Final Report

Analysis of Airborne Formaldehyde Data Over Houston Texas Acquired During the 2013 DISCOVER-AQ and SEAC^{4}RS Campaigns

AQRP Project 14-002

Submitted to:

Gary McGaughey
Project Manager
Texas Air Quality Research Program

&

Jim Smith
TCEQ Project Liaison

Prepared by:

Dr. Alan Fried
Institute of Arctic & Alpine Research
University of Colorado
1560 30th Street, Boulder, Colorado 80303
alan.fried@colorado.edu

Dr. Christopher P. Loughner
Earth System Science Interdisciplinary Center
University of Maryland, College Park, Maryland 20740
christopher.p.loughner@nasa.gov

Dr. Kenneth Pickering
Atmospheric Chemistry and Dynamics Laboratory, NASA
Goddard Space Flight Center, Greenbelt, MD

QAPP Category Number: III
Type of Project: Data Evaluation; Research or Development (Modeling)

Final Report: Version 4, February 25, 2016

QA Requirements: Audits of Data Quality: 10% Required.
All data analyzed for this project have been quality controlled and audited.

ACKNOWLEDGMENT

The preparation of this report was funded by a grant from the Texas Air Quality Research Program (AQRP) at The University of Texas at Austin through the Texas Emission Reduction Program (TERP) and the Texas Commission on Environmental Quality (TCEQ). The findings, opinions and conclusions are the work of the author(s) and do not necessarily represent findings, opinions, or conclusions of the AQRP or the TCEQ. We would like to acknowledge Rob Gilliam at the U.S. Environmental Protection Agency (EPA) for providing guidance in running WRF using the EPA iterative technique. We would also like to acknowledge Jim MacKay at the Texas Commission on Environmental Quality (TCEQ) for providing the TCEQ anthropogenic emissions inventory and Mark Estes of TCEQ for providing DNPH and auto-GC data at various ground sites and as well as his many helpful discussions. The PIs also acknowledge NASA funding for the formaldehyde measurements employed in this study during the 2013: 1) DISCOVER-AQ (Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality) campaign under grant NNX12AG85G to the University of Colorado; and 2) SEAC^{4}RS (Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys) campaign under grant NNX12AM08G to the University of Colorado.

EXECUTIVE SUMMARY

High ozone continues to be a major problem in many large U.S. cities, including Houston, Texas. Despite extensive efforts to address this problem, our understanding of the major precursors that control ozone formation is still highly uncertain and incomplete. One such major precursor is the toxic trace gas formaldehyde (CH_2O). This gas is produced in the atmosphere as an intermediate when virtually any volatile organic compound that is emitted into the atmosphere is oxidized, primarily by OH radicals. Formaldehyde subsequently rapidly decays in a matter of hours by reactions of these radicals and sunlight to produce ozone and additional radicals. These additional radicals then produce additional ozone. In Houston Texas the problem of ozone formation is particularly acute since the greater Houston-Galveston-Brazoria Metropolitan Area (HGBMA) is home to some of the largest petrochemical facilities in the United States. Highly reactive volatile organic compounds (HRVOC's) like propene and ethene, which are known to leak into the atmosphere from both normal and upset operations from these facilities, rapidly produce CH_2O and ultimately ozone.

Continued development of effective ozone control strategies requires a comprehensive understanding of the magnitude of various CH_2O sources, its photochemical production rates and pathways, and transport processes. Over the HGBMA one particularly important issue in this regard is our understanding of the relative importance of primary CH_2O sources relative to CH_2O produced from secondary photochemistry. Potential primary sources of CH_2O include any combustion process such as burning, flaring, and automotive emissions, as well as direct leaks from fugitive emissions from petrochemical facilities, to name a few sources. Photochemically produced CH_2O, which is also called secondary CH_2O, arises from the oxidation of the volatile organic sources discussed above. This also includes the oxidation of isoprene, a gas that is both emitted from certain trees as well as from petrochemical operations. Unfortunately, despite extensive efforts and advances from past studies, two competing views regarding the relative importance of primary versus secondary CH_2O sources over the greater HGBMA, which have appeared in the recent literature, have still not been resolved.

To address this critical issue as well as additional questions, a collaborative team from the University of Colorado (CU), the University of Maryland (UMD), and the NASA Goddard Space Flight Center (NGSFC) has embarked on the present project. In this study, we analyze high quality and fast airborne measurements of CH_2O measurements over Houston Texas acquired during two recent NASA airborne campaigns in 2013: Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) study; and 2) Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys ($SEAC^4RS$). The primary objectives of this study are to: 1) address the issue of CH_2O source apportionment over the HGBMA study area discussed above; 2) assess the current 2012 TCEQ emission inventories for CH_2O and its precursors; 3) assess our knowledge of the chemical mechanisms employed; 4) where possible document emission upsets; 5) identify petrochemical flaring events; and 6) confirm, where possible, the TCEQ DNPH CH_2O sampling results. The following tasks were performed (highlighted in italics) to accomplish these objectives, and the results for each task are listed with each task in boldface. Since Tasks 2 & 3 are closely related, these tasks are listed together.

1. *Prepare WRF and CMAQ input files and run the models using nested domains down to a horizontal resolution of 1 km using the 2012 TCEQ emission inventory. Once accomplished, carry out extensive model-measurement comparisons of CH_2O and other species to test the model accuracy throughout the HGBMA during the DISCOVER-AQ field campaign and assess current emission inventories where possible.*

   Modeling analysis employing the Community Multiscale Air Quality (CMAQ) model with Process Analysis, in very high-resolution mode (1 km resolution), driven by the WRF (Weather Research and Forecasting) meteorological model has been successfully developed, improved upon, and evaluated. This evaluation involved comparisons of various measured meteorological and trace chemical species concentrations (CH_2O, isoprene, CO, NO, NO_2, and O_3) with those simulated from CMAQ. Extensive CMAQ-Measurement comparisons were carried out for CH_2O and CO. Comparisons in the planetary boundary layer (PBL) and free troposphere (FT) showed reasonable daily agreement. Not considering Sept. 25, the absolute PBL biases (CMAQ-Meas.) for CH_2O and CO for all the remaining days are all relatively small. In the case of CH_2O, the average of all the daily mean PBL biases is -439 ± 392 pptv, and the average of all daily median biases is -319 ± 397 pptv. The average daily median bias percentage is -11.8 ± 15.7%. For CO, the average of all the daily mean PBL biases is -6.0 ± 14.7 ppbv, and the average of all daily median biases is -6.7 ± 14.0 ppbv. The corresponding daily median bias percentage for CO is -4.5 ± 10.7%.

   These small but persistently negative biases potentially reflect small underestimates in the emission inventories used in the calculations. However, we cannot rule out the possible contribution that CMAQ transports too much boundary layer air into the free troposphere, as has been observed on other occasions. Therefore, based on the above results, we have no firm evidence that the 2012 TCEQ emission inventory under normal conditions needs to be revised.

2. *Develop methods to identify, and provide tabulations of, time periods when sampling clearly identifiable direct emission sources of CH_2O close to their source. In this process, tabulate especially large emission sources observed from WP-3 observations and from reported petrochemical facility upsets. Where possible, estimate the magnitude of such events and provide an emission update. The CMAQ model will be re-run based on such updated emissions estimates. CMAQ output will be analyzed along the path of back trajectories to assess upstream influence. Kinematic back trajectories will be calculated from WRF model output using the WRF post-processing tool RIP (Read/Interpolate/Plot).*

3. *WP-3 observations of very large CH_2O concentrations in the 20 – 35-ppbv-range from the Sept. 25, 2013 flight during the first two circuits have identified this day as one to examine first employing the high resolution WRF-CMAQ model with updated emissions. This model will be analyzed along a forward trajectory calculated from the WRF output south to Smith Point to help in assessing the model chemistry by comparing the model and observations near and downwind of the source. Other significantly elevated time periods will be identified.*

Analysis of airborne CH_2O measurements over the greater HGBMA study area during the 2013 DISCOVER-AQ (9 sampling days over Houston) and SEAC^{4}RS (1 sampling day over Houston) campaigns over the month of September revealed that only the September 25 sampling day showed exceptional high PBL CH_2O levels in excess of 30 ppbv, levels characteristic of our past measurements over the greater HGBMA study area in 2006 and 2000. All other sampling days in 2013 showed significantly lower PBL CH_2O levels in the 2 – 10 ppbv range. We presented an observational approach based upon fast aircraft measurements of correlations between CH_2O, O_3, CO, NO_x/NO_y ratios, and propene as a means of identifying time periods revealing enhanced sources of CH_2O. In this process, we have identified a number of such plumes, which based upon strong anti-correlations of O_3 with CO and high NO_x/NO_y ratios, indicated very fresh plumes concurrent with combustion sources. Most of these plumes were found in the vicinity of petrochemical facilities. A spreadsheet with the major plumes thus identified has been supplied with this report. At present, we do not have enough information to discern if such enhanced CH_2O: 1) originates directly from the combustion sources; 2) is produced during combustion chemically from its two major precursors propene and ethene; 3) occurs simultaneously from fugitive emissions of CH_2O, propene, and ethene; or 4) some combination of the above. Likewise, we do not have enough information to even speculate on the types of petrochemical combustion sources (e.g., flaring, fluidized catalytic cracking combustion, or other potential petrochemical combustion sources) that might be responsible for our observations, and therefore efforts to correlate which petrochemical stack that might be responsible for our observations is beyond the scope of this effort.

In our plume tabulations, the largest source of enhanced CH_2O associated with petrochemical combustion occurred during the 1st circuit on 9/25/13 right over the Baytown ExxonMobil complex around 9:48 am local time. A regression analysis of fast CH_2O and CO measurements produced a CH_2O/CO slope of 82.4 ± 5.4 pptv/ppbv (r^2 = 0.83, N = 51) over this plume. Four other days were identified where we acquired CH_2O/CO slopes over this same petrochemical complex at around the same local time (9/6/13, 9/12/13, 9/13/13, and 9/24/13). The grand average for these 4-days yields a CH_2O/CO slope of 30.4 ± 12.9 pptv/ppbv and a grand median slope of 24.3 pptv/ppbv, which is a factor of 2.7 to 3.4 times lower than that on 9/25/13. Based upon the 2013 Speciated Release Inventory for CH_2O and CO under normal operating conditions (supplied as a separate spreadsheet to AQRP) one would expect a normal operating CH_2O/CO slope of ~ 12 for all three ExxonMobil facilities combined, which is a factor of ~ 2.0 to 2.5 times lower than our 4-day grand (average/median) values. However, when one considers that this 4-day grand (average/median) reflects the sum of CH_2O released as well as CH_2O produced from propene and ethene released from these same facilities under normal operating conditions, we view this factor of 2.0 to 2.5 difference as a reasonable range of values for normal operating conditions. However, the factor of ~ 7 times higher measured slope on 9/25/13 relative to the normal operating Speciated Release Inventory is considerably higher, and in our opinion, suggests enhanced emissions of CH_2O and potentially its precursors on Sept. 25 emanating from the ExxonMobil complex during the morning hours, perhaps by as much as a factor of ~ 3 relative to the other sampling days. We presented additional evidence to further support this hypothesis. We also presented counter-arguments suggesting that some or

all of these enhancements may be caused by unique meteorology on this day (strong early morning inversion with a tightly capped boundary layer ~ 0.3-km) coupled with significantly enhanced ethene and propene emissions measured on this day by TCEQ's auto-GCs during the 5 -10 am hours over the nearby Lynchburg Ferry sampling site (from unknown sources). A more definitive assessment must await additional studies based upon Lagrangian model runs employing back trajectories, and this has been identified as one of the subject areas for a future proposal.

Likewise, efforts in providing individual enhanced emission estimates for various species in moles/hour emanating from the ExxonMobil complex on Sept. 25 that could then used to compare calculated and measured CH_2O concentrations downwind at Smith Point (Task 3) turned out to be far more complicated than originally anticipated. Additional input will be required to carry this out more rigorously than our initial attempts.

Despite our partial success in arriving at firm conclusions regarding Tasks 2&3, we were successful in identifying and highlighting the uniqueness of Sept. 25 relative to other DISCOVER-AQ flight days by highlighting the extensive enhanced CH_2O levels observed throughout nearly the entire HGBM study area over most of the sampling day. Aside from the early morning measurements over the ExxonMobil complex, the majority of these enhancements were found to be coincident with elevated propene (> 5 ppbv) and arise from CH_2O that is photochemically generated from its precursors.

4. *Examine the CMAQ model output run with the Process Analysis Mode to quantify the relative importance of primary emissions and secondary photochemical production of CH_2O throughout the HGBMA study area throughout the DISCOVER-AQ field campaign.*

The CMAQ model was run in Process Analysis Mode to assess primary and secondary sources of CH_2O throughout the greater HGBMA study area throughout the month of September 2013. CH_2O from secondary production sources (Production – Destruction) is approximately a factor of 5 times higher than direct emission sources in the planetary boundary layer (PBL) over the entire month of September and approximately a factor of 7 to 8 times the direct emission source for the atmosphere over the Houston-Galveston-Brazoria Metropolitan Area up to 5-km altitude. These results were further broken down as a function of hour for the entire month of September 2013. Over the 7 am – 7 pm daylight hours, the average ratio yields a value of ~ 8/1 within the PBL. This yields a secondary CH_2O contribution of ~ 89% over the daylight hours and this agrees well with the determination from Parrish et al. [2012] of ~ 92% based upon OH reactions of ethene and propene to produce CH_2O during daylight hours. It is important to note that these results cover the entire HGBMA throughout the entire month of September in 2013 and are not restricted to times and spatial domains where measurements have been acquired. We believe these September results should reasonably represent the results for the full year. However, additional modeling studies need to be run in future studies to definitely confirm this.

5. *Tabulate optimal time periods for select comparisons of airborne CH_2O measurements with ground and mobile CH_2O measurements, focusing on overflights close to DNPH cartridge sampling sites at Clinton, Deer Park and Channelview when sampling at such sites were operative. Compare integrated DNPH measurements with 24-hour synthesized integrated airborne measurements based upon the temporal dependence calculated from the CMAQ model and the WP-3 aircraft measurements acquired at different times throughout the day. To accomplish this, the CH_2O CMAQ model output at the surface will be corrected using WP-3 CH_2O measurements for overflights close to DNPH cartridge sampling sites. The corrected CMAQ model output will then be integrated over the 24-hour DNPH sampling times and a comparison carried out.*

   We assessed the accuracy of 24-hour integrated DNPH cartridge sampling measurements for CH_2O on one occasion at the Deer Park site on Sept. 13. This was carried out by comparing *24-hour synthesized integrated airborne measurements of CH_2O,* based upon the temporal dependence calculated from the CMAQ model and the WP-3 aircraft, with the DNPH cartridge sampling measurements at Deer Park. After applying a small correction to the CMAQ results to match the observations, we determined a 24-hour integrated CH_2O value of 3.799 ± 1.9 ppb on Sept. 13 at the Deer Park sampling site, a value that is in agreement with the integrated DNPH determination of 2.673 ppbv within the precision of the CMAQ value. This 30% difference is in line with the comparison slopes reported by Gilpin et al. [1997] between diode laser measurements of CH_2O standards and those retrieved by DNPH cartridge sampling methods.

6. *As a follow-up to Task 5, employ the CMAQ model output at the surface to identify potential nighttime emissions of CH_2O and/or its precursors.*

   The CMAQ modeling results in conjunction with ground-based auto-GC measurements of propene at the Deer Park sampling site point to possible evidence of nighttime emissions of CH_2O and/or its precursors, as has been suggested by Olauger et al. [2009].

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background

Gas phase formaldehyde (CH_2O) is a ubiquitous component of the troposphere, and is typically the most abundant carbonyl compound found in the lower atmosphere. This gas is formed by the oxidation of most hydrocarbons. Over continental regions, incomplete fossil fuel combustion, biomass burning, industrial processes, and vegetative emissions are generally considered to be the major production pathways of CH_2O [Fried et al. 2003a,b and references therein]. In highly industrialized cities like Houston, Texas, which is home to some of the world's largest petrochemical refineries, significantly elevated CH_2O levels of many tens of parts-per-billion by volume (ppbv) have been observed [Wert et al., 2003, Parrish et al., 2012]. In these cases, petrochemical releases of highly reactive volatile organic compounds (HRVOC's) like ethene and propene are oxidized very rapidly in the atmosphere by OH radicals to produce CH_2O and other compounds. By contrast, in the remote atmosphere, methane (CH_4) oxidation becomes the dominant CH_2O source. Formaldehyde decomposition produces carbon monoxide (CO) and hydrogen radicals, and subsequent reaction of these radicals with nitrogen oxide compounds produce ozone (O_3). Like O_3, CH_2O is a toxic pollutant that has a dramatic effect on air quality and human health. Not only is CH_2O an irritant to the eyes and upper-airways, the International Agency for Research on Cancer [IARC, 2006, 2012] has classified CH_2O as a carcinogen due to acute and chronic exposure. The Texas Commission on Environmental Quality (TCEQ) in a report [2008] has developed both short-term and long-term Effects Screening Levels (ESLs) for outdoor CH_2O concentrations as a guide for the protection of human health and welfare. The TCEQ established short-term (1-hour) and long-term ESL heath limits of 12 ppbv, and 4.5 ppbv, respectively. Furthermore, because of its toxicity, CH_2O is considered a Hazardous Air Pollutant by the 1990 U.S. Clean Air Act.

The observed CH_2O concentrations vary quite substantially in the planetary boundary layer (PBL) over the greater Houston-Galveston-Brazoria Metropolitan Area (HGBMA). As will be shown later in this report, such CH_2O levels vary from values in the 2 – 3 ppbv range, typical of planetary boundary layer values measured in other U.S. cities, to concentrations in excess of 30 ppbv from petrochemical emission sources. Continued development of effective control strategies for CH_2O and its breakdown product O_3 therefore requires a comprehensive understanding of the attribution of CH_2O sources throughout the greater HGBMA. Formaldehyde directly emitted into the atmosphere from petrochemical and automotive sources, to name a few, are considered primary sources of CH_2O. By contrast, CH_2O photochemically produced in the atmosphere from the breakdown of HRVOCs from petrochemical facilities as well as other sources and from biogenic emissions of isoprene, are classified as secondary sources of CH_2O. Despite extensive efforts and advances from past studies, two competing views have appeared in the recent literature, which report very different primary and secondary contributions of CH_2O in the greater HGBMA. For example, based upon airborne measurements of CH_2O and its HRVOC precursors from past campaigns, Parrish et al. [2012] report that 92 ± 4% of the total CH_2O source over the HGBMA arises from secondary sources and primary emissions only account for ~ 5% of the total. By contrast, Johansson et al. [2014] employing optical remote sensing measurements close to petrochemical refineries in the Houston Ship Channel, Texas City, and Mont Belvieu areas coupled with a Lagrangian plume model determined a primary CH_2O contribution of 90% for the greater HGBMA. Despite such

divergent conclusions, both sources of CH_2O may actually be important in different regimes. Close to large petrochemical complexes direct emissions of CH_2O from flaring and other operations can indeed dominate, while further downwind secondary sources can become more important, particularly during summer months when sources of OH radicals from increased solar insolation and higher water summertime vapor levels, result in higher photochemical activity. Because of the importance of this issue, it is highly desirable to revisit the issue of CH_2O source apportionment employing new data acquired in 2013, the most up-to-date emission inventories, as well as new analysis approaches.

Emission inventories and temporal trends for CH_2O and its HRVOC precursors are additional areas requiring attention. The study by Washenfelder et al. [2010] based upon airborne measurements of various constituents (including our CH_2O measurements) carried out during the TexAQS I and II campaigns examined temporal trends for the 2000 to 2006-time period. Based upon trends in the ratios of ethene to NO_x (where NO_x is the sum of nitric oxide and nitrogen dioxide concentrations) and propene to NO_x observed for select isolated petrochemical sources in the greater HGBMA, this study reported a 30% ± 30% decrease in these ratios between data acquired in 2000 compared to 2006, with significant day-to-day and within plume variability (-50% to +100%). The median CH_2O concentration, based upon our measurements, decreased by ~ 40% for this same sampling region over this 6 year period. It is important to note that this 40% drop in CH_2O is consistent with a ~ 30 to 40% drop in CH_2O from 24-hour averaged 2,4-dinitrophenylhydrazine (DNPH) cartridge measurements over this same 6 year period. The Texas Commission on Environmental Quality (TCEQ) acquired such measurements every 6^{th} day at the Clinton, Deer Park and Channelview sites (sites very close to the Houston Ship Channel), and trends were obtained by averaging the data at each site into 3-year bin averages. Extending these measurements from 2000 out to 2011 results in an average yearly CH_2O decrease of 5.8% ± 0.6% for the 3 sites combined. It is highly desirable to further extend these temporal comparisons out to 2013 and to further validate the cartridge results employing highly accurate in situ CH_2O measurements acquired on the NASA P3 aircraft during spirals and missed approaches close to the DNPH sampling sites. This is important since past studies by Herrington and Hays [2012] and by Gilpin et al. [1997] have shown that DNPH cartridge determinations of CH_2O can contain systematic biases even when KI O_3 traps are employed.

Nighttime oxidation of emitted VOC's by O_3 and/or the nitrate radical (NO_3) may also be important sources of CH_2O that could contribute to early morning radical formation, as pointed out by Olauger et al. [2009]. Despite the fact that all WP-3 CH_2O measurements were acquired during daylight hours, it would be highly desirable to investigate the possibility of such nighttime emissions employing the analysis procedures employed in this study.

