BACKGROUND INFORMATION ON THE USE OF GC/MASS SPECTROMETRY FOR THE ANALYSIS OF ORGANIC AIR POLLUTANTS

J.M. van Niekerk & R.B. Wells
Atmospheric Sciences Division, CSIR, Pretoria

ABSTRACT

Organic pollutants are important for various reasons, such as their potential carcinogenicity or their ability to initiate photochemical smog. Modern analytical methods can isolate, identify and measure hundreds of organic compounds in our urban or industrial air. This flood of data can in itself pose problems of interpretation and reporting. Different approaches to this problem are discussed.

OPSOMMING

Organiese besoedelstowwe is van belang omdat hulle onder meer potensieel karsinogenies kan wees en tot fotochemiese rookmis aanleiding kan gee. Moderne analitiese metodes kan honderde organiese verbindings in stedelike of industriële lug isoleer, identifiseer en meet. So 'n magdom data kan wesenlike probleme by interpretasie en rapportering meebring. Verskillende benaderings tot hierdie probleem word bespreek.

INTRODUCTION

These explanatory notes were prepared for users or prospective users of the GS/MS of the Atmospheric Sciences Division. They are presented here in case they may be of wider interest to NACA members. Any questions or comments are welcome as we feel it is important to clearly explain the possibilities and limitations of such an expensive facility. Ideally, if finances allowed, we would like to strengthen the facility by adding a GC screening step, specialist sampling staff and better interpreting and reporting abilities.

WHAT DOES THE METHOD OFFER?

In essence the method can separate, identify and quantify hundreds of organic compounds present as vapours in a sample of polluted air. The method is very powerful and often gives results which could not be obtained in any other way. Nonetheless the technique is not infallible and results must be interpreted with caution. In addition costs are high and careful planning is essential.

HOW ARE SAMPLES COLLECTED?

Air is pumped through a glass tube packed with an absorbant material such as Tenax or activated charcoal. No single absorbant can be an ideal collector for all pollutant vapours and a really thorough study of a problem might require the use of several different absorbants. In spite of this reservation a single absorbant will usually yield a great deal of information about the pollutants present. Usually two tubes are connected in series in order to check for overloading by very polluted air or breakthrough of the most volatile components from the top tube to the bottom tube. Quantitative analysis would require the accurate measurement of each component in the top and bottom tubes. Moderate degrees of breakthrough can be corrected for but if similar amounts of a substance appear in the gas chromatograms from the top and bottom tubes it means that either the sampling times and rates were not suitable for the amount of pollution or that the choice of absorbant is not correct for that substance. Qualitative identification does not need to be done on both tubes and only requires GC peaks of sufficient size. Identification would most likely be carried out on the top tube, because some compounds might not penetrate to the second tube.

Before sampling even begins it is necessary for us to have detailed maps both of your chosen area of investigation and also of the surrounding area. Information on possible competing sources of organic pollution is also useful if available. In general there may be no suitable nearby source of meteorological information and our technician may also have to set up an anemometer for which a suitable site will have to be found. Practical matters such as the availability of 220V mains electricity at sampling sites, security measures, ease of access etc. will all have to be cleared up before a sampling visit.

ANALYSIS – QUALITATIVE IDENTIFICATION

Quantitative analysis is time consuming and hence more expensive and should be reserved for those compounds of special interest. Qualitative identification must therefore precede quantification of those components judged to be important enough, unless other information defines the target compounds.

Every GC peak displayed corresponds to many mass spectral scans (Fig. 1) of the separated component that emerges from the GC column into the mass spectrometer (Figs. 2). Each line in the mass spectrum corresponds to the mass of a fragment of a molecule, these fragments are produced when the sample components emerge from the GC column into the mass spectrometer source and are bombarded with electrons. The height of each line represents an intensity or court rate for fragments of that mass. Each molecule fragments differently to produce a characteristic mass spectral pattern which is used as a "fingerprint". These

