

## AQRP Monthly Technical Report

<b>PROJECT TITLE</b>	Apportioning the Sources of Ozone Production during the San Antonio Field Study	<b>PROJECT #</b>	19-025
<b>PROJECT PARTICIPANTS</b>	Aerodyne Research, Inc.	<b>DATE SUBMITTED</b>	Nov 30, 2018
<b>REPORTING PERIOD</b>	<b>From:</b> Nov 1, 2018 <b>To:</b> Nov 30, 2018	<b>REPORT #</b>	2

A Financial Status Report (FSR) and Invoice will be submitted separately from each of the Project Participants reflecting charges for this Reporting Period. I understand that the FSR and Invoice are due to the AQRP by the 15<sup>th</sup> of the month following the reporting period shown above.

---

### Detailed Accomplishments by Task

A project meeting, with all scientists in attendance, was held during this past reporting period to go over individual sub-tasks. Progress on each task was discussed. Preliminary results including those presented here were shared with the group. This meeting resulted in several good suggestions for specific analysis pathways that will further project goals.

#### *Task 1: High-Resolution (HR) Analysis*

Mass-calibrations are the first step in performing a high-quality HR fit. The full PTR-ToF mass-calibration is complete. Major progress has been made on the GC-ToF dataset, which consists of a 3-dimensional dataset with variation in sample time, chromatographic elution time, and mass spectrum. Progress on this front is documented in the “Preliminary Analysis” section. The remaining dataset to be mass-calibrated is the I-CIMS dataset.

Pseudo-HR fits (where a set of automatically-identified peak positions are used in a full HR fit) is complete for the PTR-ToF, and the resulting dataset ready for use in subsequent tasks.

#### *Task 2: PMF Analysis*

An internal training course on PMF analysis has been given by Manjula Canagaratna. Topics covered include:

- What is PMF?
- Error calculation and Q matrix
- PMF literature and history
- Application examples
- Mixing of factors
- High-resolution vs unit mass resolution
- Two-stage PMF for variability within factors
- Installation and Running
- Loading PMF Results
- Overview of the PMF Results windows
- Troubleshooting
- fPeak exploration

- SEEDs and global minima
- Convergence criteria
- Choosing the optimal number of factors

The PTR-ToF dataset was used as a training dataset in the course. Prior to and during the course, a set of internal software tools were developed to manipulate the raw mass data conveniently. Such tools include time-averaging tools which properly treat the error matrix, and mass-filtering tools to eliminate ions based either on mass or label.

Preliminary PMF results on the PTR-ToF dataset are described in the “Preliminary Analysis” section”.

### ***Task 3: 0D Box Model***

0D box model software is installed and ready for testing.

### ***Task 4: Back-Trajectory Footprint Analysis***

Back-trajectory analysis be used in this project to help understand the area where the sampled air mass originated. Progress on two fronts is outlined below:

#### Part One: Progress on the mechanics of running the model

A sequence of function scripts running in Wavemetrics-Igor Pro has been developed. The scripts are called from a workflow text file that sets up multiple CONTROL and SETUP files. The scripts are used to set the starting date/time, duration of the backtrajectory release as well as directed to accept the meteorological files. The version of the executable we have compiled (see below) is limited to 12 files. The hrrr files are 6 hours in duration (e.g. 4 files for 1 day). Due to this step size, a three day back trajectory exceeds the met file limit. An improvement in the scripting, now prioritizes using the high-resolution rapid refresh data files for the initial portion of the back trajectory, but then uses lower resolution met files as the trajectory has 'aged'. In other words, initially the particles are tracked backward from the receptor location using the high resolution met files, but the further away they travel in time, the lower resolution files are substituted in to provide temporal and spatial coverage for the (presumably) large, dispersed plume.

Additionally, we have followed the directions from NOAA to compile our own version of the hysplit executables that can engage multiple cores for faster execution. A casual analysis of the mechanism that the multiprocessor version has revealed a bottleneck. The scheme for making the multiple cores possible is a relatively easy one to transcode, however, in this implementation the different processes exchange information through disk files, thus, disk I/O limits the communication between CPUs. To take advantage of the large number of cores on our analysis computer, we now install the hysplit executables into a large virtual ramdisk. This has led to the ability to process 100's of jobs in a large batch fairly easily.

The workflow scripts described earlier generates jobs that are tested on a laptop with a low number of released particles, to verify the configuration. With only a small modification, jobs with a large number of particle releases are run on the faster machine.

