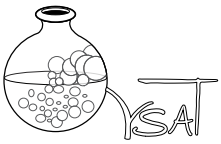
Visualization and Validation of PAMS AutoGC Data

A Presentation by Carol Meyer and
Orsat, LLC.



Orsat, LLC
1416 Southmore Ave
Pasadena, TX 77502
www.orsat.com
877-477-0171

Product specifications and descriptions in this document subject to change without notice.

Published 4/18/2019
5385-EN

INTRODUCTION

PAMS AutoGC systems collect hourly speciated VOC data which is used for developing air quality models as well as determining compliance with National Ambient Air Quality Standards (NAAQS). This data is significant because it is used over decades to compare air quality across regions as well as time. It is important that the data be collected and validated with care. AutoGC systems collect 59 VOC target species hourly and thus will generate a large amount of data. These systems have been in use for over 20 years and have proven to be able to generate consistent and reproducible data. By using frequent Quality Control checks and a good system for the generation of control charts, operators and validators may review data as it is generated and thus keep the system performance optimized to ensure the generation of easily validated data.

GUIDELINES FOR QUALITY ASSURANCE

The EPA has developed a Technical Assistance Document which addresses the recommended Quality Control (QC) checks which aimed at assuring that the AutoGC system is operating correctly[1]. Table 1 shows the scope of the data collected over the three month ozone season along with the

Originally presented by Carol Meyer on 3 April, 2019, as a technical presentation at the Air & Waste Management Air Quality Measurement Methods and Technology Conference at the Sheraton Imperial Hotel, Durham, NC, 2–4 April, 2019.

recommended QC checks. Some agencies have opted to collect this data year-round and this can only be accomplished with alternative methods for reviewing large data sets.

			Ozone Season June - August	Continuous 24/7
Sample Data		Sample Days	92	365
		Sample Hours	2,208	8760
Recommended Quality Control Checks		System Blank (SB)	92	365
		Continuous Calibration Verification (CCV)*	105	417
		Second Source Quality Control Standard (SSQC)	13	52
		Total QC Samples	210	834
Data Points Collected	Required	25	49,950	198,150
	All	59	117,882	467,634

**Includes duplicate samples for the evaluation of precision*

Table 1. AutoGC Sampling Scope

It is not practical for operators or validators to review every chromatogram so some systematic review of this large amount of data must be achieved to validate results. Some problems which occur can be corrected by simply reprocessing some data. Changes in flows, temperatures, humidity may cause analytes to shift slightly and become misidentified. Equipment failures or power interruptions occur and data must be reviewed to determine accuracy and ultimately some will be lost. While fully automated systems can collect a large amount of quality data, they can also collect a large amount of poor quality data as well. Identifying problems, implementing corrective actions and determining what is valid or invalid is still a challenge.

The EPA Technical Assistance Document for Sampling and Analysis of Ozone Precursors for the PAMS Program (currently in draft form) outlines the recommended QC checks which should be used to verify that the instrumentation is performing consistently and reproducibly[1]. Table 2 shows the recommended QC samples which should be run. This document also outlines in detail the scope of validation required for data which will be submitted to the EPA Air Quality System (AQS) database. Validation includes not only a reconciliation of QC sample failures and instrument performance but also review of ambient data for consistency with existing known relationships as well as historical data. Without a good method of visualizing and statistically evaluating ambient data gathered it would be difficult to adequately review the large amounts of data generated by the AutoGC systems. Although systems for this type of data review do exist for criteria pollutants, few if any can be configured to embrace the scope necessary to evaluate hourly ambient data for 59 targets.

