Descriptors for aircraft noise

H.M.E. Miedema

March 1994

Report title Descriptors for aircraft noise	**Distribution number**
Author H.M.E. Miedema	**Report number** disturbance 5B **Date of publication**
Institute Responsible TNO-PG PO Box 2215, 2301 AC Leiden, The Netherlands	**Report type** Disturbance Publication Series
Commissioned by: Ministry of Housing, Spatial Planning and the Environment PO Box 30945, 2500 GX The Hague, The Netherlands	**Title of research project** Noise measures for air traffic

Summary
This study contains a systematic description of the conditions which need to be met by an optimum noise exposure measure for air traffic. 'Optimum' here means the measure which best correlates with the adverse effects of noise. Because there are only limited systematic data available, this study concentrates on the effects in terms of annoyance and sleep disturbance.

Although the study is aimed primarily at aircraft noise, other noise sources are also considered. As well as the measures of aircraft noise exposure currently in use in the Netherlands - the Kosten Unit 'Ke' and the unit of noise exposure for light aircraft 'BKL' - the study also considers the 24-hour exposure measure 'L_{etm}', used in the Noise Nuisance Law, and the measure L_{dn} used in various studies in other countries.

This study does not identify a single measure as being clearly the 'best'. Although there are differences in the ways these measures are constructed, apparent major differences often turn out to be actual similarities (such as penalty factors for night-time). The measure B (used in the report to denote the Kosten units) appears to be quite satisfactory in a number of respects, the main criticism relating to the fact that it only uses the peak noise level during a flyover. Each of the other measures currently in use also suffers from one defect or another. Griefahn's curve, on the other hand, is adjudged inadequate as a measure of sleep disturbance.

The ubiquitous A-weighting in some cases gives a completely inaccurate reflection of the actually perceived loudness of a noise. An alternative frequency weighting scheme (the 'Zwicker' method) can be satisfactory in such cases. The system developed here for comparing noise exposure measures is in principle also suitable for application at the international level.

Associated reports

Supervisory Committee
J.A. Verspoor
M. van den Berg
J.F. Bakker
F. van Deventer
A.N. Lefferts
A.J. Koppert

This Guide contains a very brief and free summary of the law and regulations. You cannot refer to this publication in the event of a dispute, but should always consult the text of the laws and regulations themselves.	**Number of pages: 65**

Preface

This study addresses the question as to which is the optimum measure of exposure to aircraft noise. 'Optimum' here means that which correlates best with the adverse effects of noise. The effects considered in this study are annoyance and sleep disturbance.

Although the study is aimed primarily at aircraft noise, other noise sources are also considered. As well as the measures of aircraft noise exposure currently in use in the Netherlands - the Kosten Unit 'Ke' and the unit of noise exposure for light aircraft 'BKL', the study also considers the 24-hour exposure measure 'L_{etm}', used in the Noise Nuisance Law, and the measure L_{dn} used in various studies in other countries.

This study does not identify a single measure as being clearly the 'best'. The measure B (used in the report to denote the Kosten units) appears to be quite satisfactory in a number of respects, the main criticism relating to the fact that it only uses the peak noise level during a flyover. Each of the other measures currently in use also suffers from one defect or another. Griefahn's curve, on the other hand, is adjudged inadequate as a measure of sleep disturbance.

The ubiquitous A-weighting in some cases gives an inaccurate reflection of the actually perceived loudness of a noise. Moped noise is a case in point. Because of its frequency spectrum it can give rise to more annoyance than the A-weighted noise level would suggest. An alternative frequency weighting scheme (the 'Zwicker' method) can be satisfactory in such cases. The A-weighting is so well entrenched, however, that it is unlikely to be replaced by an alternative in the foreseeable future.

The fact that there might be better measures of exposure than those in current use does not necessarily mean that we should be thinking of replacing them. Existing measures are inextricably built into regulations already implemented, and a new measure would need to have clearly demonstrable advantages to set off against the disadvantages of a major switch. Redrawing noise zones would involve procedures with major spatial and financial consequences.

The methods developed here are very well suited for application at the international level for examining the relationships between different measures of exposure. If the European Commission should decide to undertake an initiative to harmonise noise measures and calculation methodologies, this report could serve a useful function.

Deputy Director, Noise and Traffic Department

(Dutch version signed)

H.C.G.M. Brouwer

CONTENTS

1. INTRODUCTION

This introduction begins with a statement and discussion of the objective of this study. The structure of this report is then described.

1.1 Objective

The objective of this study is to identify which noise measures can be used for setting limit values for environmental noise in residential areas, in particular from major civil aviation activities. This report gives special attention to the following measures. The reader is referred to the Annex for precise definitions of the noise measures discussed in this report.

- The L_{Aeq}(24h), defined as 10 times the logarithm of a mean over 24 hours.
- The L_{dn}, in which night noise is weighted more heavily than it is in L_{Aeq}(24h).
- The BKL, which is used for the general aviation category in the Netherlands, and is related to L_{dn}, but which is determined for a period of one year using specific rules.
- The L_{etm}, which, like L_{dn} penalises night-time noise, and also evening noise. Unlike the L_{dn}, however, the L_{etm} is based on the maximum of the contributions for the various periods in the day. The L_{etm} is used in the Netherlands for assessing noise from road and rail traffic and industry.
- B in Ke, which unlike the measures listed above, is based on the maximum sound level during an event. B is used in the Netherlands for air traffic except for the general aviation category.

The measures referred to above have been proposed as indicators of the extent to which noise causes annoyance. Griefahn's 10% awakening curve is also discussed, which has been proposed specifically for evaluating the influence of noise on sleep.

The L_{Aeq}(24h), L_{dn} and L_{etm} all relate to exposure during a 24-hour period. Griefahn's curve relates to noise exposure during one night. The definitions of B and BKL adopted in the regulations, on the other hand, relate to a year. But these measures are determined via an intermediate step in which the values for each 24-hour period are calculated. Unless otherwise indicated in this report, B and BKL are intended to refer to the 24-hour values determined in this intermediate step.

The utility of a noise measure depends primarily on how well it correlates with adverse effects. More pronounced adverse effects must be associated with higher values of the noise measure. In this study the relationship between noise measures and non-specific annoyance (part I) and with the impairment of the quality of sleep (part II) is investigated.

As well as the relationship with adverse effects, the interrelationships between the noise measures used are of importance. How, for example, can the different measures complement each other?

At present, four measures are used for aircraft noise. For major civil aviation and around military airports the noise load B, expressed in Ke, is used for aircraft take-off and landing. At Maastricht airport, as well as B, Griefahn's 10% awakening curve is calculated to estimate sleep disturbance from night flights. BKL is used for small-scale general aviation. Noise from engine testing is treated as industrial noise, and evaluated using L_{etm}. Noise from aircraft which are climbing, landing or circling (circuit flying, stacked aircraft) is referred to as overflying noise to distinguish it from noise on the ground.

In theory, a reduction in the number of measures would simplify the formulation and implementation of policy. This applies particularly to multi-source situations where several noise measures are currently used. In the event of a change to another measure, the noise contours will have to be reworked on the new basis, and this could have far-reaching practical consequences. Any modification therefore needs to be built on a sound theoretical foundation, and must be well prepared and planned. The present study examines whether a case can be made for using a single noise measure. Another reason for considering alternatives is that none of the four measures mentioned above is used elsewhere in the EC or beyond. Other measures are used in other countries. Better harmonisation of the measures used would make it easier to compare the Dutch noise standards with those in the EU and other countries. It would be sensible to relate discussions about noise measures in the Netherlands to deliberations on this subject in the EC. Aircraft noise has been under discussion within the EC for some time, supported by surveys of measurement methodologies, measures and their consequences (see, for example, ECAC/ANCAT/3, 1975; Koppert, 1991; Jonkhart, 1992; Koppert, 1993).