1.2. Overview of 2013 Airborne Campaigns Over Houston, Texas

In 2013 the University of Colorado team on this proposal deployed highly sensitive, selective and fast infrared absorption spectrometers for measurements of CH_2O during two NASA airborne campaigns over Houston, Texas from mid-August until the end of September: 1) the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) study; and 2) the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC^{4}RS). The former employed the NASA WP-3 aircraft, while the latter employed the NASA DC-8 aircraft. Although measurements from both studies are analyzed, this report focuses on DISCOVER-AQ measurements because: WP-3 measurements extensively covered the greater HGBMA study area

spanning 9 flights compared to 1 DC-8 flight. In addition, WP-3 patterns repetitively sampled the same regions over the HGBMA study area at different times of day over multiple days; and the lower altitudes (including missed-approaches) and slower aircraft speeds provided better capture of surface sources than the DC-8 aircraft. Eight surface monitoring stations were selected where the WP-3 conducted vertical spirals from near the surface up to ~ 4 km. These monitoring stations provided in situ observations of trace gases (O_3, CO, NO, reactive nitrogen species (NO_y), sulfur dioxide (SO_2)), aerosol lidar observations, and balloon soundings of ozone, NO_2, NO_x and water vapor. These eight spiral sites are shown in Fig. 1, which is color-coded by flight altitude. Typically, the WP-3 sampled these eight sites 3 times (3 circuits) each sampling day, with each circuit lasting 2.5 to 3 hours. Nominally, the start of each circuit occurred around 9 am local, 12 noon, and around 2:30 pm local. On a typical flight, the WP-3 took off from Ellington field and sampled over the following sites in the order sampled: Galveston airport, Smith Point, Moody Tower, West Houston, Conroe, Channelview, Deer Park, Manvel Croix, and back to Galveston to start the next circuit. The WP3 acquired data over the following 9 sampling days: 9/4/13, 9/6/13, 9/11/13, 9/12/13, 9/13/13, 9/14/13, 9/24/13, 9/25/13 and 9/26/13.

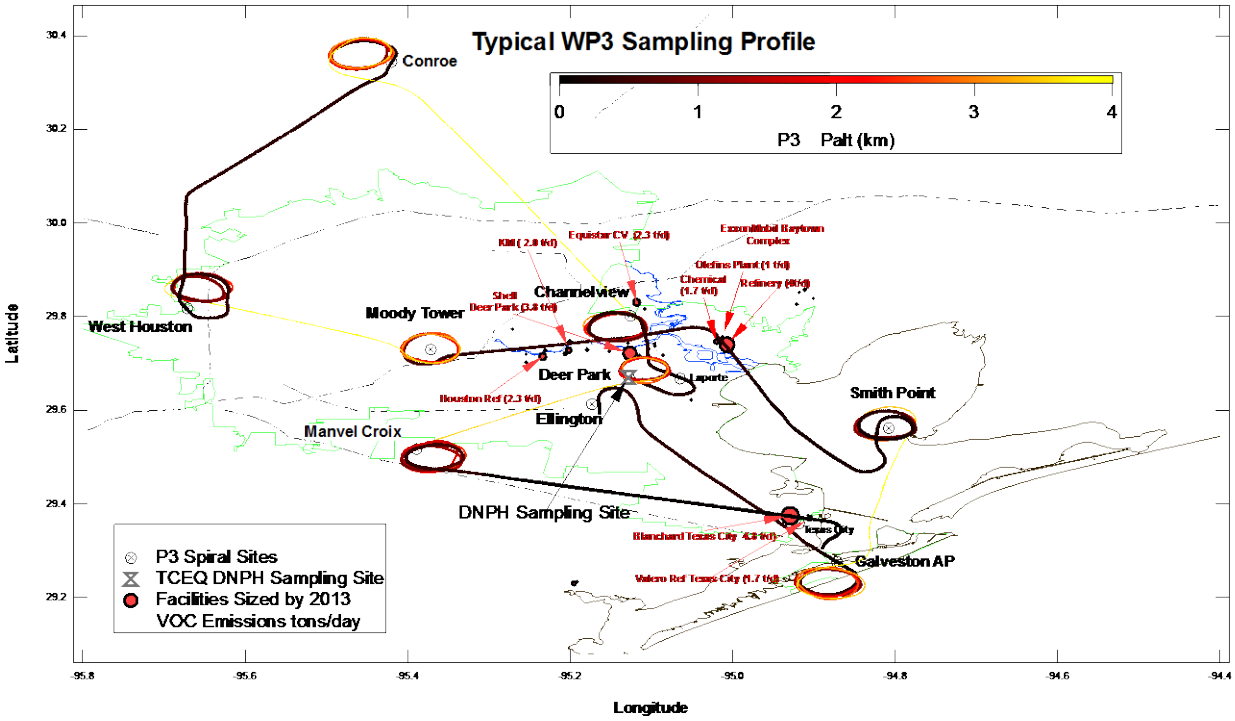

Figure 1: Typical WP-3 flight pattern (colored by altitude) over Houston, Texas, depicting the 8 spiral sites (first site, Galveston airport (AP), and the last site Manvel Croix). Also shown for reference are some of the larger petrochemical facilities under the WP3 flight track. These facilities are sized by their VOC emissions in tons/day (t/d). Also shown is the DNPH sampling site near Deer Park (to be discussed in a later section).

1.3 Project Objectives & Tasks

The primary objectives of this study are to: 1) address the issue of CH_2O source apportionment over the HGBMA study area discussed in the Introduction; 2) assess the current 2012 TCEQ emission inventories for CH_2O and its precursors; 3) assess our knowledge of the chemical mechanisms employed; 4) where possible document emission upsets; 5) identify petrochemical flaring events; and 6) confirm, where possible, the TCEQ DNPH CH_2O sampling results. The following tasks were performed to accomplish these objectives.

1. Prepare WRF and CMAQ input files and run the models using nested domains down to a horizontal resolution of 1 km using the 2012 TCEQ emission inventory. Once accomplished, carry out extensive model-measurement comparisons of CH_2O and other species to test the model accuracy throughout the HGBMA during the DISCOVER-AQ field campaign and assess current emission inventories where possible. A detailed description of the modeling method and evaluation is described in sections that follow.
2. Develop methods to identify, and provide tabulations of, time periods when sampling clearly identifiable direct emission sources of CH_2O close to their source. In this process, tabulate especially large emission sources observed from WP-3 observations and from reported petrochemical facility upsets. Where possible, estimate the magnitude of such events and provide an emission update, as described in a subsequent section. The CMAQ model will be re-run based on such updated emissions estimates. CMAQ output will be analyzed along the path of back trajectories to assess upstream influence. Kinematic back trajectories will be calculated from WRF model output using the WRF post-processing tool RIP (Read/Interpolate/Plot).
3. WP-3 observations of very large CH_2O concentrations in the 20 – 35-ppbv-range from the Sept. 25, 2013 flight during the first two circuits have identified this day as one to examine first employing the high resolution WRF-CMAQ model with updated emissions. This model will be analyzed along a forward trajectory calculated from the WRF output south to Smith Point to help in assessing the model chemistry by comparing the model and observations near and downwind of the source. Other significantly elevated time periods will be identified.
4. Examine the CMAQ model output run with the Process Analysis Mode to quantify the relative importance of primary emissions and secondary photochemical production of CH_2O throughout the HGBMA study area throughout the DISCOVER-AQ field campaign.
5. Tabulate optimal time periods for select comparisons of airborne CH_2O measurements with ground and mobile CH_2O measurements, focusing on overflights close to DNPH cartridge sampling sites at Clinton, Deer Park and Channelview when sampling at such sites were operative. Compare integrated DNPH measurements with *24-hour synthesized integrated airborne measurements* based upon the temporal dependence calculated from the CMAQ model and the WP-3 aircraft measurements acquired at different times throughout the day. To accomplish this, the CH_2O CMAQ model output at the surface will be corrected using WP-3 CH_2O measurements for overflights close to DNPH cartridge sampling sites. The corrected CMAQ model output will then be integrated over the 24-hour DNPH sampling times and a comparison carried out. Further details will be presented in a later section.
6. As a follow-up to Activity 5, employ the CMAQ model output at the surface to identify potential nighttime emissions of CH_2O and/or its precursors.

2. TOOLS & PROCEDURES DEVELEOPED

To address the above objectives & tasks, a collaborative team from the University of Colorado (CU), the University of Maryland (UMD), and the NASA Goddard Space Flight Center (NGSFC) has embarked on the present project. This study is based upon: 1) high quality and fast CH_2O measurements the CU group acquired over Houston, Texas, primarily from the NASA WP-3 aircraft, as discussed in a previous section; and 2) modeling analysis employing the Community Multiscale Air Quality (CMAQ) model with Process Analysis, in very high-resolution mode (1 km resolution), driven by the WRF (Weather Research and Forecasting) meteorological model. Near the end of this project, we discovered an initial error in the elevated point stack exit velocities, and this necessitated re-running all the CMAQ results. This in turn caused a slight delay in submitting this final report. The present results in this report have all been corrected.

2.1. WRF and CMAQ Modeling Method

Two sets of WRF model simulations were performed covering the DISCOVER-AQ Texas field deployment for Texas AQRP Project #14-004, an original and an improved set. The original WRF simulation was run with nested domains with horizontal resolutions of 36, 12, and 4 km. As part of this project, a 1 km domain was added to the improved WRF simulation. CMAQ model simulations were run with the improved WRF simulations for Projects #14-004 and this project. While Project #14-004 utilized the 4 km model results, this project used the 1 km model results. WRF modeling domains are shown in Figures 2 and 3. For this project, the 1 km domain was re-run with CMAQ with the process analysis tool to determine the contribution of direct emissions and secondary production of CH_2O. All CMAQ inputs used in running the improved CMAQ run for Texas AQRP Project #14-004, were utilized in this project. The method of running the improved WRF and CMAQ simulations, which is reported in the Final Report for Texas AQRP Project #14-004, is reviewed here.

WRF and CMAQ were initialized on August 18, 2013 to allow for adequate model spin-up time. The models were run through October 1, 2013. The WRF and CMAQ simulations employ 45 vertical levels extending from the surface to 50-mb (Table 1). The WRF simulation utilized the Multi-scale Ultra-high Resolution (MUR) Sea Surface Temperature (SST) Analysis, which has a horizontal resolution of about 1 km (available at: http://podaac.jpl.nasa.gov/Multi-scale_Ultra-high_Resolution_MUR-SST). The 12 km North American Mesoscale (NAM) model was used for meteorological initial and boundary conditions and analysis nudging. WRF and CMAQ configuration options are shown in Table 2. The WRF simulation employed observational nudging of the National Centers for Environmental Prediction (NCEP) Automated Data Processing (ADP) Global Surface (http://rda.ucar.edu/datasets/ds461.0/) and Upper Air (http://rda.ucar.edu/datasets/ds351.0/) Observational Weather Data. Observational and analysis nudging were performed on all domains. Model output was saved hourly for the 36 and 12 km domains, every 20 minutes for the 4 km domain, and every 5 minutes for the 1 km domain. The WRF model was output at higher temporal resolutions than hourly to prevent the output from being smoothed temporally. CMAQ was run to ingest the meteorology on the same temporal resolution as the WRF model output.

WRF was run straight through (i.e., was not re-initialized at all) using an iterative technique developed at the Environmental Protection Agency (EPA). The EPA successfully used the WRF iterative technique to simulate the meteorology and air quality during the DISCOVER-AQ

Maryland field deployment. Like Houston, Maryland air quality is affected by local-scale bay breeze circulations. A description of improvements and benefits to the WRF-CMAQ system by the EPA, including a description of the WRF iterative technique, is described in Appel et al. [2014]. The iterative technique involved running WRF twice. The first WRF run performed analysis nudging on all domains based on the 12 km NAM. The second WRF run performed analysis nudging on all domains based on the 12 km NAM except for 2-m temperature and humidity for the 4 and 1 km domains. The 2-m temperature and humidity from the 4 and 1 km 1st WRF iterative run was used to nudge the 2nd WRF iterative 4 and 1 km domains. This modeling technique prevented the relatively coarse NAM 12 km model from degrading the high resolution WRF modeling domains (4 and 1 km modeling domains). The 2nd iterative WRF runs were used to drive the improved CMAQ simulations.

For this project, the 4 km CMAQ was re-run for the month of September 2013 using the process analysis tool. The process analysis tool was used to determine the contribution of CH_2O due to direct emissions and secondary production.

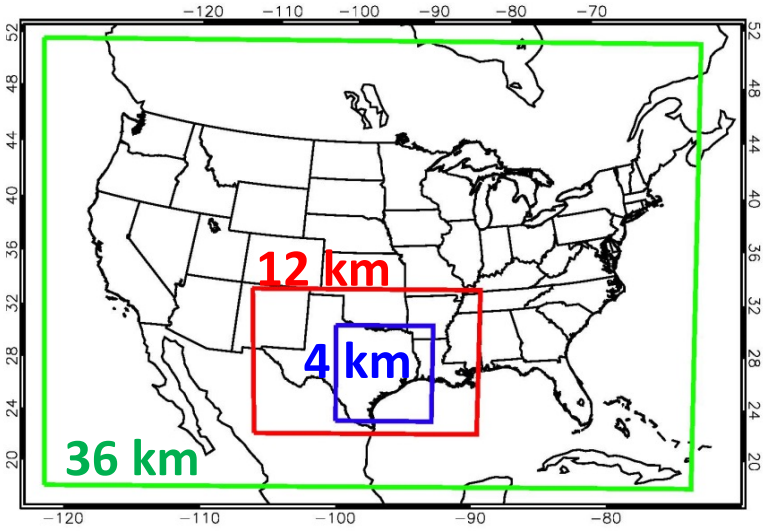

Figure 2: 36, 12, and 4 km CMAQ modeling domains.

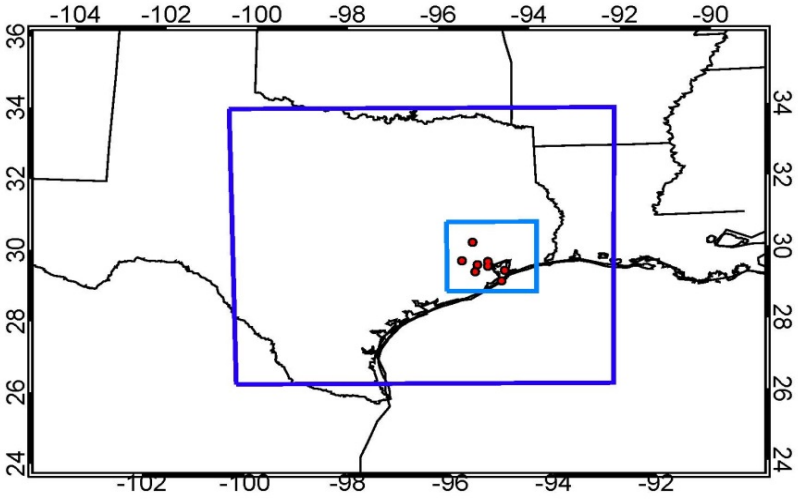

Figure 3: 4 and 1 km CMAQ modeling domains. The red dots show the NASA P-3B aircraft spiral location

η	p_h	η	p_h	η	p_h	η	p_h
1	1013.25	0.8882	905.559	0.561	590.383	0.207792	250.156
0.9974	1010.75	0.8659	884.078	0.526963	557.597	0.18447	227.691
0.994	1007.47	0.841	860.093	0.492715	524.608	0.163354	207.351
0.99	1003.62	0.82069	840.53	0.458342	491.498	0.14	184.855
0.9854	999.187	0.79947	820.089	0.4242	458.611	0.12	165.59
0.9796	993.6	0.775938	797.422	0.390373	426.027	0.1	146.325
0.9723	986.568	0.750095	772.529	0.357176	394.05	0.083	129.95
0.9635	978.091	0.721941	745.41	0.324505	362.579	0.07	117.427
0.9528	967.785	0.691895	716.468	0.292674	331.918	0.052632	100.697
0.9401	955.551	0.660275	686.01	0.262209	302.573	0.03	78.8975
0.9252	941.199	0.627918	654.842	0.233845	275.251	0	50
0.9079	924.535	0.594721	622.865				

Table 1: Terrain-following hydrostatic-pressure vertical coordinates (η) and the hydrostatic pressure (p_h) if surface pressure is 1013.25-mb for WRF and CMAQ simulations at the upper edge of each grid cell.

Weather Research and Forecasting (WRF) Version 3.6.1 Model Options	
Radiation	Long Wave: Rapid Radiative Transfer Model (RRTM) Short Wave: Goddard
Surface Layer	Pleim-Xiu
Land Surface Model	Pleim-Xiu
Boundary Layer	Asymmetric Convective Model (ACM2)
Cumulus	Kain-Fritsch
Microphysics	WRF Single-Moment 6 (WSM-6)
Nudging	Observational and analysis nudging
Damping	Vertical velocity and gravity waves damped at top of modeling domain
SSTs	Multi-scale Ultra-high Resolution (MUR) SST analysis (~1 km resolution)
CMAQ Version 5.0.2 Model Options	
Chemical Mechanism	Carbon Bond (CB05)
Aerosol Module	Aerosols with aqueous extensions version 5 (AE5)
Dry deposition	M3DRY
Vertical diffusion	Asymmetric Convective Model 2 (ACM2)
Emissions	2012 TCEQ anthropogenic emissions Biogenic Emission Inventory System (BEIS) calculated within CMAQ
Initial and Boundary conditions	Model for OZone and Related chemical Tracers (MOZART) Chemical Transport Model (CTM)

Table 2: WRF and CMAQ model options that were used in both the original and improved modeling scenarios.

2.2. WRF and CMAQ Modeling Evaluation

An evaluation of our improved WRF and CMAQ model simulations for the entire month of September 2013 are displayed in this section. Statistics used to evaluate WRF and CMAQ are described in Table 3. WRF and CMAQ statistical analyses are shown in Tables 4-6. Proposed benchmarks for evaluating WRF by Emery et al. [2001] are also shown in Table 3. The benchmarks were created to help put model results into context, not give them a passing or failing grade. For example, expectations for simulating the meteorology of a complex local-scale circulation, like sea and bay breeze circulations in and around Houston are not as high as simulating meteorology over flat terrain. Section 3.3 provides a more detailed comparison of WRF-CMAQ model comparisons with observations for select time periods as well as each day over the course of the DISCOVER-AQ mission.

Model-measurement comparisons for all chemical species were carried out over all 9-flight days of September covering the spatial domain shown in Fig. 1. For this purpose, the CMAQ model output, which represents 5-minute concentration averages in 1km grid boxes, was expanded out to 1-second data, which matched the fast species measurements acquired from the WP-3 aircraft. Comparisons were then carried out on this 1-second time basis. The merits and shortcomings of this time-coincident matching approach as it relates to comparisons of CH_2O will be discussed in a subsequent section. Compared to all hourly AQS observations (Table 5) and all WP-3 aircraft observations (combined planetary boundary layer and free tropospheric values) during the DISCOVER-AQ field deployment (Table 6), CMAQ simulated a high bias in ozone. CMAQ also simulated a low bias in carbon monoxide (CO), isoprene, nitrogen dioxide (NO_2), and nitric oxide (NO) aloft and a high bias in CH_2O and total reactive nitrogen (NO_y) compared to aircraft observations (Table 6). Figures 16a and 16b in a later section will provide further information regarding CMAQ-biases relative to WP-3 measurements of CH_2O for each sampling day in the planetary boundary layer (PBL).

An analysis and evaluation of the meteorology and air quality during an air pollution event that we focused on for this project that took place September 25 are described in the next section.

Mean Bias (MB)	$MB = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$
Normalized Mean Bias (NMB)	$NMB = \frac{\sum_{i=1}^{N}(M_i - O_i)}{\sum_{i=1}^{N} O_i} \times 100\%$
Normalized Mean Error (NME)	$NME = \frac{\sum_{i=1}^{N}\lvert M_i - O_i\rvert}{\sum_{i=1}^{N} O_i} \times 100\%$
Root Mean-Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}$
Gross Error (G)	$GE = \frac{1}{\mathrm{N}}\sum_{i=1}^{N}\lvert M_i - O_i\rvert$

Table 3: Definition of the statistics used in WRF and CMAQ model evaluations. In these equations M represents the model results, O represents the observations, and N is the number of data points.

	2 m Temperature (K)		10 m Wind Speed (m/s)		10 m Wind Direction (deg)	
	Bench mark	Model	Bench mark	Model	Bench mark	Model
MB	≤ ±0.5 K	0.7	≤ ±0.5 m/s	-0.8	≤ ±10°	33
NMB (%)		0.2		-18		26
NME (%)		0.4		38		26
RMSE		1.7	≤ 2 m/s	2.4		51
GE	≤ ± 2 K	1.2		1.8	≤ 30°	33

Table 4: Mean bias (MB), normalized mean bias (NMB), normalized mean error (NME), root mean square error (RMSE), and Gross Error (GE) of 2 m temperature, 10 m wind speed, and 10 m wind direction for the 2nd iterative 1 km WRF simulations covering all of September 2013.

	Surface Ozone (ppbv)
	Model
MB	12
NMB (%)	47
NME (%)	56
RMSE	17
GE	14

Table 5: Mean bias (MB), normalized mean bias (NMB), normalized mean error (NME), root mean square error (RMSE), and Gross Error (GE) of surface ozone for the 2nd iterative 1 km WRF simulations compared to all hourly AQS observations within the 1 km domain for the entire month of September 2013.

		O_3	CO	CH_2O	Iso	NO2	NO	NOy
Model	MB	1.6	-5.2	0.3	-0.02	-0.2	-0.2	0.6
	NMB	2.8	-4.3	-13	-6.9	-13	-52	18
	NME	15	17	37	71	70	77	62
	RMSE	12	40	1.4	0.7	2.5	2.2	4.5

Table 6: Second iterative 1 km CMAQ simulated mean bias (MB), normalized mean bias (NMB), normalized mean error (NME), and root mean square error (RMSE) of O_3, CO, CH_2O, Isoprene (Iso), NO_2, NO, and NOy covering all 1 second average measurements made onboard the NASA P-3B aircraft on all flight days during the DISCOVER-AQ field campaign.

