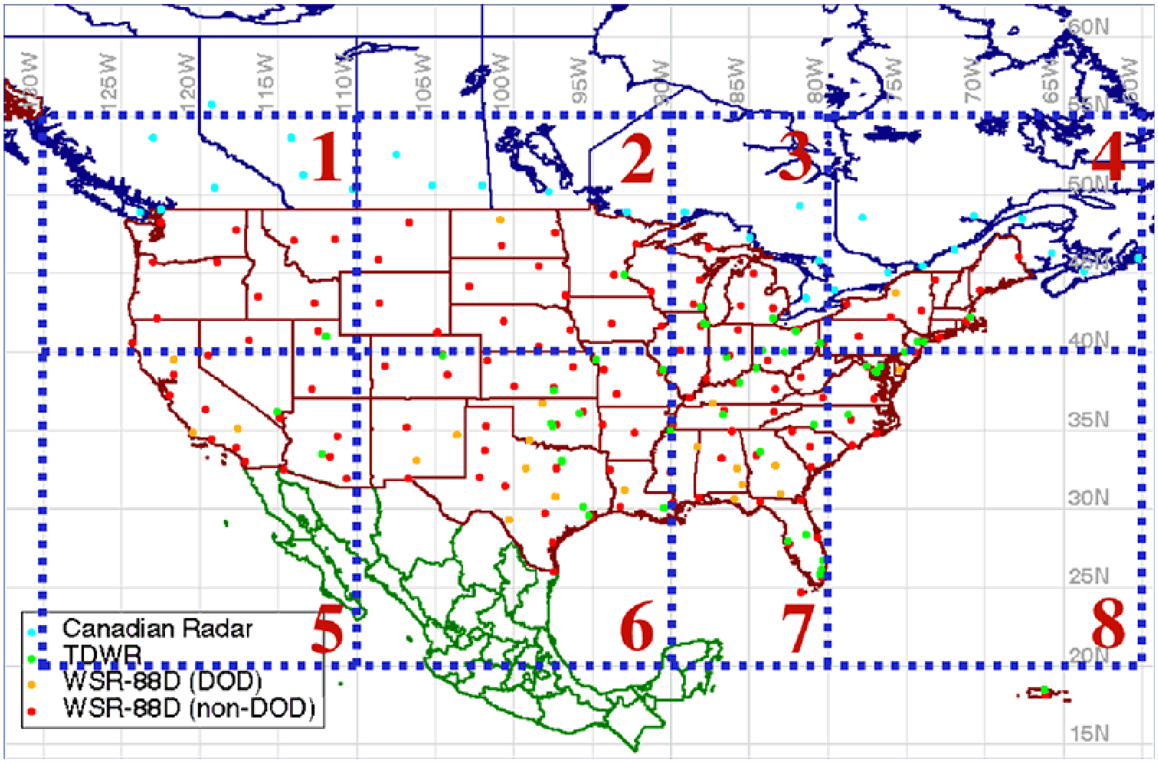


Figure 57. Map depicting the CONUS divided into 8 tiles. (From http://www.nssl.noaa.gov/projects/q2/tutorial/3dmosaic.php)

4.1. Rural area (1 million cars):

Hardware Estimate (see assumptions below):

Single node:

- 2 - Quad-core Intel Xeon 5500
- 8 GB memory
- 6 TB disk
- 1 to 2 - 100 mbit Ethernet cards

Cost: $7000

Assumptions:

Input_record_size:	40 bytes
Output_record_size:	50 bytes
Transmission_rate:	0.05 records per second (1 record per 20 seconds)
Number_cars:	1 million
Output Grid Dimension:	1km, 2000 x 2000

Output Grid Information:
Atmospheric Temperature (short),
Atmospheric Pressure (short),
Accelerometer_x (short),
Accelerometer_y (short),
Accelerometer_z (short),
Number cars (short),
Wiper State (char),
Rain Sensor (char),
Sun Sensor (char),
Headlights (char),
ABS (char),
Traction (char),
Stability (char)

Number of 2-byte Output Grids: 6
Number of 1-byte Output Grids: 7

Total: Approximately 20 bytes

Input Bandwidth = Input_record_size * Transmission_rate * Number_cars
= 40 * 0.05 * 1000000
= 2 MB / second

QC Output Bandwidth = Output_record_size * Transmission_rate * Number_cars
= 50 * 0.05 * 1000000
= 2.5 MB / second

Grid Output Bandwidth = 20 bytes per grid cell * 4000000 grid cells * 12 updates per hour
= 960 MB per hour
= 0.27 MB / second

Total I/O Bandwidth: 5 MB / second

RAM Memory Requirements in 5 minutes:

5MB * 60 * 5 = 1.5 GB

Disk Requirements for short term archive:

QC Car data:
2.5 MB * 3600 * 24 = 216 GB (1 day archive)

Grid data:
960 MB * 24 = 23 GB (1 day archive)

4.2. Urban area (1 million cars):

Hardware Estimate (see assumptions below):

Single node:
- 2 - Quad-core Intel Xeon 5500
- 16 GB memory
- 6 TB disk
- 3 to 4 - 100 mbit Ethernet cards

Cost: $8000

Assumptions:
Record_size: 40 bytes
Transmission_rate: 0.25 records per second (1 record per 4 seconds)
Number_cars: 1 million

Input Bandwidth = Input_record_size * Transmission_rate * Number_cars
= 10 MB / second

QC Output Bandwidth = Output_record_size * Transmission_rate * Number_cars
= 50 * 0.25 * 1000000
= 12.5 MB / second

Grid Output Bandwidth = 20 bytes per grid cell * 4000000 grid cells * 12 updates per Hour
= 960 MB per hour
= 0.27 MB / second

Total I/O Bandwidth: 25 MB /second

RAM Memory Requirements in 5 minutes:

7.5 GB

Disk Requirements for short term archive:

Car data:
12.5 MB * 3600 * 24 = 1.08 TB (1 day archive)

Grid data:
960 MB * 24 = 23 GB (1 day archive)

4.3. Radar/Satellite/Meteorological Data Hardware Estimate:

Hardware Estimate (see assumptions below):

Single node:

- 2 - Quad-core Intel Xeon 5500
- 16 GB memory
- 6 TB disk
- 1 - 100 mbit Ethernet card

This machine will process the following data sets:

Surface observations (ASOS, RWIS)
Radar data (NEXRAD)
Satellite Data (GOES E & W)
Climatology Monthly
Model Output Statistics

4.4. Input data record

```
/* Input data record */
 typedef struct VehicleData
{
time_t date;                     /* Includes year, month, day, hour, minute, second (4
  bytes) */
float latitude;             /* Latitude (4 bytes) */
float longitude;                 /* Longitude (4 bytes) */
short acc_x;                     /* m/s2 (2 bytes) */
short acc_y;                     /* m/s2 (2 bytes) */
short acc_z;                     /* m/s2 (2 bytes) */
short elev;                      /* Elevation (2 bytes) */
short heading;                   /* Heading (2 bytes) */
short temperature;               /* Temperature (2 bytes) */
short pressure;                  /* Pressure (2 bytes) */
unsigned char abs;               /* 1,2 or 3 (1 byte) */
unsigned char stability;         /* 1,2 or 3 (1 byte) */
unsigned char traction;          /* 1,2 or 3 (1 byte) */
unsigned char wiper;             /* 0..55 (1 byte) */
unsigned char rain;              /* 0..255 (1 byte) */
unsigned char sun;               /* 0..16 (1 byte) */
unsigned char headlights;        /* 0..4 (1 byte) */
unsigned char hours_operation;/*  Hours of operation 0..24 (1 byte) */
```

```
unsigned char latitude_zone,/*    Latitude zone 0..1 (1 byte) */
unsigned char longitude_zone; /* Longitude zone 2..3 (1 byte) */
unsigned char velocity;             /* velocity 0..255 (1 byte) */
} Vehicle_record;

enum QCFields {TEMPERATURE_FIELD, PRESSURE_FIELD, WIPER_FIELD,
RAIN_FIELD, SUN_FIELD, HEADLIGHT_FIELD, ACCELEROMETER_FIELD, ABS_FIELD,
TRACTION_FIELD, STABILITY_FIELD, NUM_QC_FIELDS};

/* Output QC data record */
typedef struct QCVehicleData
{
time_t date;                        /* Includes year, month, day, hour, minute,
  second (4bytes) */

float latitude;           /* Latitude (4 bytes) */
float longitude;          /* Longitude (4 bytes) */
short acc_x;                        /* m/s2 (2 bytes) */
short acc_y;                        /* m/s2 (2 bytes) */
short acc_z;                        /* m/s2 (2 bytes) */
short elev;                         /* Elevation (2 bytes) */
short heading;                      /* Heading (2 bytes) */
short temperature;                  /* Temperature (2 bytes) */
short pressure;                     /* Pressure (2 bytes) */
unsigned char abs;                  /* 1,2 or 3 (1 byte) */
unsigned char stability;  /* 1,2 /* 1,2 or 3 (1 byte) */
unsigned char traction;   /* 1,2 /* 1,2 or 3 (1 byte) */
unsigned char wiper;                /* 0..55 (1 byte) */
unsigned char rain;                 /* 0..255 (1 byte) */
unsigned char sun;                  /* 0..16 (1 byte) */
unsigned char headlights;           /* 0..4 (1 byte) */
unsigned char hours_operation;/* Hours of operation 0..24 (1 byte) */
unsigned char latitude_zone; /* Latitude zone 0..1 (1 byte) */
unsigned char longitude_zone; /* Longitude zone 2..3 (1 byte) */
unsigned char velocity; /* velocity 0..255 (1 byte) */
unsigned char qc_flags[NUM_QC_FIELDS]; /* up to 8 binary quality control flags for
each field
(NUM_QC_FIELDS bytes) */
} QCVehicle_record;
```