Of the three measures of overflying noise (B, BKL and Griefahn), B was formulated first, for measuring the noise from civil and military aviation. It was subsequently suggested that in certain circumstances B does not fully reflect the associated annoyance. As a result, other measures were introduced. Specifically in judging aircraft noise problems arose of having to reconcile divergent measures. Flying noise from large and small aircraft in the vicinity of the same airport, for example, are assessed by two different methods. Neither method allows for the fact that at a given location annoyance may also be caused by the other type of aircraft. We summarise below the shortcomings which have been suggested for the measure B, and which prompted the question as to whether other measures would not be more appropriate in certain situations.

Because the value of B is based on the peak noise level, it does not reflect the duration of a flyover. In the case of large civil airliners this duration is usually less than 30 seconds, depending on the definition used. In the case of light aircraft, flyover duration may be long due to circling, but training flights in large aircraft may also lead to prolonged exposure. Clearly allowance needs to be made for the fact that where flyover duration varies considerably, account will have to be taken of the contribution of this factor to the annoyance caused.

The BKL was introduced for general aviation. The value of the BKL is affected not just by the peak noise level, but also by the level during the entire flyover, so that it is partly determined by the duration of the flyover. When the BKL was introduced the question arose as to how the two measures could be combined where both small and large aircraft were involved. The present regulations do not allow for the additive effect of exposure to both light aircraft and civil airliners, but are applied for the two types of aircraft separately. Those which have the greater consequences are implemented. This approach ignores the fact that the presence of a second source type may mean that the measures based on only one source type are inadequate.

Another important respect in which BKL differs from B is the way in which exposure is aggregated over the days in a year. The BKL is quite strongly influenced by the busiest days in the busiest weeks of the year. In the case of general aviation this is generally summer weekends, because of the important recreational component in this category. Since these are the times when people have their windows open wider and more frequently, and spend more time out-of-doors, the same incident outdoor noise level can be expected to cause greater annoyance at these times. This fact should be reflected in the manner in which the noise exposure is calculated not only from light aircraft but from all aircraft, including large aircraft.

Measure B has been validated using data from questionnaires on annoyance and disturbance (including sleep disturbance) around Schiphol. Because of doubts as to whether this measure is a good indicator of sleep disturbance, and in particular of the likelihood of being awakened, for large numbers of night flights, the method of Griefahn is used around Maastricht. The introduction of this measure further complicated the question of how the various methods (B, BKL and Griefahn) can be reconciled together.

Helicopter noise, the regulations for which are based on the measure B, is a case in point currently under scrutiny. Because of the increasing civilian use of helicopters and the establishment of an airborne brigade, there is increased exposure to helicopter noise. As a result there is interest in determining which measure is a good predictor of noise annoyance from helicopters. B was designed for fixed-wing aircraft, and helicopters have different use patterns and noise characteristics. But it will in any case be clear that in practical terms the introduction of a fifth measure would be very undesirable.

The question has also arisen in the past whether B is suitable for measuring the noise around military air bases, particularly from jet fighters. De Jong [1983a] considered that it was suitable, and that no other measure was needed for military aircraft. De Jong and Groeneveld [1983] have indicated that for a given value of B annoyance in such a situation is greater than around a civilian airfield.

It has also been pointed out that since B was validated there have been changes in the aircraft fleet and in flying patterns. There are now more flights with quieter aircraft. A given value of B now reflects a different situation. It is therefore legitimate to ask whether B is still appropriate for this changed situation.

1.2 Structure of this report

The first part of this report examines the relationship between exposure to and annoyance caused by noise. The second part then looks at the relationship between exposure and awakening behaviour. General conclusions then follow in chapter 14.

Chapter 2 presents some results on the relationship between various noise measures and annoyance. A model is then discussed which indicates how the steps in the determination of noise measures for predicting annoyance can be broken down. Some existing measures can be defined in terms of steps which are permissible according to the model. Others, on the other hand, are not, and therefore do not conform with the model.
Chapter 3 deals in some detail with the first 'frequency' step in the determination of L_{Aeq}(24h). In most noise measures used the instantaneous frequency pattern is aggregated into the A-weighted sound level in this step.
Chapter 4 considers in detail the three 'time'-steps in the determination of noise measures. These involve the aggregation of the instantaneous values, determined in the first step, over 24 hours. Further integration over time, up to a value of one year, is also discussed.
The spatial variation of noise in and around the home is dealt with in chapter 5. This variation may affect nuisance.
Chapters 6 and 7 consider determinants of annoyance other than sound levels: in chapter 6 frequency-related factors such as tonal content, and in chapter 7, factors such as impulsiveness related to temporal variations in the noise.
Chapter 8 briefly outlines a procedure, based on the foregoing, for judging environmental noise. A noise measure specified on the basis of chapters 2 to 4 is an important element in this procedure.
In chapter 9 some conclusions are drawn about noise measures for rating annoyance, partly based on a comparison with the specification in chapter 8. In particular, the measures L_{Aeq}(24h), L_{dn}, BKL, L_{etm} and B are considered.
In chapter 10 the focus moves from annoyance as a criterion for a noise measure to that of sleep disturbance. For this purpose the discussion concentrates on the likelihood of being awakened. The 10% awakening curve devised by Griefahn is described and discussed.
The relationship between a standard based on this curve and one based on the measures L_{etm} and B is discussed in chapter 11.
Chapter 12 considers another indicator of the impairment of quality of sleep. The use of self-reported

sleep quality rather than awakening patterns is examined.

Finally, chapter 13 contains conclusions on the assessment of sleep disturbance due to noise, specifically in relation to the B measure and Griefahn's curve.

PART I: ANNOYANCE

2. A MODEL FOR THE INFLUENCE OF NOISE ON ANNOYANCE

The noise to which a house is exposed can be described in terms of a pattern of sounds of a given intensity, one for each moment of time and for each frequency band. This can be imagined in three dimensions. At each point in time there is a pattern of sound intensities for the different frequencies, the frequency spectrum. During a given timespan there is a sequence of such frequency spectra, which makes up the frequency time pattern.

The frequency time pattern at a location near a dwelling gives a fairly detailed picture of the noise exposure at that point, but it is not complete. It gives no information, for example, on which sound originates from which source. There may also be a spatial variation in the pattern in and around the dwelling. This last aspect is dealt with in chapter 7.

The object of this study is to indicate which noise measures should be used to set standards for environmental noise levels. A noise measure aggregates an entire frequency time pattern into a single number. Different noise measures do this in different ways.

It was observed in section 1.1 that the utility of a noise measure depends primarily on how well it correlates with the adverse effects. This part of the report deals with annoyance as the effect. A scheme for aggregating together frequency time patterns can be adjudged good if larger values of the noise measure it produces are associated with greater annoyance levels.

Two complementary approaches are considered in this chapter. In section 2.1, quantitative dose-annoyance relationships are considered for a number of relevant existing noise measures. We look, for example, at the L_{etm}-annoyance relationship and the B-annoyance relationship. These analyses do not appear to provide grounds for preferring one of the measures over the other. In section 2.1, therefore, the manner in which the frequency time pattern is aggregated when a noise measure is being calculated is considered in greater detail. Various steps are distinguished, and these different steps are considered further in the succeeding chapters.

2.1 Dose-response relationship

Schultz [1978] developed a curve for the relationship between L_{dn} and annoyance. This was one of the first studies in which the results from many dose-annoyance studies are synthesised. Schultz' article triggered off a major discussion on how the relationship between exposure to noise and annoyance can be described [Schultz, 1982; Kryter, 1982, 1983; Fidell et al, 1991]. One of the issues discussed was whether a single relationship is sufficient for the various transport modes, or whether several relationships are needed for an adequate description.

In order to clarify this matter the original data from various European noise studies were compiled and re-analysed by Miedema [1992]. There were some 13,000 evaluations of the annoyance value of noise which could be linked to an exposure level for that noise. Not only L_{dn} but also L_{Aeq}(24h) and L_{etm} were calculated. The relationship between annoyance and each of the three measures of exposure for each traffic type (aircraft, highway traffic, other road traffic and rail) is shown in figures 2.1 and 2.2. Figure 2.1 shows the percentages highly annoyed, annoyed and at least somewhat annoyed respectively as a function of the exposure.

