R&S 041349

The Application of Computer Science to Industrial Hygiene

J. E. PETERSON, H. R. HOYLE, and E. J. SCHNEIDER

The Biochemical Research Laboratory, The Dow Chemical Company, Midland, Michigan

Automatic sampling and analysis of environmental atmospheres can result in voluminous amounts of data describing exposures to chemicals. The use of a digital computer to process such data and the advantages and disadvantages of the technique are presented.

Introduction

AUTOMATIC AIR monitoring equipment is expensive. Nevertheless, if it is used to signal the need for action by operating personnel to eliminate leaks and to effect needed repairs of equipment, its cost can easily be justified in terms of reduced exposures, reduced hazards, and reduced loss of process material. Furthermore, if sample locations are properly chosen, it provides data on the concentrations of air contaminants to which men are exposed. That information, handled properly, can be used to show long-term trends of exposure which, correlated with plant operations, can result in still lower exposures. In addition, such data can be used to determine the exposures of workmen in what is, in effect, a continuing industrial hygiene survey. These data, in turn, could be correlated with medical information on effects of the exposures to confirm or deny present standards or to suggest new ones.

At The Dow Chemical Company, automatic air sampling and analysis were first used in 1950 to monitor air concentrations of carbon tetrachloride in a production plant. At that time no method was available for handling the mass of data generated by the air monitoring equipment. Even though the instrument was successfully employed by plant supervision to control exposures, full use of the data to estimate the average concentration to which men were exposed was not feasible. To illustrate: The instrument recorded air concentrations at the rate of one every six seconds; this is 432,000 times per month.

Because the first continuous air monitor was a success at day-to-day control, others were installed in several plants during the 1950's. As each monitor was put into use, even more information on workmen's exposures was "going to waste" simply because there was too much of it. The advent of digital computers suggested a solution. Taking advantage of the fact that the Computations Research Laboratory had a Burroughs B220 computer, the Environmental Research Laboratory in 1961 acquired a machine for translating air concentration data from a recorder to punched paper tape. Continuous air concentration data could then be combined with computer analysis to yield a much more complete description of inhalation exposures than had heretofore been possible.

This paper recounts some of the problems involved, from the installation of air monitoring equipment to the interpretation of the data as summarized by the computer. Actual data are used, but only for illustrative purposes, so plant, process, and even the air contaminant are immaterial. This is not a report of an environmental survey; it is a primer in the application of computer science to industrial hygiene.

Presented at the Annual Meeting of the American Conference of Governmental Industrial Hygienists, Houston, Texas, May 1965.

Obtaining Air Samples

An automatic air monitoring system must be sensitive enough to detect concentrations of significance to health. Sensitivity must range from concentrations that may be immediately hazardous to those that have little or no significance even for prolonged, repeated exposures. The instrument should be selective; it must respond only to the material(s) of interest. It must be stable; it should be unaffected by minor (or even major) variations of temperature, humidity, vibration, and line voltage. The instrument must be capable of operating unattended for relatively long periods of time. Daily attention by the operator, minor service weekly by an instrument man, and occasional performance testing should be adequate to keep it in operation. Finally, because a permanent record of air concentrations can be important, concentration data must be automatically recorded. Any continuous analyzer that satisfies these specifications can be used. So far we have used only combustion-conductivity instruments and long-path gas cell infrared spectrometers; both types produce data suitable for computer analysis.

Single-point monitoring in a production plant is seldom economical. Usually the analyzer is positioned at a central location in the plant, and air samples are brought continuously to it through probes of some sort. Tubing used for these probes must not react with, or sorb, the air contaminant of interest.

Preparing Data for Analysis

There is no method presently available for reading results automatically from multiple-point chart paper into a computer. This means that a device must be used to "digitize" the data as they are obtained. At the minimum, the digitizer must translate the analog data (usually the position of a wiper on the slide wire of a potentiometer) to digital form (for example, numbers) that can be handled by the computer.

Because data on the time of exposure can minimize programing difficulties, the digitizer includes a digital clock with an output of the day number, hour, and minute. A switch selects a "type of data" digit (from 0 to 9) which is incorporated into the digitizer output. An extra "type of data" (for example, from an infrared spectrometer or from a combustion-conductivity analyzer) digit can allow the computer to reject data that obviously do not belong with that being processed. To minimize costs, punched paper tape is used to transfer data from the digitizer to the computer.

