University of Wollongong

Assessment of benzene monitoring

EPA (Victoria) OPSIS data from Geelong
12/2002 – 3/2007

27 April 2007

ASSOC. PROF. STEPHEN WILSON, DEPT. OF CHEMISTRY, UNIVERSITY OF WOLLONGONG, NSW, AUSTRALIA, 2522.

Synopsis

Brief

EPA Victoria has requested the following information:

Provide an opinion on the accuracy and quality of benzene monitoring data generated by EPA Victoria between December 2002 and March 2007. In forming an opinion review the EPA Victoria operating systems, procedures, calibration records, and data obtained at Geelong Grammar between December 2002 and March 2007.

Report Summary

The measurements of benzene are credible. The estimates of the accuracy and uncertainty of the measurement are reasonable, although there are some minor issues associated with the documentation.

The records presented provided a good overview of the operation of the instrument, and demonstrate a clear effort by EPA Victoria to continually improve the operation, maintenance and calibration of the instrument.

Background

The OPSIS system assessed here is based around long path absorption of light (normally known as DOAS, Differential Optical Absorption Spectroscopy). With this DOAS technique, a light source is placed at a distance from the detector, and the space between the source and detector is the measurement sample. The fundamental quantity calculated by such a system is the concentration of the absorber such as benzene averaged over the total path length. In principle this average provides a much better assessment of likely community exposure when compared to measurements at a single location, which may not be representative of the wider area.

As has been illustrated elsewhere (see for example EPA Victoria publication 1019.1 or [*Volkamer, et al.*, 1998]), the DOAS technique requires careful analysis to retrieve reliable measures of concentrations of species like benzene. The absorption by benzene, for example, lies in a region where species like oxygen also absorb, and this interference needs to be carefully removed before reliable benzene concentrations can be returned [*Kourtidis, et al.*, 2000].

Reliable operation of these systems therefore relies on both good instrument (optical) performance and reliable analysis protocols. The OPSIS instrument plus software has been assessed internationally, and has been approved by a number of agencies, including the US EPA Technical Verification program and the German TÜV for benzene air quality monitoring. The operation of the OPSIS system at Geelong has been certified by NATA under accreditation 1567 (25 May 2006, documents supplied by EPA Victoria). The closed nature of the commercial instrument plus software does lead to some significant difficulties in assessing performance [E.g.*Kim*, 2004; *Xie, et al.*, 2004].

The instrument being investigated here is located at Geelong Grammar School (see EPA Victoria publication 1022 for a map.) While the OPSIS system does record and save the spectra measured by system, this information is not generally available to the end user. Instead, the software provides the results of its analysis of these spectra. These are the data that have been investigated in the current analysis.

Structure of this document

This report assesses the quality and accuracy of the results produced by the OPSIS instrument between December 2002 and March 2007 by considering the following questions:

Is the measurement technique fundamentally sound?

Has the OPSIS instrument been demonstrated to measure benzene?

Is the calibration of the instrument valid?

Are there quality control measures in place for the data produced by the instrument?

The answers to these questions will be used to justify the overall assessment.

OPSIS Data Accuracy and Quality (12/2002 – 12/2006)

Material Assessed

For this assessment EPA Victoria provided the following records of instrument performance for the period under investigation:

- **CARMS File 60388-1 OPSIS Daily Reports Corio 2002**
  (12/ 2005 – 12/2006) Includes daily reports, service visit log sheets, multipoint bench calibrations, in path span checks, monthly validation summaries, and monthly reports.

- **CARMS File 60387-1 OPSIS Logs Corio 2002**
  (20/8/2002 – 22/12/2006.) Instrument Log sheets.

- **CARMS File 60401-1 OPSIS Calibration and Performance Checks Corio 2002**
  (28/4/1998 – 11/11/2005) Historical calibration and performance checks.

- **CARMS File 60401-2 OPSIS Calibration & Performance checks Corio 2002**
  (6/6/2005 – 29/11/2006) Multipoint bench calibrations, in path span checks.

- **Documents EC 343_1 through EC 343_13**
  Operational methods for the OPSIS DOAS instrument.