Figure 2.2 plots the annoyance score against noise exposure. A score of 0 indicates no annoyance amongst the exposed population and a score of 100 that every individual in the exposed population is highly annoyed. Coefficients of correlation of 0.49, 0.50 and 0.50 were obtained for the relationships between annoyance and L_{Aeq}(24h), L_{etm} and L_{dn} respectively. The differences between these values are marginal and do not provide a basis for preferring one of these measures to the others. The great similarity between correlation coefficients may be due to the strong relationship between the sound levels in the various periods of the day. The corresponding correlation coefficients for air traffic only between L_{Aeq}(24h), L_{etm} and L_{dn} and annoyance are 0.53, 0.54 and 0.53 respectively.

In interpreting these correlations it is important to bear in mind that they are calculated at the individual level: the noise loading is correlated with the reported annoyance on a per person basis. Correlation coefficients of the order of 0.8 and 0.9 are sometimes reported. These are calculated at an aggregated level: for each noise loading range, the mean noise and the mean annoyance are calculated, and these pairs are then correlated with one another. Correlations calculated on this basis obviously say little about the spread in the original data.

The correlation coefficients mentioned above relate to figure 2.2. More detailed information can be found in the source referred to, from which the figures are taken.

In addition to the three measures considered, BKL and B are also of particular interest for the Dutch situation. The relationship between the BKL and annoyance was not analysed quantitatively. The way in which, in the case of the BKL, a 'representative' day is determined from the days of the year is very specific to that measure, but the manner in which the value is determined for that 'representative' day displays a strong similarity with the L_{dn}. The main difference is that the BKL distinguishes three rather than two periods in the day. The day-time period of the L_{dn} (7.00 - 22.00) is further subdivided into a day-time (7.00 - 19.00) and an evening (19.00 - 22.00) period, and a penalty of 5 dB(A) is applied to the evening. Leaving aside the way in which the 'representative' day is determined, the analogy with the L_{dn} (day-night average level) could be brought out by designating it the L_{den} (day-evening-night average level).

It is difficult to compare the correlations between L_{Aeq}(24h), L_{dn}, L_{etm} and annoyance with the correlation for B because there are no studies in which both B and one or more of the other measures are related to annoyance. There are studies which analyze the relationships between the measures themselves, but these do not give any information on the relationship with annoyance (ECAC/ANCAT/3, 1975; NATO/CCMS, 1989; Jonkhart, 1992).

It is therefore difficult to make a readily interpretable comparison between the various noise measures, and in particular between B or BKL and L_{Aeq}(24h), L_{etm}, or L_{dn} on the basis of their correlation with annoyance. In the following section we concentrate instead on comparing the way in which the various noise factors are quantified in the light of knowledge about how they affect annoyance. We first consider what seems to be the best way of quantifying the various factors, and this is then compared with the way these are actually quantified in the various measures under consideration.

2.2 A model

The frequency-time pattern for a single source is considered here. How well this frequency-time pattern is summarised by a noise measure depends on the manner in which the sound intensity for each frequency-time combination which makes up the pattern affects annoyance. What reduction in the sound intensity for a given frequency and time, for example, will offset an increase in intensity for another frequency-time combination such that the annoyance remains unchanged? Considerations of this nature are referred to as 'trade-offs' between contributions, in this case contributions to the annoyance.

Miedema [1995] has formulated a model which describes qualitatively the properties of such trade-offs. By specifying the characteristics precisely, it is possible to deduce which noise measures display these characteristics. These are only certain measures, namely those which can be defined in terms of a so-called hierarchical power-sum.

A hierarchical power-sum is calculated by repeated application of the following rule:

$$y = [\ \sum_i (b_i x_i)^{a_i}\ ]^{1/a}$$

where x_i is either an initial value, i.e. a sound intensity for a given frequency-time combination, or the output of a previous application of the rule. The parameters a, a_i and b_i are positive.

The following points in regard to this combination rule should be noted:

- If a_i = a and these tend to infinity (∞) then y tends towards the maximum of all the $b_i x_i$.
- The parameter a is in fact redundant except in the case referred to in the preceding point is not performed. Unless the maximum of the $b_i x_i$ is being taken in a step, it is henceforth assumed that a = 1.
- It is primarily the order of the numbers ultimately obtained which is of significance. The order of the numbers indicates the rank in terms of annoyance of the situations for which they were calculated. After the last step it is possible, if desired, to carry out certainly strictly monotonic transformations, such as positive power transformations or logarithmic transformations. Since these transformations are not essential [Miedema, 1995] we shall sometimes for reasons of simplicity, refer to measures obtained from one another by such a transformation as the same.

A number of noise measures, such as percentile measures (for example L_{A10} and L_{A90}), the traffic noise index (TNI) and noise and number index (NNI) are not consistent with the model: they cannot be written in terms of hierarchical power sums. A definition of B or another measure in which values are only included in the calculation of the measure if they exceed a certain threshold value, are not consistent with the model either. In the following the noise load B (or another measure) refers to the version in which no such threshold is applied.

Measures such as L_{Aeq}(24h), L_{dn} and BKL are consistent with the model. Some measures, for which a maximum is set for one or more of the steps, such as L_{etm} and B, are 'limit cases'. They do not strictly accord with the model, but measures which are arbitrarily close to them are in accordance with the model.

The various noise measures which are consistent with the model or to which measures which are consistent can be made arbitrarily close, can be described by specifying the power sums of which they are composed. For L_{Aeq}(24h), L_{dn}, BKL, L_{etm} as well as B, the sound intensities per frequency-time combination, expressed in $10^{-12}W/m^2$, are the base data which are aggregated together by means of power sums. The steps by which the base data are aggregated together are broadly comparable for these measures. There is one 'frequency step' followed by three 'time steps'. It is only the parameters used in the power sums which differ for the different measures.

In the frequency step the intensities for the various frequency bands at a given time are aggregated into an instantaneous value. Then follow the three time steps. First of all the instantaneous values for an event are combined to give an event value. The values for the separate events during part of a day are then aggregated together to give a value for that part of the day. Finally the values for the parts of the day are aggregated together to give a single value for the noise during 24 hours. In practice some of these steps can be combined for some measures, but we will not go into details, to avoid complicating matters.

The main difference between the various noise measures described lies in the parameters used in each step for the power sum. Once the values of these parameters have been specified, then a measure is largely defined. All that remains is, for example, the precise definitions of the different periods in the day.

The following tables show the weights b_i and the exponents a_i for each of L_{Aeq}(24h), L_{dn}, BKL, L_{etm} and B.

L_{Aeq}(24h)

frequency	time within event	time of event	period of the day
A-weights; 1	1; 1	1; 1	1; 1

L_{dn}

frequency	time within event	time of event	period of the day
A-weights; 1	1; 1	1; 1	1, 10; 1

BKL

frequency	time within event	time of event	period of the day
A-weights; 1	1; 1	1; 1	1, 3.16, 10; 1

L_{etm}

frequency	time within event	time of event	period of the day
A-weights; 1	1; 1	1; 1	1/12, 3.16/4, 10/8; $\rightarrow \infty$

B

frequency	time within event	time of event	period of the day
A-weights; 1	1; $\rightarrow \infty$	1; 2/3	4,1,2,3,4,6,10; 1

The first column corresponds to the first step, the second column to the second step, etc. The figure(s) before the semi-colon in each cell indicate the weights. If these are all the same, only one figure is shown. The figure(s) after the semi-colon indicate the exponents for that step; again, only one is shown if they are all the same. The symbol $\rightarrow \infty$ means that the measure is approached by taking vary large values of a = a_i. (As mentioned earlier, a = 1 in all other cases). Figure 2.3 contains the same information as the above summary. This figure indicates, in the form of a tree structure, how the various noise measures can be obtained by taking different values for the weights and the exponents in the various steps.