For each datum the digitizer punches two "words" on the tape. The first "word" contains the day number, hour, and minute, as well as the "type of data" identification number. The second "word" contains the probe number and the actual datum which, in this case, is a number from 0 to 999, proportional to the concentration of the air contaminant. This information is obtained at a rate that may vary from two times a minute to once in 2.5 minutes. Because collecting each datum generated is not always necessary, the digitizer can be programed to skip the collection of some data. In every case, once data from probe No. 1 are punched onto the tape, data from the other probes are obtained in serial order, but the digitizer can ignore all but every second, fifth, or tenth sets of analyses.

Programing the Computer

We decided that the minimum time in the data summary should be an eight-hour shift. For this period the Dow Computations Research Laboratory programed the computer to calculate the mean concentration at each location, the standard deviation of these data, the percentage of time that the concentration was above several preselected levels, and the appropriate time-weighted averages.

Computers are versatile. For instance, if one probe is located to sample air outside the building, the computer can be programed to correct automatically for deviations in the "background." It can take into account peculiarities in the calibration curve of the instrument used for air analysis and can recognize several kinds of errors in the data and use only "good" data for calculations.

Computer Output

The basic computer output consists of the mean concentration, standard deviation, number of analyses recorded, and percentage of time the concentration was above preselected levels, all at each location for each shift during the survey. Shift or daily averages can also be obtained over any selected time interval such as a week or a month. These data, further identified as to department, air contaminant, etc., are permanently stored on magnetic tape.

The computer has been programed to reject poor data such as that caused by a sticking punch or by a faulty encoder. If poor data are being obtained, this fact is often signaled first by an unexplained decrease in the number of analyses used by the computer to obtain the shift averages. This has become the normal signal for nonroutine maintenance of the digitizing equipment.

In addition to the basic output, the computer calculates time-weighted averages of two kinds. The first kind is the "usual" time-weighted average concentration to which men are exposed. It is obtained by combining the air analysis data with "job analysis" information on the percentage of time spent by men in the vicinity of specific analyzer probes. Job analysis information is also combined with the data on the percentage of time during which concentrations exceed the preselected levels. This results in a second kind of average which is the "time-weighted percentage of time" exposures exceeded the preselected levels.

An example of how the time-weighted percentage of time is calculated is given below. Assume that the following information is true for an operator in a plant:

Data Location Number	Percentage of Time Spent at That Location
1	12.5
7	31.3
4	18.8
12	25.0
Unexposed	12.4
	100.0

Furthermore, assume that, over the time period of interest, concentrations at these data locations behaved in this manner:

% of Time Concentration Exceeds Selected Values at Specified Locations

Concentration greater than	Data Location Number 1	7	4	12
25 ppm	26.63%	28.68%	93.44%	31.55%
50	16.39	22.13	20.08	23.77
100	6.55	10.65	17.62	11.47
250	1.63	2.04	16.39	3.68
500	0.00	0.00	0.00	0.00

The time-weighted percentage of time spent by operators in concentrations above 25 ppm will then be:

$$\frac{(12.5)\ (26.63) + (31.3)\ (28.68) + (18.8)\ (93.44) + (25.0)\ (31.55) + (12.4)\ (0)}{100} = 38.92$$

This means that *on the average,* during the time period of interest, men in the operators classification encountered concentrations above 25 ppm 38.92% of the time.

Similar calculations show that these men encountered concentrations above 50 ppm 28.24% of the time; above 100 ppm 10.95% of the time; above 250 ppm 4.00% of the time; and above 500 ppm 0.00% of the time.

Data Analysis

Automatic air monitoring reveals that in an industrial situation the variation of concentration with time can be large despite the use of a rather large time base. Figure 1 is a plot of daily time-weighted average concentrations to which operators were exposed on each shift over one week. Each point on the graph is the mean of several hundred determinations spaced equally over an eight-hour period. These are actual plant data.