#### Part Two: Progress on the model results

A full high-resolution data set has been generated. The data consists of three things. 1) for each hour of the project (t), a lat/lon grid depicting the results of particle release run backwards for each t - n hours (n from 1 to 72) is archived. 2) the sum of all grids for each (t) is archived (this is the 'footprint') and 3) an image of the footprint is bound into a movie frame. This final product

will enable the research team to quickly get a snapshot view of where the air was likely to have come from for every hour. The actual digital footprint can be used in more advanced analysis yet to come.

## Preliminary Analysis

**Mass calibrations across GC chromatograms.** The main challenge in performing HR analysis on mass spectra obtained from GC separation is finding species that span an appropriate mass range to use for the mass calibrations across chromatogram retention times. Our goal was to find a series of compounds that are eluted constantly into the detector (EI-ToF) from our GC, and thus would be subjected to the temperature ramping and any pressure changes that occur, due to the GC oven temperature ramp, which has been shown to affect mass drift.

The solution that we found to account for this was to use compounds that were being emitting from the GC column itself, and to use their fragments as the mass calibration ions. These include ions like  $C_8H_7O_2Si_2^+$  and  $C_5H_{15}O_3Si_3^+$  which come from the stationary phase of the column that was used during the study. Using this approach, we were able to generate a unique peak list of mass calibration ions that allowed us to obtain  $\pm 10$  ppm mass calibrations across the chromatograms for the data that came from this campaign.

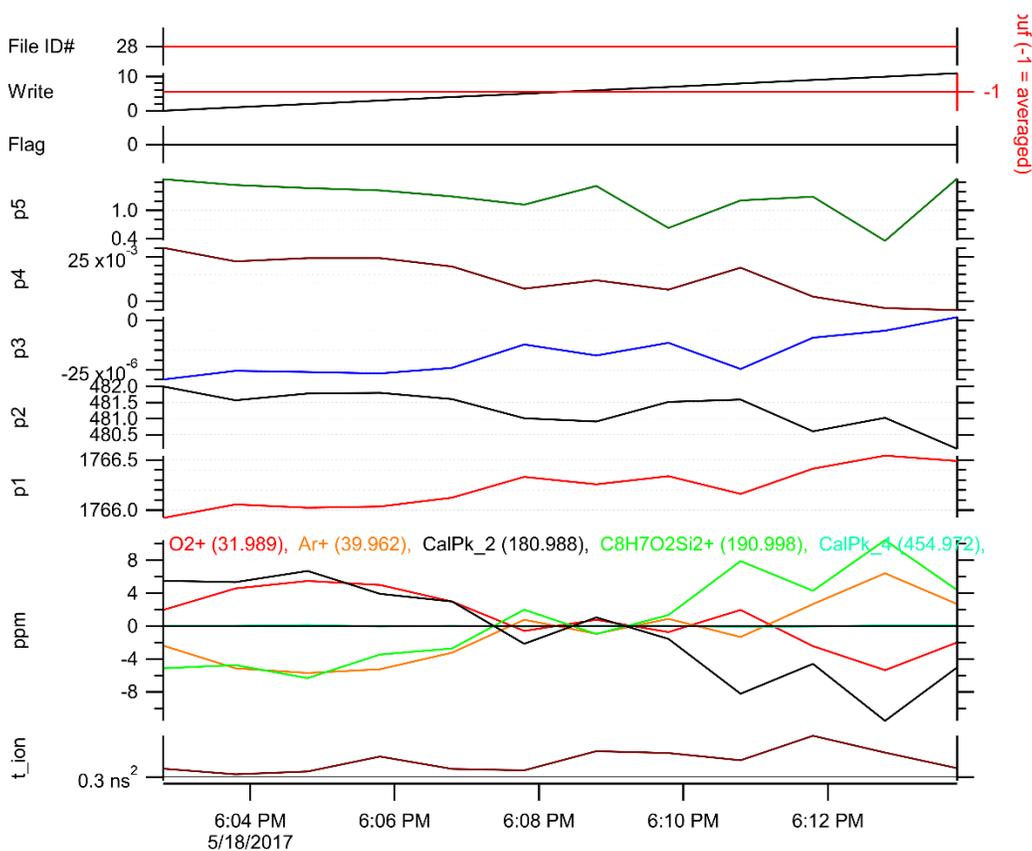


Figure 1. Mass calibration result for the GC-ToF instrument.

With the HR mass calibrations now complete for the campaign chromatograms, we have generated a peak list of analyte ions present in our chromatograms for identification and quantification. One of the large benefits of conducting this HR analysis is that we have significantly reduced the backgrounds of our ion trances. An example of this is shown below for the ion  $m/z$   $C_3H_7^+$  resolved using the HR analysis from its unit mass trace  $m/z$  43. Aside from reducing baselines, using the HR analysis we are able to distinguish between ions of different

molecular formulas but with the same unit mass. This is shown in the second panel, where a section of the chromatogram has been selected to show that while there are ions eluting with  $m/z$  43, only a subset of those chromatographic peaks can be associated with the  $C_3H_7^+$  ion.