This can be illustrated by considering the sumss by which the L_{etm} is built up. The first step involves aggregating together the intensities I_i for the various -octave or octave bands i for each point in time, using the standard A-weights [IEC 225] and exponent 1:

$$I_A = \sum_i A_i I_i$$

It should be noted that the A-weighted sound level L_A, is equal to $10\log I_A$. In the second step, the I_A value at time t is denoted by $I_A(t)$. In this step all the weights and the exponents are equal to 1, so that the sum per event, I_{AX}, is given by:

$$I_{AX} = \sum_t I_A(t)$$

The A-weighted 'sound exposure level', L_{AX}(=SEL), is equal to $10\log I_{AX}$. In the third step, the I_{AX} value for event j is denoted by $I_{AX}(j)$. In this step all the weights and the exponents are again equal to 1, so that the sum for the relevant part of the day, I_{Aeq}, is given by:

$$I_{Aeq} = \sum_j I_{AX}(j)$$

The A-weighted equivalent sound level for the part of the day, L_{Aeq}, is equal to $10\log I_{Aeq}$. In the fourth and final step, the I_{Aeq} values for the three parts of the day distinguished are denoted by $I_{Aeq}(7\text{-}19)$, $I_{Aeq}(19\text{-}23)$ and $I_{Aeq}(23\text{-}7)$ respectively. In this step the weights are equal to 1/12, 3.16/4 and 10/8. The exponents $a = a_i$ are taken as very large values. I_{etm} is therefore given by:

I

$$I_{etm} = \lim_{a \to \infty} \{ [I_{Aeq}(7 - 19) /$$

The so-called 24-hour level, L_{etm}, is equal to $10\log I_{etm}$ - 10log3600.

The model is based on a few general notions only. But more information is available, and this can be used to further constrain the alternatives generated by the model, of which the five measures referred to above belong. This information is presented in the following chapters.

Figure 2.1 Percentage highly annoyed (top row), annoyed (bottom row) and at least somewhat annoyed (overleaf) as a function of exposure (A = air traffic, H = highway traffic, O = other road traffic, R = rail traffic, I = impulse sources).

Figure 2.2 Annoyance score as a function of exposure (lines from top to bottom are for air traffic, highway traffic, other road traffic and rail traffic).

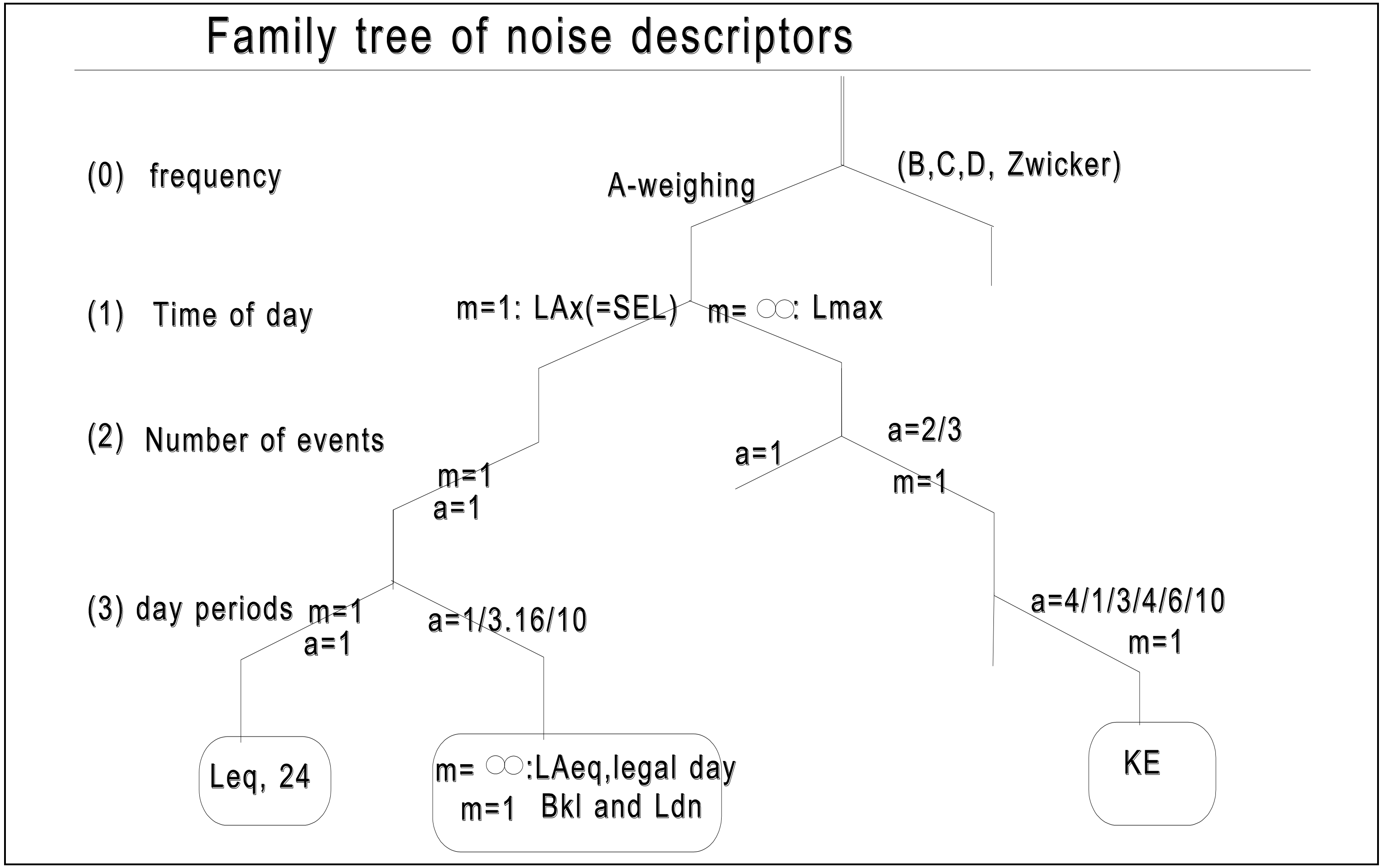
Figure 2.3 Tree structure describing the composition of noise measures used in the Netherlands. For each of the steps shown the same type of aggregation rule is applied (see text) but the parameters may be different for each step. These parameters are indicated in the tree structure. Starting with a complete picture over time of noise intensities per frequency band, the steps are as follows:

Frequency step: values for the separate frequency bands are aggregated for each time point in time.

Time step 1: values for the separate time intervals are aggregated for each event.

Time step 2: values for the separate events are aggregated for each relevant part of the day.

Time step 3: values for the separate parts of the day are aggregated for a 24-hour period.



3. LOUDNESS (AT A SINGLE MOMENT)

This chapter examines the first step in calculating noise measures such as L_{dn}, L_{Aeq}(24h), L_{etm} and B. In this first step, the intensities for the various frequency bands at a given point in time are aggregated to give a value for the instantaneous loudness. For the noise measures mentioned, this is done by calculating the A-weighted sound level. But there are alternatives to this.

Other measures for the instantaneous loudness are for example the B, C or D-weighted sound level or the loudness calculated according to the procedures of Stevens or Zwicker. These measures are discussed in this chapter. Like the A-weighted sound level, they can also be used as a basis for assessing noise situations over 24 hours.

At the end of this chapter the so-called integration time and the decay pattern for instantaneous noise are considered. Many meters used to measure A-weighted sound levels allow a choice from three options for the integration time and decay rate, designated as 'slow', 'fast' and 'impulse'. In his loudness model, Zwicker allows for the way integration time and decay pattern influence the instantaneous koudness.

The ideas discussed in this chapter appear to have only limited practical implications, at least in the shorter term; the instantaneous A-weighted sound level is firmly entrenched in present practice, and forms the basis of most noise measures used in the Netherlands and elsewhere. Some cases will be indicated in this chapter for which this approach does not predict annoyance optimally. This means that there are limits to the improvements which can be achieved if the present noise measures based on instantaneous A-weighted sound levels are replaced by other noise measures also based on the A-weights.

An issue which is considered in this chapter and which is currently the subject of discussion in the Netherlands is the integration time used for calculating instantaneous levels. The choice of integration time is important for measures based on the maximum A-weighted level per event, in particular where the events have a short, relatively high peak, i.e. impulse sounds and close, rapidly moving sources (low-flying aircraft). In estimating B for air traffic, a meter setting of 'slow' has been chosen. In general it appears, as we shall see in section 3.5, that a setting of 'fast' would be a better choice. For most types of traffic (road, rail and air) the choice between 'fast' and 'slow' has very little effect on the calculated maximum A-weighted level. This notwithstanding, the use of a single setting for all measurements for regulatory purposes, based on the way noise is actually perceived, would be simpler and more theoretically satisfactory.

