Med. Lavoro, 4, 412-419, 1982

Clinica del Lavoro "L. Devoto"
La medicina del lavoro
Via San Barnaba 8, Milano, Italy

A STATISTICAL EVALUATION OF THE POLLUTION LEVEL IN WORK ENVIRONMENTS

A. Borroni [1,2], G.A. Gino [1], B. Mazza [1,2] and G. Nano [1,2]

« Valutazione statistica del livello di inquinamento degli ambienti di lavoro ». - La caratterizzazione degli ambienti industriali dal punto di vista degli inquinanti presenti pone alcuni problemi di corrispondenza tra valori reali e valori misurati, essendo necessariamente limitato nello spazio e nel tempo il numero dei prelievi che è possibile effettuare. Da questa incertezza deriva la difficoltà di stabilire con sicurezza, una volta fissati per i vari inquinanti degli standard di accettabilità, ad es. i TLV, se i valori reali siano sempre e comunque inferiori agli standard. La trattazione statistica di un limitato numero di campioni è in grado di fornire, con un livello di fiducia predeterminato, indicazioni coerenti sia sulla concentrazione reale dell'inquinante che sulla sua pericolosità, in relazione al TLV considerato, per la salute dei lavoratori. Nello studio qui riportato di un ambiente di lavoro caratterizzato dall'emissione di polveri, si sono misurate le concentrazioni dell'inquinante distribuite sia spazialmente che temporalmente; le misure sono state ripetute con le stesse modalità a distanza di quattro mesi. I risultati mostrano che l'ambiente di lavoro in esame è caratterizzato da una deviazione geometrica standard delle concentrazioni (indice di dispersione intorno alla mediana nella distribuzione lognormale considerata) pressoché costante e che è possibile dare con un buon livello di fiducia un giudizio sulla pericolosità o meno di tale ambiente per la salute dei lavoratori.

INTRODUCTION

The measurement of contaminants in a work environment may be conducted with any one of the following objectives: 1) test and inspection of industrial plants and equipment; 2) description of the work environment; 3) evaluation of the degree of employees' exposure. Consequently, before undertaking any sampling program, the proposed objective must be clearly defined, in order to select the most appropriate instruments and techniques.

Fixed temporal samplings are used in testing machines, and fixed temporal and spatial samplings are used in evaluating work environments. To evaluate the degree of exposure of a single worker, we use personal samplings which average in time and space the worker's length of exposure in various areas *.

In every case, we must consider the problem of discrepancies between real and measured situations, since random and systematic errors are always committed and the

[1] Istituto di Chimica-fisica, Elettrochimica e Metallurgia del Politecnico di Milano - Piazza Leonardo da Vinci, 32 - 20133 Milano.

[2] Gruppo per gli Studi sui Sistemi di Produzione e sul Lavoro del Politecnico di Milano - Piazza Leonardo da Vinci, 32 - 20133 Milano.

Financial support to this work by Comune di Sesto S. Giovanni is acknowledged.

* The results obtained with this technique cannot be the basis for conclusions regarding the entire work environment, but rather only the exposure of the individual worker under study.

contaminant in a work environment is not characterized by homogeneous temporal and spatial distribution. It is therefore useful to assign to the measured data a certain degree of variability which defines the interval in which the real concentrations can be found with a given probability. Statistical analysis of the data can satisfactorily meet this requirement.

Statistical Analysis of the Data

In order to reduce the existing disparity between the real environment and our representation of it, the number of samplings can be increased; however, the improvement in accuracy obtainable by each additional sampling decreases as the total number of samplings increases. It is therefore important to determine the appropriate number of samplings that will allow us to identify with a good degree of accuracy the characteristics of the environment being studied. It can be said that accuracy decreases rapidly if less than three samples are taken, while it slowly improves if more than eight samples are used.

In order to introduce metters, let us assume that we are referring to an environment in which a contaminant exists in a constant and uniform concentration: repeated samplings would give different concentration values X_i, with deviations $X_i - \bar{X}$ distributed symmetrically (normal distribution) around the mean value $\bar{X}$ *, which is an estimate of the real concentration C.

*
$$\bar{X} = \frac{\sum_{i=1}^{N} X_i \; t_i}{\sum_{i=1}^{N} t_i}$$

where t_i is the duration of the i^{th} sampling and N the number of samplings.