3. SUMMARY OF RESULTS BY TASK

3.1. Overview Summary of Results

Two initial tasks were carried out in parallel by the CU and UMD/NGSFC teams. The CU team initiated their efforts by identifying WP-3 and DC-8 aircraft sampling periods arising from clearly identifiable sources **(Tasks 2 & 3).** These periods were intended to provide a focus for further study by WRF-CMAQ analysis. Because of the large and dynamic pollution levels trapped in a shallow boundary layer, the CU team identified the Sept. 25 DISCOVER-AQ flight for the initial analysis. This team started this analysis by quantifying CH_2O/CO slopes from the final DISCOVER-AQ data for 4 specific CH_2O source regions where: 1) petrochemical refinery emissions were dominant over the Baytown ExxonMobil petrochemical complex; 2) biogenic isoprene emissions were dominant near Conroe; 3) where CH_2O photochemical production downwind of the Baytown and Deer Park petrochemical complexes were dominant over Smith Point; and 4) where automotive and urban sources mixed with residual CH_2O from the previous night together with transported CH_2O were dominant over the center of Houston over Moody Tower. The 4 regions, which are shown in Fig. 1, were meant to illustrate the different source characteristics and highlight the regions for further analysis. Figure 4 below illustrates these 4 regions employing a regression plot of 1-second CH_2O concentration data as a function of 1-second CO data acquired from the WP-3 aircraft during the first circuit (C1) and second circuits (C2) of September 25, 2013.

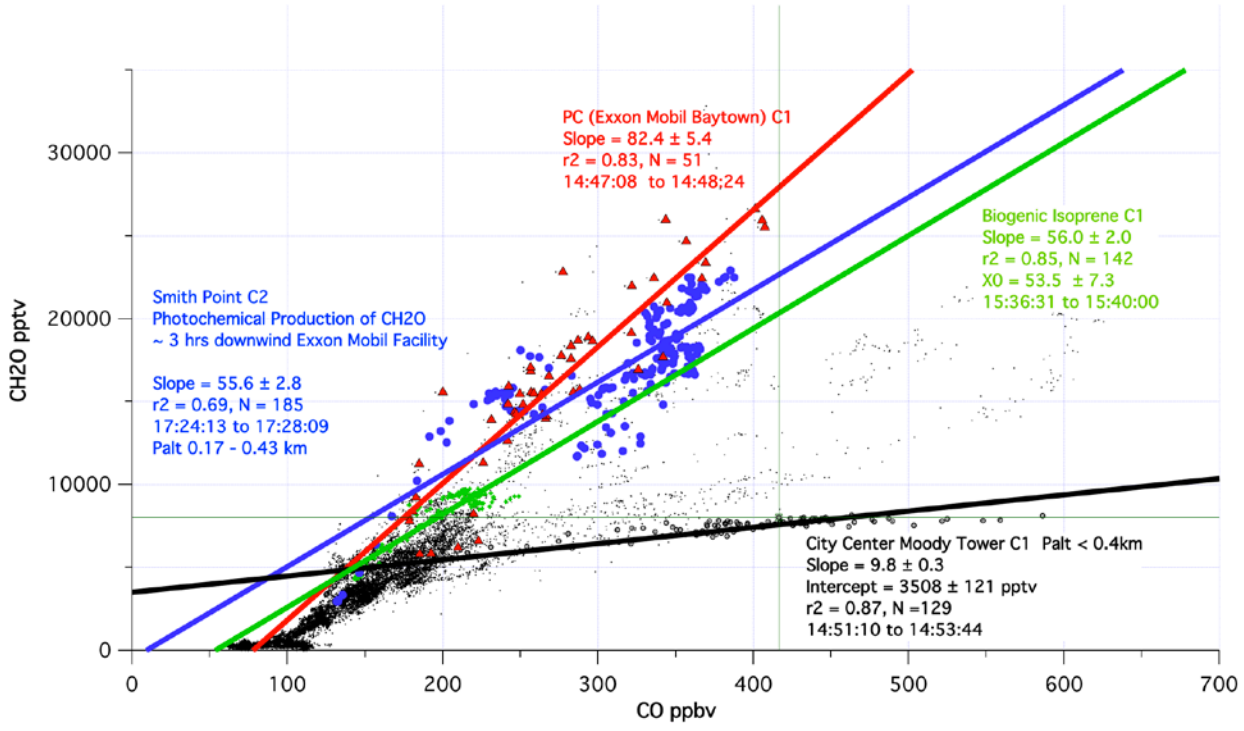

Figure 4: Linear regression slopes of CH_2O versus CO (units of ppt/ppb) for 4 specific CH_2O sources observed from the WP-3 aircraft on Sept. 25, 2013.

The 4 different regions shown in Fig. 4 are: 1) petrochemical refinery emissions near the Baytown ExxonMobil petrochemical complex during the 1st circuit (red PC C1 regression, the 51 regression points in this fit are shown by red triangles); 2) biogenic isoprene emissions (green Biogenic regression C1) near Conroe; 3) secondary photochemical production of CH_2O downwind of the Baytown and Deer Park petrochemical complexes (dark blue C2 regression) observed during the Smith Point spiral; and 4) automotive and urban emission sources over the center of Houston over Moody Tower mixed with residual CH_2O from the previous night as well as transported sources from other regions (black C1 regression). These regions were respectively identified by elevated: propene (PC trace); isoprene (biogenic trace); low propene (reacted away) but favorable wind direction from the Baytown and Shell Deer Park complexes to Smith Point and the shallow boundary layer (<0.5 km from ozonesonde and CH_2O measurements) that confines the emissions and their products (dark blue trace). The local times are 5 hours earlier than the GMT times indicated here.

Figures 5a – 5c further highlight the extreme CH_2O concentrations observed by the CU team from their WP-3 measurements acquired on Sept. 25 relative to other sampling days throughout September in 2013 (Fig. 5d). In all figures, we display the spiral sites and overlay the CH_2O distributions measured on the WP-3 aircraft colored and sized by these concentrations. The CH_2O scale has been restricted in all cases to 20,000 pptv (20-ppbv) to better contrast the levels observed on all days. In Fig. 5a, we plot the CH_2O distributions observed on the WP-3 aircraft during the 1st circuit. A maximum CH_2O concentration of 26.6 ppbv was measured over the ExxonMobil Baytown petrochemical complex at 9:48 am local time. Other significantly elevated levels were observed about 2 minutes earlier over the Bay (22.2 ppbv) and over the Deer Park Shell complex (20.7 ppbv) at 11:18 am local time

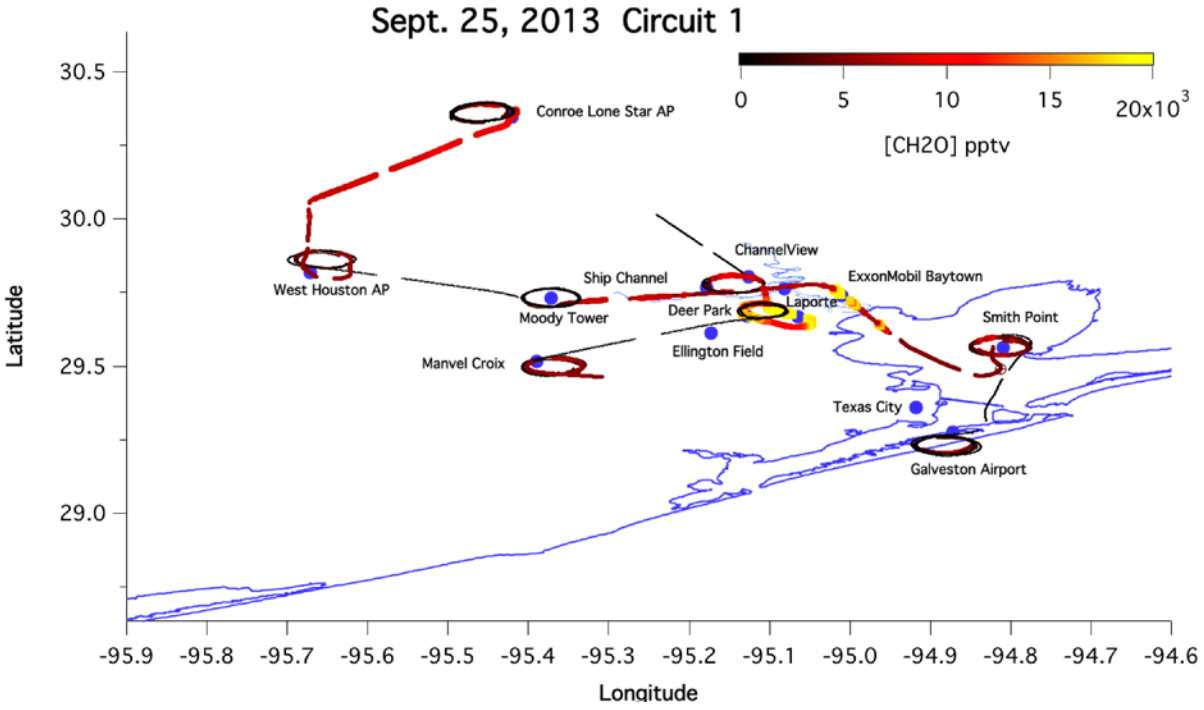

Figure 5a: CH_2O distributions measured on the NASA P3 aircraft during the 1st circuit of the Sept. 25 flight. The color scale here has been restricted to 20 ppbv. The start of this circuit began ~ 9 am local time.

Figure 5b shows the CH_2O distributions observed during the 2nd circuit (starting around noon local time). The measured CH_2O over the ExxonMobil Baytown petrochemical complex at 12:34 local time dropped to 14.5 ppbv but the levels over Galveston Bay and at Smith Point attained values as high as 32.8 ppbv (12:30 pm local) and 22.9 ppbv (12:26 pm local), respectively. In a latter section we will display back trajectories showing that the enhanced CH_2O levels observed over Smith Point during the 2nd circuit passed over the ExxonMobil complex several hours earlier. During the 3rd circuit (Fig. 5c) the CH_2O further dropped to 4.2 ppbv over the ExxonMobil complex at 2:56 pm local and 10.3 ppbv at 2:50 pm local over the Bay just near Smith Point.

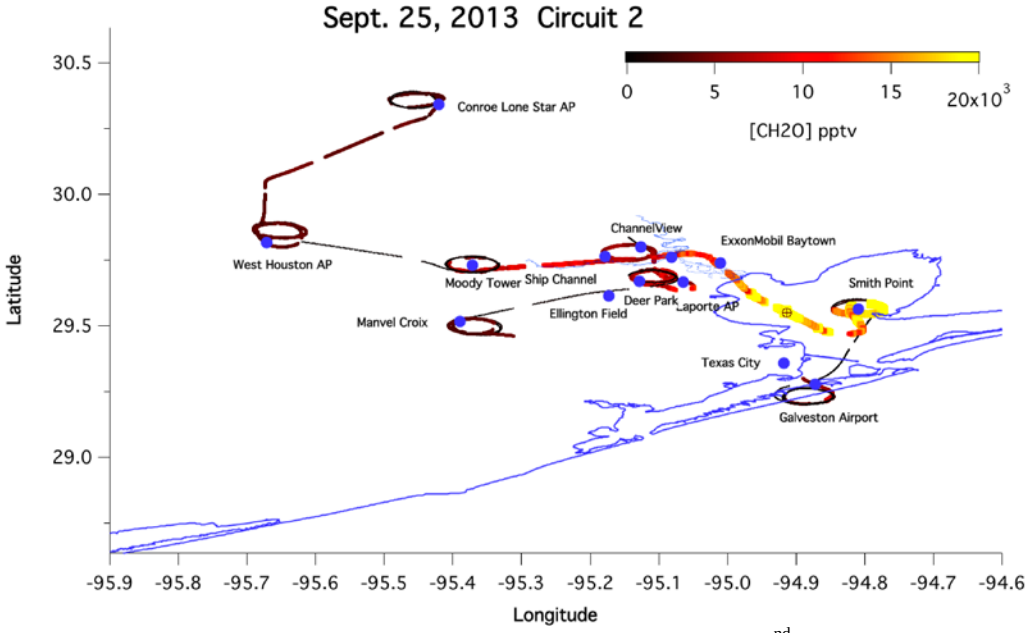

Figure 5b: CH_2O distributions measured on the NASA P3 aircraft during the 2nd circuit of the Sept. 25 flight. The start of this circuit began ~ 12 noon local time.

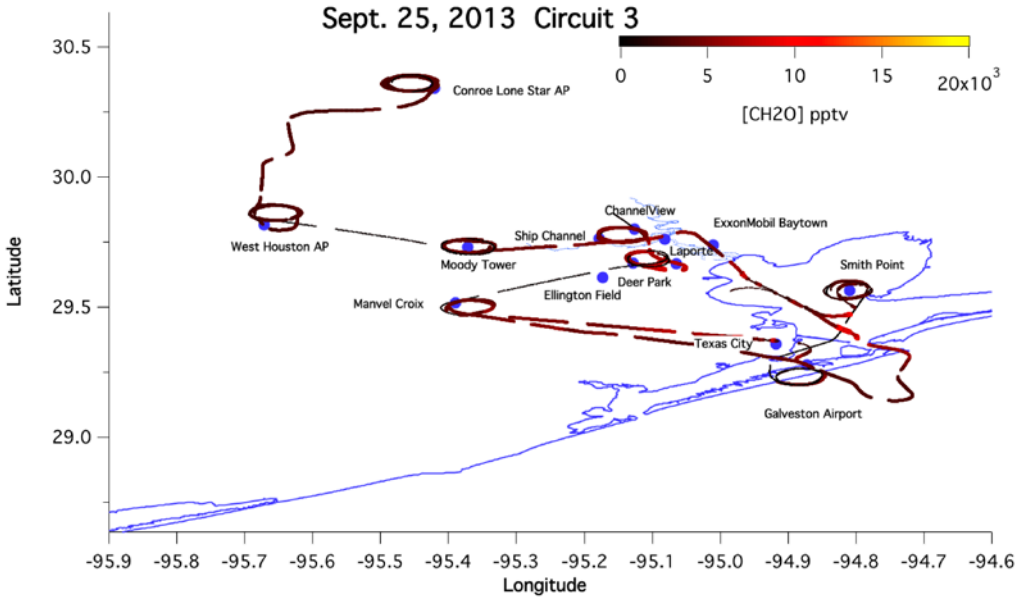

Figure 5c: CH_2O distributions measured on the NASA P3 aircraft during the 3rd circuit of the Sept. 25 flight. The start of this circuit began ~ 2:30 pm local time.

Fig 5d shows the composite of our WP-3 CH_2O measurements observed on all days spanning the time period from Sept. 4 to Sept. 26 with the exception of the 3 circuits on Sept. 25 for comparison. As can be seen, the 1st two circuits on Sept. 25 clearly standout to the rest of our observations. At pressure altitudes less than or equal 0.5-km, we tabulate the following CH_2O concentrations for Fig. 5d: average = 3145 ± 1485 pptv, median = 3064 pptv, 25% value = 2010 pptv, 75% value = 4053, 90% value = 5008 pptv, and maximum value = 14437 pptv. Although many different sources of CH_2O are no doubt responsible for these observations, **it is our opinion that the ExxonMobil Baytown facility plays some role in these extreme events on Sept. 25. However, the significantly elevated CH_2O levels observed on Sept. 25 during the 2nd circuit over Smith Point also has a major contribution from the Deer Park-Channelview region.** Evidence in support of these assertions will be presented in a later section of this report. As a result, Sept. 25 cannot be used in the assessment of normal operating emission inventories nor can it be included in our model validations. However, the extreme events on Sept 25 may serve as an excellent opportunity in assessing the veracity of the model chemistry downwind of large emission sources (**Task 3).** Back trajectories from the CMAQ model indicate that the WP-3 sampled the large ExxonMobil emissions approximately 3 hours downwind at Smith Point during the 2nd circuit, and comparisons of WP-3 CH_2O measurements during the 2nd circuit at Smith Point with CMAQ model results using normal and enhanced emissions provides an excellent opportunity to assess the above topics. However, this aspect, which was included in our original final report, turns out to be far more complicated than originally anticipated and must await additional analyses that are not part of the present study. Therefore, in this report we only discuss our evidence supporting what we believe to be enhanced emissions from the ExxonMobil facility during the 1st circuit of Sept. 25.

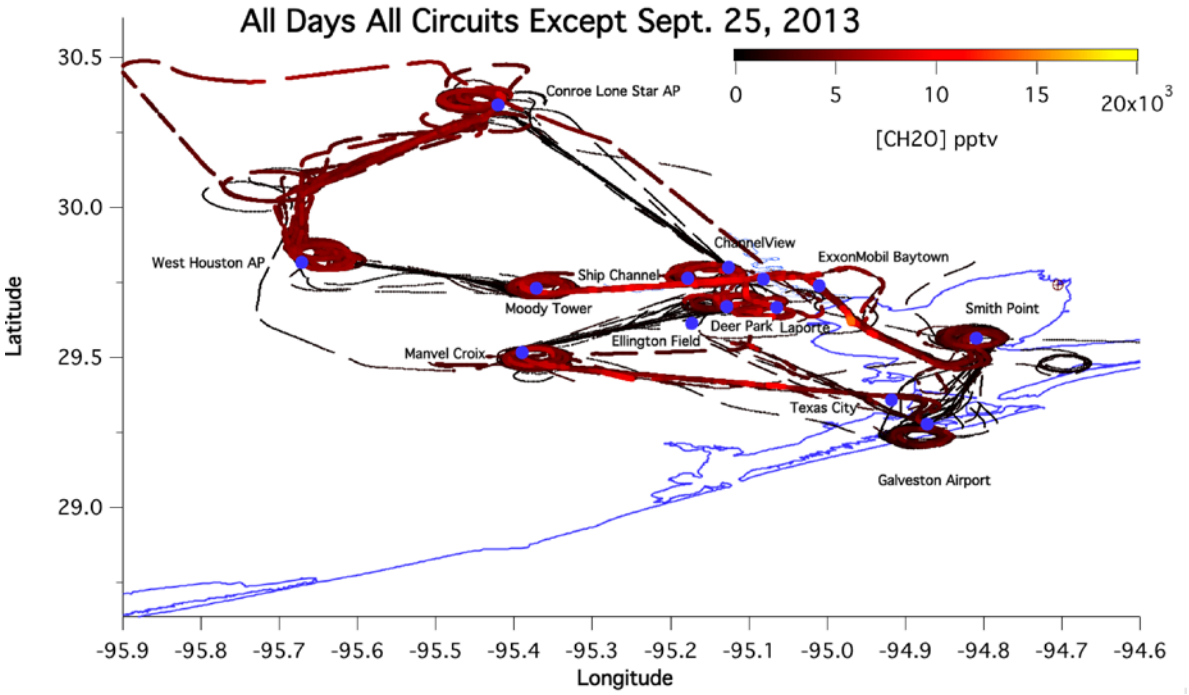

Figure 5d: CH_2O distributions measured on the NASA WP-3 aircraft during the entire DISCOVER-AQ study from Sept. 4 to Sept. 26 with the exception of the 3 circuits on Sep. 25.

In parallel with the efforts above, the UMD/NASA GSFC team updated their modeling technique (**Task 1)** to begin WRF simulations down to a horizontal resolution of 1 km based on preliminary WRF and CMAQ simulations run down to a horizontal resolution of 4 km. A comparison between observations and the original WRF simulation run down to a horizontal resolution of 4 km reveals that the model simulated weaker sea and bay breezes than observed (Figure 6). For September 25, the model simulated the sea breeze front just onshore over Galveston, while observations revealed the bay breeze front pushing farther inland. Recognizing the importance of correcting this problem, the UMD/NASA GSFC team carried out a new WRF simulation to improve the model representation of sea and bay breezes using a new modeling technique, higher resolution meteorological initial and boundary conditions (North American Mesoscale, NAM, 12 km model), and the inclusion of a 1 km horizontal resolution domain. This group performed observational nudging on all model domains and ran WRF iteratively. For the iterative simulation, WRF was first run performing analysis nudging based on the NAM 12 km, and then re-run performing analysis nudging based on the previous WRF simulation. This modeling technique prevented the relatively coarse NAM 12 km model from degrading the high-resolution (4 km and 1 km) WRF modeling domains. Meteorology-Chemistry Interface Processor (MCIP) was run to create meteorological input files for CMAQ for all four domains (36, 12, 4 and 1 km). Figure 7 shows the resulting improved WRF model output alongside temperature and wind velocity observations, and the improvements are further tabulated in Table 7.

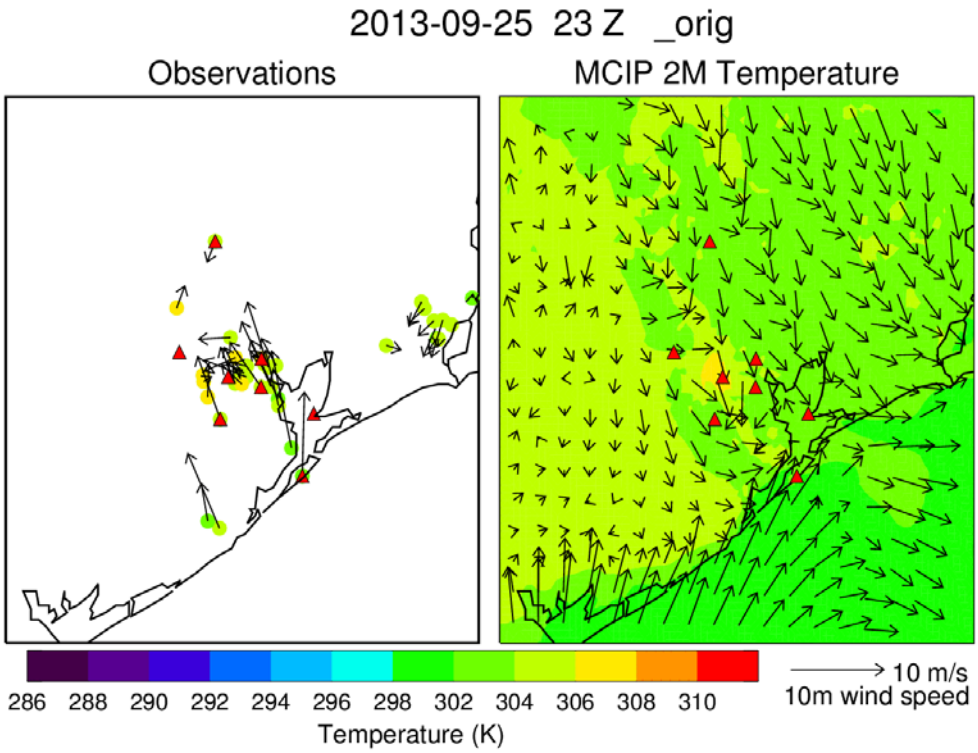

Figure 6: Observed (left) and WRF diagnosed (right) 2-m temperature and 10-m wind velocity at 23 UTC 25 September 2013 from the original 4 km WRF simulation. Strength of WRF simulated bay and sea breezes are weaker than observations.