3.1 A, B, C and D-weighting

The A, B, C, and D-weighted sound levels are defined as 10 times the log of the weighted sum of the sound intensities per frequency band. The weighting scheme is different for each of these measures.

These measures have quite a long history. The A-weighted sound level was calculated in a noise survey in New York dating back to 1929. A tentative standard for calculating the A-weighted sound level was published in the Journal of the Acoustical Society of America in 1936.

In formulating this measure the premise seems to have been that it should be (the logarithm of) a weighted sum of the intensities per frequency band x_i. In summing the intensities together, each of them was weighted with a factor b_i. The general form of such a measure is therefore $y = \sum_i b_i x_i$. Research by Fletcher and Munson [1933] showed clearly that in such an approach, the weighting scheme which is appropriate depends on the absolute level of the intensities in the different frequency bands. The A-weighting is correct for noises which sound about as loud as a 1000 Hz tone of 30 dB.

The other weighting schemes are intended for louder sounds.

3.2 Zwicker

It is possible, in addition to applying multiplicative weights to the band intensities ($b_i x_i$), to apply a power transformation to give an aggregation scheme of the type $y = \sum_i (b_i x_i)^{a_i}$. The intensity in each frequency band is exponentiated, and the results are aggregated together to produce the total loudness. This aggregation scheme conforms to the model for annoyance referred to in section 2.2.

This more general aggregation rule allows an empirical aspect of noise to be described which is not catered for by the use of a weighted sum as mentioned in the preceding section. Setting the exponents a_i to less than 1 implies that a given sound intensity will produce more loudness if it is distributed evenly over several frequency bands rather than being concentrated in a single frequency band.

This is the approach taken in Zwicker's model to allow for this phenomenon. Zwicker's procedure, however, does not take the intensities per frequency band as its base data, but the excitations of different sections of the basilar membrane corresponding to the frequency bands. Fletcher and Munson [1933, 1937] and Fletcher [1940, 1953] had already described loudness in terms which in the most important respects corresponded to the later model of Zwicker. Broadly, excitation levels are first determined for each critical band. Transformations are then applied to convert them into specific loudnesses, which are summed to give the total loudness.

The first version was described by Zwicker and Feldtkeller [1955], and later in English by Zwicker and Scharf [1965]. A complete description of the model, incorporating the results of the study on the time integration of loudness (see section 3.5), is given by Zwicker [1982]. There is a computer program for calculating loudness from ⅓-octave band spectra [Paulus and Zwicker, 1972; Zwicker, Fastl and Dallmayr, 1984], and a meter has been developed which measures specific and total loudness directly [Zwicker and Fastl, 1983].

3.3 Stevens

The model of Stevens [1955] is less an extension of the earlier developments than was the work of Zwicker. His 1955 paper contains an addendum in which he considers the differences between his own model and that of Zwicker and Feldtkeller [1955], which he only saw after completing his own paper. In this addendum he identified two important differences between his model (including the six later, adjusted versions) [Stevens, 1956, 1957, 1961, 1971] and the various versions of Zwicker's model.

One difference was the use of octave (and later ⅓-octave) bands instead of critical bands. On this point he stated that, although there were probably better grounds for using critical bands, it remained to be seen whether these could be determined precisely. He later stated [Stevens, 1957], referring to a paper of Zwicker, Flottorp and Stevens [1957], that it would indeed be better to start with critical bands, but considered that the fact that the use of octave and ⅓-octave bands had become established to be a decisive argument for retaining them.

The main difference from Zwicker's method, which Stevens also mentioned in the addendum, is the method of aggregating the contributions from the various frequency bands to the total loudness. Zwicker sums together transformations of the excitations while Stevens sums the largest contribution and a fraction of the differences between the other contributions and that maximum.

3.4 Evaluation

Zwicker's procedure incorporates information about the characteristics of perceived loudness, such as that concerning the aggregation of contributions from the different frequency bands and the importance of critical bandwidths, better than do the methods of Stevens or the A-weighted sound level. All three methods have been standardised (ISO 532, IEC 225).

An important question is whether these differences mean that Zwicker's method predicts better which noises will be perceived as louder than others. In experiments in which *loudness* was evaluated the Zwicker procedure indeed gave somewhat better results than that of Stevens, and both gave better results than the A-weighted sound level [Scharf and Hellman, 1980: a summary of the literature; Van Wyck, 1981: a round robin test; Zwicker, 1980; Fastl, 1981; Zwicker 1982a: 142-143; Zwicker, 1982b: Darbietung 5.6; Fastl, Markus and Nitsche, 1985].

Broadly speaking, however, and with some exceptions, the differences between the results of Zwicker's procedure and the A-weighted level are not great.

A study of the procedures shows which pairs of sounds exhibit the greatest difference between matching on the basis of the A-weighted level and matching in the basis of Zwicker's loudness value. Suppose that sounds are matched with pink noise. The difference between the two methods is great for sound which is highly concentrated in a narrow frequency band and for strong low-frequency sound. It transpires that for both of these types of sound, Zwicker's method correctly indicates when they sound of equal loudness to the pink noise. The A-weighted level, on the other hand, underestimates the sound level needed to give equal loudness in the case of a sound which is highly concentrated in a narrow frequency band, and overestimates it for the strong low-frequency sound. Examples of both of these types of sound are given.

Fastl [1981] shows, for example, that Zwicker's method indicates very accurately when a 1000 Hz tone is as loud as broad band white noise, whereas using the A-weighted level results in an underestimation of 15 dB(A) of the level required for matching. Zwicker [1986] gives a practical example of this. He shows that a passing moped which registers 87 dB(A), and which has a clearly peaked frequency spectrum, is less loud, using his method, than one recording 84 dB(A), but with a flatter spectrum. The equally loud 1000 Hz tone was 96 dB(A) for the first moped and 99 dB(A) for the second. The first moped gives a value using the A-weighted level 3 dB(A) higher than the second, whereas the equally loud 1000 Hz tone is 3 dB(A) lower.

Other practical examples relating to noise from aircraft, goods vehicles and mopeds, and to the effect of double glazing, are given by Zwicker [1985] and Zwicker et al [1985].

3.5 Integration time

Two temporal factors which play a role in determining perceived loudness are the integration of excitation over time and the gradual decay of the excitation, even if the external noise stops abruptly. The way these temporal aspects are dealt with is important in choosing between the settings 'slow', 'fast' and 'impulse' available on many sound level meters. The A-weighting can be used with any of these settings. The integration times associated with each of these settings (1000, 125 and 35 ms. respectively) affects the maximum reached for sounds of short duration, whereas there is no difference for constant sounds. For an impulse lasting less than 35 ms, an integration time of 35 ms will record a maximum 5.5 dB(A) higher than an integration time of 125 ms.
Stevens [1961, 1971] states that his method is only applicable to 'steady' noise. He does not consider the use of his procedure for strongly fluctuating or impulsive sounds.

Zwicker [1977] does consider temporal aspects in his model. He points out that a decision must be made as to which is done first: the aggregation over frequency bands or the integration over time. He demonstrates [Zwicker, 1969, 1974], with the help of experimental data, that aggregation over the frequency bands should precede the integration over time. Account is taken of the attenuation?? for each frequency band, before the integration over time. The question of time integration is now considered in greater detail.

Zwicker bases the integration time used in his model on two types of experiment. In the first type he investigated how the duration of step-function impulses affects their loudness. Increasing this duration up to a certain limit (the integration time) proved to increase loudness but above this level there was no further increase in loudness. In the second type of experiment the effect of an increase in the frequency of the impulse on loudness was investigated, using for example a 2 kHz impulse of 5 ms duration. It transpired that loudness only increased with impulse frequency over a certain value (the reciprocal of the integration time). Zwicker concluded, on the basis of both types of experiment, that the integration time is approximately 100 ms.