QC Parameter	Frequency	Description	Acceptance Criteria
Initial Calibration (ICAL)	Prior to sampling	Average RF of at least 3 levels from 1–25 ppbC	$R^2 \geq 0.99$ x-intercept ±0.5 ppbC each level ± 30% RSD ≤ 10%
Continuing Calibration Verification (CCV)	Daily	Humidified standard within range of ICAL	All Target VOCs within ± 30%
System Blank (SB)	Daily	Humidified zero air	Less than MDL or 0.5 ppbC
Second Source QC Standard (SSQC)	Weekly	Humidified second source standard within range of ICAL	All Target VOCs within ± 30%
Retention Time Standard (RTS)	Weekly	Humidified sample containing all targets ~2–60 ppbC or less	All Target VOCs within retention time windows identified correctly
Precision Check	Weekly	Replicate CCV samples	Absolute RPD for each target ≤ 25% Ongoing comparison of CCV to have RSD ≤ 15%

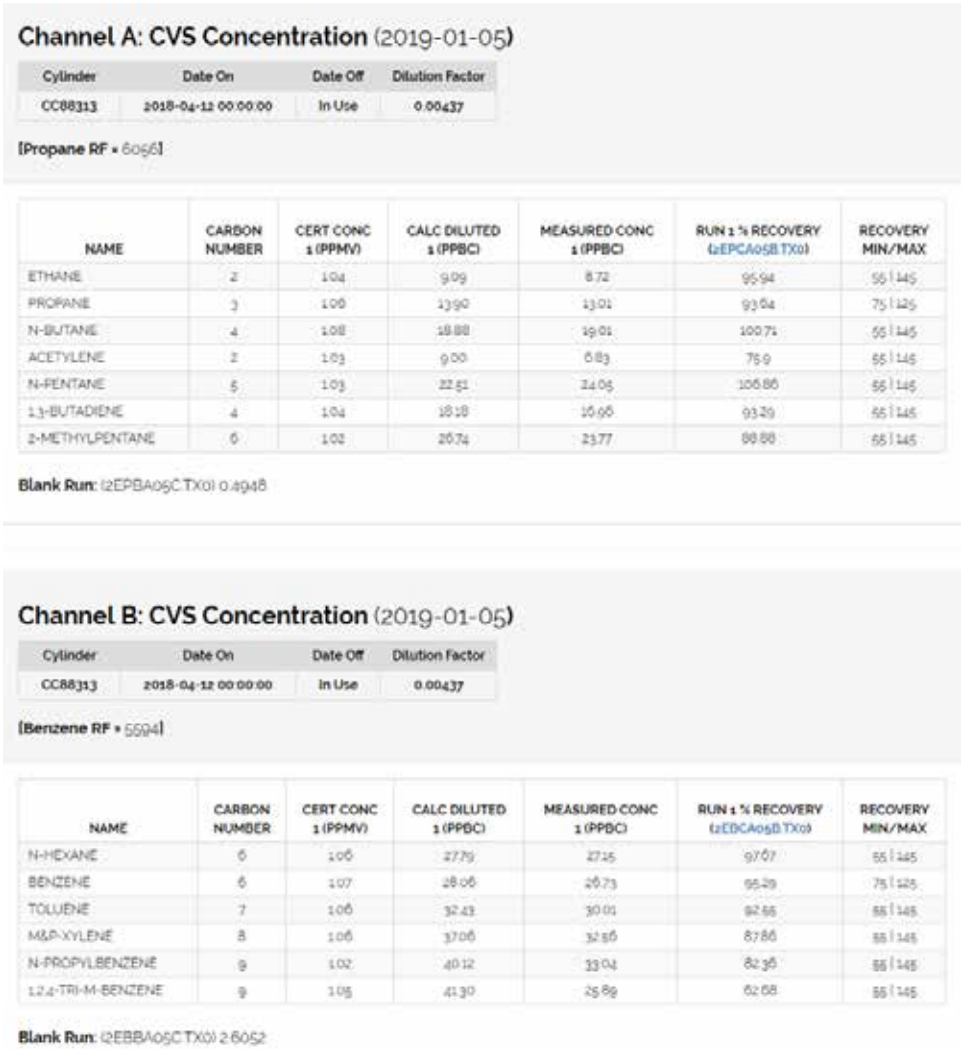
Table 2. TAD Quality Control Parameter Summary

VISUALIZATION OF PAMS DATA USING THE MERLIN AUTOGC XPLORER (MAX) FOR DATA VERIFICATION AND INITIAL VALIDATION

The Merlin AutoGC Xplorer (MAX) is a cloud-based system developed specifically for AutoGC data from Agilent and PerkinElmer AutoGC Systems. It allows the review of all quality control data, application of control limits, and review of ambient data for commonly occurring issues such as peak retention time shifting which will allow users to keep their systems generating quality data and complete the initial data review. The initial steps required to verify the instrument calibration drift and correct identification of all peaks correctly would involve a tedious review of chromatography and check sample control charts and possibly some reprocessing of data and additional review. MAX allows the initial verification of data completeness, calibration drift and target identification with tools to help data validators locate specific areas which need attention and will allow reprocessed data to be reloaded and reviewed to insure issues originally noted have been corrected.