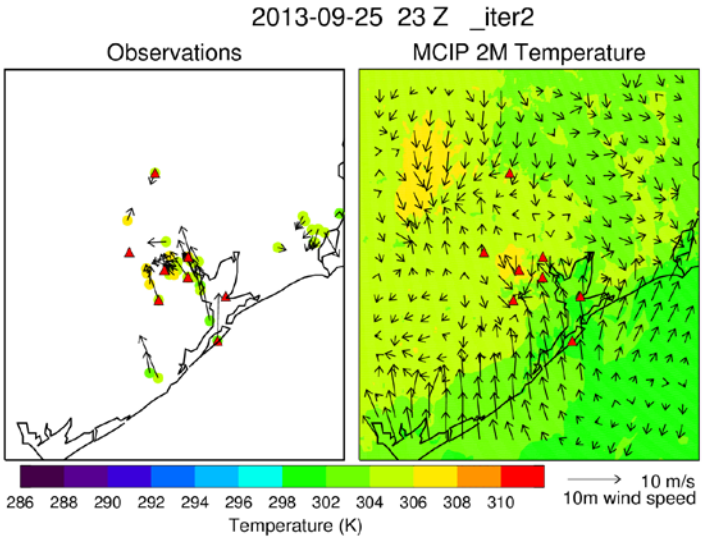

Figure 7: Observed (left) and WRF diagnosed (right) 2-m temperature and 10-m wind velocity at 23 UTC 25 September 2013 from the new 1 km WRF simulation. Strength of WRF simulated bay and sea breezes are in better agreement than in the original simulation.

	2 m Temperature (K)		10 m Wind Speed (m/s)		10 m Wind Direction (deg.)	
	Orig (4km)	Iter 2 (1 km)	Orig (4km)	Iter 2 (1 km)	Orig (4km)	Iter 2 (1 km)
MB	-0.8	-0.1	-0.8	-0.4	56	38
NMB (%)	-0.2	-0.04	-21	-12	39	26
NME (%)	0.4	0.4	50	43	39	26
RMSE	1.6	1.4	2.2	1.9	73	55
GE	1.3	1.1	1.7	1.5	56	38

Table 7: Mean bias (MB), normalized mean bias (NMB), normalized mean error (NME), root mean square error (RMSE), and Gross Error (GE) of 2 m temperature, 10 m wind speed, and 10 m wind direction for the 2nd iterative 1 km WRF simulations covering September 24-26, 2013.

On September 25, CMAQ simulated a low ozone bias compared to surface observations at Smith Point and near the Ship Channel around Baytown (Fig. 8). **Smith Point is of interest since it is the receptor of significant pollution plumes originating from the ExxonMobil Baytown and Shell Deer Park petrochemical facilities on this day.** As will be shown, large CH_2O enhancements were observed by the WP-3 near Baytown and Deer Park on this day, **which were not included in the emissions inventory used by CMAQ**.

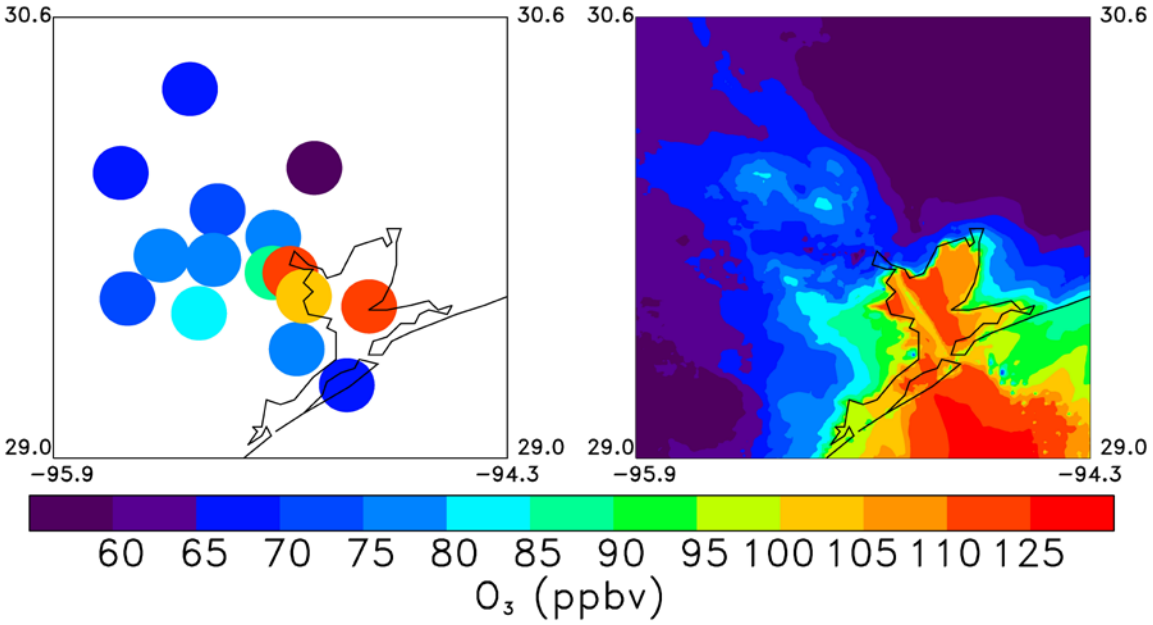

Figure 8: Eight-hour average ozone maximum from observations (left) and new 1 km CMAQ simulation on 25 September 2013.

Like ozone, CH_2O comparisons of CMAQ with measurements acquired from the WP-3 aircraft during the Sept. 25 flight (**Tasks 2, 3)** reveals a low model bias near the ExxonMobil Complex during the 1st and 2nd circuits, downwind of the ExxonMobil Complex over Smith Point, and in all three of the Deer Park spirals (Figure 9).

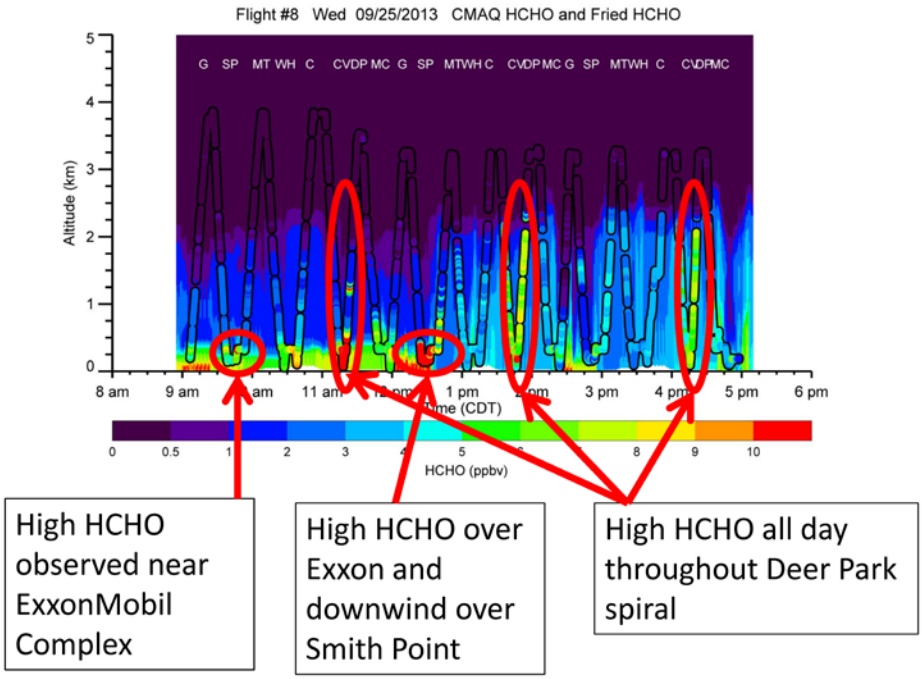

Figure 9: Annotated curtain of CMAQ CH_2O concentrations (background) along the NASA WP-3 flight track (overlay) shows high CH_2O concentrations near the Exxon-Mobil Complex during the first circuit, downwind of the Exxon-Mobil Complex over Smith Point (SP) during the 2nd and 3rd circuits, and during all three of the Deer Park (DP) spirals. The other abbreviations at the top of this figure are: Galveston (G), Moody Tower (MT), West Houston (WH), Conroe (C), Channelview (CV), and Manvel Croix (MC). Figure 1 shows these locations on a map of the HGBMA study area.

In addition, CMAQ exhibited a low bias for CO and isoprene near the Exxon-Mobil Complex, downwind of the ExxonMobil Complex, and over Deer Park; a low bias for NO over the Exxon-Mobil Complex and Deer Park; a low bias for NO_2 downwind of the ExxonMobil Complex over Smith Point; **and high biases in ozone over combustion sources near the ExxonMobil complex from ozone titration with NO (to be shown).** The causes of these biases are now explained. The low model biases in CH_2O, CO, and NO near the ExxonMobil Complex and Deer Park are all consistent, and suggests a low bias in model emissions estimates in these areas, as previously mentioned. This behavior is shown over the ExxonMobil Complex in Fig. 10 for all the gases mentioned above. Although, even if the CMAQ emissions input files included the emissions from this emissions enhancement (to be discussed in connection with Fig. 12a), the model is not expected to capture the peak concentrations in the plume, and minimum for ozone due to titration, and due to the lack of plume in grid treatment within the model. Figure 11 shows a similar time-series for all of the gases mentioned above downwind of the ExxonMobil complex near Smith Point showing the low model bias downwind of the enhanced emissions event.

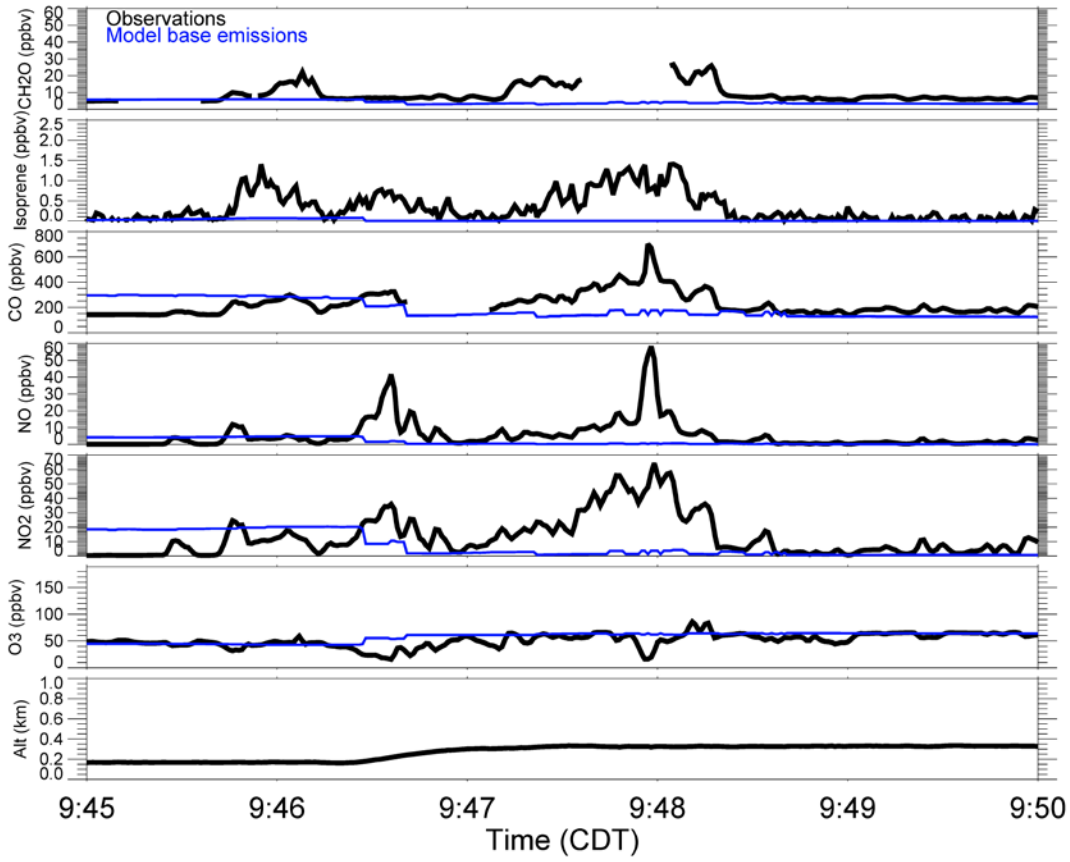

Figure 10: Time-series comparisons of WP-3 (black) and CMAQ CH_2O, CO, NO, and O_3 (red) near the Exxon-Mobil Complex during the 1st circuit on September 25. The inverse relationship exhibited by O_3 around 9:48 CDT relative to NO, CO, and CH_2O is indicative of a combustion event from the Exxon-Mobil facility. This will be further shown in another plot in Section 3.2.

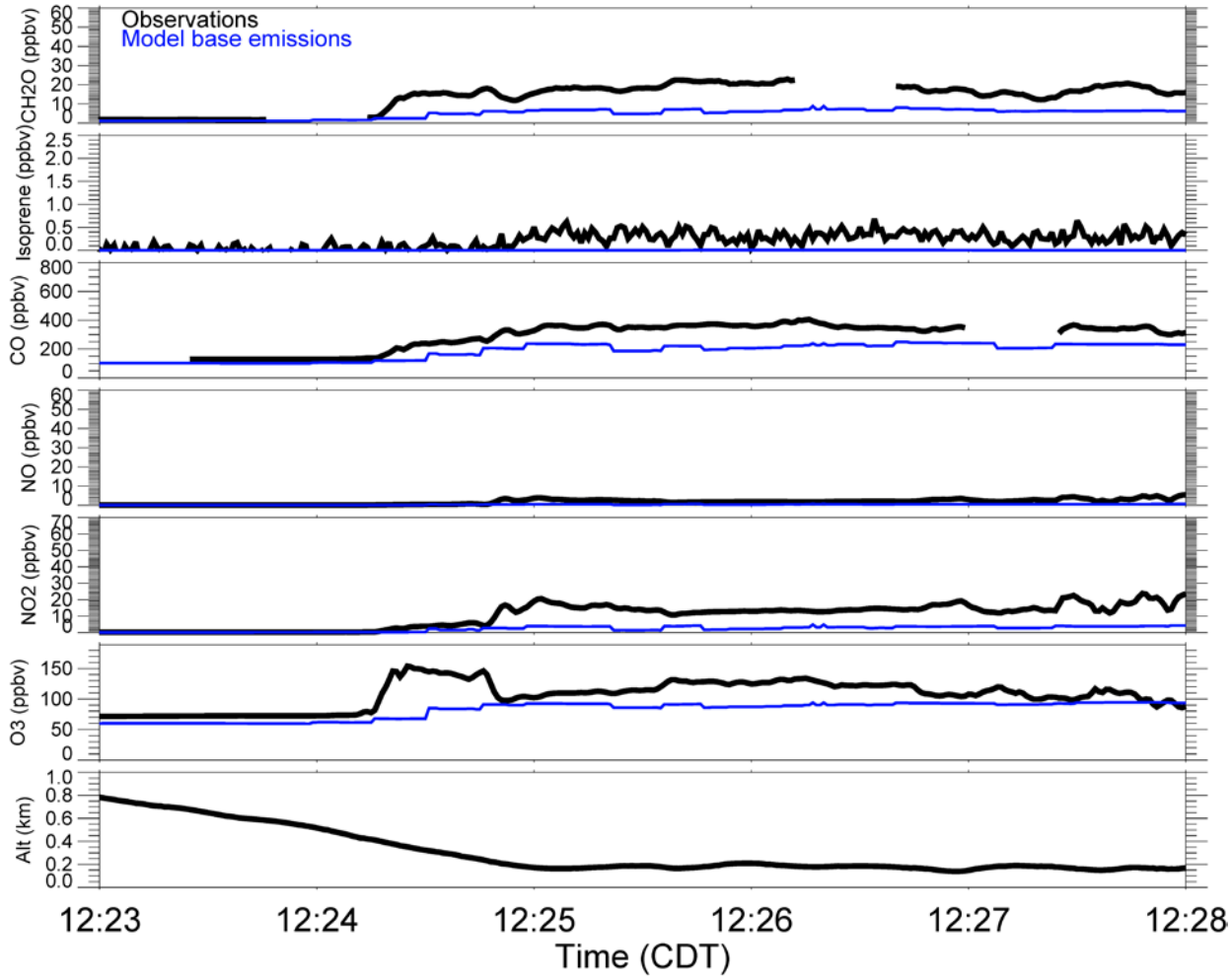

Figure 11: Time-series comparisons of WP-3 (black) and CMAQ CH_2O, CO, NO, and O_3 (red) downwind of the Exxon-Mobil Complex near Smith Point during the 2nd circuit on September 25.

3.2. Measurement Characteristics of Direct Emission Events (Tasks 2 & 3) & Evidence for Enhanced CH_2O Emissions From the ExxonMobil Complex on Sept. 25

Events where we observe major direct emissions of CH_2O and other gases from the WP-3 over petrochemical facilities, possibly from flaring events, in the greater HGBMA study area has received a great deal of interest from AQRP and TCEQ. Although there is some discussion as to whether or not our direct emission observations reflect petrochemical flaring operations, the evidence we present below shows enhanced CH_2O and CO emissions from petrochemical sources without further characterizing the nature of the specific sources. **As we will show, we observe enhanced CH_2O emissions concurrent with combustion sources.** However, at this point we do not have enough information to discern if such enhanced CH_2O: 1) originates directly from the combustion sources; 2) is produced during combustion chemically from its two major precursors propene and ethene; 3) occurs simultaneously from fugitive emissions of CH_2O and propene and ethene; or 4) some combination of the above. In the discussions below, reference to *enhanced direct CH_2O emissions* refers to all 4 possibilities. Likewise, we do not have enough information to even speculate on the types of petrochemical combustion sources (e.g., flaring, fluidized catalytic cracking combustion, or other potential petrochemical combustion sources) that might be responsible for our observations, and therefore efforts to correlate which petrochemical stack that might be responsible for our observations is beyond the scope of this effort.

The overall topic regarding the presence of large and sporadic direct CH_2O emission events from petrochemical facility upsets in the greater HGBMA area has also received a great of interest and at the same time has been a source of some controversy among researchers studying Houston. On one hand, Parrish et al. [2012 and references therein] provide evidence based upon TexAQS 2000 and 2006 studies employing our aircraft CH_2O measurements, which sampled numerous individual petrochemical plumes, that such upset direct events were not evident. By contrast, Olauger et al. [2009] argue that such direct primary CH_2O emissions from flares could explain some of their observations and follow up studies regarding the observation of flaring emissions [Johansson et al., 2014, as one example] supported the importance of such events. Therefore, an effort in this study was devoted to revisiting this issue by identifying and documenting time periods whenever the WP-3 sampled major (to be defined) direct emission sources of CH_2O, either directly from combustion or concurrent with such sources over the greater HGBMA study area.

Although the discussions below focus on the largest enhancement event we observed in this study (the 1st circuit of Sept. 25 over the ExxonMobil complex), this is not meant to convey that this was the only direct CH_2O emissions event we observed. In fact, as shown in a spreadsheet provided to AQRP, we observed a number of other direct emissions sources of CH_2O concurrent with combustion, and the time series plot for the entire Sept. 25 day (to be shown) reveals elevated CH_2O over large regions of the study area. This will be further discussed.

Figure 12a below provides a more detailed view of the results shown in Fig. 10, showing the temporal profiles for CH_2O, CO, propene, O_3, altitude, and the NO_x/NO_y ratio for the first WP-3 circuit on Sept. 25, 2013 near the Baytown ExxonMobil Complex. The times in this figure, which are GMT, correspond to the CDT times in Figs. 10 & 11 by subtracting 5 hours. As noted in Fig. 1, this complex consists of 3 individual facilities in close proximity. As will be discussed in Fig. 12b, this complex is not only one of the largest sources of VOC emissions in the greater HGBMA study area but also is somewhat isolated from the numerous other sources in and around the Ship Channel. In our analysis of Fig. 12a, we employ the ratio NO_x/NO_y since it provides an indication of the degree of photochemical processing. Fresh emission plumes, such as from flaring and other combustion processes exhibit ratios in the 0.9 to 1.0 range, where nearly all of the nitrogen-oxides are in the form of NO_x ($NO+NO_2$). As the air mass ages, the NO_x undergoes oxidation to form species such as HNO_3, PAN, alkyl nitrates and other species. In addition, in fresh combustion, O_3 is titrated by the emitted NO, resulting in highly anti-correlated (negative) CH_2O-O_3 and CO-O_3 slopes. The time over which such anti-correlation persists will vary somewhat depending upon several photochemical variables, the most important of which is the photolysis frequency of NO_2 in producing NO and $O(^3P)$, which then goes on to reproduce O_3 downwind. Mueller et al. [2015] employing the Master Chemical Mechanism (v3.3) Chemistry 0-D photochemical box model in simulating a small fire plume over Georgia, indicates that such anti-correlation will last for ~ 10 minutes of processing, after which net formation of O_3 starts. **Thus, such anti-correlations will reflect highly localized combustion events, and this together with NO_x/NO_y ratios near unity provide an indication when localized combustion has been sampled.**

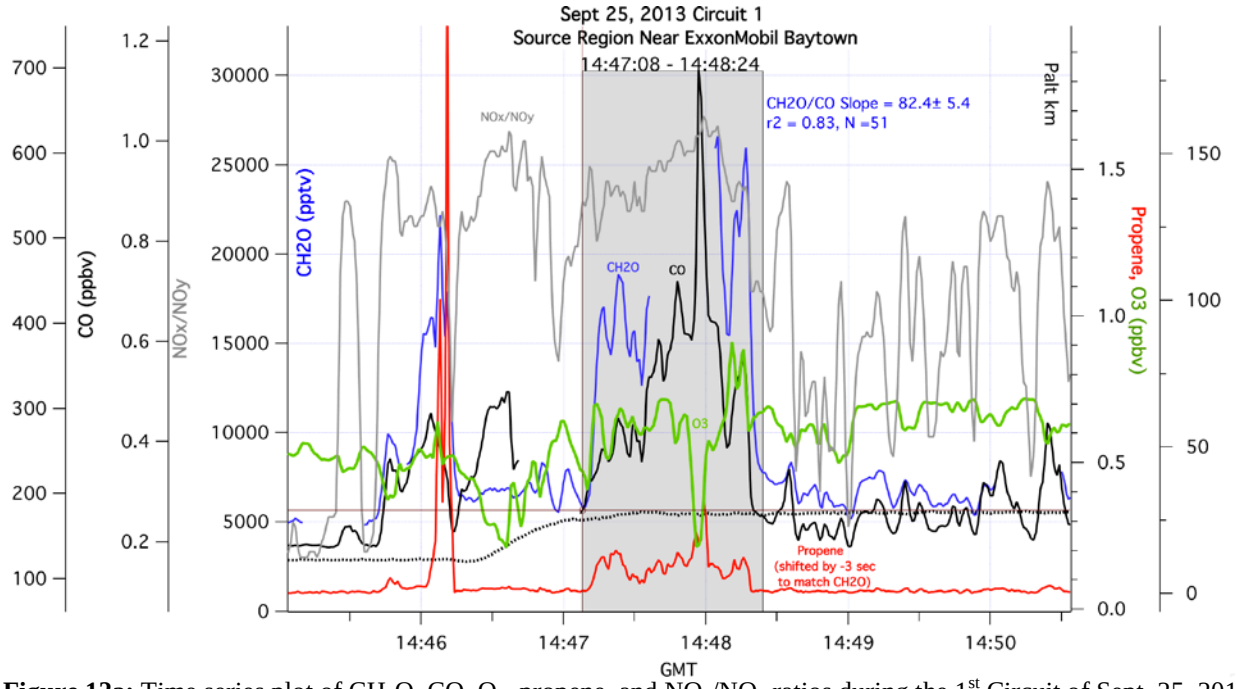

Figure 12a: Time series plot of CH_2O, CO, O_3, propene, and NO_x/NO_y ratios during the 1st Circuit of Sept. 25, 2013 near the Exxon-Mobil Baytown complex**.** Note the time axis here spans the same range as Fig. 10, only the axis here is GMT (CDT + 5 hours).