Smoorenburg [1979: 18-21] reviews the literature on this first type of experiment. He summarises the findings with regard to the integration time t with the following remarks:

> 'The values of t vary between 10 and 230 ms, and the increase in the sound level needed to compensate for a reduction in the duration by a factor of 10 to give the same loudness varies between 5 and 15 dB. These latter values are distributed around the theoretical value of 10 dB for the integrator model. No clear effect of the frequency or bandwidth on t was observed. Some researchers report that t is affected by the sound level, however: as the latter increases, t is reduced. There is also some evidence (particularly from Stevens and Hall) that when this level has fallen to just over the threshold, t rises more strongly.'

The further discussion by Smoorenburg and the data adduced by him from a very extensive worldwide round robin?? study reveals considerable uncertainty about the integration time. Some integration times outside the range 10 - 200 ms turned up in the round robin??. If a best value had to be chosen, that was 80 ms

In his book on psycho-acoustics, Moore [1982], opts for an integration time of 150 ms He points to a number of factors, such as a low level just above the threshold, which influence the integration time.

4. LOUDNESS: TEMPORAL FACTORS

The preceding chapter dealt with alternatives to the A-weighted sound level as a measure of instantaneous loudness. We now consider the aggregation of these instantaneous values over time.

As already described in section 2.2, three time steps can be distinguished. These are discussed in turn in the following sections. Section 4.1 considers the aggregation of the instantaneous values for a single event. The aggregation of the results of this over part of a day is dealt with in section 4.2. Finally section 4.3 considers the aggregation needed to give a single 24-hour value.

4.1 Event

Many exposures to noise are made up of separate sound events. In the case of traffic these consist of the passage of a car, a train, an aircraft, etc. Impulses caused by hammer blows, shooting, etc. from various stationary sources may form clearly recognisable events. At some distance from a highway, however, separate events are not, with a few exceptions, generally observed. The same applies to noise from a number of stationary sources, such as refineries, ventilators and generators. In these cases the noise during each of a series of consecutive time intervals can be regarded as the events.

The question considered here is how the instantaneous values for a single event are to be aggregated together.

In the L_{Aeq}(24h), L_{dn} and L_{etm}, the instantaneous values per event are added together. An alternative to summation, also widely used, is to take the maximum of the instantaneous values during an event. This is what happens when B is evaluated for aircraft noise.

Whether it is better to sum or to take the maximum has been studied in laboratory experiments in which test subjects are asked to compare events. The picture which emerges is that when loudness is being compared it is primarily the maximum level during an event which matters. But if questions are asked about 'noisiness' then the non-maximum levels also play a role. The instructions given to the test subjects about the aspect which is to be compared determines the measure which best predicts the result of the comparison. It is therefore conceivable, for example, that the non-maximum levels might play an even more important role than they do for noisiness if the subjects are asked to compare the quiet in two periods during which events occur.

It is therefore important to know what determines annoyance over a longer time: the loudness of events, their noisiness, or perhaps something different? It is known that very loud events, such as a noisy moped or motorcycle passing by, cause more annoyance than would be supposed from their contribution to the L_{Aeq}(24h). But experiments carried out by Kryter and Pearsons [1965] and simple examples suggest that it is not only the maximum which is of importance. As an extreme example, take an event for which the maximum level persists for a period. An event in which the plateau lasts twice as long will cause more annoyance, despite the fact that the maximum is unchanged. As was mentioned in the introduction, it was precisely considerations of this nature which caused the measure BKL to be introduced for light aviation. Unlike B, BKL is not only determined by the maximum values.

The most likely hypothesis would seem to be that the maximum levels during an event have a greater influence on annoyance than is implied by their contribution to the L_{Aeq}(24h), but not to the point that it is only the maximum value which counts. A family of measures has been proposed in the literature [see Schultz, 1982], denoted by Q, which describe this. A parameter is set such that the instantaneous values are raised to a power greater than 1, before being summed for the event. This means that the non-maximum values affect the result in a manner intermediate between a straight summation (as for example in the L_{Aeq}(24h)) and no influence at all as where only the maximum is taken (as for example in B). The aggregation rule tends to summation as the exponent tends to 1, and to the maximum as the exponent tends to infinity.

4.2 Time of day

The previous section considered the calculation of a single value per event. We now consider the second time step, the aggregation of a number of events to give a single value for all the events in a part of the day.

In the case of L_{Aeq} (24h), L_{dn} and L_{etm}, the values arrived at per event are summed over all the events within the part of the day in question, and, in the case of L_{etm}, are divided by the number of hours in the period. The most interesting alternative approach, of the measures actually used in practice, is adopted for B. In this case it is not the values for the separate events from the first time step which are summed together, but the values raised to the power □. One consequence of this is that less loud events influence the result more strongly than would be the case with simple summation.

B is based on the maximum per event. Butabove this procedure could equally be used in the second time step if event values are based on the sum of the instantaneous values during the event, as in the case of L_{Aeq}(24h). The summation of the per event values raised to the power □ would make the result more sensitive to lower values than the L_{Aeq}. In the case of the measure Q, referred to in section 4.1, the choice of a parameter leads to the per event values being raised to a power, for example □, before they are summed. The formulation of Q does not allow for different exponents to be used for, on one hand, aggregating together the instantaneous values per event and, on the other hand, aggregating the results of this over a part of a day, however. But in section 4.1 it was stated that an exponent greater than 1 appears to be the best choice for the first step, whereas in the second step, as will be explained below, an exponent less than 1 appears indicated. This is because the sub-maximum instantaneous values within an event appear to have quite a minor role, whereas when events are being aggregated together it is the other way round, and the sub-maximum events play a disproportionately large role.

The fact that sub-maximum levels cause more annoyance than is implied by L_{Aeq}(24h) has been pointed out by Finke et al [1980]. Reducing levels which lie 10 dB(A) below L_{Aeq}(24h) hardly has any effect on the L_{Aeq}(24h) value. But if, for example, the L_{Aeq}(24h) is 65 dB(A), a situation in which there are significant pauses during which it is really quiet will cause less annoyance than one in which a sound level of 55 dB(A) persists during these pauses. This favourable effect of pauses can be allowed for by applying exponents less than 1 in the second time step.

4.3 The 24-hour period

People are generally assumed to be more sensitive to night-time than day-time noise. The way in which this is expressed in noise measures can be described by looking first at L_{Aeq}(24h). For this measure, the instantaneous values per event are first summed together, and the totals per event thus obtained are then summed together for the events occurring within the relevant parts of the day. Finally, the values for the parts of the day are summed to give a single value and divided by the number of seconds in 24 hours. The preceding two sections considered the first two time steps. We now discuss the third step.

The L_{dn} makes a distinction between a day-time period (7.00 - 22.00 hours) and a night-time period (22.00 - 7.00 hours). The overall value is the weighted sum of the values for the two periods. Weights of 1 and 10 are used for the day and night values respectively.
The L_{etm} is used for most sources in the Netherlands. This distinguishes day-time (7.00 - 19.00 hours), evening (19.00 - 23.00 hours) and night-time (23.00 - 7.00 hours) periods. The values for these three periods are weighted, and the maximum taken. The weights used are 1, 3.16 and 10 respectively. (If we were working with 10 times the logarithm of values obtained for the various periods in the day, then

this would be equivalent to adding 0 (= 10 log 1), 5 (= 10 log 3.16) and 10 (= 10 log 10)). As indicated earlier, this maximum can be approximated by raising the weighted values for the three periods to a high power, summing them, and then taking the inverse power of the result. (In the case of road traffic, L_{etm} is defined somewhat differently, but this does not affect the present argument.)

In calculating B, nine separate periods in the day are distinguished. Just as for L_{dn}, a weighted sum of the results for the various periods is calculated. The weight for the period covering the middle of the day is 1, and that for most of the night is 10. The weights for the other periods of the day lie between these two values.

The available analyses of the relationship between the different measures and annoyance do not throw up one system of weighting the different periods of the day as being the best. Nor do these analyses demonstrate that the result of applying such weighting schemes accounts better for annoyance than simply leaving the weights out of the calculation, as for L_{Aeq}(24h) [Fields, 1986; Miedema 1992].

This result may be caused by the fact that there is a high correlation between the noise loads in the different periods of the day. But occasionally there may be large differences. If for example a source only emits noise at night then the L_{dn} will be 10 dB(A) higher than the L_{Aeq}(24h).