The dispersion of data can be characterized by the standard deviation S:

$$S = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N-1}}$$

The accuracy of the estimate depends upon the pre-established confidence level that one has decided to work at. In general, a confidence level $(1 - \alpha) = 0.95$ (or $\alpha = 0.05$) is used. Once the confidence level has been established, the real concentration C will be found with $(1 - \alpha)$ probability in an interval symmetrical to the average value, according to the formula:

$$\bar{X} \pm t_{(1-\frac{\alpha}{2},\, N-1)} \cdot \frac{S}{\sqrt{N}}$$

where $t_{(1-\frac{\alpha}{2},\, N-1)}$ is a function of $(1 - \frac{\alpha}{2})$ and of $(N - 1)$ which can be determined from probability tables (Student's t-distribution). Such an interval is called the two-sided confidence interval, because it includes the upper and lower limits that bound the value of the real concentration C.

When the number of samples is kept constant, the confidence interval widens as the confidence level increases; on the other hand, when the confidence level is kept constant, the confidence interval is reduced as the number of samples is increased.

If we want to compare the measured value with a pre-established standard (for example, TLV), it is not sufficient to compare TLV with $\bar{X}$, which is only an estimate of C. Instead, once the confidence level has been established, we must determine if C is above or below the TLV. In this case, we are interested in an one-sided confidence limit for the real concentration C, since only the lower limit or the upper limit is pertinent.

If, for example, $\bar{X}$ is less than the TLV, and we want to be $(1 - \alpha)$ confident that C is also less than the TLV, we must verify that:

$$\bar{X} + t_{(1-\alpha,\, N-1)} \cdot \frac{S}{\sqrt{N}} < TLV$$

On the other hand, if $\bar{X}$ is greater than the TLV and we want to be $(1 - \alpha)$ confident that C is greater as well, we must verify that:

$$\bar{X} - t_{(1-\alpha,\ N-1)} \cdot \frac{S}{\sqrt{N}} > TLV$$

Even with $\bar{X} < TLV$, if

$$\bar{X} + t_{(1-\alpha,\ N-1)} \cdot \frac{S}{\sqrt{N}} > TLV$$

we cannot assert with $(1 - \alpha)$ confidence that C is less than the TLV, i.e., we are in a condition of uncertainty or nondecision. In this case, we can choose to increase the number of samplings or reduce the confidence level (however, this is not advisable when dealing with particularly dangerous contaminants).

In real work environments, the concentration of contaminants is not constant, nor is it uniformly distributed; rather, it is variable according to time and space but with values that can only be $\geqslant 0$. In addition, errors in determining the concentration of contaminants increase as the concentration level itself decreases. This is reflected in an asymmetrical distribution of the deviations X_i-$\bar{X}$ around the average $\bar{X}$. A lognormal distribution is the type which best describes the distribution of contaminants in a work environment.

In order to simplify our procedure, the data for concentrations X_i taken from repeated samplings can be plotted on lognormal probability paper against the respective cumulative percentage (which is obtained by dividing the number of samplings accumulated up until the desired concentration by the total number of samplings plus one *). In this way, the data are approximatively arranged on an interpolative straight line ** from which we can obtain graphically a central value and a dispersion index which correspond respectively to the median X_m and the geometric standard deviation GSD, both of which effectively characterize the series of of samplings taken. The median is defined as the concentration corresponding to a cumulative percentage of 50%. The geometric standard deviation is defined as the ratio between the concentration corresponding to a cumulative percentage of 84% and the concentration corresponding to a cumulative percentage of 50%.

Once these quantities have been obtained, it is possible to statistically evaluate the samplings as previously illustrated for the normal distribution, by substituting:

$\ln X_m$ for $\bar{X}$
ln GSD for S
ln TLV for TLV

Application

The described method was applied in the study of work environment of a plate rolling mill fed with thick slabs and ingots. Figure 1 shows the layout of the department and the location of the sampling positions. In point A, eight successive samples were taken, each for a duration of 30 minutes; a one-hour sample was simultaneously taken in points A, B, C, D. Ferric oxide dust was the contaminant being measured. All of the samplings were repeated four months later in the same points and with the same methods.

Table 1 shows the results of the two series of temporal samplings taken in point A; the results are in order of increasing concentration levels and are shown in relation to the respective cumulative percentages. In figure 2 the same data have been plotted on lognormal probability paper, revealing that for the first series of samplings: $X_m = 2.67$ ($\ln X_m = 0.98$) and GSD = 1.68 (ln GSD = 0.52); and for the second series: $X_m = 2.00$ ($\ln X_m = 0.69$) and GSD = 1.50 (ln GSD = 0.41).