On the second shift there was little variation during the first four days, but over the last three days of this period the shift average varied by a factor of almost four. If the threshold limit value (TLV) for this material were 50 ppm, an industrial hygienist taking air samples during the first three days on the second shift would probably have declared that the hazard to health was low or nonexistent. On the other hand, if he had

ısume that, over the time
, concentrations at these
ıved in this manner:

ntr Exceeds Selected
ʾpecified Locations

	Data Location Number		
	7	4	12
⁶	28.68%	93.44%	31.55%
	22.13	20.08	23.77
	10.65	17.62	11.47
	2.04	16.39	3.68
	0.00	0.00	0.00

ɪ percentage of time spent
ʿentrations above 25 ppm

) (0)
—— = 38.92

ı *the average,* during the
ʿest, men in the opera-
ːountered concentrations
:% of the time.

ːs show that these men
rations above 50 ppm
above 100 ppm 10.95%
ʾ50 ppm 4.00% of the
ɔpm 0.00% of the time.

itoring reveals that in
ı the variation of con-
an be large despite the
me base. Figure 1 is a
ȝhted average concen-
ators were exposed on
ek. Each point on the
ɜeveral hundred deter-
lly over an eight-hour
ıal plant data.

there was little varia-
ur days, but over the
ɜriod the shift average
almost four. If the
TLV) for this mate-
dustrial hygienist tak-
he first three days on
ǀ probably have de-
to health was low or
her hand, if he had

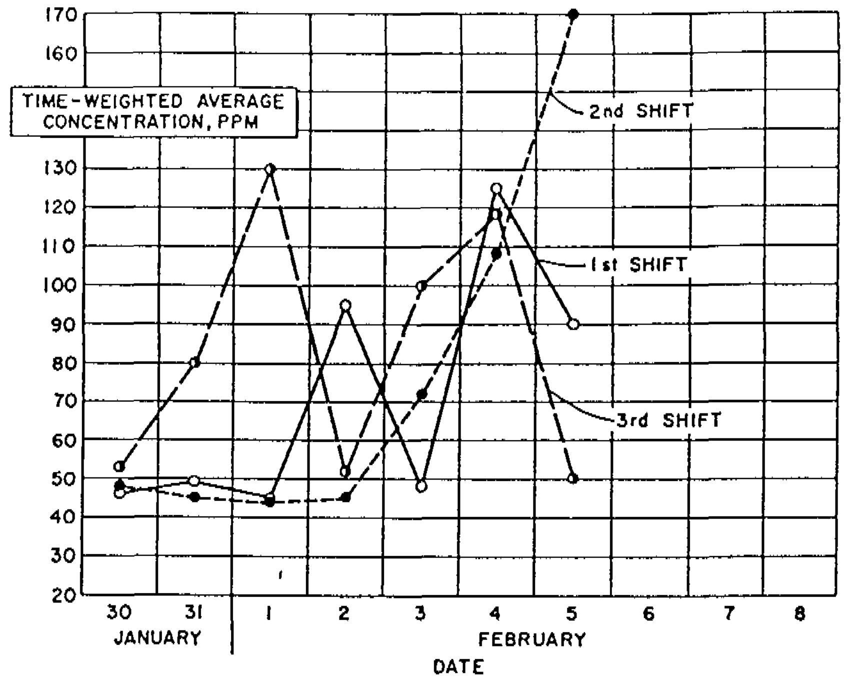

FIGURE 1. Daily time-weighted average exposures for each shift from January 30 to February 5. Time-weighted average for the week (all shifts) was 82.32 ppm.

sampled on the last day of this period (on the second shift) he could have experienced a strong inclination to "push the panic button."

Variation of concentration between shifts can also be striking, as illustrated by a comparison of the third shift with the first or second shifts during the first three days.

Figure 2 places Figure 1 in a broader context. Data for Figure 2 are also time-weighted average concentrations to which operators were exposed. The rather extreme variation shown in Figure 1 for a time basis of one shift can be repeated for a time basis of one week. The information contained in Figure 2 is based upon hundreds of thousands of individual air samples.

Average concentrations and time-weighted average concentrations are not the only data summaries provided by the computer; average concentrations are accompanied by the standard deviation of the data. Having the standard deviation and using it quantitatively are two different things, however. We do not yet know how to use such information efficiently.

On the other hand, the spread or variability of the data is indicated in a more meaningful manner by the percentage of time the concentration exceeded certain levels. Figure 3 is a plot of the time-weighted percentage of time above these levels on log-probability paper. The interval during which data were gathered and the operational classification are identical to those in Figure 2. This graph shows that the median concentration to which men on all shifts in this classification were exposed was 40 ppm and that they were exposed to 500 ppm or higher 2% of the time.