Taking L_{Aeq}(24h) as the measure for annoyance implies that a situation with an L_{Aeq} of 60 dB(A) for the night period (22 - 6 hrs) and quiet at other times is equivalent in annoyance terms to a situation with an L_{Aeq} of 60 dB(A) for the day period (9 - 17 hrs) and quiet at other times. Taking L_{dn}, however, the former situation is equivalent to a situation with an L_{Aeq} of 70 dB(A) for the day period and quiet at other times. Further analyses, with highly contrasting time distributions for the noise exposure are needed in order to ascertain whether different weights are required for the different parts of the day, and if so, what these weights should be.

The foregoing considers the weighting scheme for the periods of the day. We now turn to the exponents to be used in aggregating the contributions from different periods. An example can be taken to make it clear that simply taking the maximum of the values for the different parts of the day, as is done for L_{etm}, does not always correctly reflect annoyance. A situation X with values for L_{Aeq} for day, evening and night-time of 70, 65 and 60 dB(A) respectively will cause more annoyance than a situation Y with the same day-time level, but quiet in the evening and night. L_{etm} is equal for both of these situations, however, i.e. 70 dB(A). In the case of L_{Aeq}(24h), L_{dn} and B, the weighted sum of the contributions from the various parts of the day are taken, and situation X scores higher than situation Y. The L_{Aeq}(24h) values for the two situations are 67.7 and 67 dB(A) respectively, and the L_{dn} values are 69.5 and approximately 67 dB(A) respectively, depending to some extent on the L_{Aeq} between 22 and 23 hrs and between 6 and 7 hrs. L_{Aeq}(24h) and L_{dn} therefore treat X as causing more annoyance than Y.

4.4 Week, season and year

A distinction is made not only between the different periods of the day, but also between weekdays and the weekend, and between the summer and winter months.

Distinctions of this kind are justified by the differences in *personal exposure* in the home which occur at these various times. At present, exposure is as a rule expressed in terms of a measure at the external face of the house experiencing the greatest loading. But many people are at home more in the evenings and at night than in the day-time and more in the weekend than on working days. Noise in the evenings and at weekends therefore contributes more to the exposure at home. The summer months will make a considerable contribution to the personal exposure in the home because windows are kept open wider and more frequently at that time, and because more time is spent out-of-doors, for example in the garden. A better indicator of the personal exposure in the home could be obtained by combining the behaviour over time of the sound level on the external wall of the house with the highest noise loading with a 'standard behaviour pattern' which indicates when people are outside, when inside with the windows open and when inside with the windows closed.

Certain periods could be distinguished on the grounds that human 'activities' carried on at certain times are *more sensitive* to noise than at other times. The assumption that sleep is more sensitive to noise than day-time activities is the rationale for making a distinction between day and night. Differences in sensitivity are much more difficult to allow for quantitatively. Penalty factors are incorporated into L_{etm} and B to reflect differences in sensitivity during different periods of the day. It will be difficult to theoretically underpin penalty factors of this kind for different periods in the week and different seasons.

Apart from matters of sensitivity and personal exposure, there is a question related to the trade-off between exposure on different days. In other words whether, for a given total personal exposure within a period with a given sensitivity, the way the noise is distributed affects annoyance. For example is the annoyance greater, the same or less when it is concentrated on a single work-day as opposed to being spread out over the whole week?

Vos and Geurtsen [1992] recently reported on published data and their own work on this topic. They contend that a noise measure should not be calculated only for the period that the noise source (in their case a firing-range or air-base) is active, because this would suggest that concentrating activities would increase the annoyance caused. They suggest that, on the contrary, the concentration of noise within a limited number of days could diminish annoyance. This concentration must not go too far, however. It appears difficult to be precise about the relationship between concentration and annoyance.

In this section, three factors have been discussed which are of importance in aggregating together noise from different periods. These are: differences in the extent to which the noise contributes to the personal exposure in the home, differences in the sensitivity of activities during the different periods and the trade-off between exposure during different periods. The first factor appears to be the easiest to quantify. If this factor is brought into the calculation then sound levels during the weekend and in the summer, as measured on the external wall experiencing the greatest loading, are likely to be weighted more heavily, relative to sound levels on weekdays and in the winter (which will therefore be weighted less heavily) than is the case at present in a measure such as B.

5 FACTORS OTHER THAN LOUDNESS: FREQUENCY-RELATED FACTORS

The L_{Aeq}(24h) is based on the A-weighted sound intensities, which give an indication of instantaneous loudness. As we saw in chapter 2, there are also other measures of instantaneous loudness. Furthermore instantaneous aspects of the sound other than its loudness can affect the annoyance caused. Results from Berglund et al [1988] suggest that other aspects, in the relevant experiment the sharpness of the noise, are particularly important when loudness is low and are less important for louder noises.

Section 5.1 discusses the concept of 'noisiness'. The term is used both for instantaneous noise and for noise events. It is used to designate the aspects of the instantaneous noise or of events, in their entirety, which affect the annoyance caused by the noise, judged over a longer period. Section 5.2 deals with some of these instantaneous aspects. The temporal aspects are considered in chapter 6.

5.1 Noisiness

Loudness is not the only factor which affects annoyance. The description of noisiness given above makes it clear that loudness and noisiness are generally not the same. Given this fact, the confusion which the introduction of the concept of noisiness has caused in the world of applied psycho-acoustics is rather surprising. One source of confusion is that in evaluating loudness models, not only loudness rating experiments but also noisiness rating experiments are used. A second point, discussed further below, is that it is not clear whether Stevens' model relates to loudness or noisiness.

Stevens' original model, published in 1956, unquestionably dealt with loudness. He also stuck to the notion of loudness in Stevens [1961], in which the sixth version of the model was described. In this he refers to an article by Kryter [1959], which proposed an adjustment to the second version of Stevens' model. The proposal was simply that the specific loudness curves be replaced by so-called 'noy curves' (the 'noy' from 'annoyance'). There seems to be little theoretical basis for this modification. It is not clear what significance the idea underlying the loudness model should have for noy curves.

Stevens rejected this proposed modification in 1961, referring to results which show that loudness is somewhat different from noisiness. In Stevens' view, the 'Perceived Noise Level' (PNL, logarithmic equivalent: PNdB), as Kryter called the result of his procedure, should not be confused with loudness. Kryter [1960], and then following Stevens' 1961 paper Kryter and Pearsons [1965], claimed that their procedure was suitable for calculating loudness, however.

The confusion was further compounded when Stevens [1971] presented the seventh and last version of his model as a model for both loudness and noisiness, and adopted the designation PNdB for the calculated result, even though he did not use Kryter's noy curves in the calculations.

Experimental data show what was to be expected, namely that noisiness and loudness are sometimes the same and sometimes not, and therefore are in general not the same. This is illustrated by results of, on the one hand, Berglund, Berglund and Lindvall [1976] and Fastl and Widmann [1990], which show that for overflying aircraft there is a fixed relationship between loudness and noisiness, and on the other hand, of Lubke and Mittag [1965], who considered that the relationship between loudness and noisiness depends on the presence of high tonal components in the noise. Based on other studies, Stevens concluded the same: "Some studies show a difference, some do not." His conclusion, one sentence later, comes as something of a surprise: "In sum, the evidence suggests that a single composite weighting function should prove adequate to the needs of noise evaluation", by which he meant that the same procedure can be used for determining loudness and noisiness.
Kryter [1985: 167], on the other hand, reconsidered his earlier claim, and stated that: "The attribute of noisiness includes, in addition to loudness, some perceived effects of impulsiveness and duration of sound upon their unwantedness that are not generally perceived as a part of the attribute of loudness".

Elsewhere in his book, Kryter discusses corrections for the two factors mentioned, impulsiveness and duration, which can be incorporated into the PNL. It is noteworthy that the influence of tonality is not mentioned here, whereas Kryter had earlier proposed the Effective Perceived Noisiness Level (EPNL), in which a correction for tone was applied to the PNL for tonality (see section 5.2).

