INTERNATIONAL STANDARD

ISO 9359

First edition
1989-09-15

Air quality — Stratified sampling method for assessment of ambient air quality

Qualité de l'air — Échantillonnage stratifié pour l'estimation de la qualité de l'air ambiant

Reference number
ISO 9359 : 1989 (E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 9359 was prepared by Technical Committee ISO/TC 146, *Air quality*.

Annexes A, B and C of this International Standard are for information only.

International Organization for Standardization
Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

Introduction

The ambient air quality at a particular location or region is generally variable with time, this variation being caused by a number of factors, especially meteorological conditions, topography and patterns of emissions.

Such circumstances may require that a large number of measurements be made over a long interval of time to ensure that a sufficiently wide range of conditions is covered. Stratified sampling is one method which reduces the number of measurements needed to assess certain aspects of ambient air quality. This technique has been applied for example in ambient air quality surveys and in noise surveys [1] (see examples given in annex B).

The basic aim of stratification and stratified sampling is to reduce the number of measurements needed to obtain results with a desired precision, or to improve the precision of the results without increasing the number of measurements.

To do this, information is necessary on the conditions which are likely to give rise to high, low or intermediate values of the ambient air quality characteristic of interest in the area being studied. This information is used to introduce a stratified sampling scheme in which the total number of measurements made is distributed among the different strata in such a way that the variance of the data obtained, within a given stratum, is reduced compared with the overall variance.

The reliability of the stratification scheme selected will depend upon the extent and validity of a priori knowledge, covering emission sources and the influences of topography and meteorological conditions on atmospheric dispersion. The use of results from previous measurement surveys or from specially mounted pilot surveys can be extremely helpful in the choice of strata [2], as can the application of ambient air quality models. Data from existing ambient air quality monitoring stations which are sited to be representative of the area under examination can be used in the method.

The main body of this International Standard outlines principles to be applied for obtaining a meaningful stratification scheme. Annex A outlines the method to be used for carrying out the calculations.

This page intentionally left blank

Air quality — Stratified sampling method for assessment of ambient air quality

1 Scope

This International Standard specifies a method for the assessment of certain aspects of ambient air quality in terms of percentiles and means using the principle of stratified sampling.

This is by estimating percentiles and means of the frequency distribution of measurements of ambient air quality characteristics. The application to the estimation of means, however, is restricted to cases where certain assumptions about the frequency distribution of the ambient air quality characteristic can be made using a priori knowledge or when a sufficient number of statistically independent measurements are available (see ISO 2854 and ISO 2602).

The results may be used to assess ambient air quality during the period of the measurement survey. (For length of period see also ISO 7168.) By using information on the longer-term occurrence of the various strata, an assessment for a longer period can be obtained using the same database.

Thus, although meteorological conditions have a profound effect on the concentration and distribution of air pollutants, stratified sampling enables results which are independent of the actual meteorological conditions prevailing during the interval of time of measurement to be calculated for a longer term.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 2602 : 1980, *Statistical interpretation of test results — Estimation of the mean — Confidence interval.*

ISO 2854 : 1976, *Statistical interpretation of data — Techniques of estimation and tests relating to means and variances.*

ISO 3534 : 1977, *Statistics — Vocabulary and symbols.*

ISO 7168 : 1985, *Air quality — Presentation of ambient air quality data in alphanumerical form.*

3 Definitions

For the purposes of this International Standard, the following definitions apply.

3.1 percentile: The value X_P which separates the range of the population of a parameter into two groups at the level of the percentage P.

3.2 fractile; quantile: The value X_P which separates the range of the population of a parameter into two groups at the level of the fraction $f = P/100$, where P is a given percentage.

3.3 stratified sampling: Of a population which can be divided into different sub-populations (called strata), sampling carried out in such a way that specified proportions of the sample are drawn from the different strata. [ISO 3534]

3.4 stratum: Sub-population of a population characterized by certain features.