SYSTEM STABILITY AND CALIBRATION VERIFICATION

Calibration stability is easily reviewed for both the daily continuing calibration verification (CCV) check samples and the weekly second source quality control (SSQC) check standard. The database is configured with the CCV and SSQC concentration information which allows the automatic calculation of recoveries for QC samples. Figure 1 shows the QC summary from the daily QuickLook page from MAX which is designed to show the daily data collected along with recoveries for any QC samples which have run. Information on target concentrations in daily blanks and recoveries for daily checks are



Channel A: CVS Concentration (2019-01-05)

Cylinder	Date On	Date Off	Dilution Factor
CC88313	2018-04-12 00:00:00	In Use	0.00437

[Propane RF = 6056]

NAME	CARBON NUMBER	CERT CONC 1 (PPMV)	CALC DILUTED 1 (PPBC)	MEASURED CONC 1 (PPBC)	RUN 1 % RECOVERY (2EPCA05B.TX0)	RECOVERY MIN/MAX
ETHANE	2	104	9.09	8.72	95.94	55 | 145
PROPANE	3	100	13.90	13.01	93.64	75 | 125
N-BUTANE	4	108	18.88	19.01	100.71	55 | 145
ACETYLENE	2	103	9.00	6.83	75.9	55 | 145
N-PENTANE	5	103	22.51	24.05	106.86	55 | 145
1,3-BUTADIENE	4	104	18.18	16.96	93.29	55 | 145
2-METHYLPENTANE	6	102	26.74	23.77	88.88	55 | 145

Blank Run: (2EPBA05C.TX0) 0.4948

Channel B: CVS Concentration (2019-01-05)

Cylinder	Date On	Date Off	Dilution Factor
CC88313	2018-04-12 00:00:00	In Use	0.00437

[Benzene RF = 5594]

NAME	CARBON NUMBER	CERT CONC 1 (PPMV)	CALC DILUTED 1 (PPBC)	MEASURED CONC 1 (PPBC)	RUN 1 % RECOVERY (2EBCA05B.TX0)	RECOVERY MIN/MAX
N-HEXANE	6	106	27.79	27.15	97.67	55 | 145
BENZENE	6	107	28.06	26.73	95.29	75 | 125
TOLUENE	7	106	32.43	30.01	92.55	55 | 145
M&P-XYLENE	8	106	37.06	32.56	87.86	55 | 145
N-PROPYLBENZENE	9	102	40.12	33.04	82.36	55 | 145
1,2,4-TRI-M-BENZENE	9	105	41.30	25.89	62.68	55 | 145

Blank Run: (2EBBA05C.TX0) 2.6052

Figure 1. Quality Control Summary from MAX QuickLook.

available daily for review. MAX can be configured to load data continuously so operators and validators have near-real-time access to site performance. Because the data is loaded into the database hourly, the chromatographic data system be revised where issues occur.

Figure 2 shows the resulting control chart of CCV recoveries. Calibration checks sometimes fail as a result of some error in collection or operator error and this data can be removed or flagged so control charts do not reflect non-representative control samples. Occasionally target components in the nightly check will be misidentified and may need to be reviewed, reprocessed and reloaded.