The shaded region in Fig. 12a between 14:47:08 and 14:48:24 depicts the overpass over the ExxonMobil complex during the 1st circuit on Sept. 25, and this is further shown on the map in Fig. 12b. A subsection of this region (14:47:44 – 14:48:11) shows the aforementioned strong anti-correlations as well as NO_x/NO_y ratios near unity and enhanced propene (one of the two major highly reactive CH_2O precursors) measured by A. Wisthaler's group employing a PTR-MS instrument onboard the WP-3. As can be seen, on both sides of this subsection, CO-O_3 correlations are evident rather than anti-correlations. The regression slopes and r^2 values of Fig. 12a are: CH_2O/CO (82.4 ± 5.4 pptv/ppbv, r^2 = 0.83, N = 51); O_3/CO (-0.132 ± 0.019 ppbv/ppbv, r^2 = 0.64, N = 28 for the subsection 14:47:44 – 14:48:11); and propene/CO (39.6 ± 4.2 pptv/ppbv, r^2 = 0.54, N = 77). The corresponding range for the NO_x/NO_y ratio is 0.88 to 1.05. Unfortunately, we do not have CH_2O data at the peak of the CO observations at 14:47:57 since the instrument was in a zeroing mode. Based on the CH_2O/CO slope of 82.4 pptv/ppbv and the peak CO value of 702.5 ppbv, one would expect a peak CH_2O value of 57.9 ppbv at this time.

We can compare the CH_2O/CO slope during the 1st circuit of Sept. 25 with similar measurements carried out over this same facility at approximately the same time on other days to further highlight what we believe to be enhancements on Sept. 25. Figure 14 shows one example. Here we show measurements acquired over the ExxonMobil facility during the 1st circuit of Sept. 13, and this plot shows exactly the same profile behavior as Fig. 12a only with continuous measurements of CH_2O at the peak CO and a significantly reduced CH_2O/CO slope. We thus assert that the missing CH_2O data at the peak CO in Fig. 12a, although inconvenient, does not invalidate our CH_2O-CO regression analysis. In total, we have measurements of CH_2O/CO slopes on 4 days in addition to Sept. 25 over the ExxonMobil complex at around the same local time during the 1st circuit of the WP-3. These days and the resulting regression slopes are: (1) 9/6/13 (slope = 49.7 pptv/ppbv ± 5.2, r^2 = 0.65, N = 52); (2) 9/12/13 (slope = 24.3 pptv/ppbv ± 5.5, r^2 = 0.54, N = 19); (3) 9/13/13 (slope = 24.3 pptv/ppbv ± 1.3, r^2 = 0.94, N = 22); and (4) 9/24/13 (slope = 23.2 pptv/ppbv ± 1.8, r^2 = 0.89, N = 23). This analysis is predicated on the assumption that CH_2O, propene, and CO are co-emitted from a common source, and hence co-vary, and that additional sources of these gases from other locations have a minimal impact. The

latter topic will be further discussed. This approach has been employed by Herndon et al. [2012] to arrive at mole fraction ratios in flaring plumes.

We consider the above 4 days as *normal operating days* for the ExxonMobil facility. More comprehensive details of these and other time periods where we observe clear direct emissions of CH_2O associated with combustion events can be found in the spreadsheet supplied to AQRP with this document. It is important to note that in some cases the fast measurements of one or more species had to be shifted by several seconds in order to line up distinct structural features before slopes were calculated for this table. The -3 seconds shift in propene in Fig. 12a is one example of this. The grand average for the above *normal operating days* yields a CH_2O/CO slope of 30.4 ± 12.9 pptv/ppbv and a grand median slope of 24.3 pptv/ppbv, which is a factor of 2.7 to 3.4 times lower than the 82.4 slope measured on Sept. 25. Based upon the 2013 Speciated Release Inventory under normal operating conditions (this document is included with this report as a separate spreadsheet) one would expect a normal operating slope CH_2O/CO slope of ~ 12 for all three ExxonMobil facilities combined, which is a factor of ~ 2.0 to 2.5 times lower than our 4 day grand (average/median) values. However, when one considers that our 4-day grand (average/median) measured slopes reflect the sum of CH_2O released as well as CH_2O produced from propene and ethene released from these same facilities under normal operating conditions, we view this factor of 2.0 to 2.5 difference as a reasonable range of values for normal operating conditions. However, the factor of ~ 7 times higher measured slope on Sept. 25 relative to the normal operating Speciated Release Inventory is considerably higher**,** and in our opinion, suggests enhanced emissions of CH_2O and potentially its precursors on Sept. 25 emanating from the ExxonMobil complex during the morning hours, perhaps by as much as a factor of ~ 3. What's interesting is that in contrast to the CH_2O/CO slope comparisons, the propene/CO slope between the 4-day normal operating average (30.3 ± 1.9 pptv/ppbv) is in the ballpark of the 39.6 ± 4.2-pptv/ppbv value measured on Sept. 25. This may imply that CH_2O and ethene emission enhancements might be more dominant than propene enhancements or that propene may be destructively removed nearly equally on each of the sampling days from flaring operations while CH_2O is not.

We present below additional arguments to support our hypothesis above as well as potential caveats that might argue against this. In our previous version of this final report, we attempted to provide individual enhanced emission estimates for various species in moles/hour emanating from the ExxonMobil complex on Sept. 25 that were then used to calculate concentrations downwind at Smith Point. Unfortunately, other than our factor of ~ 3 deduced enhanced CH_2O plus alkene emissions from this complex, there are too many unknowns at the present time to further justify this exercise. This will have to await further analysis under a separate proposal.

Fig. 12b provides further information relating our measurements during the 1st circuit of Sept. 25 relative to the major VOC sources in the region. Figure 12b shows the WP-3 flight track colored and sized by the measured CH_2O during the first circuit of Sept. 25. The shaded region of Fig. 12a (14:47:08 to 14:48:24) over the ExxonMobil complex is highlighted here in red. Like Fig. 1, we show the VOC emissions (tons/day) from the largest facilities reported in the 2013 Speciated Release Inventory for normal operations, and these are sized by the emissions levels. The wind vectors measured on the WP-3 at the start, end, and at the maximum CH_2O level (14:48:05, 26.6 ppbv) are shown by the red vector arrows. As can be seen, the highest CH_2O levels occur right over the ExxonMobil complex, almost immediately downwind of the chemical facility and no other large emission sources upstream are evident. The next largest

source, the Shell Deer Park facility is ~ 11 km upwind and not in the general direction of the wind flow during the overpass time. There are also a number of smaller VOC emission sources in this same direction as the Shell Deer Park facility. The only close large VOC emission source to the northeast of ExxonMobil is the Mount Belvieu complex at 67 degrees approximately 16 km away. Wind vectors at the Lynchburg Ferry site (approximately 7 km northwest of the ExxonMobil complex at a compass bearing of ~ 282 degrees from this facility, see Fig. 12b) during the early morning hours between 4 and 7 am indicated airflow from the general sector of Lynchburg Ferry (294 to 323 degrees, wind speeds of 1.2 to 1.5 m/s) to the ExxonMobil complex. Between 8 and 10 am the winds at Lynchburg Ferry shifted to northeasterly flow (15 to 34 degrees, 2.0 to 3.1 m/s). Based on this information, it is not immediately obvious that highly localized sources of CH_2O, ethene, and propene affect our ExxonMobil observations at 9:48 am. However, we will discuss this further at the end of this section.

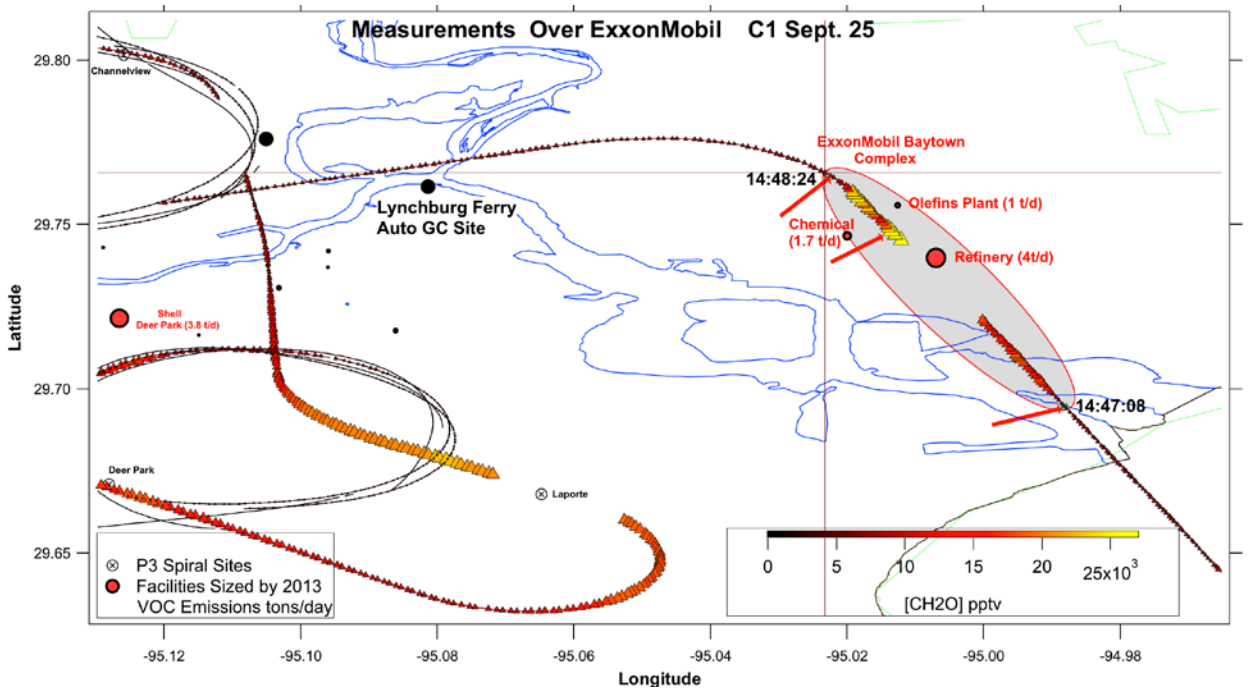

Figure 12b: WP-3 flight track colored and sized by the measured CH_2O concentrations during the 1st circuit of Sept. 25. Also shown for reference are some of the larger petrochemical facilities under the WP3 flight track. These facilities are sized by their VOC emissions in tons/day (t/d).

Figure 13 below shows contrasting behavior to that depicted in Fig. 12a. Here we plot the same series of measurements acquired over the ExxonMobil complex (shaded region) only during the 2nd circuit approximately 3 hours after the 1st circuit. As can be seen, CH_2O, CO, and O_3 are all highly correlated (not anti-correlated as in Fig. 12a) and the NO_x/NO_y ratio resides around 0.14. The O_3/CO slope is + 0.242 ± 0.009 ppbv/ppbv (r^2 = 0.89, N = 93), the CH_2O/CO slope is 67.6 ± 5.2 pptv/ppbv (r^2 = 0.72, N = 68), and the propene/CO slope is 11.8 ± 3.2 pptv/ppbv (r^2 = 0.13, N = 93). All these observations are consistent with the prevalence of photochemical production of CH_2O in the atmosphere and not enhanced CH_2O emissions concurrent with combustion sources shown in Fig. 12a. Hence, arguments suggesting that the results of Fig. 12a reflect the dominance of CH_2O produced from stagnant recirculating air is not supported by our observations.

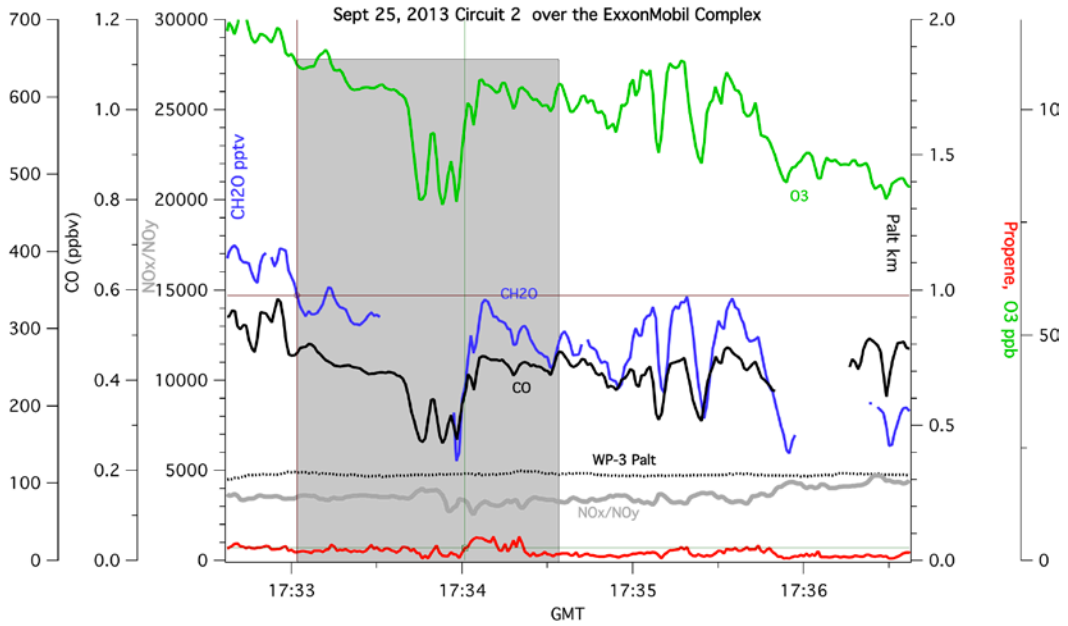

Figure 13: WP-3 measurements during the 2nd circuit of Sept. 25 over the ExxonMobil complex.

As mentioned previously, Fig. 14 below provides additional information regarding our Sept. 25 analysis. Here we show measurements acquired over the ExxonMobil facility during the 1st circuit of Sept. 13, which is one of the 4 days we have considered as *normal operating days*. This plot shows the exact same behavior as Fig. 12a only with a factor of 3.4 lower CH_2O/CO slope.

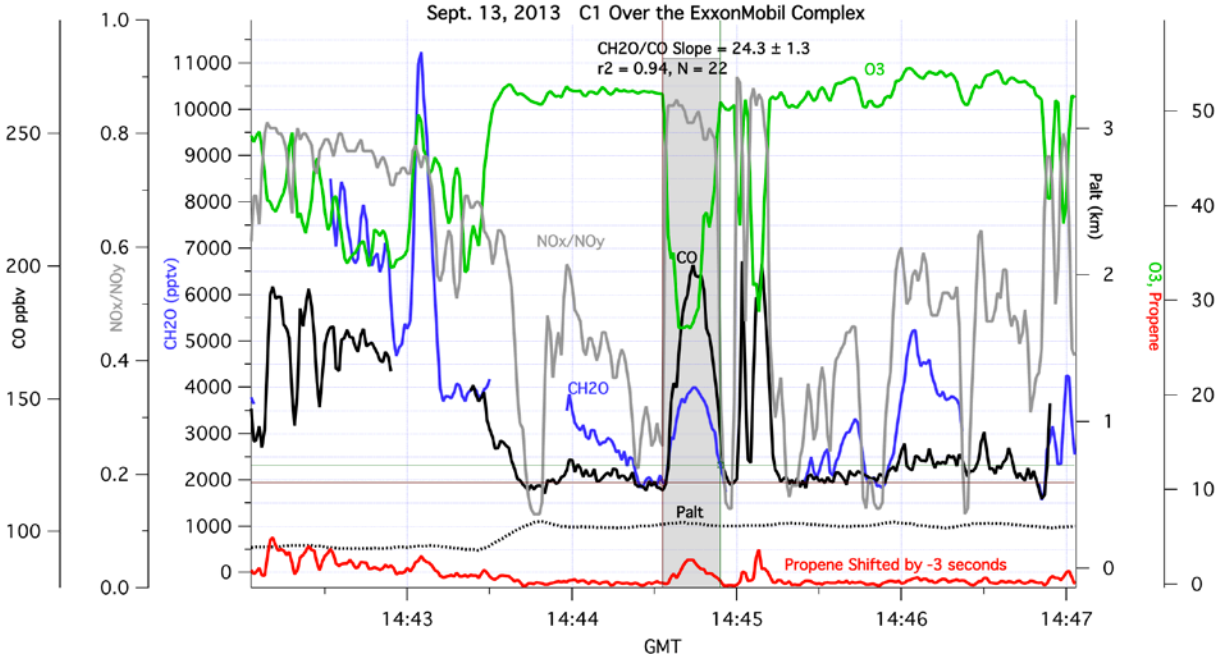

Figure 14: Time series plot over the ExxonMobil facility during the 1st circuit on Sept. 13 in the same format as Fig. 12a

The meteorology on Sept. 25 was somewhat unique relative to the other sampling days. Aircraft profile measurements on this day over Moody Tower reveals a strong inversion and a thin boundary layer (PBL height~ 0.3 km). Any emissions trapped in this thin layer will result in higher mixing ratios relative to emissions emitted into a thicker boundary layer. It is our belief that our analysis, which relies on CH_2O/CO regression slopes, should be not be affected by this unless enhanced CH_2O production simultaneously occurs due to elevated mixing ratios of its precursors in this thin layer from large unknown nearby sources of these VOCs. Upwind measurements taken immediately after the ExxonMobil overpass in Fig. 12b past the Lynchburg Ferry site and over Moody Tower reveals a typical urban CH_2O/CO slope of 9.8 ± 0.3 (Fig. 4), suggesting that such unknown large VOC sources are not present, at least around the time of our ExxonMobil measurements. However, we cannot eliminate the possibility that large VOC sources several hours earlier may have an effect on our profiles shown in Fig. 12a by coalescing with the freshly emitted combustion plume, thus affecting our CH_2O/CO slope. Lynchburg Ferry TCEQ auto GC measurements of ethene and propene in the early morning hours from 5 am to 10 am in fact reveal the highest ethene and propene levels on Sept. 25 relative to all other days during the DISCOVER-AQ study period. The average 5 to 10 am ethene mixing ratios on Sept. 25 were 3 to 15 times higher than the corresponding values during the 4-*normal operating days.* The corresponding propene averages were 2 to 10 times higher on this day compared to the 4-*normal operating days*. A preliminary box model calculation with the enhanced observed Lynchburg Ferry ethene and propene levels indicate substantial CH_2O production, which if transported to the ExxonMobil site at or within a few hours of the time of our overpass, this could explain some or all of the enhanced Sept. 25 CH_2O/CO slope. However, we would expect that such behavior should change the shape of the profiles observed in Fig. 12a to profiles looking more like Fig. 13, where photochemical production of CH_2O is evident. At present, the exact extent of transported-photochemically produced CH_2O relative to direct emissions needs further study, and must await more detailed Lagrangian model runs employing back trajectories, and this will be the subject of a future proposal.

In this section we have provided evidence suggesting that enhanced emissions of CH_2O and/or its alkene precursors emanating from the ExxonMobil Baytown complex in the early morning hours of Sept. 25 may possibly explain our observations. We estimate enhancements by a factor of ~ 3 relative to other sampling days. However, unique meteorology and observations of significantly enhanced ethene and propene levels measured by TCEQ's auto-GCs during the 5 -10 am hours over the nearby Lynchburg Ferry sampling site (from unknown sources) could explain some or all of these enhancements.