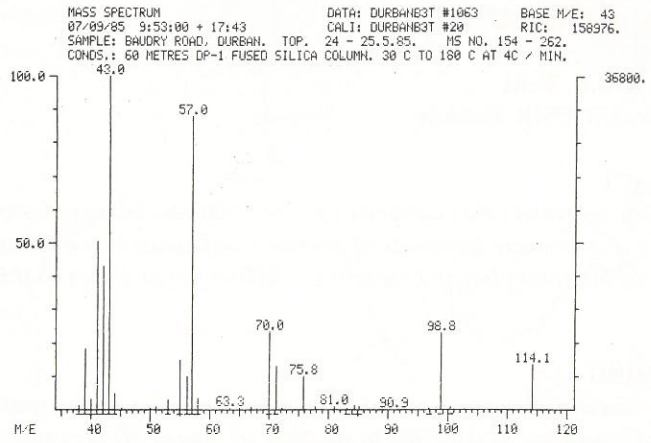
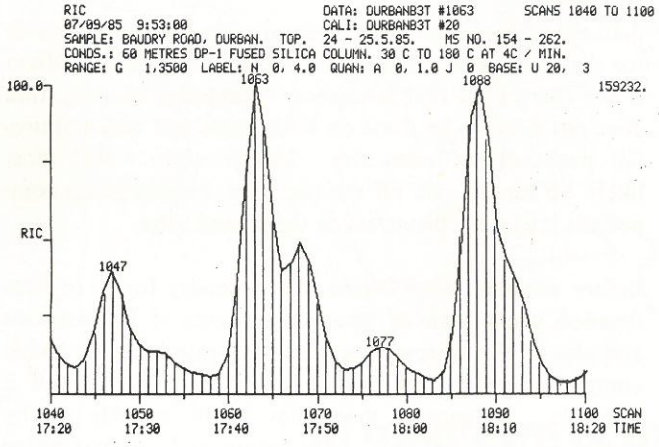


FIGURE 1



NOTE:
Each line here is the sum of all the intensities of the lines in a scan such as shown in Figure 1.
The envelope of these give the G.C. spectrum as in Figure 5 or 6.

FIGURE 2

mass spectral patterns are stored by the data system and then identified by comparison with a library of some 38 000 stored mass spectra. Without going into detail the data system offers great versatility in eliminating background, carrying out different search modes, separating overlapping GC peaks and picking out specified compounds. In the print-out called "library search" (Fig. 3) three parameters are given, namely, PUR, FIT, RFIT. In all cases 1 000 is the maximum value, anything over 850 in the cases of "FIT" and "RFIT" denotes that the compound was most probably correctly identified. Between 600 and 800 the identification is tentative and below 600 it is not reliable. A useful visual aid in comparing sample spectra with library spectra is the difference display shown in Fig. 4. Because the GC/MS instrument which recorded the library spectrum may differ slightly from ours the spectra from one compound will never totally correspond. If the FIT and RFIT values exceed 950 the chance of correct identification is very good,

LIBRARY SEARCH DATA : DURBANB3T 1063
BASE M/E : 57
07/0985 9:53: 00 + 17:43 CALI : DURBANB3T 20
RIC : 124671
SAMPLE : BAUDRY ROAD, DURBAN TOP. 24–25.5.85
MS NO. 154 – 262
CONDS : 60 METRES DP–1 FUSED SILICA COLUMN. 30 C TO 180 C AT 4C / MIN.

RANK IN NAME

1. 2085 Hexane, 2,5-dimethyl
2. 2086 Heptane, 2-methyl-
3. 3340 Hexane, 3-ethyl-4-methyl
4. 4854 Oc-ane, 2,7-dimethyl-
5. 10042 Heptane, 3-bromo-

RANK	FORMULA	M.WT	B.PK	Purity	FIT	RFIT
1.	C8. H18	114	43	930	954	938
2.	C8. H18	114	43	908	935	949
3.	C9. H20	128	43	795	877	844
4.	C10. H22	142	43	790	87	848
5.	C7 H15.BR	178	57	786	864	827

This printout ranks the 5 closest fitting spectra in the library for one mass spectral pattern as in Figure 1.

FIGURE 3

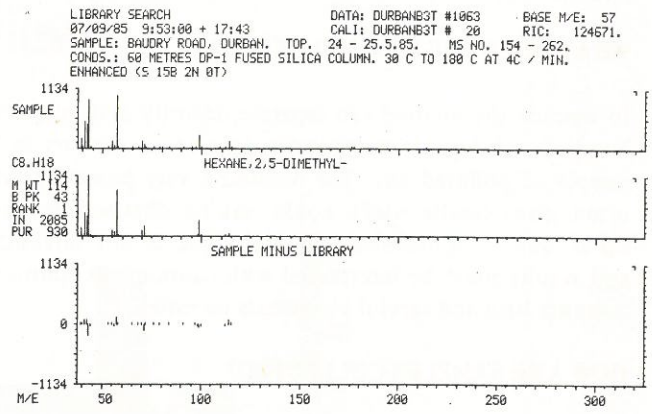


FIGURE 4

however, if the mass spectra of compounds are very similar as in the case of aliphatic hydrocarbons the FIT may be very good but the correct isomer may not have been found. Only by running a calibration to obtain the mass spectra and retention times of pure standards can one approach full certainty of identification. The "bought" library of spectra might not include some compounds important to you, but this can be rectified by creating your own supplementary library by analysing standards. The drawbacks to this can be the high cost and the long delay in obtaining pure commercial standards.