* The various samplings should be classified in order of increasing concentration. The cumulative percentage of the m^{th} datum in N is : $m/N + 1 \cdot 100$.

** Drawn by eye.

Table 2 shows the results (once again in relation to the respective cumulative percentages) of the two series of spatial samplings taken simultaneously in points A, B, C, D. In figure 3 the same data have been plotted on lognormal probability paper, revealing that for the first series: $X_m = 3.38$ ($\ln X_m = 1.22$), GSD $= 1.58$ (ln GSD $= 0.46$); and for the second series: $X_m = 1.61$ ($\ln X_m = 0.47$), GSD $= 1.51$ (ln GSD $= 0.41$).

If we want to estimate the risk of the work environment to employees' health, we can assume a confidence level of 95% and a TLV for ferric oxide of 5 mg/m³. We would then have to consider the following term:

$$\ln X_m + t_{(.95,\, N-1)} \cdot \frac{\ln \mathrm{GSD}}{\sqrt{N}}$$

which must be compared to ln TLV.

For the temporal samplings taken in position A, we obtain a value of 1.33 for the first series, and 0.86 for the second series. For the spatial samplings taken simultaneously in points A, B, C, D, we obtain a value of 1.76 for the first series, and 0.95 for the second series. When compared to the standard of 5 mg/m³ (ln TLV = 1.61), we find that in all but one case, these values are lower than the standard.

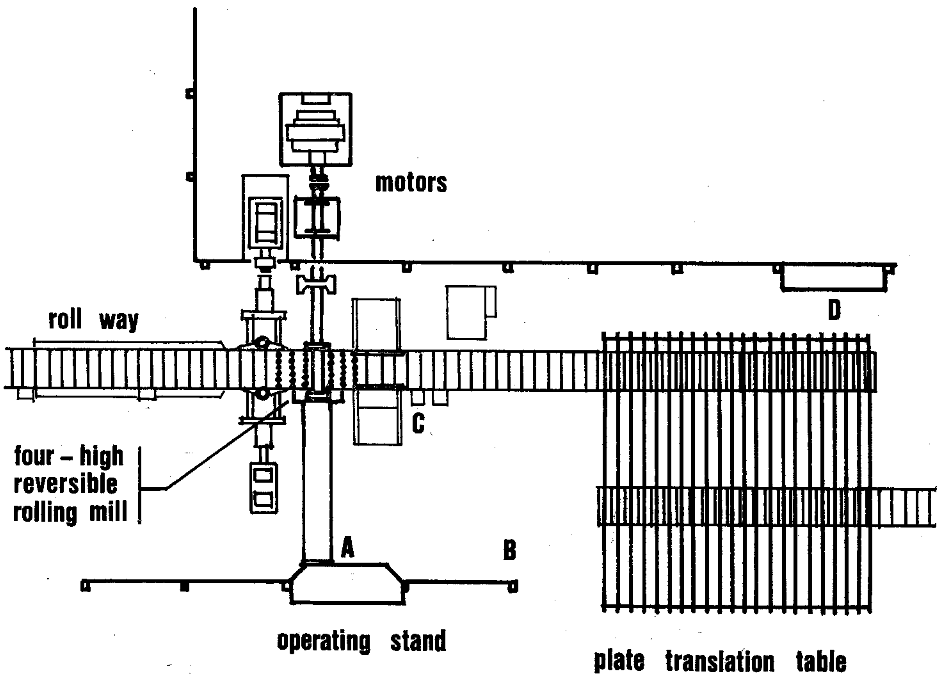

Fig. 1 - Layout of the department studied (plate rolling mill) indicating the points where samplings were taken.