- **Method Validation report for the measurement of benzene, toluene, NO$_2$ and SO$_2$ by DOAS using OPSIS**
  (1/2006) Summary of the performance of the OPSIS DOAS instrument

- **Measurement Uncertainty Determination Report for the Measurement of benzene, toluene, NO$_2$ & SO$_2$ by DOAS using OPSIS method.**
  (2/2006) Documentation of the uncertainty estimation calculations.

- **Benzene Air monitoring in Corio 2003 – 2005**
  EPA Victoria Publication 999, June 2005

- **Benzene Air monitoring in Corio 2003 – 2005**
  EPA Victoria Publication 1022, March 2006

- **OPSIS 2007 Corio**
  Logs, check cals and operational communications for the OPSIS instrument in Corio.

The period of analysis starts with the installation of the instrument at this site.

Is the measurement technique fundamentally sound?

The DOAS technique measures a fundamental property of the target molecule, its ability to absorb light. This is altered by the temperature (and slightly by the pressure) of the molecules. Provided the absorption of the molecule of interest can be separated from the absorption features of other molecules in the sample, the measurements are highly reliable.

In practice, several key issues need to be addressed. In order to determine the absorbance of any of the species, the spectrum of the illuminating lamp needs to be characterized (often known as the background or reference spectrum.) The resulting spectrum then needs to be carefully analysed to ensure that the absorbing molecules are independently determined. [Eg. *Volkamer, et al.*, 1998]

While the basic principles of the OPSIS implementation of these retrieval methods can be determined, the details are not readily accessible. However, the OPSIS system has been tested by a number of agencies worldwide for the purpose of measuring hydrocarbons, including the US EPA Technical Verification program and the German TÜV for benzene air quality monitoring.

The operation of the OPSIS system at Geelong has been certified by NATA under accreditation 1567 (25 May 2006, documents supplied by EPA Victoria).

Has the Geelong OPSIS instrument been demonstrated to measure benzene?

Tests outlined in the method validation report clearly show three important features.

When benzene is placed on a test bench the instrument detects the presence of benzene, and shows good linearity between the reported concentration and the amount estimated based on a calibration gas. The test range of this determination (0 – ~250 $\mu g/m^3$, 0 – 80 ppbv) spans a range much larger than that normally encountered in atmosphere in Geelong, and so provides a good indication that the retrievals are performing reliably, with respect to benzene.

The second test (Method Validation) is the assessment of the cross sensitivity of the measurement to the other species present in the calibration standards. This test is performed by increasing the amount of one gas in the path during normal operation (via the use of an in-path gas cell). The retrieved amount of benzene was found to be insensitive to the amount of the other gases (toluene, SO_2 and NO_2), indicating good selectivity.

The third indication comes from the routine in-path calibration checks for benzene. These checks are subject to real variations in the amount of benzene in the path during the measurement, and so results are expected to be more scattered than those of the test bench. (Results presented in Figure 3). The relative constancy of the retrievals reported in these tests indicates that there is no significant interference from other unknown chemical species, at least under the conditions where these tests are performed.

Is the calibration of the instrument valid?

Fundamentally, for an instrument with a linear concentration response (as discussed above), the calibration of the instrument consists of determining the instrument zero (what does the instrument measure when there is no benzene in the path), and the response to a known concentration of benzene in the path. These will be considered separately as they present different challenges to the instrument and operations. The precision (repeatability) of the measurements made by this instrument also needs to be considered.

Zero determination

The zero of the instrument is determined by the background spectrum (sometimes called the reference) that is being used. Every time the background spectrum is altered, there will be a change in the zero of the instrument. This can be quite difficult to validate, as the only means of knowing the concentration of benzene at the time of measuring the background spectrum is to use prevailing meteorological conditions (and hence an assumption of the source of benzene) and the OPSIS reported benzene concentration, which depends on the amount of benzene present when the previous background was measured. It is not clear from the documentation provided when the background is altered.