REVIEW OF PEAK IDENTIFICATION

In the process of checking the calibration drift, validators may indeed determine periods of time where the chromatographic system was not identifying all targets correctly. Failures in the weekly retention time check

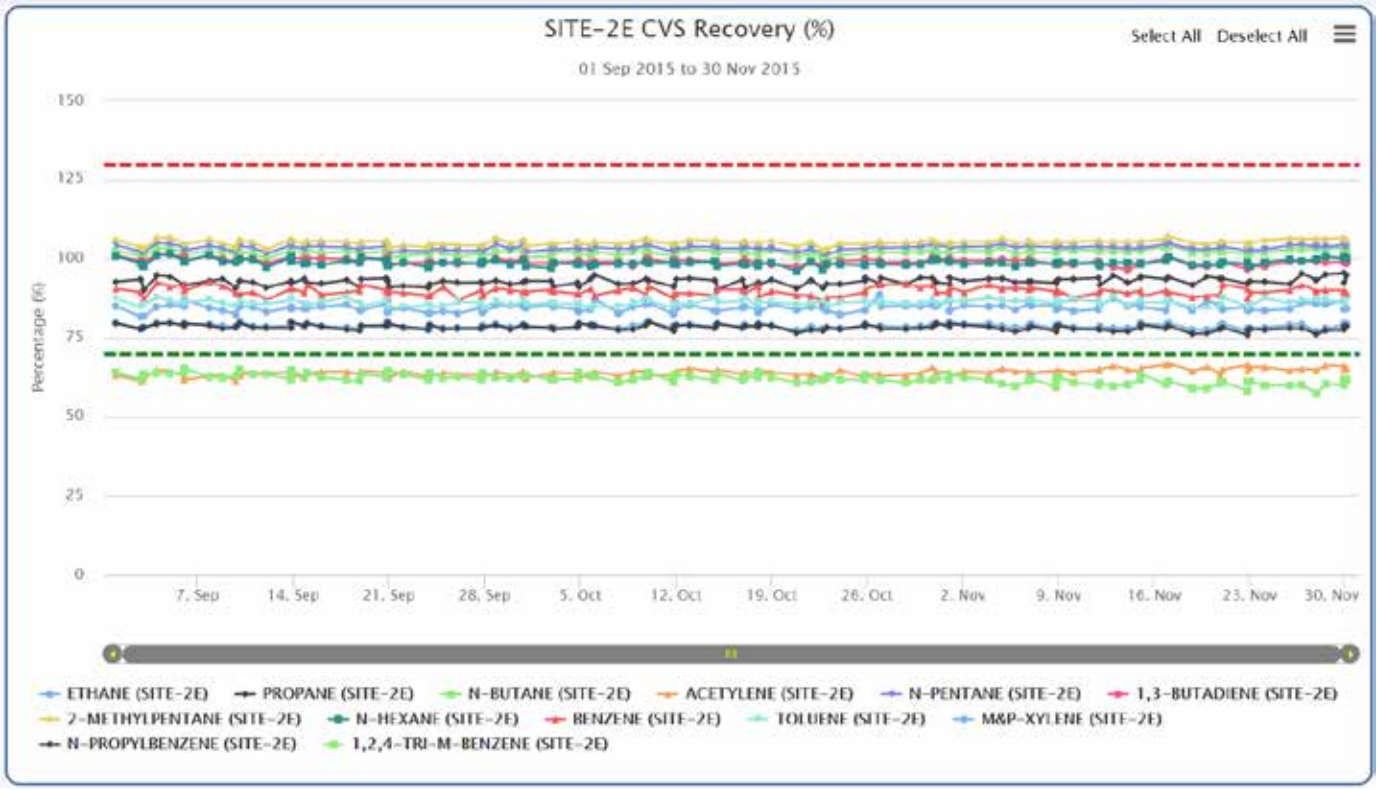


Figure 2. CCV Quality Control Chart

sample may require a lot of data review to determine where exactly targets are misidentified. MAX allows the plotting of retention times in ambient and this can show more precisely where retention time drifting occurs. This allows targeted review of chromatographic data and thus reduces the amount of chromatographic review required. Validators may resubmit individual hourly data which have been reprocessed to capture misidentified targets. Once re-submitted the same retention time plots can confirm that areas of misidentification have been corrected

Figure 3 shows the retention time plot of three closely eluting targets on the end of the PLOT column chromatogram which are a good indicator of retention time shifts which result in misidentification of peaks on the PLOT chromatogram. Drift is easily recognized allowing operators feedback on optimizing methods to maintain good peak identification. Rollover identification of individual data points allows validators to evaluate further by showing the data file name and concentration. Concentrations below the detection limit may be more difficult to identify accurately and may not warrant significant attention or reprocessing.