4 Symbols

Symbol	Meaning
f	weighted fraction of the values of f_i
f_+	upper confidence limit for the weighted fraction
f_-	lower confidence limit for the weighted fraction
f_i	fraction of the ith stratum being below (or above) a given value
k	number of strata
n	total number of measurements
n_i	number of measurements in the ith stratum
m_i	number of measurements having a value below a given value in the ith stratum
P_i	percentage of the ith stratum being below (or above) a given value ($P_i = 100 f_i$)
$s^2(f)$	estimate for the variance of f
$s^2(\bar{x})$	estimate for the variance of $\bar{x}$
$s_i^2(f_i)$	estimate for the variance of f_i
$s_i^2(x_{ij})$	estimate for the variance of x_{ij}
$t_{\nu;1-\alpha}$	tabulated value of the t-distribution for the one-sided test at the significance level α and for ν degrees of freedom (for tables see ISO 2602)
$u_{1-\alpha}$	tabulated value of the standardized normal distribution for the one-sided test at the significance level α
$\bar{x}$	weighted mean of the values of $\bar{x}_i$
$\bar{x}_+$	upper confidence limit for the weighted mean
$\bar{x}_-$	lower confidence limit for the weighted mean
$\bar{x}_i$	arithmetic mean of measurements within the ith stratum
x_{ij}	jth measurement in the ith stratum
w_i	probability of occurrence of the ith stratum given as a weighting factor
X_P	percentile; fractile (or quantile)
α	significance level
$1 - \alpha$	confidence level
Δ	margin of error
μ	mean of population
μ_i	mean of the ith stratum
ν	number of degrees of freedom
$\sigma_i^2(f_i)$	variance of f_i
$\sigma_i^2(x_{ij})$	variance of x_{ij}

5 Guidelines for stratification

The stratification scheme should be designed in such a way that the strata means, μ_i, are different from each other and the variances are smaller than the variance of the population. The probability of occurrence of each stratum, w_i, should be known in advance (see 5.1 to 5.6). To calculate final results it is necessary to use weighting factors, w_i, which refer to the time period for which the assessment of ambient air quality is being made.

If a long-term assessment is needed, based on only a relatively short interval of time of measurement, then weighting factors, w_i, appropriate to the long-term situation should be used to weigh the strata data, and not weighting factors, w_i, for the interval of time of measurement. In a similar way, it is possible to use the method for prospective assessments of ambient air quality — as might be required, for example, in relation to projected increases in traffic flow — and it would then be necessary to use predicted weighting factors, w_i.

Often, there will be some uncertainty in the weighting factors, w_i, and the effect this has on the calculated percentiles or means will need to be determined by using equations (A.6) and (A.7) (see annex A).

For a satisfactory stratification, a priori information is required relating the magnitude of the ambient air quality characteristic of interest to those factors which affect it or result from it. This information is used to estimate μ_i, σ_i, and w_i. The factors on which this information is usually based are the temporal and spatial patterns of emissions, transport and dispersion, associations with other ambient air pollutants, and effects of the ambient air pollutant of interest. Examples of ways in which these factors could be used to set up stratification are given below.

5.1 Patterns of emissions

Certain emissions show obvious variations with time or in their spatial pattern.

EXAMPLES

1 Emissions of sulfur dioxide and other combustion products produced as a result of space heating have a strong seasonal variation. It could be appropriate to define strata covering different periods of the year, e.g. summer and winter, if the assessment of ambient air quality being made is likely to be affected by these emissions.

2 Road traffic exhaust emissions usually have a strong diurnal variation; their source can be considered to be linear in the case of a major road or areal in the case of an urban area. If airborne lead levels were being assessed in the vicinity of a main highway, then the strata used could be based on distance from the highway and time of day in relation to peak traffic flow periods. A spatial or temporal stratification could be used to assess carbon monoxide levels in an urban area.

5.2 Transport and dispersion

When selecting criteria for a stratification scheme related to the transport and dispersion of air pollutants in the atmosphere, it is necessary to decide whether long-range transport of air pollution, governed by synoptic factors, or nearby sources, meteorological effects and topography, is the dominant influence on the ambient air quality characteristic of interest. Stratification may then be based on factors such as

- local topography;
- air temperature;
- wind speed and wind direction;
- atmospheric stability;

— mixing height;

— solar radiation;

— weather type;

— air mass type;

or on results of dispersion models. These models using emission and meteorological data predict the temporal and spatial pattern of ambient air quality which may be used to devise stratification schemes.

EXAMPLES

1 Considering the effects of a single emission source at some distance from the area under examination, a stratification based on wind speed and wind direction may be useful (see also clause B.1).

2 An assessment of ambient air quality may be required in an urban-industrial basin area with many emission sources. Here, a variety of parameters could be examined for their suitability, e.g. atmospheric stability, seasonal effects, wind speed and wind direction (see also clause B.2).

3 If concentrations of carbon monoxide need to be assessed at street level in an area with tall buildings, e.g. in the central part of a large city, then wind speed and wind direction may be used along with the time of day (see also 5.1, example 1).