One method of looking at the change this produces is to look at the 25th percentile of the reported benzene measurements. This measure is chosen to ignore most pollution events, and so to provide a proxy for the zero of the instrument. For daily data sets this is shown in the figure below. There appear to be jumps in the data, represented in the graph by the arrows. These jumps are most likely due to changes in the background spectrum being used. It should be noted that these changes are small in comparison to the 75 $\mu g/m^3$ benzene intervention level specified in Victoria's State Environment Protection Policy.
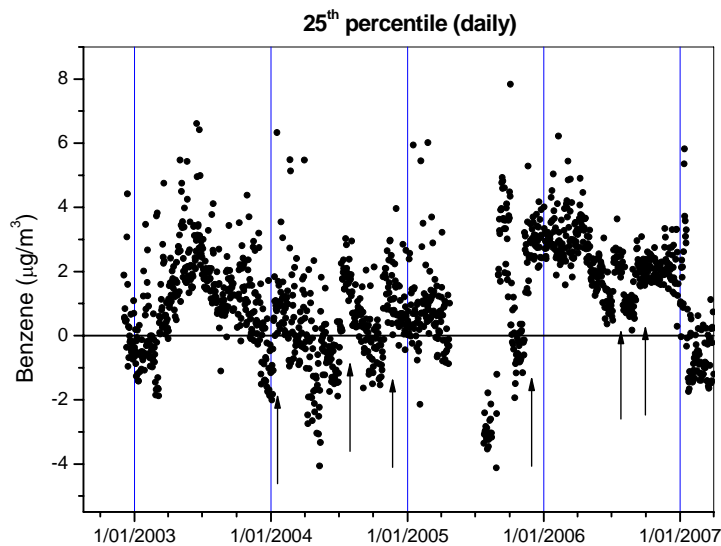
- Figure 1 Benzene concentration of the 25th percentile based on daily data sets. Some points where the calculated benzene concentration appears to change are highlighted with arrows.

The zero therefore appears to be well defined, with a standard deviation of around 0.5 $\mu g/m^3$ estimated from the last portion of the 2006 data. This value is consistent with the estimations based on zero air gas measurements (Method Validation). However, the systematic variation in zero appears to be of the order of $\pm$ 3 $\mu g/m^3$ throughout the period. While not limiting the detection of pollution events, it is an issue that would need to be resolved for the reliable measurement of much lower concentrations. These variations are within the uncertainty estimates ascribed to the data (see EPA Victoria Publication 999).

Calibration at a known concentration

The quantities reported by the OPSIS instrument are based on reference spectra of the target chemical compounds (pre)-recorded under well controlled conditions. For operation in the field this is then compared with standard gas mixtures to determine whether the calibration remains valid. A significant disagreement between the existing calibration and calibration generated by gas mixtures indicates a problem with either the calibration procedure or a change in instrumental performance.

The calibration tests involve two different configurations, the monthly in-path span check and the 3 monthly multi-point calibrations. The multi-point calibration is deemed to be the same as the previous calibration if the slope lies within ±20% (Method Validation). At the Victoria's State Environment Protection Policy Intervention level this corresponds to approximately a 20% uncertainty in the retrieved value.[1]

The slopes retrieved during the period in question are shown in the figure below. This clearly demonstrates that the ±20% is clearly met for benzene throughout the period. From the documentation provided it therefore appears that the initial calibration provided with the OPSIS has been used for the entire period.

Based on the calibration data slopes, the standard gases used imply that the calibration used by the OPSIS is 7 $\pm$ 2% (95% confidence) low for the period.

---

[1] Note that the zero offset and slope will be correlated in the retrieval, and so the actual variation at the retrieved concentration could be smaller.

OPSIS - Corio, benzene
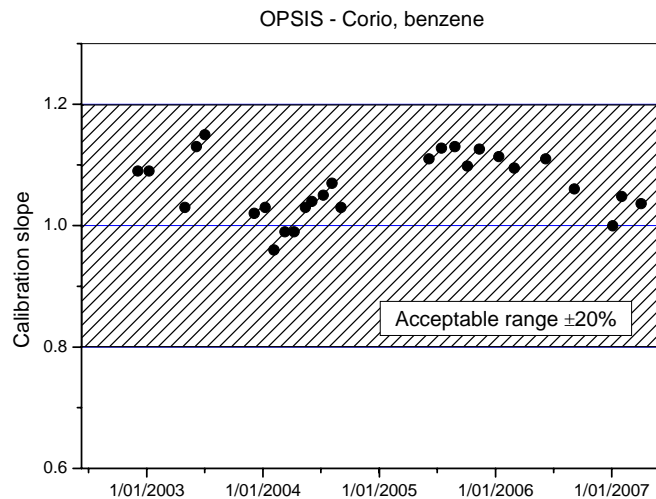
- Figure 2 Plot of the calibration slope determined for benzene as part of the routine calibration procedure.