AFTER IDENTIFICATION, THEN WHAT?

The primary reason for identifying and later measuring organic compounds in the air is their toxicity to human or other life. The Atmospheric Sciences Division has available toxicity data from various sources such as the American Conference of Governmental Industrial Hygienists and the US National Institute for Occupational Safety and Health. In addition overseas data banks are accessible via a computer link up. Although some comments on toxicity may be offered with GC/MS reports it would probably be best for you to informally discuss the report findings with us before embarking on any large scale search for toxicity data.

Other reasons for wanting quantitative measurements could be that a substance may be a precursor of photochemical smog or that it indicates a malfunction of an industrial plant.

QUANTITATIVE MEASUREMENT

As stated above this will require analysis of known pure standards. It will also mean the use of internal standards and a high ratio of calibration checks to analytical GC runs. Because a sample may take between 1 and 4 hours to finish emerging from the GC column only a few samples per day can be analysed. Even though the pollutants of interest may have been measured in the first hour of the gas chromatogram run the next sample cannot be injected until the components of the previous run are clear of the column. This time will be unknown until the qualitative study is done and correct costing of a quantitative study cannot be done until it is known. In compensation the information content of each analytical run may be very high. The cost of running and manning the GC/MS is currently R120/hour. Clearly quantitative sampling and analysis needs greater than usual care in choosing a sampling strategy. It is important to remember that exposed sampler tubes should not be stored longer than 3 weeks before analysis to prevent losses or the formation of artifacts. In spite of this caveat it is often useful to collect perhaps 2 or 3 times the number of samples that you finally intend to analyse, especially if long distances must be travelled to obtain the sample or if the event observed is infrequent. It is unfortunate that the high cost of GC/MS work precludes the analysis of large number of samples covering a long time. This is because general experience of air pollution studies has shown that short term sampling may give information about non-typical periods and lead to erroneous conclusions. If measurements are for example planned to determine organic pollution near a given factory, the factory should be in the desired mode of operation and weather conditions should also be defined. There must be no chance that a factory will close for routine maintenance just when the technician arrives to collect samples. Good liaison is therefore essential.

DATA PRESENTATION – QUALITATIVE IDENTIFICATION

The gas chromatogram reconstructed by the data system from all the mass spectral scans (Fig. 5) will be given with the identified peaks being numbered (Fig. 6) and tabulated as in Table 1. The peaks not identified may either fall below some predetermined peak height or area or they may on investigation fail to yield a meaningful mass spectrum after the background signal has been subtracted. The numbered peaks will be listed together with their identification. If a peak is of special interest the mass spectra of the sample and of the closest fitting compounds in the library may be given. It is also sometimes informative to display the difference between the sample mass spectrum and the closest spectra in the library. For the sake of brevity in reporting this will not be done for all peaks unless specially requested. In cases where a peak could result from column bleed or unavoidable slow absorbant decomposition this will be noted. Users should not automatically dismiss such peaks but should consider carefully if they might nonetheless be genuine components of their samples. The list of closest fit identifications may sometimes show duplicates, this indicates the presence of very similar if not isomeric compounds.