Data which is missing, out of hourly collection windows or which has been reprocessed and reloaded are clearly identified within the database. This allows review and confirmation of reprocessing activities.

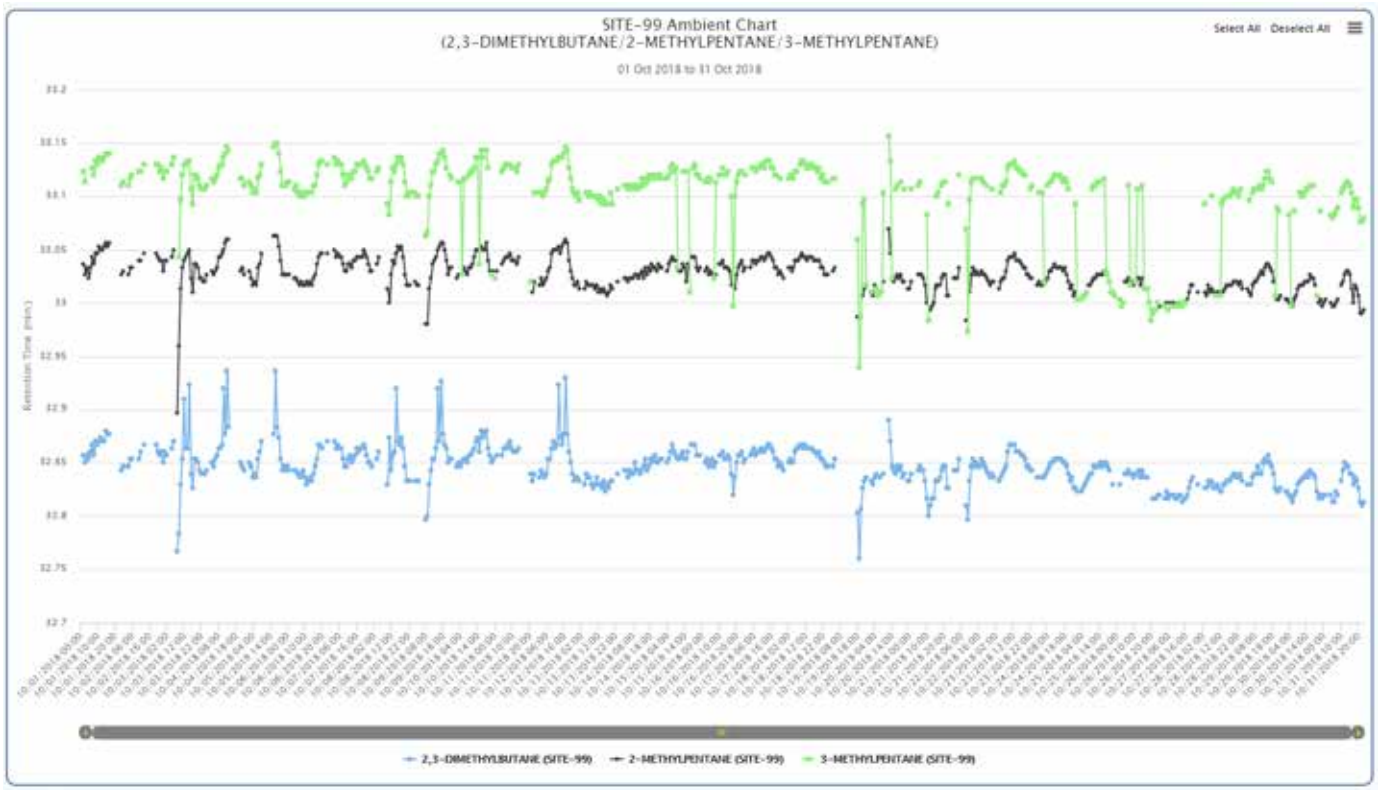


Figure 3. Retention Time of 2,3-Dimethylbutane, 2-Methylpentane and 3-Methylpentane on the PLOT column.