The other technique used for assessing the calibration, the in-path calibration, has only recently been introduced. The results from this calibration are shown below, with a scale similar to Figure 2. Prior to June 2006 the two calibration methods agreed in the magnitude of the difference between the gas standard and the internal OPSIS calibration. Post June 2006 there is a divergence, which is exaggerated by the greater number of measurements during this period using the in-path method. The documentation shows that the period at the end of 2006 was recognized as a time when there were instrumental problems. This will be discussed in more depth later.

The in-path checks are an important component of the overall performance assessment. They permit more regular checking of instrumental performance than the span checks listed above, and also provide the only check of the impact of interference in the measurement of the target species.

OPSIS Corio In-Path checks
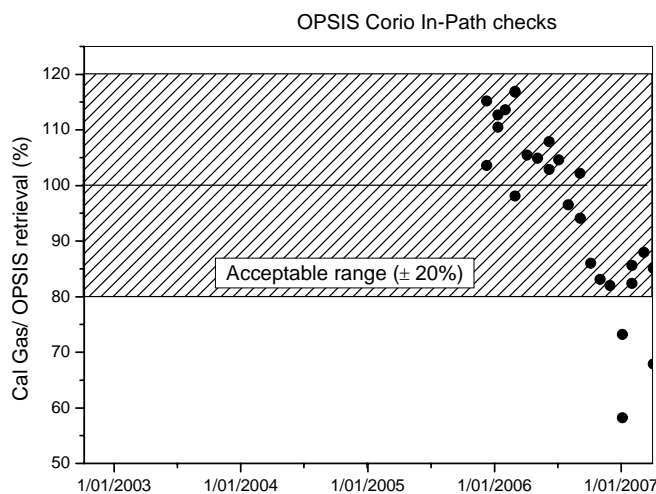
- Figure 3 In-path calibration results for the OPSIS system till the end of 2006. The supplied data has been redrawn to match the scale of Figure 2. For a number of days in the later part of the record there are two measures of the retrieval, made before and after system alterations were made.

Uncertainty estimates

The estimation of the uncertainty is documented at length in the Measurement Uncertainty Report. The model used for estimating the uncertainty in the measurement (Figure 1, in the Measurement Uncertainty Report), has some minor problems. For example, it would be useful to clearly delineate between the processes of calibration and those of routine operation.

There is one apparent error in the use of °C for the uncertainty estimate, whereas the quantity used in the calculation itself is K.

The uncertainty model used assumes that there is a linear relationship between the various listed sources of uncertainty and the reported concentration and this is not necessarily true. With temperature, for example, the molecular absorption will change with temperature, but it is not clear how this will impact upon the retrieved quantity, as it should be fundamentally non-linear, and from the information available it is not known whether the OPSIS software considers this. The temperature is also used in the determination of the retrieved concentration ($\mu g/m^3$) and mixing ratio (ppbv), and in this situation the linear model is appropriate. Similar issues exist also for pressure.

The error budget has not been recalculated for this assessment. However, the overall magnitude of the uncertainty appears reasonable and the underlying assumptions used are appropriate given the information provided with the instrument. The magnitude of the uncertainty estimate is in agreement with estimates of others. [*Xie, et al.*, 2004]

Are there quality control measures in place for the data produced by the instrument?

The documentation provided shows that a series of measures have been put in place to assess the quality of the data and to report both issues of instrument failure and high benzene concentrations. The clear picture arising from these records is that the procedures for operation of the instrument have been significantly improved with time, with more carefully defined procedures carried out which parallel earlier practice, but which make it more likely that the processes are adhered to. This includes both more automated data collection systems and more streamlined documentation of the methodologies.

There is a clear record of issues being identified and being resolved. For example, problems were identified with the change to Eastern Daylight Saving Time, and that this was subsequently resolved by ensuring that all measurements were made in Eastern Standard Time.