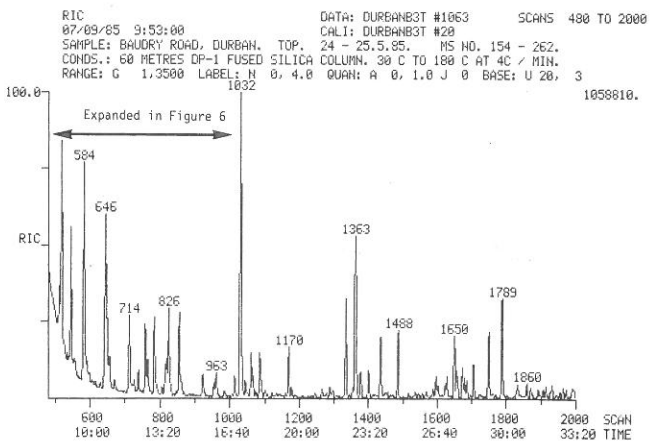


FIGURE 5

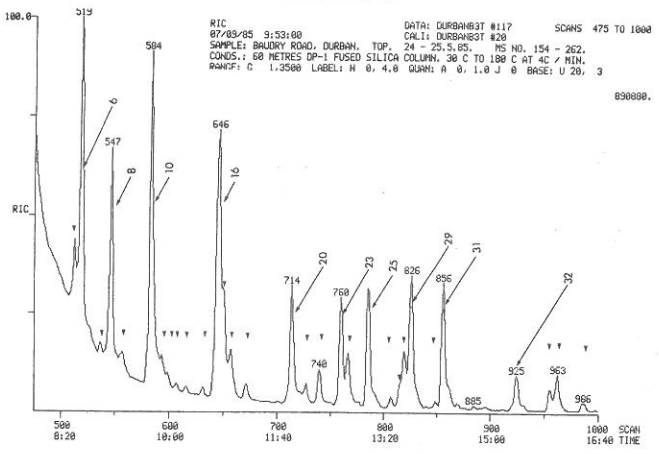


FIGURE 6

TABLE 1

DATA : DURBANB3T CALI : DURBANB3T
07/09/85 9 : 53 : 54
SAMPLE : BAUDRY ROAD, DURBAN. TOP.
24 – 25.5.85. MS. NO. 154–262

NO.	SCAN NO.	COMPOUND NAME	FIT
1	206	Air	
2	222	Formicacid, methylester	850
3	340	Butane, 2-methyl-	978
4	379	Pentane	956
5	513	Butane, 2,3-Dimethyl-	949
6	519	Pentane, 2-methyl-	972
7	537	2-Butanone	897
8	547	Pentane, 3-methyl-	978
9	558	1-Hexene	953
10	584	Hexane	987
11	594	2-Hexene-(E)-	915
12	599	1 Butene, 2,3-dimethyl-	914
13	607	Not identified	
14	616	2-Hexene-(z)-	945
15	631	2-Pentene, 3-methyl-,(z)-	955
16	646	Octafluorotoluene Internal Standard	1000
17	650	Cyclohexane	924
18	658	Pentane, 2,4-Dimethyl-	966
19	672	Ethane, 1,1-dichloro-1-nitro-	950
20	714	Benzene	1000
21	727	Methane, Tetrachloride	856
22	740	1-Pentene, 2-Methyl-	979
23	760	Hexane, 2-Methyl-	986
24	767	Pentane, 2,3-dimethyl-	955
25	785	Hexane, 3-methyl-	991
26	807	Cyclopentane, 1,3-dimethyl-, Trans	983
27	815	1-Hexanol, 5-methyl-	893
28	819	Ethene, trichloro-	961
29	826	Pentane, 2,2,4-Trimethyl-	979
30	849	2-Hexene, 2-Methyl-	949

DATA STORAGE

When your samples are analysed the initial data will be stored on disks which will be copied onto other disks at intervals to provide a backup. The Division is at present unable to store data permanently on magnetic tape so this data will ultimately have to be wiped off the disks. It is therefore important that users study the reports submitted to them and raise any queries as soon as possible. For example a compound reported by us may have a significance to you not apparent to us at the time of analysis and you may want detailed information on that GC peak.

FEEDBACK

In order for us to make the best use of the expensive GC/MS facility and in order to improve our reporting of the data your comments on our reports are essential. Organisations differ greatly in the availability of expertise within them to interpret reports. It will be useful to both parties if our analysts know from the beginning of the project who will have to evaluate the results.

COST FACTORS

These factors have been mentioned under various headings but are listed fully here. Actual rates are not given because they change periodically, however, certain factors are easily overlooked.

Factors	Comments
Initial visit to site, transport and consultation time, study of feasibility.	Should not be omitted, will reduce ultimate costs and avoid misunderstandings.
Purchase of standards, chemicals new GC-columns.	May not be needed, but takes time if needed.
Sampling visit: transport, technician and equipment. Accommodation and S & T costs, overtime.	Overtime may be needed for 24-hour sampling.
GC/MS analysis time. NB this includes standard runs as well as analytical runs.	Present cost + R120/hour (1985)
Data interpretation.	Unpredictable, time depends on sample complexity and client's needs.
Data presentation, report writing, checking, possible toxicity, literature studies.	