3.3. Enhanced CH_2O Levels Observed Over Broad Areas Throughout Sept. 25

Although the above discussions center around potential enhanced emissions of CH_2O and/or its precursors from the ExxonMobil complex on Sept. 25, elevated CH_2O levels were observed in many other locations throughout much of the entire DISCOVER-AQ sampling day on Sept. 25. As we discussed previously in connection with Fig. 5d, the composite of our WP-3 CH_2O measurements observed on all sampling days spanning the time period from Sept. 4 to Sept. 26 over the HGBMA at pressure altitudes ≤ 0.5-km, excluding the 3 circuits on Sept. 25, (comprising 52,589 measurements) yield the following CH_2O values: average = 3.145 ± 1.485 ppbv, median = 3.064 ppbv, 25% value = 2.010 ppbv, 75% value = 4.053 ppbv, 90% value = 5.008 ppbv, and maximum value = 14.437 ppbv. In Fig. 15 we plot time series measurements for

CH_2O and propene measured on the WP-3 aircraft throughout Sept. 25 over each of the 8 spiral sites during each of the 3 circuits. The elevated CH_2O levels up to 26 ppbv can be seen over the ExxonMobil facility during the 1st circuit (shown in the gap between Smith Point (SP) and the Moody Tower (MT)). In addition, significantly elevated CH_2O levels (greater than the 90th percentile of all other sampling days) of 23 ppbv can be seen during the 1st circuit boundary layer measurements over Channelview (CV) and Deer Park (DP), and up to 33 ppbv during the 2nd circuit over Smith Point. Second circuit Channelview-Deer Park measurements up to 12.5 ppbv can be seen as well as many additional elevated measurements in the 5 – 10 ppbv range throughout the boundary layer in the 2nd and 3rd circuits. The majority of these enhancements, which are coincident with elevated propene (> 5 ppbv), arise from CH_2O that is photochemically generated from its precursors. NASA Langley box model calculations by Jason Schroeder and James Crawford at NASA Langley indicate that CH_2O can be photochemically generated from enhanced propene and ethene in time periods as short as 15-minutes. We will be working closely with this group to further study CH_2O and O_3 production using their box model. However, what's unclear and beyond the scope of this present project are the sources of these enhanced highly reactive alkenes. Further, in this report we only focus on CH_2O distributions and their enhancements and not on the related subject of enhanced O_3 formation throughout the greater HGBMA. Although the latter is clearly related to the former, meteorology also plays a large role here and this is the subject area of other studies.

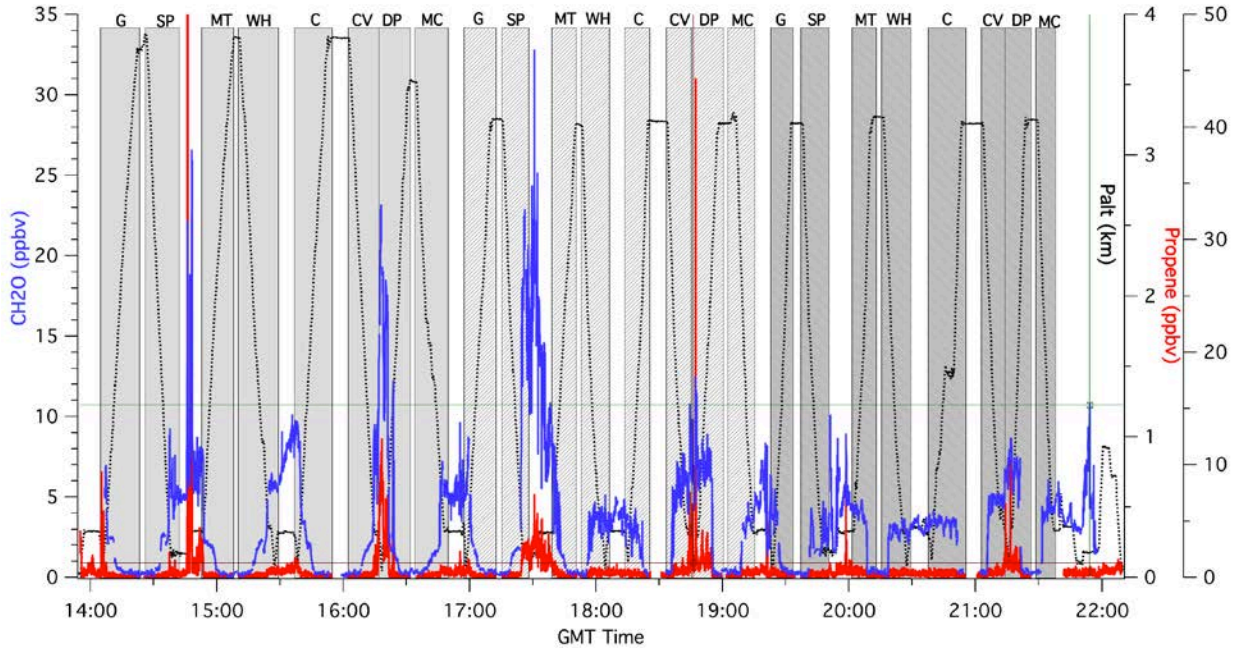

Figure 15: Time series plot of CH_2O (blue) and propene (red, from Armin Wisthaler's PTR-MS measurements) over the 8 spiral sites (G = Galveston, SP = Smith Point, MT = Moody Tower, WH = West Houston, C = Conroe, CV = Channelview, DP = Deer Park, and MC = Manvel Croix) during each of 3 circuits on Sept. 25. The propene scale on the right has been restricted to 50 ppbv to visualize many of the enhanced levels. The large off-scale propene spike at 14:46 is 197 ppbv.

3.4. Evaluation of Detailed CH_2O CMAQ-Measurement Comparisons (Task 1)

As part of our effort in completing **Task 1,** we evaluated CMAQ CH_2O results by comparing the model with measurements for all the DISCOVER-AQ flights. Figures 16 & 17 present

detailed temporal and spatial (CMAQ-Measurement) comparisons for CH_2O for two select days: September 13, 2013 and September 25, 2013. The former exemplifies typical results for many of the days studied, while the latter illustrates results from the extreme event just discussed. Figure 16 shows this comparison for the 2nd WP-3 circuit on September 13. It is important to note that in contrast to model results, which calculate relatively constant 5-minute average CH_2O concentrations in 1km grid boxes, the 1-second WP-3 CH_2O measurements often reflect large changes in airmasses as the WP-3 traverses ~ 0.1km each second. This results in large measurement variances compared to model results. To facilitate measurement-model comparisons, the model results are calculated at each 1-second aircraft time period, even though the model results are temporally and spatially averaged over longer periods. Although the true variance is smoothed out in the model, one can still compare results for select time periods for a given spatial region with common sources at relatively constant altitudes. In Figures 16 & 17, these time periods are denoted by dark horizontal traces, where the blue and red lines respectively represent the average CH_2O measurement and CMAQ model results. These traces are only meant to graphically show the overall biases for the select time periods. The actual (CMAQ-Measurement) biases are determined by point-to-point differences over the select time periods and these values along with their standard deviations are given in the plots. As can be seen, although there are differences, the measurements and model values follow the same overall trends. Fig. 16 shows 4-free troposphere (FT) and 3-planetary boundary layer (PBL) time periods for comparison.

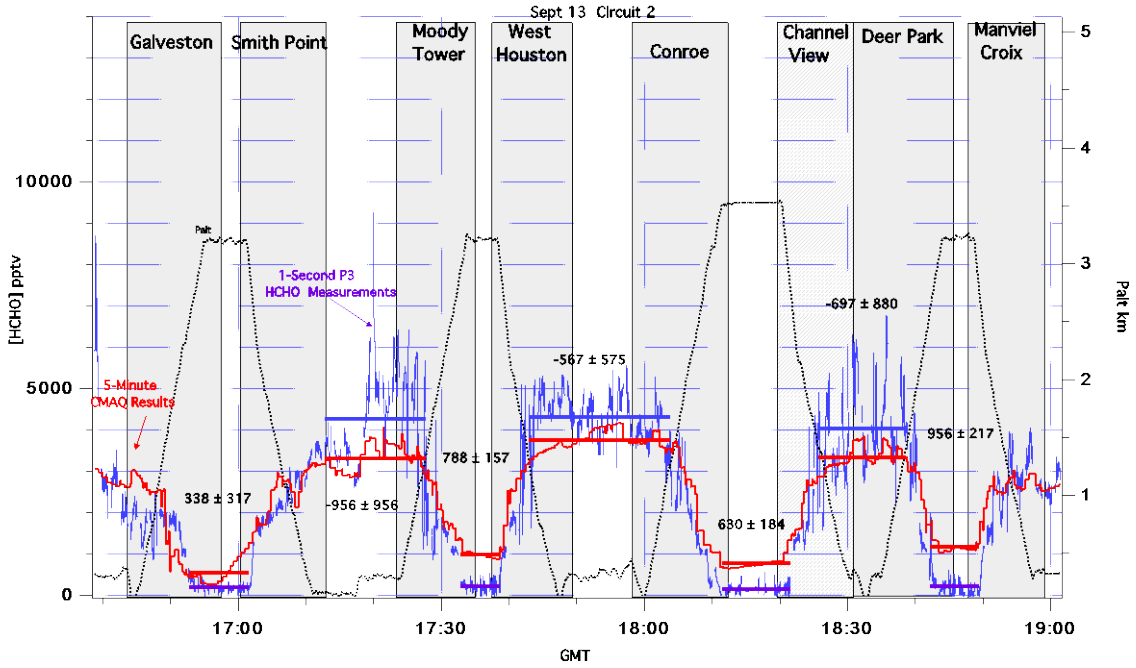

Figure 16: WP-3 CH_2O measurements (1second data, blue trace) and 5-minute CMAQ model results (red trace) for the 8 spiral sites during the 2nd circuit of the September 13, 2013 DISCOVER-AQ flight. The dark horizontal blue and red traces are averaged values over the select PBL and FT time periods, with the resultant point-by-point average (CMAQ-Measurement) differences in units of pptv and standard deviation given for each period.

For the purposes of this study we base our classifications of FT and PBL airmasses by comparing the WP-3 flight pressure altitudes (Palt) with the CMAQ determined boundary layer heights. As can be seen in Fig. 12a there are occasions where the latter are lower than the

indicated aircraft altitudes by as much as ~ 0.1 to 0.3 km yet large plumes are still observed. During the 1st circuit of Sept. 25 during the spiral down over Smith Point, we arrive at an estimate of 0.36-km for the PBL height based upon altitude profiles of various meteorological and trace gases and this compares to a CMAQ-determined boundary layer height of 0.13-km (Meas.-CMAQ = 0.23-km). To account for this discrepancy, we operationally define the PBL in this study whenever the WP-3 pressure altitudes are less than or equal to 0.3-km above the CMAQ-determined boundary layer heights.

Figure 17 shows the corresponding plot for the 1st circuit of Sept. 25, which is the day with the high-observed concentrations of CH_2O discussed previously. As can be seen, the results for the select periods shown in Fig. 17 for this day are significantly different than Sept. 13. Obviously, the discrepancies in CMAQ-PBL heights have some bearing on this discrepancy, as the calculated emissions have not reached the aircraft altitudes. Enhanced emissions of CH_2O and its precursors not reflected in the CMAQ modeling input also play a role here.

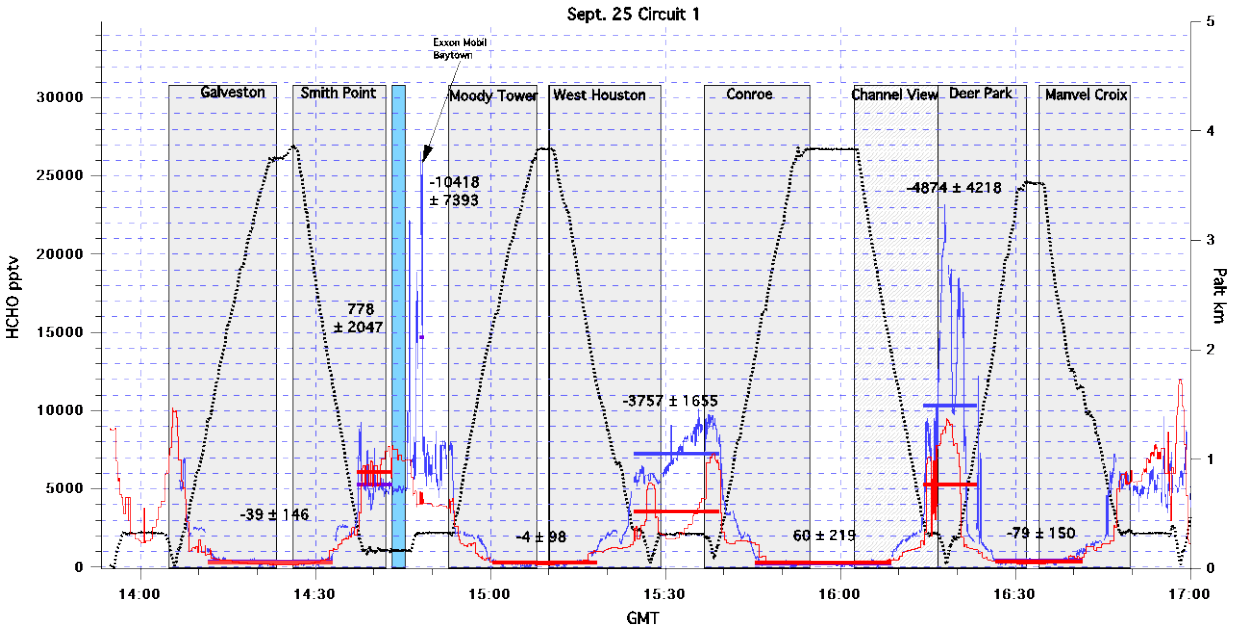

Figure 17: WP-3 measurements and CMAQ model results for Sept. 25, 2013 in the same format as Fig 14.

Figures 16 & 17 are only meant to show the differences between these two sampling days. Rather than focus on these small sample sizes, we now present comparisons for all the PBL legs for each of the 9 DISCOVER-AQ measurement days **(Task 1).** Figure 18 shows such plots in the PBL for the entire month of September in 2013 based upon all the point-by-point 1-second comparisons. The top trace shows the mixing ratios (mean ± 1σ and median) for both the measurements and modeled values, while the bottom trace shows this information for the daily mean and median biases (CMAQ-Measurement) along with the bias percentages (CMAQ-Measurement/Measurement) on the right axis. Like the absolute biases, the bias percentages were calculated from the point-by-point comparisons of 1-second data. Figure 18b shows this same comparison for CO for all the PBL legs.

In the discussion that follows we do not consider Sept. 25 for the reasons discussed previously. The absolute PBL biases for CH_2O and CO for all the remaining days are all relatively small. In the case of CH_2O, the average of all the daily mean PBL biases is -439 ± 392

pptv, and the average of all daily median biases is -319 ± 397 pptv. The average daily median bias percentage is -11.8 ± 15.7%. For CO, the average of all the daily mean PBL biases is -6.0 ± 14.7 ppbv, and the average of all daily median biases is -6.7 ± 14.0 ppbv. The corresponding daily median bias percentage for CO is -4.5 ± 10.7%.

These small but persistently negative biases in Fig. 18 potentially reflect small underestimates in the emission inventories used in the calculations. However, we cannot rule out the possible contribution that CMAQ transports too much boundary layer air into the free troposphere (as suggested by the high FT biases in Fig. 16) and the related error of low biases in CMAQ-determined PBL heights, which as discussed, underestimates the observed concentrations at the WP-3 aircraft altitudes. **Therefore, based on the above results, we have no firm evidence that the 2012 TCEQ emission inventory under normal conditions needs to be revised (Task 1).**

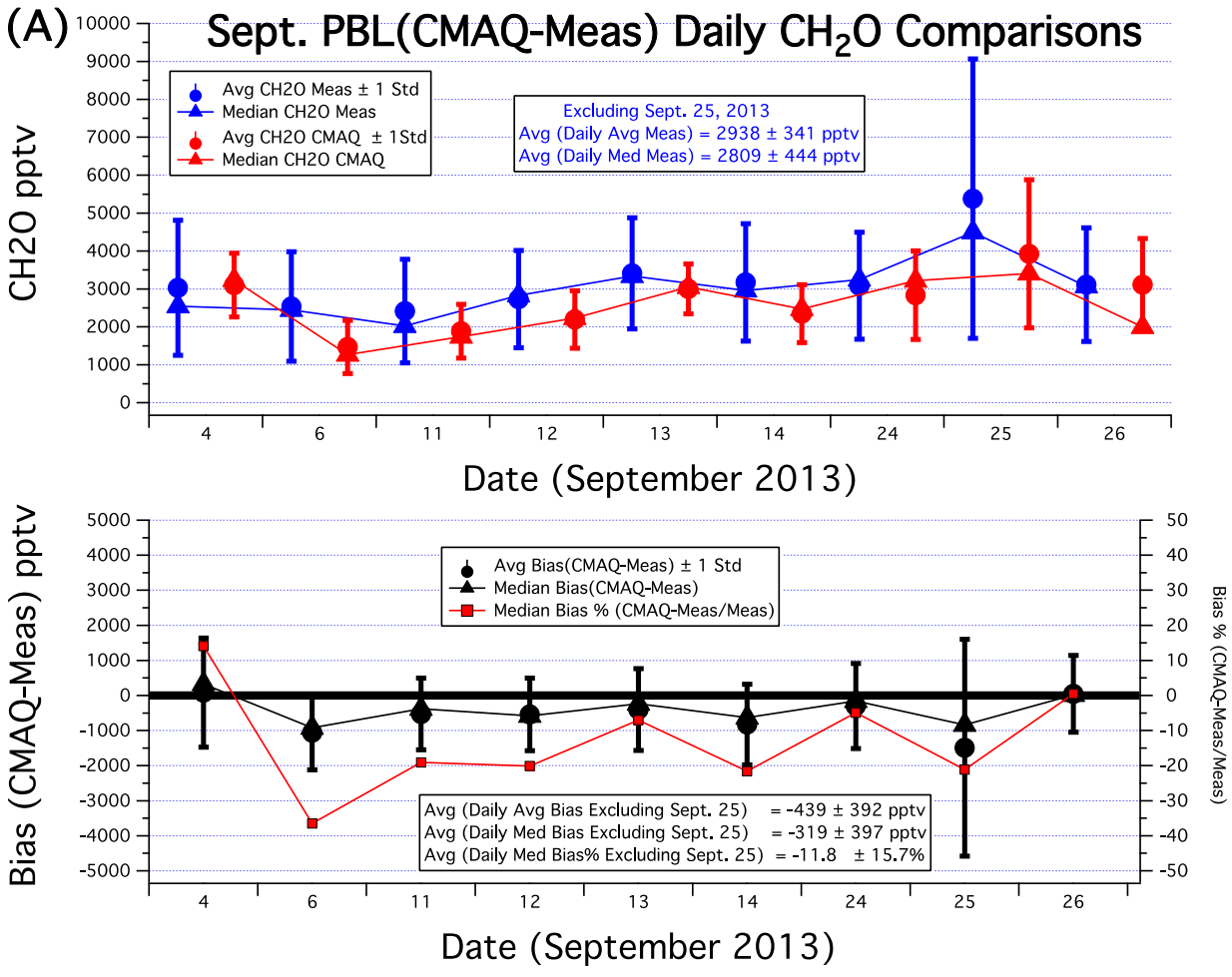

Figure 18: (A) Comparison of daily CH_2O mixing ratios from the WP-3 observations and the 1km CMAQ model results based upon point-by-point comparisons data in the planetary boundary layer (PBL).

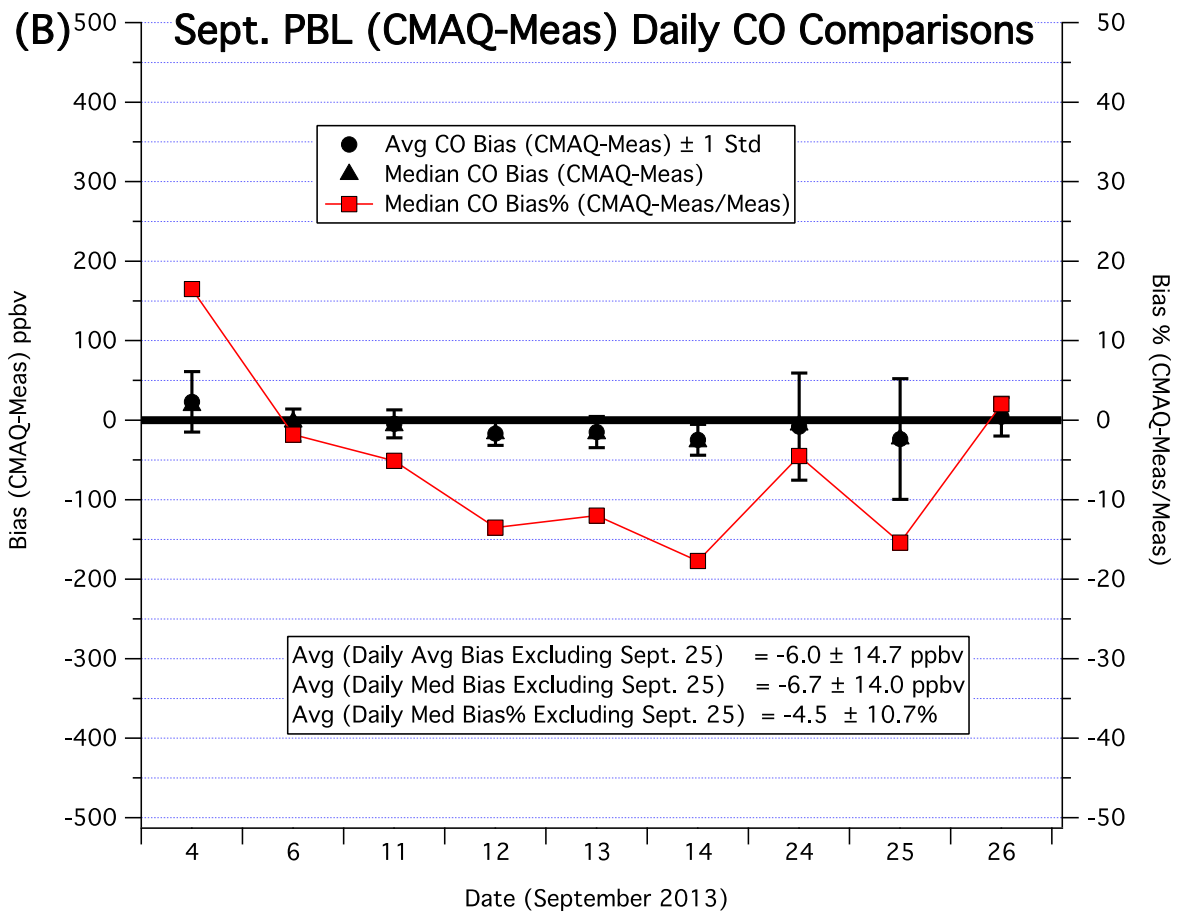

Figure 18: (B) Similar comparisons for CO measurements and model values.