Chapter 5 Conclusions

Specifically, this report focused first on assessing how many of the temperature and pressure observations pass quality-control tests, and whether the pass rates vary by day, time of day, vehicle speed, ambient temperature, or precipitation class. Secondly, for the observations passing QCh, this report compared the observations to a known standard (KDTW Automated Surface Observing System (ASOS)) and computed the correlation coefficient, mean absolute error, bias, and t-test for significance of the difference between the vehicle data and KDTW. Lastly, this report outlined the hardware requirements for processing vehicle and ancillary data.

The following are the major conclusions of this report:

- The SRT, CRT, and NST, finalized in the CAT, appear to provide a robust QCh set. The MAT and NVT may be valuable, but will need additional tuning. The ROT was not investigated here.
- For DTE09, the Jeep Cherokees proved superior to the Ford Edges and Nissan Altima. For all three data sets, temperature measurements appear to be far superior to pressure measurements.
- There is some indication that environmental conditions (precipitation, temperature) affect the QCh pass rates, but vehicle characteristics (speed) and time of day do not appear to play a large role. The effects of environmental conditions on these datasets were not statistically significant for the most part, and physically the differences were small. Additional measurements in future experiments should focus on the influence of environmental conditions on QCh pass/fail rates.
- For the temperature observations that passed QCh in DTE09, the resulting statistics indicate that the vehicle data is very similar to KDTW. Some of the differences may be related to the distance from KDTW to the DTE, and future work should include better surface validation sensors, such as a mobile RWIS.
- Storage of vehicle and ancillary data will require considerable disk space.

Chapter 6 References

Dion, François and Ralph Robinson, 2009: VII Data Use Analysis and Processing (DUAP) Project Report. 141pp.

Drobot, S.D., and 8 others, 2009: IntelliDrive[SM] road weather research & development - the Weather Data Translator. *Proceedings of the 2009 Intelligent Transportation Society of America Annual Conference*. 13 pp.

NCAR (National Center for Atmospheric Research), 2008: VII Requirements: Vehicle Probe Data, Ancillary Data, and Potential Quality Checking Algorithms. 24 pp.

NCAR, 2009a: Task 3: Creation of VII-based Derived Observations Recommendation Report. 37pp.

NCAR, 2009b: Detroit IntelliDrive[SM] Experiment – Winter 2009 Detailed Test Plan. 26pp.

Petty, Kevin R. and William P. Mahoney III, 2007: Weather Applications and Products Enabled Through Vehicle Infrastructure Integration (VII). 119 pp.

APPENDIX A. List of Acronyms

ABS	Anti-lock Braking System
ASOS	Automated Surface Observing System
CAT	Combined Algorithm Test
CONUS	Contiguous United States
CPU	Central Processing Unit
CRT	Climatological Range Test
DTE09	Development Testbed Environment 2009
DUAP	Michigan Department of Transportation VII Data Use Analysis and
FHWA	Federal Highway Administration
GOES E	Geostationary Operational Environment Satellite East
GOES W	Geostationary Operational Environment Satellite West
I/O	Input/Output
KDTW	Detroit Metropolitan Wayne County Airport
MAE	Mean Absolute Error
MAF	Mass Air Flow
MAP	Manifold Absolute Pressure
MAT	Model Analysis Test
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NEXRAD	Next-Generation Radar
NST	Neighboring Surface Station Test
NVT	Neighboring Vehicle Test
OBE	On-board Equipment
PoC	Proof of Concept
QC	Quality Control
QCh	Quality Checking
RAM	Random Access Memory
RITA	Research and Innovative Technology Administration
ROT	Remote Observation Test
RSAS	Rapid Update Cycle Surface Assimilation System
RSE	Road Side Equipment
RUC	Rapid Update Cycle
RWIS	Road Weather Information System
SRT	Sensor Range Test
UTC	Universal Time Coordinate
VII	Vehicle Infrastructure Integration

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free “Help Line” 866-367-7487
www.its.dot.gov

FHWA-JPO-10-062
NTL # 14968