REVIEW OF SYSTEM PERFORMANCE AND AMBIENT DATA

Once the QC data have been evaluated and the analytical system is determined to be capable of generating correctly identified valid ambient data, the next level of validation involves evaluation of the ambient data itself for internal consistency. This requires comparisons of target species and observations which are best made using a database capable of allowing the review of ambient data graphically in time series, scatter plots and bar graphs to facilitate rapid review and highlight outliers.

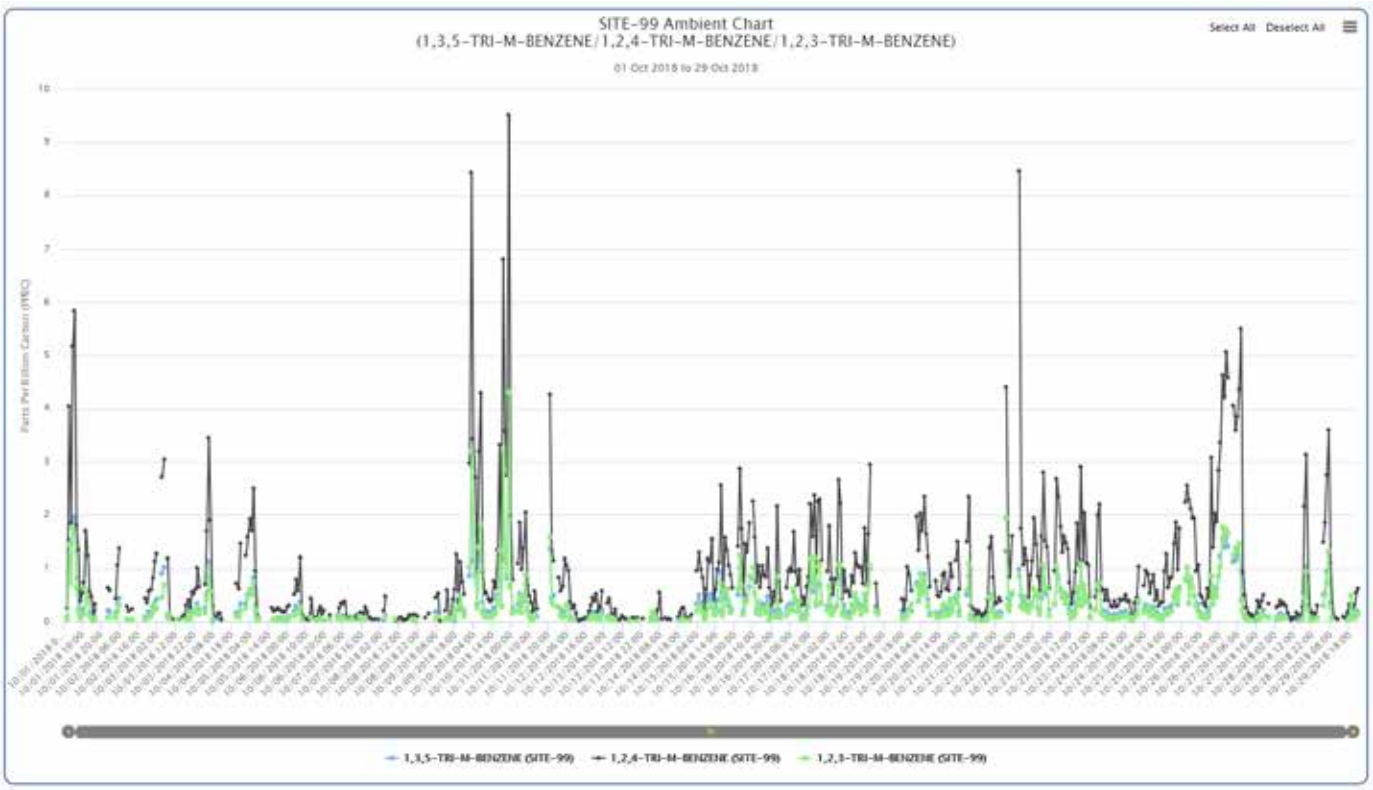
Ambient data review should include

1. Checks for known abundant species, evaluation of historical and representative concentrations and ratios of TNMHC to TNMTC
2. Consistent relationships among known targets such as:
   a. m & p-xylene > o-xylene
   b. propane > propylene
   c. 1,2,4-trimethylbenzene > 1,2,3-trimethylbenzene or 1,3,5-trimethylbenznene
3. Spacial and temporal characteristics
   d. Isoprene – diurnal and seasonal patterns
   e. Relation of benzene or toluene to totals