3.5. Assessment of TCEQ DNPH Sampling Results (Tasks 5 & 6)

Having established the general level of agreement between CMAQ-modeled-CH_2O and WP-3 measurements in the PBL, we next address **Task 5:** the comparison of *24-hour synthesized integrated airborne measurements*, based upon the temporal dependence calculated from the CMAQ model and the WP-3 aircraft, with DNPH cartridge sampling measurements from the Clinton, Deer Park and Channelview TCEQ sites. As stated previously, this objective is important since the 24-hour DNPH sampling results have been employed every 6th day over a number of years to collect averaged CH_2O levels at both Deer Park and Clinton, and these data have been used to infer decreasing yearly trends in CH_2O. Unfortunately Sept. 13 at Deer Park was the only day where the DNPH sampling system was operational during a P3 flight day. Fortunately, as shown in Fig. 19, the WP-3 overflights passed very close to the TCEQ sampling site at low altitudes. Figure 19 shows the WP-3 flight track, colored by altitude, near the DNPH sampling site. Although this figure only plots the 2nd WP-3 circuit, each of the 3 circuits passed close to this sampling site over ~ a 6-hour time span.

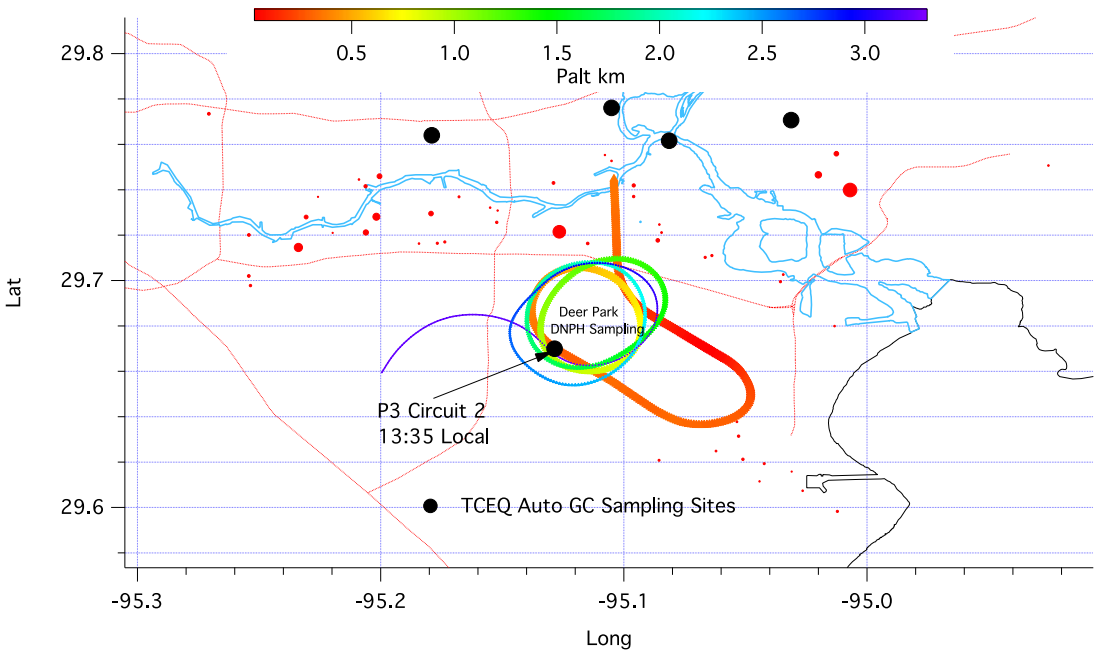

Figure 19: P3 flight track during the 2nd Circuit, colored by altitude over the DNPH TCEQ auto-GC Deer Park sampling site on Sept. 13, 2013.

The traces of Fig. 20 below show the results of comparing WP-3 measurements with ground-based results from the Deer Park TCEQ sampling site. In the top trace, we plot the WP-3 measurements for both CH_2O (blue points) and propene (red points) at the point of closest approach to the Deer Park sampling site for each of 3 circuits at the indicated local sampling times. The CH_2O measurements also include the total uncertainty (systematic plus random), and we indicate the sampling altitude. The 24-hour averaged DNPH measurements acquired by the TCEQ system at Deer Park are shown by the solid blue line spanning the 24-hour time period. The agreement in WP-3 propene measurements from the Wisthaler's group PTRMS instrument with the ground-based TCEQ auto-GC measurements collocated with the DNPH sampling system indicate that the WP-3 and ground sampling site are in the same airmass for all 3 circuits. Without any further information it would be impossible to tell if the significantly elevated propene measured by the auto-GC sampler at around 4 am produces elevated CH_2O. The major photochemical pathway involving the production of CH_2O from its HRVOC precursors involves OH, which is produced only during daylight hours by photolysis of O_3 and reaction of the subsequently formed $O(^1D)$ radicals with water vapor. Since there is no significant OH at night to initiate oxidation of propene to CH_2O, one would expect no corresponding CH_2O increase in the dark unless ozone reacts with propene in the dark to produce CH_2O. This would only occur if the elevated propene does not simultaneously occur with simultaneous large emissions of NO from flaring, which would titrate down the O_3.

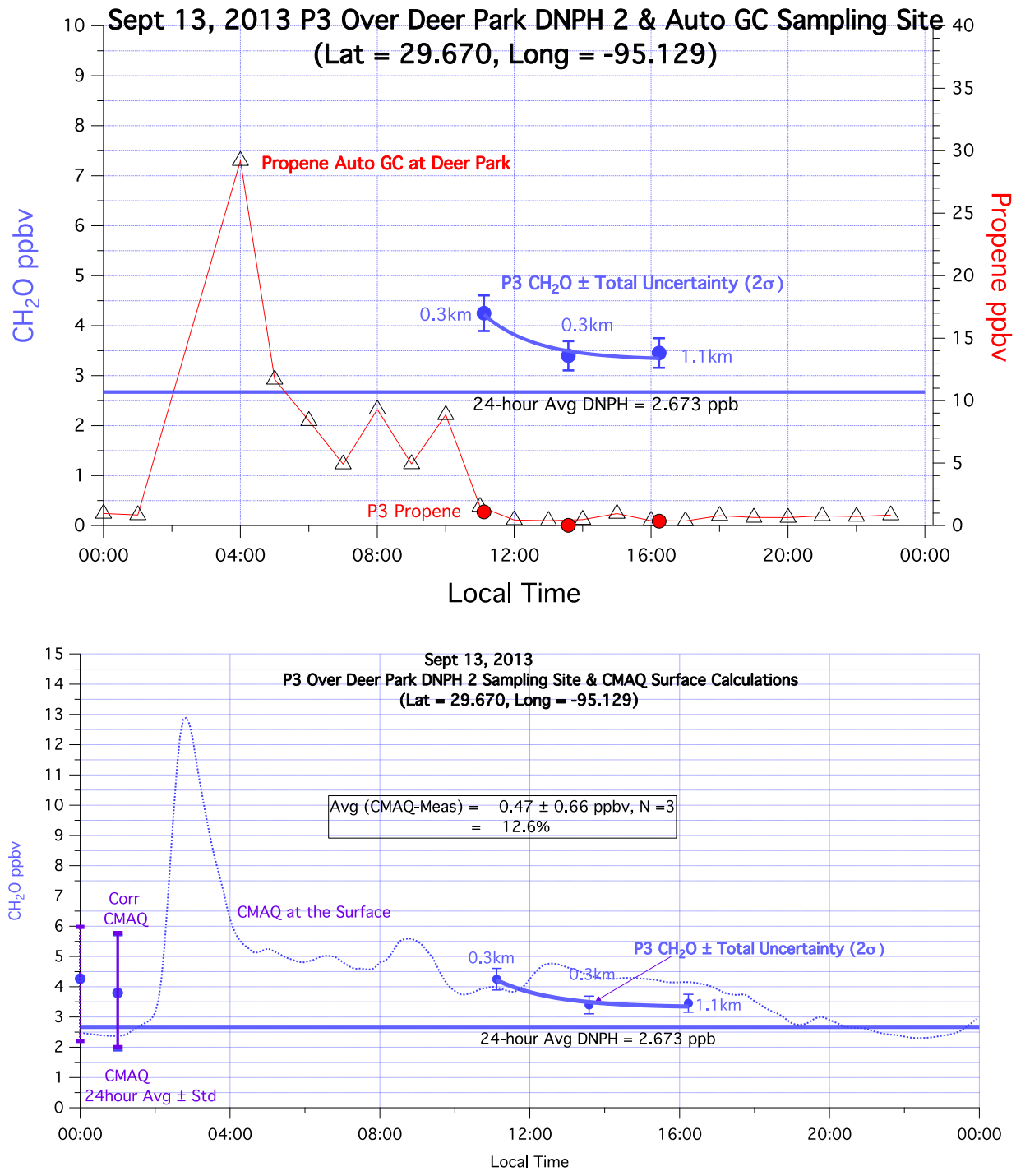

Figure 20: (Top trace) Comparisons of WP-3 propene measurements with the auto-GC measurements at Deer Park. (Bottom trace) Comparison of WP-3 measured CH_2O concentrations with CMAQ calculations at the surface of Deer Park and the 24-hour averaged DNPH results as well as the 24-hour integrated CMAQ results (blue points near time 00:00 with error bars).

It is interesting to note that the 24-hour CMAQ temporal profile of CH_2O at the Deer Park surface site shown in the lower trace of Fig. 20 (dashed blue line) indicates a large increase in calculated CH_2O at around 4 am. In fact, the CMAQ temporal CH_2O profile follows the measured propene profile. Since CMAQ employs an average emission inventory for each hour for a given season, the apparent coincidence in elevated calculated CH_2O and measured propene on. Sept. 13 in the early morning hours implies that early morning propene spikes at the surface at Deer Park should be a daily occurrence. Throughout the month of September in 2013 the hourly Deer Park auto-GC measurements in fact show such propene spikes on most days between the hours of 4 am and 7am, with typical levels in the 10-30 ppbv range and a maximum value of ~ 90 ppbv. This would in turn imply that elevated surface CH_2O at Deer Park might be a regular occurrence from perhaps fugitive emissions and subsequent reactions of O_3 with propene or perhaps ethene in the dark unless O_3 is simultaneously titrated by flares. This latter process, however, would directly release CH_2O. This interesting observation hints at the potential importance of nighttime emissions of CH_2O and/or its precursors, as suggested by Olauger et al. [2009]. Dedicated round the clock ground-based measurements of CH_2O at one or all of the 3 auto-GC sampling sites would be important to carry out in the future to resolve this and eliminate the possibility that such propene enhancements are simply a result of a compressed boundary layer. Addressing this might also help in our efforts to further understand the Sept. 25 results and the connection to our analysis over the ExxonMobil complex.

Aside from the interesting time dependence and associated speculation just discussed, the 24-hour CMAQ surface modeled CH_2O at Deer Park shown in the lower trace of Fig. 20 can be used in conjunction with the WP-3 CH_2O observations to assess 24-hour DNPH results. As can be seen, the CMAQ model results at the surface at Deer Park agrees with the W-P3 measurements to within 468 pptv, which is similar to but opposite in sign with all our previous comparisons. Averaging the CMAQ CH_2O results over the 24-hour DNPH sampling period yields the 4.267 ± 1.9 ppbv value shown at the left on the Y-axis. Applying a small - 468-pptv correction to match the CMAQ results with the 3 WP-3 measurements yields the 3.799 ± 1.9 ppb result shown with the Corr-CMAQ point. This value is in agreement with the averaged DNPH results (2.673 ppbv) within the precision of the CMAQ mean value. The difference is 30%. It is interesting to note that the DNPH value is in line with our averaged daily mean and median CH_2O values of **2.938 ± 0.341 ppbv, and 2.809 ± 0.444 ppbv, respectively, for the composite PBL shown in Fig. 18.** It is also interesting to note that this 30% level of agreement is in line with the comparison slopes reported by Gilpin et al. [1997] between diode laser measurements of CH_2O standards and those retrieved by DNPH cartridge sampling methods. **Based on these limited observations, the Deer Park DNPH sampling system should accurately reflect 24-hour integrated surface CH_2O levels at this site (Task 5).** Clearly more comparisons should be carried if the opportunities arise in the future. In particular, CH_2O measurements with our IR spectrometer located at the Deer Park and Clinton DNPH sites, sampling for at least 1-month each, would provide extremely valuable information. In addition to providing more substantial comparisons with DNPH results, such observations would help to address the nighttime questions just discussed.

3.6. CMAQ Model Analysis Run in Process Analysis Mode to Assess Primary & Secondary Sources of CH_2O Throughout the HGBMA Study Area (Task 4)

The next step in our analysis is **Task 4:** Employ the CMAQ model output runs in Process Analysis Mode to quantify the relative importance of direct emissions and secondary

photochemical production of CH_2O. This was initially carried out with the existing 2012 emission inventories. Figure 3 shows the 4 and 1-km domains. Only the 1-km domain calculations are used in our source attribution assessments. As can be seen, this domain focuses on the Houston-Galveston-Brazoria Metropolitan Area. As can be seen in Figure 3, the 1 km modeling domain covers the Houston-Galveston-Brazoria Metropolitan Area.

CMAQ run with Process Analysis can be used to quantify the contributions of individual physical and chemical processes to changes in model pollutant concentrations [Byun and Ching, 1999]. It can be used to determine the relative importance of processes (i.e., emissions vs. chemical production) impacting changes in pollutant concentrations. In relation to regulatory decisions, Process Analysis can be used to help determine what type of emissions should be controlled in order to most effectively improve air quality.

CMAQ run with Process Analysis for formaldehyde yielded information on how formaldehyde concentrations changed for each grid cell due to emissions, chemistry, transport (horizontal and vertical advection and diffusion), dry deposition, and cloud processes. We examined the relative importance of emissions and chemistry on changes in formaldehyde concentrations within the 1 km domain. Figure 21 shows the results of our daily average calculations for the entire month of September in 2013 averaged over the entire 1 km domain.

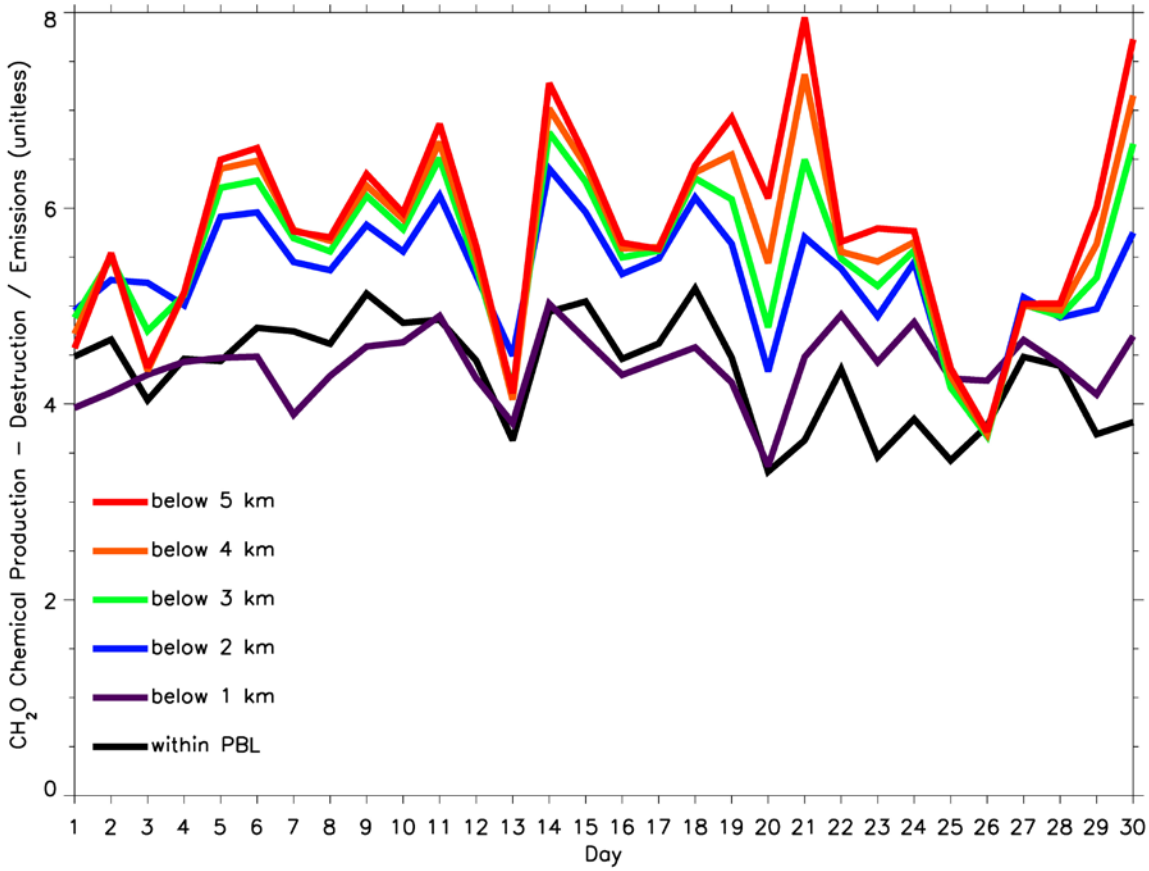

Figure 21: 1-km CMAQ model ratios of CH_2O from secondary production sources (production-destruction) relative to direct emission sources for the entire month of September in 2013 over the Houston-Galveston-Brazoria Metropolitan Area.

As can be seen, CH_2O from secondary production sources (Production – Destruction) is approximately a factor of 5 times higher than direct emission sources in the planetary boundary layer (PBL) over the entire month of September and approximately a factor of 7 to 8 times the direct emission source for the atmosphere over the Houston-Galveston-Brazoria Metropolitan Area up to 5-km altitude. These results are in agreement with our qualitative assessment presented in one of our reports based upon our fast CH_2O-O_3-NO_x/NO_y correlations.

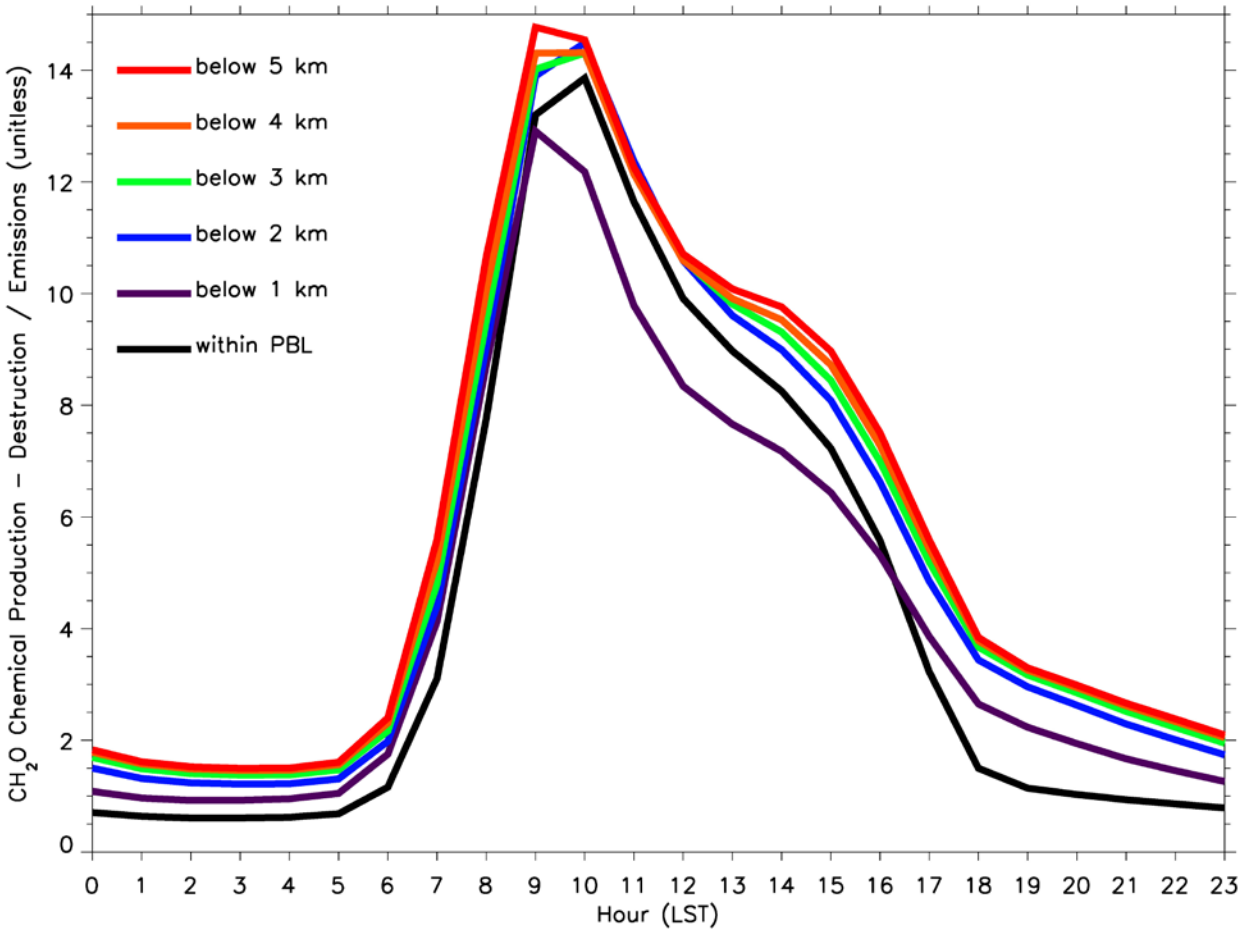

Figure 22: CMAQ model results from Fig. 19 broken out by hour of day for the entire month of September 2013.