The conventional plot (Figure 2) shows a maximum concentration of about 135 ppm, but it is one of seven-day averages, whereas Figure 3 is a summation of instantaneous values. Both kinds of graph have advantages. A conventional plot illustrates better how exposures vary with time, and any trends become readily apparent. With this kind of graph, however, the only usable index of exposure is the time-weighted average, a number of limited utility because it cannot reflect

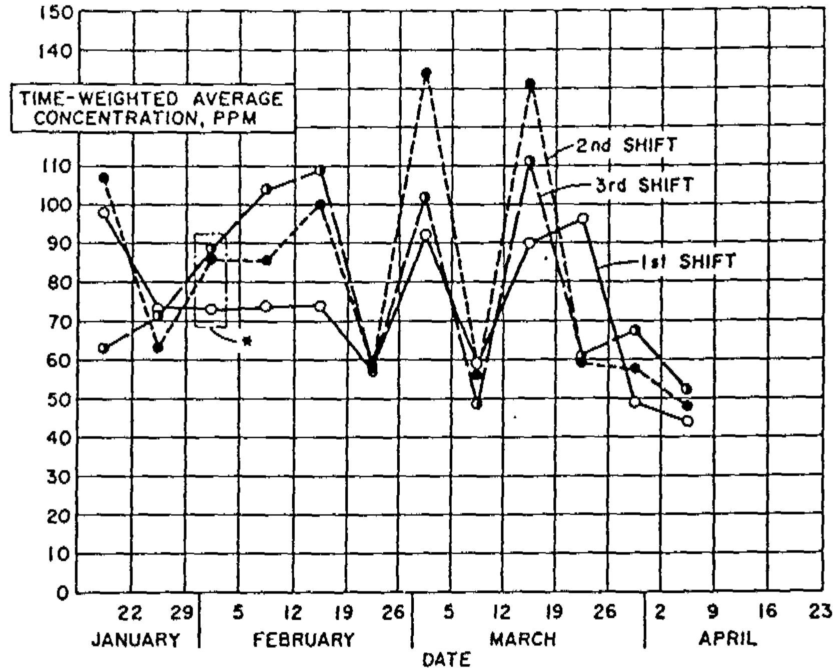

FIGURE 2. Weekly time-weighted average exposures for each shift from January 16 to April 10. Time-weighted average for the whole period (all shifts) was 79.09 ppm. *These points are expanded in Figure 1.

concentration variations.

A log-probability plot such as that shown in Figure 3 is not very useful as a trend indicator, but the fact that time-weighted *exposure* data plot as a straight line on this graph paper offers exciting possibilities. In our limited experience all time-weighted percentage of time data have been best fitted by a straight line on log-probability paper. That straight line is a complete summary of exposure information, including measures both of the average concentration and of the concentration variability.[3]

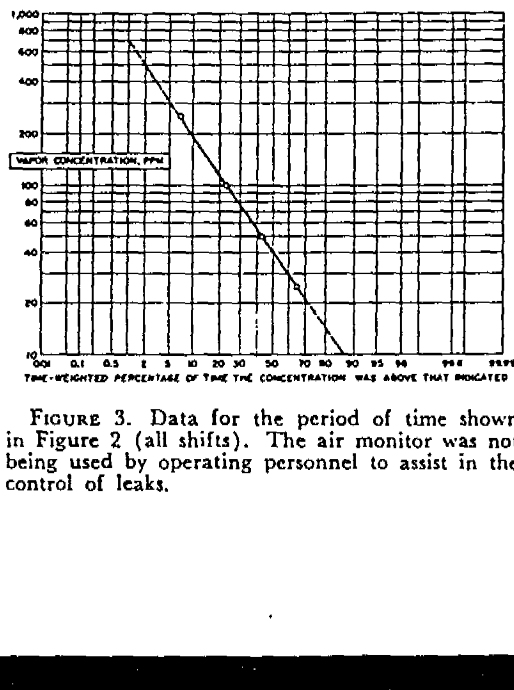

FIGURE 3. Data for the period of time shown in Figure 2 (all shifts). The air monitor was not being used by operating personnel to assist in the control of leaks.