The use of the light intensity as a flag of either questionable performance ($< 20\%$) or rejection of data ($< 5\%$) (e.g. EPA Victoria publication 1022) is appropriate. The changes in the standard deviation of the retrieved benzene as a function of light intensity are shown in Figure 4, using raw measurement data from 12/2002 through 5/2005. The standard deviation in the data will increase with increased concentration. However, it clearly also increases significantly as the lamp intensity decreases.
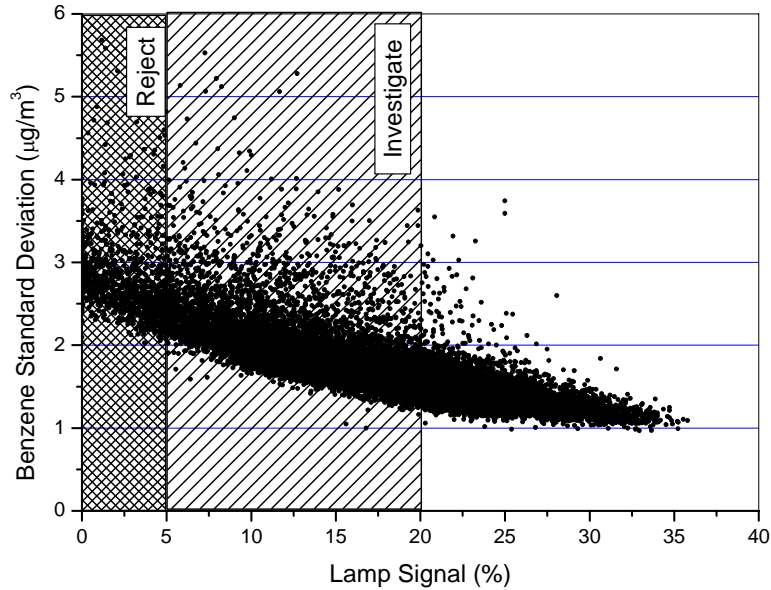
- Figure 4 Plot of the reported standard deviation for the hourly benzene values as a function of lamp intensity. At high intensities the reported standard deviation is consistently small. As the lamp signal decreases, the standard deviation of the retrieved benzene concentration increases, and the variability in the standard deviation climbs significantly. Data from 12/2002 – 5/2005.

Two specific episodes in the data have been investigated to assess the data handling procedure, chosen as there are clear gaps or concerns expressed regarding the data.

2005: From the 22 April 2005 through 1 September 2005 there is very little data (See EPA Victoria publication 1022). An instrumental fault (rather than problems with trees (suspected 21 April 2005), are identified in the operational log as of 6 June 2005). The data record for this period shows a steady loss of intensity consistent with the operational logs (see Figure 5). EPA Victoria rejected most of the data from the second half of April due to low signal intensity. The data is clearly suspect for the last part of the April before the instrument was shut down and needed to be rejected from the data record due to low light intensity.

It would seem that the updated procedures implemented in 2006 for data quality assessment should pick up these changes more rapidly.
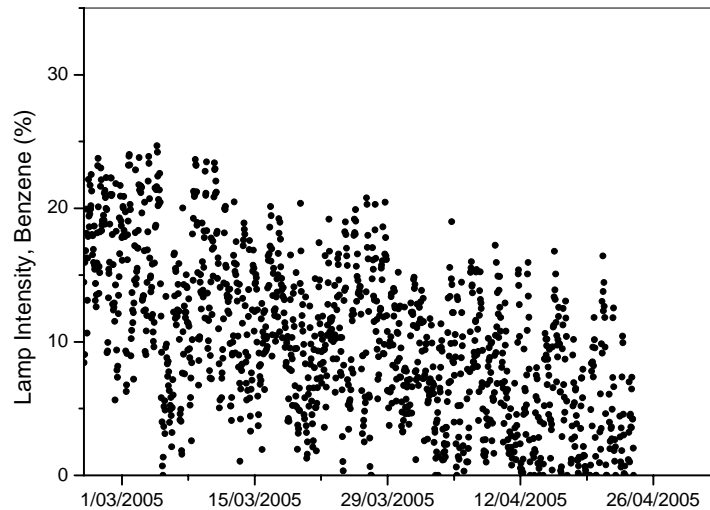
Figure 5 Lamp intensity in the first part of 2005, Benzene retrieval