4 If oxidants are being assessed, a stratification based on solar radiation, wind direction and temperature may be useful.

5 For inter-regional or long-range transported air pollutants, the air mass concept may be used.

5.3 Associations with other ambient air pollutants

Some ambient air quality characteristics are indicative of atmospheric conditions or are associated with other ambient air quality characteristics of interest. The concentrations of certain ambient air pollutants may be closely correlated and stratified sampling for the ambient air pollutant of interest may then be carried out on the basis of the level of the indicator ambient air pollutant.

EXAMPLE — The concentrations of the ambient air pollutant of interest, as measured for example at a fixed, continuous monitoring station, can be used to define strata within which random sampling can be performed (see also clause B.3).

5.4 Effects

The effects of ambient air pollution may lend themselves to establish a stratification.

EXAMPLES

1 Effects on plant growth or crops could lead to an areal or temporal stratification.

2 Frequency and occurrence of complaints from the public about odours, for example, could also be used.

3 Plants and growth periods.

5.5 Pilot surveys

If it is not possible to devise a stratification on the basis of existing knowledge of ambient air quality and the factors governing its variation in the area of interest, then it may be necessary to set up a pilot survey of ambient air quality, or to carry out calculations using ambient air quality models, which are themselves based on stratification.

5.6 Ambient air quality models

Ambient air quality models, using emission and meteorological data to predict the temporal and spatial emission patterns of ambient air quality, may be used to devise stratification.

6 Measurement guidelines

6.1 Number of strata and number of measurements per stratum

Having decided on a stratification criterion, it is then necessary to decide how many strata are to be used, and how many measurements are to be made within each stratum to achieve a desired confidence limit in the assessment.

Experience with the stratified sampling technique has shown that the reduction in variance obtained by increasing the number of strata, k, soon falls off, and that $k = 2, 3$, or 4 is usually sufficient.

If the probability, w_i, and the estimate of the variance of each stratum, s_i^2, are known from a priori knowledge (see 5.1 to 5.4) or a pilot survery (see 5.5), the total number of measurements, n, for a given margin of error, Δ, is given by equation (1):

$$n = \left(\frac{2\, t_{\nu;\,1-\alpha}}{\Delta} \right)^2 \left(\sum_{i=1}^{k} w_i s_i \right)^2 \qquad \ldots (1)$$

The theory of stratified sampling shows that, once the total number of measurements to be made has been decided upon, it is possible to allocate these among the different strata so as to achieve a minimum variance in the results calculated for the population.

If the arithmetic mean is determined, then this optimum allocation is achieved when

$$n_i = n \frac{w_i s_i}{\sum_{i=1}^{k} w_i s_i} \qquad \ldots (2)$$

And if fractions are determined, then

$$n_i = n \frac{w_i \sqrt{f_i (1 - f_i)}}{\sum_{i=1}^{k} w_i \sqrt{f_i (1 - f_i)}} \qquad \ldots (3)$$

In other words, a large number of measurements should be carried out in a particular stratum if that stratum has a higher probability of occurrence, as given by w_i, or a large variance, as given by s_i^2 or f_i $(1 - f_i)$.

To calculate the number of measurements in the ith stratum, n_i, precisely, knowledge of both w_i and σ_i (or f_i) is needed. While the weighting factors, w_i, may be established before the measurements begin, it is likely that there is little or no information on σ_i (or f_i). Consequently, it is advisable to refine n_i as the measurements proceed, by calculating s_i (or f_i) and then making use of these in equations (2) or (3).

NOTE — The validity of some of the equations given in annex A will depend upon the number of measurements made in a particular stratum. Thus equation (A.9) requires that $n_i > 15$ since it is based on the approximated binomial test; if this is not satisfied, then exact theory must be used. The equation giving confidence limits for the weighted mean [equation (A.10)] is always valid if $n_i \geqslant 5$. n_i may be less if the frequency distribution of the air quality characteristic in the ith stratum is gaussian.

In cases where information about the variance is not available, proportional allocation to w_i may be suitable.

6.2 Independence of measurements

In order to apply the equations and methods of calculation specified in this International Standard, it is necessary for the measurements to be carried out in a manner such that they may be assumed to be independent.

NOTE — Ambient air quality measurements at a particular measurement site are often highly auto-correlated and it may thus be necessary to ensure intervals of time of sufficient length between the measurements. For example, strong auto-correlations have been observed in central Europe for periods up to six days. When stratified sampling is employed, independent measured values could well be obtained at shorter intervals of time, because correlation effects due to changes from one stratum to another have been eliminated.