The data in Figure 3 were obtained by a survey using the equipment described. How are these data to be interpreted? If we still assume a TLV of 50 ppm, is the exposure represented by Figure 3 a hazardous one or not? These data are among the most complete ever gathered for men industrially exposed to a potentially hazardous vapor; what kind of animal experiments should be conducted to interpret these exposures? We cannot yet answer these questions, but we know that they must be answered eventually.

Despite this void in our knowledge, the data from this type of survey have several uses. Any information about the intensity of exposures helps to interpret medical findings. Even without complete interpretation, ranking of exposures is possible. From graphs such as Figure 3 one can judge if the exposure is obviously hazardous or nonhazardous. For instance, in Figure 4 the concentrations to which the men were exposed are shown to exceed 15 ppm 50% of the time. They exceed the assumed TLV of 50 ppm only about 14% of the time. Without any further knowledge, the exposure represented by this curve would probably be judged nonhazardous by most industrial hygienists, provided that contacts with high concentrations were more or less randomly distributed, of short duration, and of little or no consequence in themselves.

The data in Figure 4 were obtained in the same plant, for the same group of men, and by the analyzer using the same sample points as for Figure 3. During the period of time when the information in Figure 2 was obtained, the air analyzer was being virtually ignored by operating personnel in the plant. On the other hand, the data in Figure 4 were obtained when the concentrations recorded by the air analyzer were used to prompt corrective action. Even plants with a large number of leaks and spills can be brought from a potentially hazardous situation to one that is without hazard by using the automatic air analyzer as an operational tool. This is the way an automatic air monitor should be used; it will not justify its expense unless it is so used.

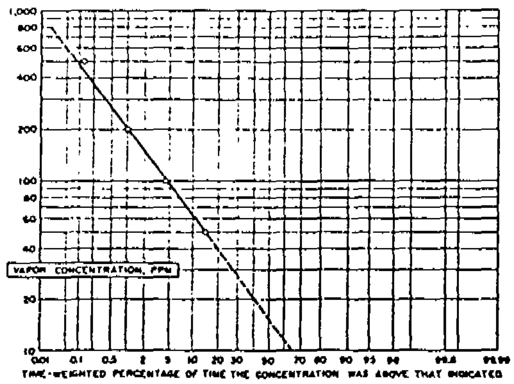

FIGURE 4. Exposures for the same people as shown in Figures 2 and 3, but for the period January 5 to June 18 when the air monitor was used to signal the need for maintenance. The time-weighted average concentration for this period was 28.83 ppm.

Advantages of the Technique

The advantages of automatic air analysis coupled with a computer analysis of the data are many. First, air monitoring gives a true picture of the manner in which air concentrations of the contaminant vary with time. Coupled with a good analysis of the percentage of time men spend in the various locations, this technique can give the best estimate presently obtainable of the actual exposures encountered. Second, this technique enables a correlation of air concentrations with varying plant activities. This type of correlation can point out "trouble spots" quite readily and, when used properly, can lead to dramatic reduction of concentrations to which men are exposed and a corresponding reduction in materials lost. Third, automatic air monitoring is the only technique so far devised that truly shows the air concentration variation between shifts. Too often, air sampling is done only on the day shift, and the afternoon and midnight shifts are either ignored or assumed to be similar to the day shift. Finally, the records obtained can be stored on magnetic tape for indefinite periods of time in a minimum of storage space. Access to records on magnetic tape is simple and easy. These are important considerations, because proof of past accomplishments can be of considerable value.

Conclusions

Use of automatic air sampling and analysis equipment, coupled with a computer analysis of the data so obtained, can be an extremely potent weapon in the fight of industrial hygienists against inhalation hazards. Mere installation of an air monitor, however, does not end the battle. Data obtained by the monitor must be used by operating personnel in much the same way that monitors of other operating variables are used, namely, to assure control. Computer analysis of the data obtained by an air monitor can give the best estimate presently obtainable of the concentrations of air contaminants to which men are exposed. In fact, the environmental analysis is so thorough that it poses a challenge to toxicologists to devise animal experiments that will provide a better basis for judgment than does the present threshold limit value.

References

1. LITCHFIELD, J. T., JR., and F. WILCOXON: A Simplified Method of Evaluating Dose-Effect Experiments. *J. Pharmacol. Exptl. Therap. 96:* 99 (1949).