Figure 22 further shows this breakdown as a function of hour for the entire month of September 2013. Over the 7 am – 7 pm daylight hours, the average ratio yields a value of ~ 8/1 within the PBL. **This yields a secondary CH_2O contribution of ~ 89% over the daylight hours and this agrees well with the determination from Parrish et al. [2012] of ~ 92 ± 4% based upon OH reactions of ethene and propene to produce CH_2O during daylight hours.** ***These results and the findings in Parrish et al. [2012] imply that emissions control efforts should not focus on primary CH_2O emissions, but rather HRVOC emissions that are precursors to CH_2O. This is a major finding of the present study.*** It should be mentioned that our results are based upon present emission inventories for ethene and propene and span the entire HGBMA study area throughout the entire month of September in 2013 and are not restricted to times and spatial domains where measurements have been acquired. We believe

these September results should reasonably represent the results for the full year. However, additional modeling studies need to be run in future studies to definitely confirm this.

4.0. PROJECT SUMMARY

The following is a summary of the primary conclusions from the analysis in this study.

(1) Analysis of airborne CH_2O measurements over the greater HGBMA study area during the 2013 DISCOVER-AQ (9 sampling days over Houston) and SEAC^{4}RS (1 sampling day over Houston) campaigns over the month of September revealed that only the September 25 sampling day showed exceptional high PBL CH_2O levels in excess of 30 ppbv, levels characteristic of our past measurements over the greater HGBMA study area in 2006 and 2000. All other sampling days in 2013 showed significantly lower PBL CH_2O levels in the 2 – 10 ppbv range. Enhanced CH_2O levels were observed throughout the day and over nearly the entire HGBMA study area. Aside from the early morning measurements over the ExxonMobil complex, the majority of these enhancements were found to be coincident with elevated propene (> 5 ppbv) and arise from CH_2O that is photochemically generated from its precursors.

(2) We developed an observational approach based upon fast aircraft measurements of correlations between CH_2O, O_3, CO, NO_x/NO_y ratios, and propene as a means of identifying time periods revealing enhanced sources of CH_2O. We have identified a number of such plumes, which based upon strong anti-correlations of O_3 with CO and high NO_x/NO_y ratios, indicated very fresh plumes concurrent with combustion sources. Most of these plumes were found in the vicinity of petrochemical facilities. At present, we do not have enough information to discern if such enhanced CH_2O: 1) originates directly from the combustion sources; 2) is produced during combustion chemically from its two major precursors propene and ethene; 3) occurs simultaneously from fugitive emissions of CH_2O, propene, and ethene; or 4) some combination of the above. Likewise, we do not have enough information to even speculate on the types of petrochemical combustion sources (e.g., flaring, fluidized catalytic cracking combustion, or other potential petrochemical combustion sources) that might be responsible for our observations, and therefore efforts to correlate which petrochemical stack that might be responsible for our observations is beyond the scope of this effort.

(3) In our plume tabulations, the largest source of enhanced CH_2O associated with petrochemical combustion occurred during the 1st circuit on 9/25/13 right over the Baytown ExxonMobil complex around 9:48 am local time. Regression analysis of fast CH_2O and CO measurements on this day were compared to 4 other days (9/6/13, 9/12/13, 9/13/13, and 9/24/13) where we acquired CH_2O/CO slopes over this same petrochemical complex at around the same local time. These 4 days produced CH_2O/CO slopes within a factor of ~ 2 of that determined from the 2013 Speciated Release Inventory for CH_2O and CO under normal operating conditions for all three ExxonMobil facilities combined. Since our measurements reflect the sum of CH_2O released as well as CH_2O produced from propene and ethene released from these same facilities under normal operating conditions, we view this factor of 2.0 to 2.5 difference as a reasonable range of values for normal operating conditions. However, the factor of ~ 7 times higher measured slope on 9/25/13 relative to the normal operating Speciated Release Inventory is considerably

higher, and in our opinion, suggests enhanced emissions of CH_2O and potentially its precursors on Sept. 25 emanating from the ExxonMobil complex during the morning hours, perhaps by as much as a factor of ~ 3 relative to the other sampling days. We also presented counter-arguments suggesting that some or all of these enhancements may be caused by unique meteorology on this day (strong early morning inversion with a tightly capped boundary layer ~ 0.3-km) coupled with significantly enhanced ethene and propene emissions measured on this day by TCEQ's auto-GCs during the 5 -10 am hours over the nearby Lynchburg Ferry sampling site (from unknown sources). A more definitive assessment must await additional studies based upon Lagrangian model runs employing back trajectories, and this has been identified as one of the subject areas for a future proposal.

(4) Likewise, efforts in providing individual enhanced emission estimates for various species in moles/hour emanating from the ExxonMobil complex on Sept. 25 that could then used to compare calculated and measured CH_2O concentrations downwind at Smith Point (Task 3) turned out to be far more complicated than originally anticipated. Additional input will be required to carry this out more rigorously than our initial attempts.

(5) Modeling analysis employing the Community Multiscale Air Quality (CMAQ) model with Process Analysis, in very high-resolution mode (1 km resolution), driven by the WRF (Weather Research and Forecasting) meteorological model has been developed, improved upon, and evaluated. This evaluation involved comparisons of various measured meteorological and trace chemical species concentrations (CH_2O, isoprene, CO, NO, NO_2, and O_3) with those simulated from CMAQ.

(6) Extensive CMAQ-Measurement comparisons in the PBL and FT showed reasonable daily agreement. Not considering Sept. 25, which in addition to possible enhanced CH_2O emissions had potential complications caused by unique meteorology, the absolute PBL biases (CMAQ-Meas.) for CH_2O and CO for all the remaining days are all relatively small. In the case of CH_2O, the average of all the daily mean PBL biases is -439 ± 392 pptv, and the average of all daily median biases is -319 ± 397 pptv. The average daily median bias percentage is -11.8 ± 15.7%. For CO, the average of all the daily mean PBL biases is -6.0 ± 14.7 ppbv, and the average of all daily median biases is -6.7 ± 14.0 ppbv. The corresponding daily median bias percentage for CO is -4.5 ± 10.7%.

(7) These small but persistently negative biases potentially reflect small underestimates in the emission inventories used in the calculations. However, we cannot rule out the possible contribution that CMAQ transports too much boundary layer air into the free troposphere, as has been observed on other occasions. Therefore, based on the above results, we have no firm evidence that the 2012 TCEQ emission inventory under normal conditions needs to be revised.

(8) We assessed the accuracy of 24-hour integrated DNPH cartridge sampling measurements for CH_2O on one occasion at the Deer Park site on Sept. 13. This was carried out by comparing *24-hour synthesized integrated airborne measurements of CH_2O,* based upon the temporal dependence calculated from the CMAQ model and the WP-3 aircraft, with the DNPH cartridge sampling measurements at Deer Park. After applying a small correction to the CMAQ results to match the observations, we determined a 24-hour integrated CH_2O value of 3.799 ± 1.9 ppb on Sept. 13 at the Deer Park sampling site, a value that is in agreement with the integrated DNPH determination of 2.673 ppbv within

the precision of the CMAQ value. This 30% difference is in line with the comparison slopes reported by Gilpin et al. [1997] between diode laser measurements of CH_2O standards and those retrieved by DNPH cartridge sampling methods.

(9) In the process of investigating (8) above, the CMAQ modeling results in conjunction with ground-based auto-GC measurements of propene at the Deer Park sampling site point to possible evidence of nighttime emissions of CH_2O and/or its precursors, as has been suggested by Olauger et al. [2009].

(10) The CMAQ model was run in Process Analysis Mode to assess primary and secondary sources of CH_2O throughout the greater HGBMA study area throughout the month of September 2013. CH_2O from secondary production sources (Production – Destruction) is approximately a factor of 5 times higher than direct emission sources in the planetary boundary layer (PBL) over the entire month of September and approximately a factor of 7 to 8 times the direct emission source for the atmosphere over the Houston-Galveston-Brazoria Metropolitan Area up to 5-km altitude. These results were further broken down as a function of hour for the entire month of September 2013. Over the 7 am – 7 pm daylight hours, the average ratio yields a value of ~ 8/1 within the PBL. This yields a secondary CH_2O contribution of ~ 89% over the daylight hours and this agrees well with the determination from Parrish et al. [2012] of ~ 95% based upon OH reactions of ethene and propene to produce CH_2O during daylight hours. It should be mentioned that our results are based upon present emission inventories for ethene and propene and span the entire HGBMA study area throughout the entire month of September in 2013 and are not restricted to times and spatial domains where measurements have been acquired. We believe these September results should reasonably represent the results for the full year. However, additional modeling studies need to be run in future studies to definitely confirm this.

5. REFERENCES

Appel, K.W., R.C. Gilliam, J.E. Pleim, G.A. Pouliot, D.C. Wong, C. Hogrefe, S.J. Roselle, and R. Mathur: *Improvements to the WRF-CMAQ modeling system for fine-scale air quality simulations,* EM, September 2014, 16-21, 2014.

Byun, D., and J. K. S. Ching, 1999: *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*. U. S. Environmental Protection Agency Rep. EPA-600/R-99/030, 727 pp. [Available from Office of Research and Development, EPA, Washington, DC 20460.]

Emery, C., E. Tai, and G. Yarwood: *Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes, prepared for the Texas Natural Resource Conservation Commission*, prepared by ENVIRON International Corporation, Novato, CA, 2001.

Fried, A., Y. Wang, C. Cantrell, B. Wert, J. Walega, B. Ridley, E. Atlas, R. Shetter, B. Lefer, M.T. Coffey, J. Hannigan, D. Blake, N. Blake, S. Meinardi, B. Talbot, J. Dibb, E. Scheuer, O. Wingenter, J. Snow, B. Heikes, and D. Ehhalt, *Tunable Diode Laser Measurements of Formaldehyde During the TOPSE 2000 Study: Distributions, Trends, and Model Comparisons*, J. Geophys. Res., **108 (D4),** doi: 10.1029/2002JD002208, 2003a.

Fried, A., J. Crawford, J. Olson, J. Walega, W. Potter, B. Wert, C. Jordan, B. Anderson, R. Shetter, B. Lefer, D. Blake, N. Blake, S. Meinardi, B. Heikes, D. O' Sullivan, J. Snow, H. Fuelberg, C. Kiley, S. Sandholm, D. Tan, G. Sachse, H. Singh, I. Faloona, C. N. Harward, and G.R. Carmichael, *Airborne Tunable Diode Laser Measurements of Formaldehyde During TRACE-P: Distributions and Box-Model Comparisons*, J. Geophys. Res., **108 (D20),** 8798 doi: 10.1029/2003JD003451, 2003b.

Fried, A., J. Walega, J.R. Olson, J.H. Crawford, G. Chen, P. Weibring, D. Richter, C. Roller, F. Tittel, B. Heikes, J. Snow, H. Shen, D. O'Sullivan, M. Porter, H. Fuelberg, J. Halland, and D. Millet, 2007: Formaldehyde over North America and the North Atlantic during the Summer 2004 INTEX Campaign: Methods, Observed Distributions, and Measurement-Model Comparisons, *J. Geophys. Res.* **113 (D10302),** doi: 10.1029/2007JD009185, 2008a.

Fried, A., J.R. Olson, J.G. Walega, J.H. Crawford, G. Chen, P. Weibring, D. Richter, C. Roller, F.K. Tittel, M. Porter, H. Fuelberg, J. Halland, T.H. Bertram, R.C. Cohen, K. Pickering, B.G. Heikes, J.A. Snow, H. Shen, D.W. O'Sullivan, W.H. Brune, X. Ren, D.R. Blake, N. Blake, G. Sachse, G.S. Diskin, J. Podolske, S.A. Vay, R.E. Shetter, S.R. Hall, B.E. Anderson, L. Thornhill, A.D. Clarke, C.S. McNaughton, H.B. Singh, M.A. Avery, G. Huey, S. Kim, and D.B. Millet, Role of convection in redistributing formaldehyde to the upper troposphere over North America and the North Atlantic during the summer 2004 INTEX Campaign, *J. Geophys. Res.* **113 (D17306),** doi: 10.1029/2007JD009760, 2008b.

Fried, A., C. Cantrell, J. Olson, J.H. Crawford, P. Weibring, J. Walega, D. Richter, W. Junkermann, R. Volkamer, R. Sinreich, B.G. Heikes, D. O'Sullivan, D.R. Blake, N. Blake, E. Apel, A. Weinheimer, D. Knapp, A. Perring, R.C. Cohen, H. Fuelberg, R.E. Shetter, S.R. Hall, K. Ullmann, W.H. Brune, J. Mao, X. Ren, L.G. Huey, H.B. Singh, J.W. Hair, D. Riemer, G. Diskin, and G. Sachse, Detailed comparisons of airborne formaldehyde measurements with box models during the 2006 INTEX-B and MILAGRO campaigns: potential evidence for significant impacts of unmeasured and multi-generation volatile organic carbon compounds, *Atmos. Chem. Phys.,* **11,** 11,867-11,894, 2011.

Gilpin, T., E. Apel, A. Fried, B. Wert, J. Calvert, Z. Genfa, P. Dasgupta, J.W. Harder, B. Heikes, B. Hopkins, H. Westberg, T. Kleindienst, Y.N. Lee, X. Zhou, W. Lonneman, and S. Sewell: Intercomparison of six ambient [CH_2O] measurement techniques, J. Geophys. Res. **102,** 21161 – 21188, 1997.

Herndon, S.C., D.D. Nelson, E.C. Wood, W.B. Knighton, C.E. Kolb, Z. Kodesh, V.M. Torres, and D.T. Allen, A, *Application of the Carbon Balance Method to Flare Emissions Characteristics,* Ind. Eng. Chem. Res., **51,** 12577 – 12585, 2012.

Herrington, J.S. and M.D. Hays: Concerns regarding 24-h sampling for formaldehyde, acetaldehyde, and acrolein using 2,4-dinitrophenylhydrazine (DNPH)-coated solid sorbents, Atmos. Environ., **55,** 179 – 184, 2012.

International Agency for Research on Cancer (IARC), Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Lyon, France, 2006.

International Agency for Research on Cancer (IARC), Formaldehyde. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Lyon, France, 2012.

Johansson, J., J. Mellqvist, J. Samuelsson, B. Offerle, J. Moldanova, B. Rappenglück, B. Lefer, and J. Flynn: Quantitative measurements and modeling of industrial formaldehyde emissions in the greater Houston area during campaigns in 2009 and 2011, J. Geophys. Res., **119,** 4303 – 4322**,** doi:10.1002/2013JD020159**,** 2014.

Müller, M., B. Anderson, A. Beyersdorf, J.H. Crawford, G. Diskin, P. Eichler, A. Fried, J. W. Walega, F.N. Keutsch, T. Mikoviny, A.J. Weinheimer, M. Yang, R. Yokelson, and A. Wisthaler, In situ measurements and modeling of reactive trace gases in a small biomass burning plume, in review Atmos. Chem. Phys. Discussions, 2015.

Olaguer, E. P., B. Rappenglück, B. Lefer, J. Stutz, J. Dibb, R. Griffin, W.H. Brune, M. Shauck, M. Buhr, H. Jeffries, W. Vizuete, and J.P. Pinto: Deciphering the role of radical precursors during the second Texas air quality study, J. Air Waste Manage. Assoc., **59**, 1258–1277, doi: 10.3155/1047- 3289.59.11.1258, 2009

Texas Commission on Environmental Quality (TCEQ), *Formaldehyde. Development Support Document*, August 7, Austin, TX, 2008.

Parrish, D.D., T.B. Ryerson, J. Mellqvist, J. Johansson, A. Fried, D. Richter, J.G. Walega, R.A. Washenfelder, J.A. de Gouw, J. Peischl, K.C. Aikin, S.A. McKeen, G.J. Frost, F.C. Fehsenfeld, and S.C. Herndon: Primary and secondary sources of formaldehyde in urban atmospheres: Houston Texas region, Atmos. Chem. Phys., 12, 3273 -3288, 2012.

Washenfelder, R. A., M. Trainer, G. J. Frost, T.B. Ryerson, E.L. Atlas, J.A. de Gouw, F.M. Flocke, A. Fried, J.S. Holloway, D.D. Parrish, J. Peischl, D. Richter, S.M. Schauffler, J.G. Walega, C. Warneke, P. Weibring, and W. Zheng: Characterization of NO_X, SO_2, ethene, and propene from industrial emission sources in Houston, Texas, J. Geophys. Res., **115**, D16311, doi: 10.1029/2009JD013645, 2010

Wert, B.P., Fried, A., and Henry, B.: Evaluation of inlets used for the airborne measurement of formaldehyde, J. Geophys. Res., **107,** doi: 10.1029/2001JD001072, 2002.

Wert, B. P., M. Trainer, A. Fried, T.B. Ryerson, B. Henry, W. Potter, W.M. Angevine, E. Atlas, S.G. Donnelly, F.C. Fehsenfeld, G.J. Frost, P.D. Goldan, A. Hansel, J.S. Holloway, G. Hubler, W.C. Kuster, D.K. Nicks Jr., J.A. Neuman, D.D. Parrish, S. Schauffler, J. Stutz, D.T. Sueper, C. Wiedinmyer, and A. Wisthaler, *Signatures of terminal alkene oxidation in airborne formaldehyde measurements during TexAQS 2000*, J. Geophys. Res., **108**, 4104, doi: 10.1029/2002JD002502, 2003

6. AUDITS OF DATA QUALITY

6.1 Quality Assurance Checks on the Airborne CH_2O Measurements

Our extensive measurements of CH_2O (both ground-based and airborne) have a long heritage where we have developed and implemented approaches and tests to ensure measurements with high and verifiable accuracy. The numerous references by Fried and his collaborators in the reference section document the extensive calibrations, crosschecks, interference testing, zeroing, inlet testing, and intercomparisons with other techniques that have been carried out by this group in this regard. The technique intercomparison studies by Gilpin et al. [1997] and the extensive aircraft inlet testing of Wert et al. [2002] are but two of the many papers that highlight the extreme care attention to detail in our CH_2O measurements.

During the 2013 DISCOVER-AQ and SEAC^{4}RS campaigns we carried numerous QA/QC checks, prior to the missions, during the missions, and post mission. A comprehensive discussion of these procedures can be found online in our DISCOVER-AQ data submission. However, for completeness, we briefly outline the salient features of these procedures here. Calibrations for both campaigns relied on aircraft permeation calibration systems, whose absolute calibrations were determined from multiple approaches, including direct absorption employing the Beer-Lambert Law relationship. Laboratory tests have been carried using multiple calibration standards to ensure linearity. In the field we generally employed a 3-point calibration during flight. Calibration standards were added multiple times each flight on top of both ambient air (true standard addition) and on top of zero air. Addition of calibration standards as well as well as background zero air (air scrubbed of CH_2O) was added to nearly the entire inlet to check for inlet and cell line losses and outgassing in flight. Zero air was generated in flight employing two scrubbers in series (a heated Pd/Al_2O_3 catalytic converter followed by a room temperature SnO_2 scrubber). Extensive tests have been carried out over the years to ensure that this scrubber pair completely removes ambient CH_2O. Background zero air additions were carried out every 1 to 2 minutes in flight by overflowing the inlet with zero air. This frequent zeroing procedure very effectively captures and removes optical spectrometer noise as well as residual outgassing from inlet line and cell contaminants. Eliminating this latter effect is very important for high measurement accuracy, particularly when sampling low CH_2O concentrations in the upper atmosphere that were proceeded by very high boundary layer levels. The paper by Wert et al. [2002] highlights the importance of these procedures.

6.2 Data Quality Audits

More than 10% of the WRF and CMAQ model input and output files, scripts, and analysis products were reviewed for quality assurance purposes by Loughner. Model inputs and outputs, model evaluation statistics, and graphics generated for this project are being stored and will continue to be for at least three years after the completion of the project at NASA GSFC or the University of Colorado. In addition, all model inputs, outputs, and post-processing analyses will be sent to the University of Texas after the completion of the project.

7. RECOMMENDATIONS OF FUTURE WORK

In the process of addressing the various tasks in this study we have identified a number of topics for future studies, and these include:

1. The one comparison of our highly accurate infrared (IR) spectrometer CH_2O measurements with DNPH sampling at the Deer Park site was highly successful. However, as we point out it would be extremely important to more directly validate these very useful multi-year DNPH time trends employing our measurement systems co-located at multiple DNPH sampling sites for some extended period of time. The University of Colorado team is presently working on the development of smaller, lighter, and autonomous instruments for CH_2O that either could be readily transported between sites or provide multiple simultaneous measurements at various sites employing multiple instruments. In addition, because of the extreme flexibility of the IR measurements, one could develop instruments for fast in situ detection of many other hydrocarbon gases of importance like: propene, ethene, methanol, methane, and propane. We have already deployed instruments to simultaneously measure CH_2O and ethane, and we are working on a fast ethane/methane ratio spectrometer.
2. Our fast CH_2O measurements offer extreme flexibility in that these measurements can be integrated over 24-hours for direct DNPH comparisons. Since the 1-second data are still retained, one can further investigate potential suspected nighttime emissions of CH_2O and/or its precursors in conjunction with auto-GC measurements.
3. As we discussed, the CMAQ model results run in the Process Analysis mode revealed the predominance of secondary sources of CH_2O compared to primary emission sources throughout the HGBMA study area for the entire month of September in 2013. However, as we indicated it would be highly desirable to confirm this result for the entire year, and perhaps over multiple years.
4. **It would be highly desirable to further confirm the validity of our CH_2O/CO regression analysis by additional ground-based measurements at one or more of the auto-GC measurement sites. Our analysis to identify the presence of petrochemical combustion plumes, potentially from flaring operations, has generated a lot of interest because of past studies on this topic. It would be highly desirable to continue and further verify the methods developed in the present study with additional observations.**
5. In addition, Lagrangian model runs employing back trajectories on data acquired during the 1st circuit of Sept. 25, 2013 over the ExxonMobil complex should be carried out to address the issues raised regarding the relative importance of enhanced emissions from this facility compared to photochemically-generated CH_2O from the transport of highly reactive VOCs.
6. **The potential enhanced ExxonMobil facility emissions coupled with favorable meteorology on September 25 provides an excellent opportunity to further study the production of CH_2O downwind at Smith Point. Unfortunately, this task turned out to be far more complicated than originally anticipated in the present study. It would be highly desirable to continue this effort by employing a Lagrangian model that tracks the pollution plume from both the ExxonMobil complex and the Shell Deer Park facility back to Smith Point employing measured inputs at the time of the petrochemical facility overpasses.**