TABLE 1 - *Results of dust samplings taken in position A (temporal distribution).*

Sample number m	1st series (•)		2nd series (x)	
	Concentration (mg/m^3)	Cumulative percentage (%)	Concentration (mg/m^3)	Cumulative percentage (%)
1	1.70	11	0.89	11
2	1.80	22	0.91	22
3	2.00	33	2.19	33
4	2.29	44	2.20	44
5	2.30	56	2.30	56
6	3.30	67	2.80	67
7	4.50	78	3.00	78
8	6.60	89	3.20	89

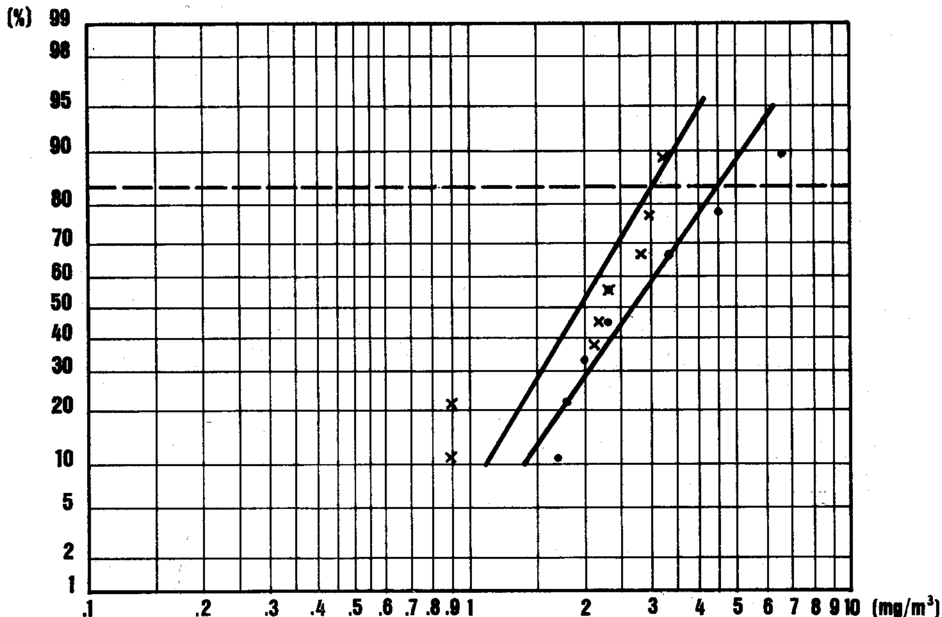

Fig. 2 - Data from Table 1 (dust concentrations and respective cumulative percentages) plotted on lognormal probability paper in order to determine X_m and GSD.
• 1st series; x 2nd series.
(Keuffel and Esser Co. 46 8040 probability paper).

TABLE 2 - *Results of dust samplings taken (simultaneously) in positions A, B, C, D (spatial distribution).*

Sample number m	1st series (•)		2nd series (x)	
	Concentration (mg/m³)	Cumulative percentage (%)	Concentration (mg/m³)	Cumulative percentage (%)
1	2.3	20	0.8	20
2	3.1	40	1.6	40
3	3.2	60	2.0	60
4	6.4	80	2.6	80

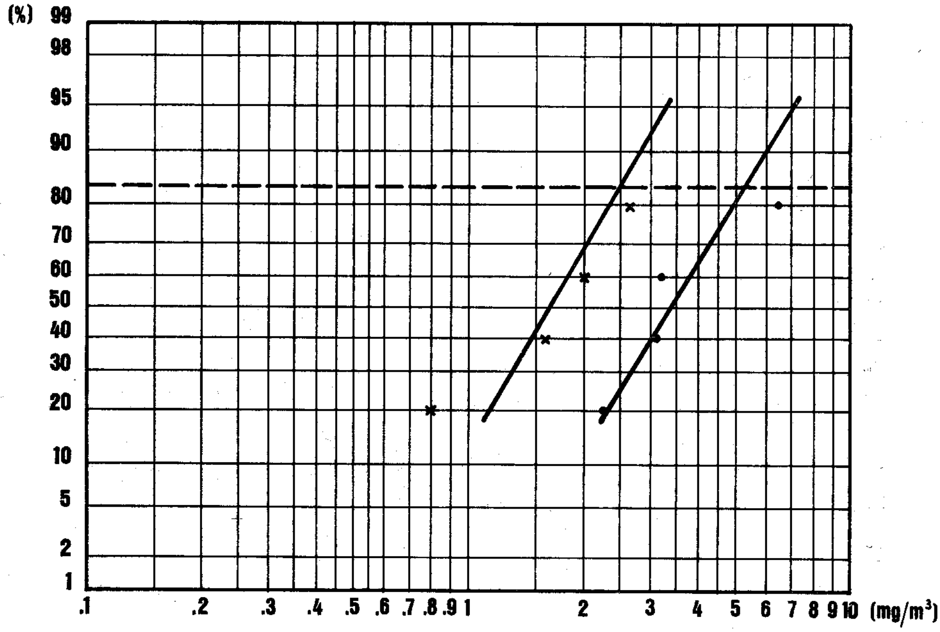

Fig. 3 - Data from Table 2 (dust concentrations and respective cumulative percentages) plotted on lognormal probability paper in order to determine X_m and GSD.
· 1st series; x 2nd series.
(Keuffel and Esser Co. 46 8040 probability paper).