Outliers should be identified and a determination made if the target species is misidentified and data can be corrected by reprocessing, or it may

be qualified or invalidated. It should be noted that relationships between species can be site specific depending on the actual sources identified as contributing to the VOC totals. Figures 4 and 5 show example plots of ambient data which can be used to access system performance and locate outliers. MAX also has multiple output formats which allow data to be extracted from the database in comma-delimited or Excel™ formats for additional statistical evaluation. In addition to daily review of system performance, the system incorporates the necessary features to facilitate the flagging and generation of the necessary data for submission to the EPA Air Quality System. The secure cloud-based system facilitates data review by both operations and quality control managers. The secure cloud website allows the configuration of multiple users and permissions levels.

Figure 4. Graph of Selected Ambient Target Species.



SUMMARY

According to the Photochemical Assessment Monitoring Stations Implementation Manual 2, the objective of the PAMS program is to "provide an air quality database that will assist air pollution control agencies in evaluating, tracking the progress of, and if necessary refining control strategies for attainment of the ozone NAAQS." The objective of the validation process is to produce data of known quality to be added to this greater database. Data can be characterized as one of three categories; valid, valid but qualified or invalid. The PAMS data completeness requirements suggest that 75% of the measured values should be in the first two categories.

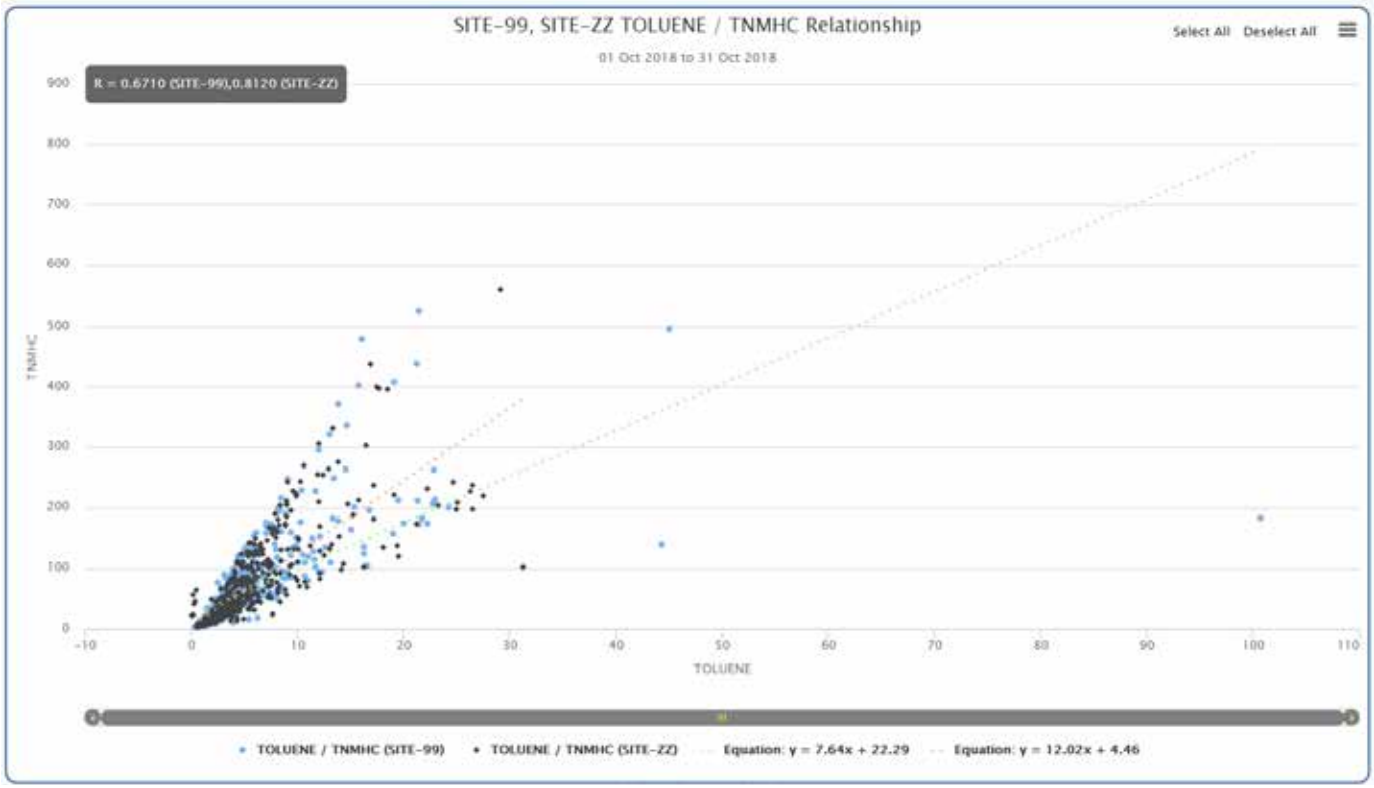


Figure 5. Relationship of Toluene to TNMHC in Co-located PerkinElmer and Agilent PAMS Analyzers.

The validation of PAMS AutoGC data requires both the evaluation of the instrumentation performance as well as the consistency and integrity of the ambient data. The use of quality control checks and the development of a clear pass/fail criteria will help validators to evaluate the data collected by these systems and insure that the AutoGC system is performing well. The ability to update the dataset with corrections will reduce data losses due to misidentification and increase data capture. While it is important to have a database capable of allowing the visualization of the quality control checks to allow validators to determine the system performance, it is also important to have the tools to assess the consistency and integrity of the ambient data. The review of ambient data includes