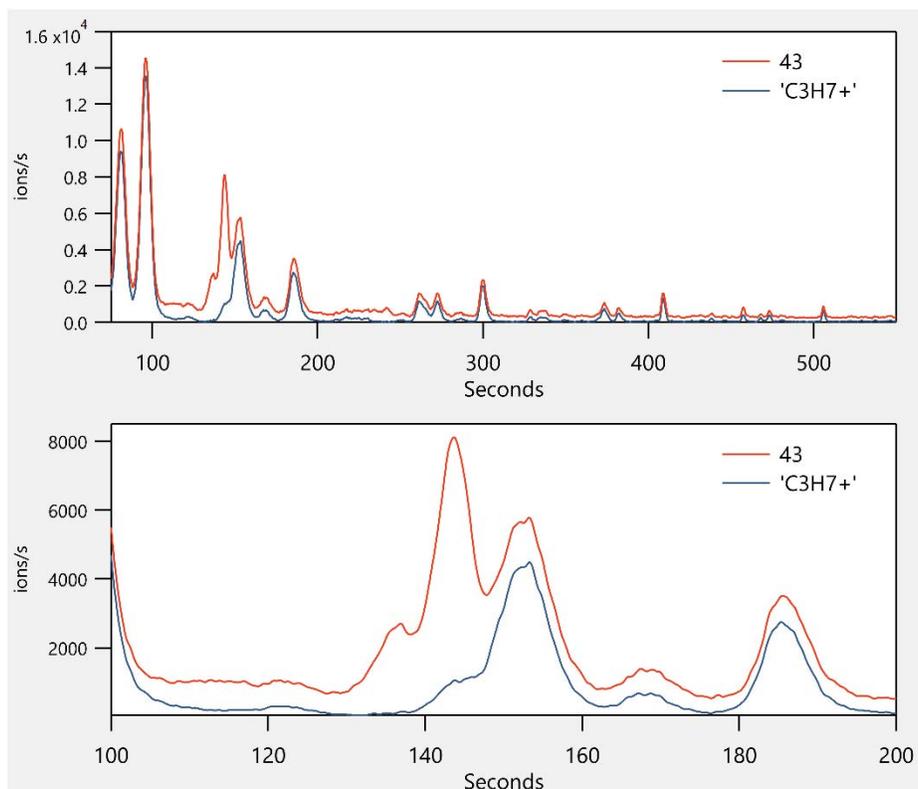


Figure 2. A full chromatogram from the GC-EI-ToF instrument (top) showing two ions with a mass-to-charge ratio of  $m/z=43$ . A zoomed-in view (bottom) is also shown.

With the mass calibrations complete, and an analyte peak list in place, our GC-EI-ToF data set is ready to be used for PMF analysis, and to be taken out of the Tofwerk software and into TERN for retention time shifting, peak fitting, and quantification of the chromatographic peaks.

**Preliminary PMF Results on PTRMS Dataset.** Proton transfer mass spectrometer data (EC instrument) was considered, using mass data from the 10-second pseudo-HR fit (see previous monthly report). The pseudo-HR fit took a small set of known ions and added an extensive set of automatically fit peak positions.

Data is in raw counts, and has not been zero-corrected or calibrated. Error from un-baselined data was used. Data was post-averaged onto a 5-minute time-base. Only stationary periods were considered. Any QA-ed periods in original delivered PTR data have also been excised.

The mass list has been aggressively paired down to tailor the PMF analysis on ions of greatest potential atmospheric importance. This list included no negative mass defect ions (usually associated with the chlorine-containing mass calibrant), and focused instead on ions that showed decreases with additions of zero air. Data with  $S/N < 2$  was downgraded by 2. Only fPeak 0 was considered for this first exploration.

A 7 factor PMF solution is shown below. This solution was chosen over the fewer-factor solutions because it is the first that separates aromatics into two distinct factors.