6.3 Interval of time and site of measurements

Having established a stratification, the number of strata to be used, the number of measurements to be carried out and the way in which these are to be allocated among the different strata, it is then necessary to decide on a scheme for obtaining measurements within each stratum which are random in time and space, respectively. The interval of time of measurement should be less than the expected duration of the stratum condition.

NOTES

1 Randomness in time can be achieved simply by using random number tables, but for purely practical reasons it may be necessary to limit intervals of time of measurement to, for example, normal daytime working hours. In this case, great care is needed to ensure that bias is not introduced to the measured values due to a temporal pattern in the air quality characteristic being assessed, e.g. the diurnal variation in traffic exhaust emissions. A temporal form of stratification should be decided upon if considerable influence is to be expected. This will necessitate measurements being carried out outside normal working hours.

2 If randomness in space is used, it can be achieved by identifying a large number of measurement sites using a grid, from which the accessible measurement sites are then listed and numbered. The measurement sites to be used for measurements can then be selected by using random number tables.

An alternative method of measurement site selection may be used if sufficient information is available from previous studies to make possible the identification of representative measurement sites, which may allow a reduction in the number of measurements.

7 Methods of calculation

Estimates for fractiles, means, and fractions of the population are calculated as sums of the corresponding weighted values for the various strata. The variances of these sums are obtained by calculating the respective weighted sums of the variances of the strata. These are usually less than or equal to the variances which would be obtained by random unstratified sampling [2].

Appropriate formulae to be used are set out in annex A. Examples of how this International Standard can be applied, together with calculated results, are given in annex B.

Annex A
(informative)

Mathematical equations

All the symbols used in this annex are as defined in clause 4.

The fraction below the value of the fractile X_P is determined from equation (A.1):

$$f_i = \frac{m_i}{n_i} \qquad \dots \text{(A.1)}$$

The arithmetic mean of the measurements in the ith stratum is obtained from equation (A.2):

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \qquad \dots \text{(A.2)}$$

The estimate for the variance in terms for the fraction in the ith stratum is given by equation (A.3):

$$s_i^2(f_i) = f_i\,(1 - f_i) \qquad \dots \text{(A.3)}$$

and the estimate for the variance of x_{ij} is given by equation (A.4):

$$s_i^2(x_{ij}) = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \qquad \dots \text{(A.4)}$$

The estimate for the fraction from the population is found by calculating the sum of the weighted fractions from equation (A.5):

$$f = \sum_{i=1}^{k} w_i f_i \qquad \dots \text{(A.5)}$$

where $\sum_{i=1}^{k} w_i = 1$

The mean of the population, μ, is estimated by the sum, $\bar{x}$, of the weighted arithmetic means from equation (A.6):

$$\bar{x} = \sum_{i=1}^{k} w_i \bar{x}_i \qquad \dots \text{(A.6)}$$

where $\sum_{i=1}^{k} w_i = 1$

The variance for the weighted fraction is estimated using equation (A.7):

$$s^2(f) = \sum_{i=1}^{k} \frac{w_i^2\, s_i^2(f_i)}{n_i} \qquad \dots \text{(A.7)}$$

where $\sum_{i=1}^{k} w_i = 1$

The confidence limits for the weighted fraction are given by equations (A.8):

$$\left.\begin{aligned} f_+ &= f + \left(s(f)\, u_{1-\alpha} + \frac{1}{2\,n} \right) \\ f_- &= f - \left(s(f)\, u_{1-\alpha} + \frac{1}{2\,n} \right) \end{aligned}\right\} \qquad \dots \text{(A.8)}$$

The variance for the weighted mean is estimated using equation (A.9) (see also note to 6.1):

$$s^2(\bar{x}) = \sum_{i=1}^{k} \frac{w_i^2\, s_i^2(x_{ij})}{n_i} \qquad \dots \text{(A.9)}$$

where $\sum_{i=1}^{k} w_i = 1$

The confidence limits for the weighted mean are given by equations (A.10) (see also note to 6.1):

$$\left.\begin{aligned} \bar{x}_+ &= \bar{x} + s(\bar{x})\, t_{\nu;\,1-\alpha} \\ \bar{x}_- &= \bar{x} - s(\bar{x})\, t_{\nu;\,1-\alpha} \end{aligned}\right\} \qquad \dots \text{(A.10)}$$