the evaluation of species which should trend together, the review of diurnal patterns in species which exhibit these patterns as well as comparisons of species across sites and time are all important evaluations to insure that data is valid. The ability to review large datasets for consistency is integral to generating the monitoring data necessary for good ambient air models and for predicting and assessing the trends in ambient levels of VOCs over years and decades. Consistent data quality objectives will make these data useful for many years to come.

REFERENCES

1. United States Environmental Protection Agency. Revised Draft Technical Assistance Document for Sampling and Analysis of Ozone Precursors for the Photochemical Assessment Monitoring Stations Program. Office of Air Quality Planning and Standards (C304-06): Research Triangle Park, NC, March 2018.

2. EPA- Air Monitoring Technology Information Center (AMTIC), PAMS Data Analysis Workbook; see https://www3.epa.gov/ttn/amtic/pamsworkbook.html

25 years

Carol Meyer *has been actively involved with PAMS monitoring since the initial 1993 Coastal Ozone Assessment for Southeastern Texas (COAST) study. Orsat, LLC has provided ambient monitoring services to both state and industry, and currently operates 25 PAMS AutoGC monitoring sites in Texas for the TCEQ, UT CEER and AECOM. Using both the PerkinElmer Ozone Precursor system and the Agilent GC System with a Markes Unity 2 thermal desorber, Orsat has developed a fully automated application for the continuous monitoring of 56 NMHCs hourly.*

Fully Automated, Round-the-Clock, Photochemical Assessment Monitoring Stations (PAMS)

Orsat has customized the installation of hardware and software to produce a robust application for the measurement of VOCs in the ambient air known as the AutoGC. Orsat has been involved with continuous unattended ambient air VOC monitoring since its earliest implementation in the State of Texas Coastal Oxidant Assessment for Southeast Texas (COAST) program in 1994. Today, Orsat's services encompass all aspects of site operation from deployment to operator training and application assistance in topics ranging from Microsoft Windows operation to gas chromatographic theory. Over the past three decades, Orsat has deployed over 40 sites and currently maintains over 35 sites in the state of Texas for both public and private industry. In particular, Orsat has worked closely with the Texas Commission on Environmental Quality (TCEQ) to monitor air quality in the Barnett and Eagle Ford Shale Formations.

Orsat, LLC
1416 Southmore Ave
Pasadena, TX 77502
www.orsat.com
877-477-0171

Published 4/18/2019
5385-EN