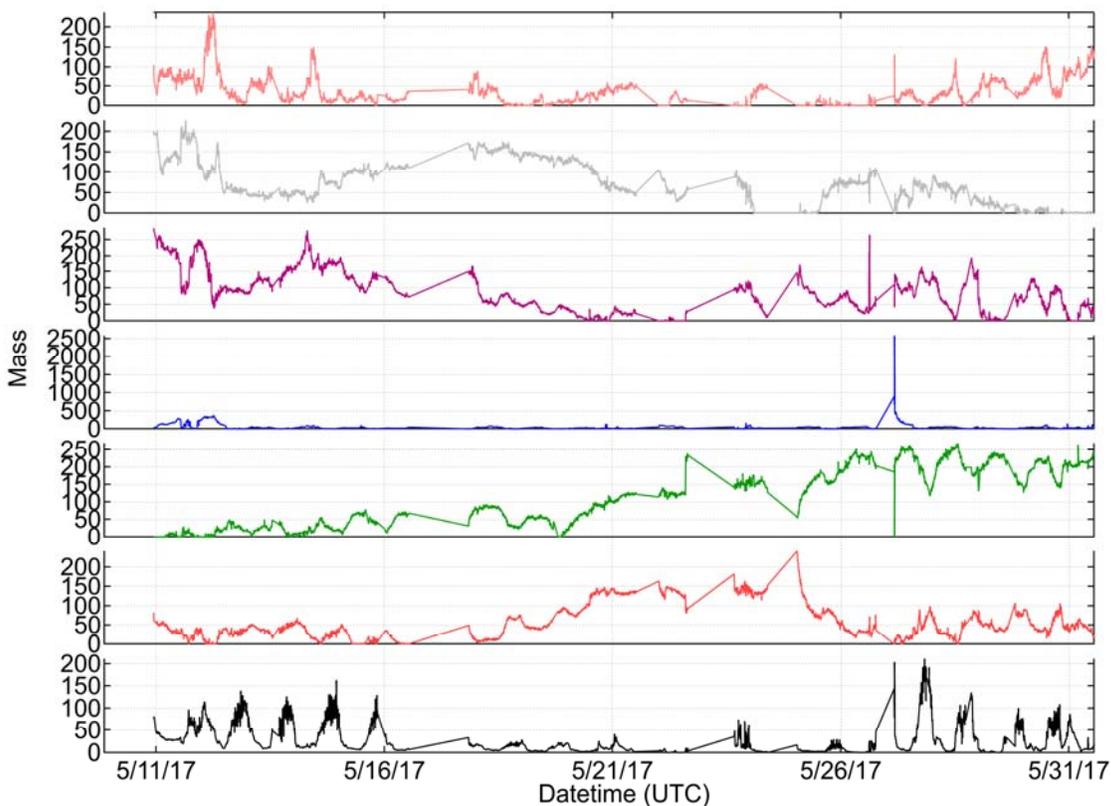


Figure 3. PMF factor time traces for a 7-factor solution using a pared-down pseudo-HR fit from the PTRMS instrument.

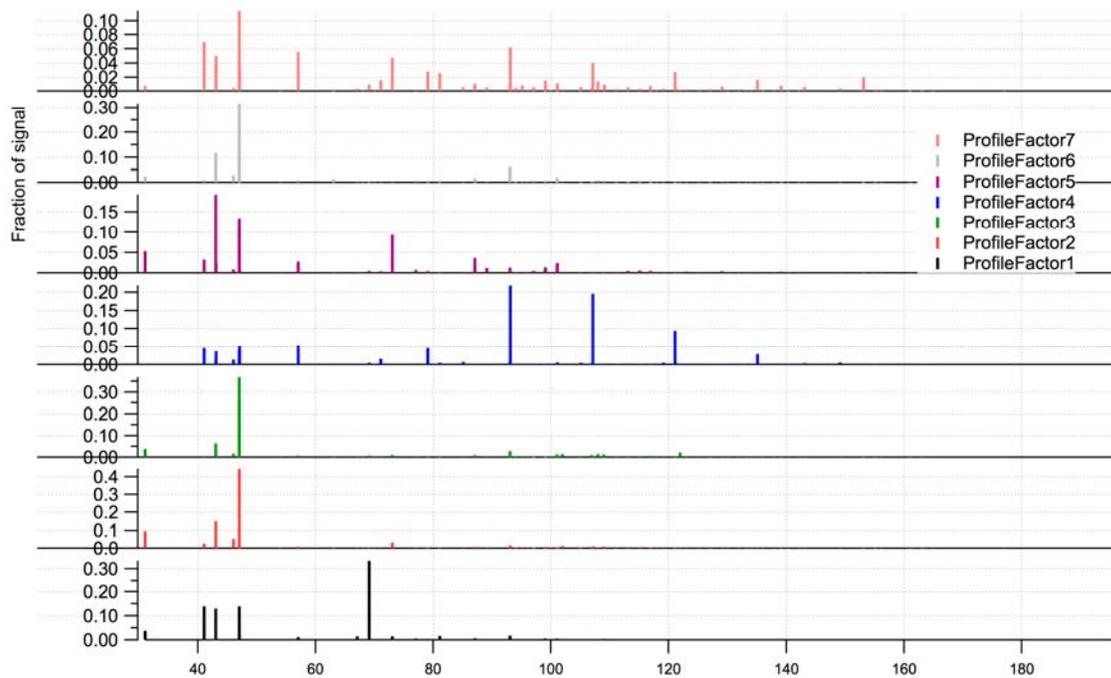


Figure 4. PMF factor mass spectra for the 7-factor solution shown in Figure 3.