Conclusions

Under the present conditions, it is impossible to maintain with 95% confidence that the concentrations of the contaminants are always below the pre-established standard of 5 mg/m^3 in all areas of the department. On the other hand, however, it is also impossible to argue the contrary.

Instead, we can state (with 95% confidence) that in position A the concentration of the contaminant is always lower than the pre-established standard.

As far as the total work environment is concerned, the description of the combined positions A, B, C, D indicates that this is a "questionable" area. The work environment itself must therefore be kept under control when rolling operations are conduced on steel plates which have a particular chemical composition, since these operations involve a great number of passes in order to obtain thin sheets *.

We should point out that if a statistical method of evaluation had not been used, and if the number of samplings had been more limited in time and space, completely different conclusions could have been drawn, depending upon the data taken into consideration. The above statement regarding position A is particularly relevant, despite the measurement of a value of concentration which was above the pre-established standard (see Table 1).

We should also emphasize another result of this study: though the median concentration values X_m were considerably different in the two series of samplings conducted at a distance of four months, the GSD dispersion indexes proved to the comparable. We can take this to be an indication of the constancy of the law of spatial distribution of the contaminant in the environment under study. This is a law of distribution which depends upon both the modality of emission of the contaminant (in relation to the production process, the type of plant, the operating conditions, etc.) and the contaminant dispersion model which is characteristic of that particular environment (in relation to its form and volume, the presence of air currents or ventilation, etc.).

In conclusion, by the help of the statistical analysis we have performed, we can be relatively certain in the judgement made regarding the risk of the environment under study to the health of workers.

* These are the conditions which characterized the first series of samplings, for which the term:

$$\ln X_m + t_{(.95,\, N-1)} \cdot \frac{\ln GSD}{\sqrt{N}}$$

exceeds ln TLV.

SUMMARY

The quantitative identification of the contaminants present in industrial environments poses certain problems regarding discrepancies between real and measured values, since the number of samplings that can be taken is necessarily limited in time and space. This uncertainty makes it difficult to establish with certainty if the real values are always below the acceptable standard that has been established, for example, the TLVs. A statistical analysis of a limited number of samples can, with a predetermined level of confidence, provide coherent indications regarding both the real concentration of the contaminant and the risk it poses, in relation to the established TLV, to the health of workers.

In this study of a work environment characterized by dust emissions, we conducted spatial and temporal measurements of the pollutant to determine its concentration; four months later, the measurements were repeated using the same techniques. The results show that the work environment under consideration is characterized by an almost constant geometric standard deviation of the concentrations (a measure of relative dispersion of the lognormal distribution taken into consideration). We can be relatively certain in the judgement made regarding the risk of such an environment to the health of workers.

REFERENCES

1. ANG A.H.S., TANG W.H.: *Probability Concepts in Engineering Planning and Design,* Vol. 1, *Basic Principles.* John Wiley & Sons, New York, chapters 5 and 6, appedix A (1975).
2. LEIDEL N.A., BUSCH K.A., LYNCH J.R.: *Occupational Exposure Sampling Strategy Manual.* NIOSH, Cincinnati, chapter 4, appendices I, J and M (1977).

Accettato il 18/2/1982.

3. MELWIN W.F.: Air Sampling and Analysis for Contaminants in Work Places. In *Air Sampling Instruments* (5th Ed.). ACGIH, Cincinnati, p. A-1 (1978).
4. PETERSON C.M.: Aerosol Sampling for Particle Size Analysis. In *Air Sampling Instruments* (5th Ed.). ACGIH Cincinnati, p. F-1 (1978).
5. SOULE R.D.: Industrial Hygiene Sampling and Analysis. In Clayton G.D., Clayton F.E. Eds.: *Patty's Industrial Hygiene and Toxicology.* Vol. 1, *General Principles.* John Wiley & Sons, New York, p. 707 (1978).