2006-2007: The second period to be considered is at the end of 2006- 2007, when the in-path calibration started to decrease (see Figure 3), a feature not observed in the calibration bench test. This appears not to be due to a loss of lamp intensity (Figure 6), where the intensity is normally above 20% throughout the period.
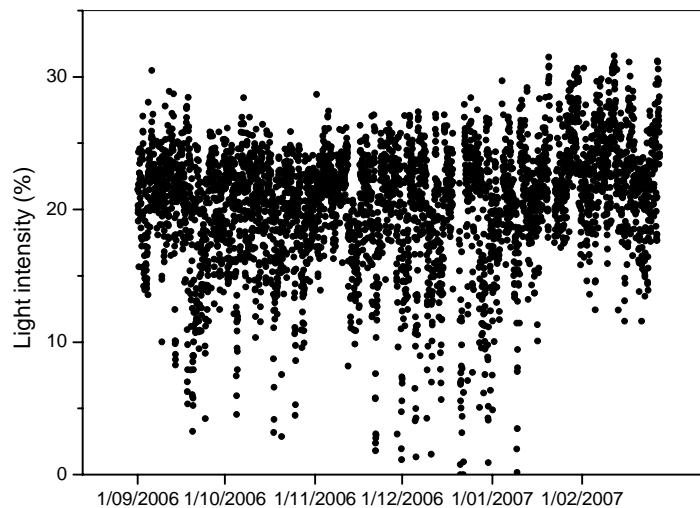
Figure 6 Lamp intensity reported for benzene retrievals.

A plot of the standard deviation versus lamp intensity for September 2006 through February 2007 is shown in Figure 7. It appears that the standard deviation follows the trend of previous data sets (Figure 4). Indeed, if anything the standard deviation appears to lie on the bottom edge of the envelope from the earlier period, suggesting the instrument is performing well. Therefore, from the experimental data presented it seems likely that the instrument is performing acceptably. It seems more likely that the deviation between the calibrations is due to a problem with the calibration itself. This is consistent with the findings of the final tests reported in the documentation presented which reports problems with the calibration cell. It can therefore be concluded that for this period the measured data are highly likely to be acceptable

and it is the calibrations that are suspect, which appears to be the conclusion reached by the Victorian EPA.
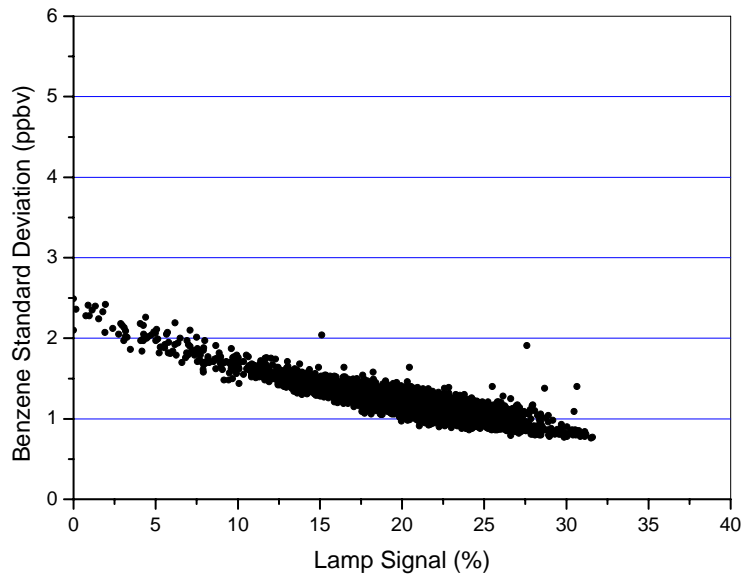
- Figure 7 Standard deviation for benzene retrieved for Sept 2006 – February 2007 inclusive.

Conclusions

The DOAS measurement technique is an internationally accepted method for long-path (fence line) measurements of a range of important atmospheric pollutants, including benzene.

The OPSIS system, as operated by EPA Victoria at Geelong, has been certified by a number of agencies internationally, and the operation of this particular instrument has been assessed and NATA certified. The calibration of the instrument is routinely checked by traceable standards. The calibration data suggests that there has been a bias in the overall calibration over the period of operation, although this bias is less than the uncertainty ascribed to the overall method.