A preliminary investigation of these factors has been attempted. From top down:

- Factor 7 (peach): This factor contains peaks corresponding to both aromatics, alkane fragments and oxygen-containing species.
- Factor 6 (grey): unknown. Broad baseline signal
- Factor 5 (purple): unknown. Looks similar to factor 6, but includes more diurnal component.
- Factor 4 (blue): This factor contains strong aromatic components.
- Factor 3 (green): Possible instrumental function. Some diurnal component can be seen. This factor drifts up over the course of the campaign. At least one negative mass-defect ion that escaped the original data purge is present in this factor and not in others.
- Factor 2 (red): Another possible instrumental function. Opposite diurnal behavior as green. Many ions are similar to those present in Factor 3 (green).
- Factor 1 (black): Biogenic. Dominated by isoprene

Factors 7 and 4 (peach and blue) are two factors that warrant additional investigation and exploration. The mass spectra associated with these factors will be investigated for characteristic features that might indicate specific source categories, such as fuel vapors, aged or oxygenated compounds indicative of transport, or others.

### **Data Collected**

No data will be collected as part of this project. However, data will be generated after completion of Task 1, HR analysis.

### **Identify Problems or Issues Encountered and Proposed Solutions or Adjustments**

During the PMF analysis workshop, we identified that the source datasets were too large to be analyzed easily and quickly with current PMF software tools. This presented a significant challenge since much of PMF analysis relies on exploring a wide solution space, requiring multiple runs of the software. One possible solution was to divide the month-long dataset into sub-periods. However this might obscure variability between different measurement locations, for example. Instead, an averaging algorithm was developed that can quickly average the mass spectral results, while properly treating their associated error matrices, allowing the 10-second PTR-ToF dataset to be turned into a 5-minute dataset, which was adequately small in size. This algorithm takes only a few seconds and bypasses the need to re-fit the pseudo-HR dataset on a slower time base (several hours/days). Additional mass-filtering algorithms described previously also help in moderately reducing the size of the PMF input dataset.

### **Goals and Anticipated Issues for the Succeeding Reporting Period**

In the next reporting period, there are several goals:

- Task 1: Make progress on I-CIMS dataset mass calibrations and generation of a pseudo-HR dataset.
- Task 2: Hold a mini-meeting to discuss the preliminary PMF mass spectra from the PTR-ToF dataset, focusing on identifying specific mass signatures and potential source categories.
- Task 2: Develop a method for doing PMF analysis on the GC-ToF dataset.
- Task 3: Test the 0D box model with the provided test case.
- Task 4: Investigate HYSPLIT results for dependence on release height.

No issues are anticipated, besides the delay in project start date, discussed previously.

## **Detailed Analysis of the Progress of the Task Order to Date**

Now that the inaugural project meeting has taken place, progress is ramping up in earnest. Each scientist has a list of to-dos and all team members are on the same page as to the task at hand.

During this past reporting period, significant progress has been made on Task 4, HYSPLIT back-trajectories. This task is mostly independent of the other tasks, and as a result could be begun earlier than planned. Progress was aided by technical know-how developed as part of unrelated projects.

Significant progress was also made on the PMF analysis task, aided by an internal PMF seminar that used real project data as a test case.

**Do you have any publications related to this project currently under development? If so, please provide a working title, and the journals you plan to submit to.**

Yes       No

**Do you have any publications related to this project currently under review by a journal? If so, what is the working title and the journal name? Have you sent a copy of the article to your AQRP Project Manager and your TCEQ Liaison?**

Yes       No

**Do you have any bibliographic publications related to this project that have been published? If so, please list the reference information. List all items for the lifetime of the project.**

Yes       No

**Do you have any presentations related to this project currently under development? If so, please provide working title, and the conference you plan to present it (this does not include presentations for the AQRP Workshop).**

Yes       No

**Do you have any presentations related to this project that have been published? If so, please list reference information. List all items for the lifetime of the project.**

Yes       No

---

Submitted to AQRP by      Dr. Tara Yacovitch  
Principal Investigator