The uncertainty determined for the method is somewhat difficult to determine, as a result of a number of factors. However, the magnitude of the uncertainty estimate at the "intervention level" for benzene is reasonable. The zero calibration is problematic with this technique, as it is not possible to ensure that there is no benzene present when the background is defined. This appears to lead to changes in the measured low concentrations. While not a problem given that the primary role of the instrument is to assess air quality with respect to the intervention level specified by the Victorian State Environmental Protection Policy, any future requirements to quantify significantly smaller concentrations may need to consider this issue in more depth, or to use other methods of determination, such as using canisters (Victorian EPA publication 999) which provide better sensitivity, but in general less spatial coverage and less time resolution.

References

Kim, K. H. (2004), Comparison of BTX measurements using a differential optical absorption spectroscopy and an on-line gas chromatography system, *Environmental Engineering Science*, *21*(2), 181-194.

Kourtidis, K., I. Ziomas, C. Zerefos, G. Achilleas, D. Balis, and P. Tzoumaka (2000), Benzene and toluene levels measured with a commercial DOAS system in Thessaloniki, Greece, *Atmos. Env.*, *34*(9), 1471-1480.

Volkamer, R., T. Etzkorn, A. Geyer, and U. Platt (1998), Correction of the oxygen interference with UV spectroscopic (DOAS) measurements of monocyclic aromatic hydrocarbons in the atmosphere, *Atmos. Env.*, *32*(21), 3731-3747.

Xie, P. H., W. Q. Liu, Q. Fu, R. B. Wang, J. G. Liu, and Q. N. Wei (2004), Intercomparison of NOx, SO_2, O_3, and aromatic hydrocarbons measured by a commercial DOAS system and traditional point monitoring techniques, *Advances in Atmospheric Sciences*, *21*(2), 211-219.

Appendix

The purpose of this section is to highlight issues that are peripheral to the brief. They are listed here as comments and suggestions, and primarily relate to the documentation process but do not affect the conclusions of the main document.

Units: Data is presented in both "ppbv" and $\mu g/m^3$. I recognize that the Victorian "State environment protection policy (Air Quality Management)" uses units related to both. I also note that in some places "ppm" is used in the documentation where ppmv is meant (See 343_08, for example). For the purposes of the documentation it would be good if a single unit system could be used, and I suspect that $\mu g/m^3$ is likely to be the most useful. I gather that the instrument reports in ppbv, but maybe OPSIS could alter this? I note also that the $\mu g/m^3$ concentrations are at 25° C and at an absolute pressure of one atmosphere (101.325 kPa) (according to the policy), although I did not see that mentioned in the documentation. This is important when looking at the error budget in particular, as the quantity derived initially by the software will be the concentration at ambient temperature and pressure, and this is then adjusted.

An example of the unit confusion lies in the file "Rawhourly 4-12-2002 - 24-05-2005.xls". The units are not given, but a comparison with the file "Validated Data.xls" indicates that the benzene values are given in $\mu g/m^3$. However, when the validation data files from 2006 are investigated, the units are clearly given in ppbv. The relationship between signal intensity and standard deviation for the Raw Hourly file listed above (see Figure 4) and the file for December 2006 (see Figure 7) would suggest that the concentration units are the same for both files. There was no way I could see of resolving this.

Bad flag data – light intensity: The published cut off for poor light intensity is 5% (See EPA Victoria publication 1022, monthly reports such as Dec 2006). In the manual EC 343_11 the cut off listed is 10%. *(This has now been dealt with; I had an old version of the document).*

Definition of intercept recorded. EC343 Section 8 page 3 states offset factors are (-intercept values), although the recorded values are always (as far as I can tell) the intercept values.

Linearity: The discussion of linearity in the Measurement Uncertainty Determination is presumably a remnant from another document, as the mention of metals is not relevant in this context. More relevant is the measured linearity using standards, which clearly justifies the use of linear regression (as does the claimed linearity range for the measurement, (343_01).

Recovery: I find the term "recovery" misleading. As far as I am aware, it usually means the amount of material "recovered" following a procedure. In this instance it is the value returned during the calibration process for one of the standards. The returned value by the instrument has an assumed calibration built into the software (scaled by "1" as far as I can tell). Therefore the "recovery" is:

$$\frac{\text{value from internal calibration}}{\text{value based on calibration standards}}$$

If this value is significantly different from 1.0 either the calibration standards are in error or the internal calibration is in error. The term "bias" that is then used to assess this quantity is testing whether the value is significantly different from 1. This is more relevant terminology, as it is using the multipoint calibration gas to assess the internal calibration.