Despite the confusion which there has been it will be clear that there are instantaneous and temporal factors other than loudness which can affect annoyance. The instantaneous factors are discussed in the following section, and temporal factors are considered in chapter 6.

5.2 Tonality, high and low sound, sharpness

Broadly speaking, sound is said to be more tonal in character the more the energy is concentrated in a narrow frequency band. There are no known experiments in which test subjects' assessments of the tonal nature of sound is itself related to a measure of the concentration of sound energy. Such measures of the tonal character of sound have been related to judgements of its noisiness or annoyance value.

The results of these experiments are not conclusive. Scharf and Hellman [1980] summarise the research which has been carried out in which test subjects compared sounds with and without tonal components with each other. They concluded: "When the attribute judged is either loudness or noisiness, tonal components do not seem to alter the subjective magnitude of noise for sounds below 80 dB SPL. Above 80 dB SPL, tonal components slightly increase the noisiness of sounds".

With regard to the procedures designed to correct for the effect of tonal components, Scharf and Hellman state "None of the examined procedures designed to correct for the presence of tonal components improved the effectiveness of the descriptors to which they were applied; the variability and discrepancy between calculated and judged level either remained the same or increased". This related to the tonality correction procedures FAR 36 [FAA, 1969] and that of Kryter and Pearsons [1965], both to be applied to the PNL, and an unpublished procedure of Stevens which can be combined with his 1972 method of calculating loudness. The FAA procedure is based on Little [1961], and was developed by Sperry [1968] [see Kryter 1985: 135].

The conclusion suggested by the laboratory experiments that tonality has no effect on annoyance or noisiness runs counter to the dominant view that tonal components can increase annoyance. The results of a number of field studies suggest a relatively strong annoyance due to tonal noise, for example from trams or braking buses [see Miedema and van den Bergh, 1985, and Ericz et al, 1986]. But the review carried out by Scharf and Hellman indicates that it is difficult to demonstrate such an effect in the laboratory. This might be explained, at least in part, by the inadequacy of the measures of tonality used. Robinson [1992] recently proposed a new measure to describe the effect of tonality on annoyance, depending on a whole complex of factors, including the frequency of the tonal components.

It has been clearly demonstrated in a number of laboratory studies that the presence of a single tonal component in random noise affects noisiness or annoyance, the relationship depending on the frequency of that component [Kryter and Pearsons, 1965; Pearson, 1968: 135; Robinson, 1991]. Broadly speaking, the effect of a tonal component of constant magnitude increases as its frequency rises. This could indicate that it is more the sharpness of the noise rather than its tonality which is important. Zwicker [1982] indicates how the sharpness of a noise can be calculated. Sharpness can also play a role in non-tonal noises. Zwicker's formula indicates that the sharpness of a 'hissing' white noise is greater than that of pink noise. It is possible that this measure of the sharpness of the noise is a better predictor of the extra annoyance than measures of its tonality.

Generally speaking, sharpness increases as the high frequency components in the noise become more dominant. But low components can also be important for annoyance. It has been found that the A-weighted sound level does not predict annoyance from sonic booms [Galloway, 1990] and heavy artillery [Buchta, 1990] well. Both of these authors advocate the use of the C-weighted sound level for such noises. In the C-weighting, low-frequency components weigh more heavily than in the A-weighting scheme. It is very likely that it would be unnecessary to use a different measure if Zwicker's procedure for calculating loudness were adopted. We have already seen above that the A-weighting underestimates the loudness of the low-frequency noise of goods trucks, whereas Zwicker's procedure indicates loudness correctly.

Apart from its effect on loudness, sometimes underestimated by the A-weighting, low-frequency noise can have another effect. Low-frequency noise and inaudible vibration may cause vibrations in buildings. Vibration in the home may either be felt directly, or may lead to perceptible noises (rattling windows or cutlery). These vibrations are caused, for example, by sonic booms and heavy artillery, but also by heavy surface transport, such as trucks and goods trains, and by aircraft.

6. FACTORS OTHER THAN LOUDNESS: TEMPORAL FACTORS

Loudness during a 24-hour period is described by measures such as the L_{Aeq}(24h), L_{dn}, L_{etm} and B. But there are temporal factors which affect annoyance and which are not expressed in these measures. One such factor is sudden increases in noise. It is not only the level of loudness which is important, but also its derivative, in particular its rate of increase. This is discussed in stion 6.1. Another temporal factor, the roughness of the noise, is discussed in stion 6.2.

6.1 Increases in loudness

Pure impulse noise, impulsive noise from trams, low-flying jet fighters and sonic booms from aircraft cause more annoyance, for a given L_{Aeq}(24h), than noise from, for example, road traffic. In all these cases this can be attributed for a large part to rapid rates of increase in the noise level. High peak levels can also contribute to this (see stion 4.1), while sonic booms are also of very low frequency (see stion 5.2).

The difference in annoyance from road traffic and from impulsive noises has been demonstrated in various studies. Miedema [1992] presented certain relationships between annoyance and L_{Aeq}(24h) for road traffic and for impulse noises in a comparable form. It could be inferred that annoyance due to impulse noise is approximately equal to that due to road noise with a L_{Aeq}(24h) level 15 dB(A) higher. Trams negotiating points cause more annoyance than on the straight (Miedema and van den Berg, 1985). Miedema [1992] presented certain relationships between annoyance and the L_{Aeq}(24h) for road traffic and for impulsive noise from trams in a comparable form. For this type of tram noise the relationship is weak. There is clear evidence, however, that annoyance due to impulsive tram noise is greater than that due to road traffic noise with the same L_{Aeq}(24h) level. On the basis of this and the results referred to above for 'purer' impulse noise, it can be expected that the annoyance caused by rail and also road traffic noise will increase when there is impulse noise involving for example 'banging' and 'crashing'.

Some evidence of the difference in annoyance for low-flying jet fighters was given by De Jong [1986]. He investigated the annoyance experienced by residents living under low flight paths. Although no noise data were collected, it was clear that, in view of the small number of overflights, that the L_{Aeq}(24h) under these low flight paths was much lower than the L_{Aeq}(24h) from road traffic associated with the same annoyance.

Because aircraft are no longer permitted to break the sound barrier over the Netherlands, sonic booms will not be considered here.

Schultz [1982] discusses four measures which take account of fluctuation in the noise level, and seven measures which incorporate the rate of change.

The Noise Pollution Level (NPL) is an example of the first type. This is a weighted sum of the L_{Aeq}(24h) and the standard deviation of the instantaneous values. But there is no clear relationship between the degree of fluctuation and annoyance. This can be made clear with simple examples, without further research. This is done specifically for the NPL (and other measures) by Miedema [1985].

A model of the sond type was proposed by Muller. He proposed an additive combination of the L_{Aeq}(24h) and the root mean square of the derivative of the (A-weighted) sound level. This seems a more promising approach. Some qualifications are, however, that rate of increase appears to be more important than the rate of decrease, that the rate of increase below a certain value should have no effect, and above a certain value should have a constant effect, and that the measure is in practice not easy to calculate.

Laboratory research by Yaniv et al [1982] showed that L_{Aeq}(24h) is a better predictor of annoyance than, for example, either of the measures mentioned above.

Plotkin et al [1987] have suggested a correction for the contribution of noise events to the total L_{Aeq}(24h) which depends on the rate of increase. For increase rates below 15 dB/s there is no correction, between 15 and 30 dB/s a variable correction which depends on the rate of increase, and above 30 dB/s a constant correction of 5 dB(A). This correction does not suffer from the drawbacks mentioned above. The correction is based on the rate of change and not on the variability. Furthermore the correction applies specifically to the rate of increase, only applies over 15 dB/s, and is constant over 30 dB/s. One difficulty, however, is that the exact rate of increase in the range 15 - 30 dB/s will in practice be difficult to determine.

The approach just described is therefore attractive, but suffers from the practical problem referred to, and the question also arises whether the limits have been correctly set. As far as the latter is concerned, research by Harris [see Galloway, 1988] is relevant. This shows that the rate of increase only becomes important above 30 dB/s. By inspecting the temporal pattern of various noise events, Miedema [1992] provisionally and non-exhaustively distinguished three categories: < 10 dB/s., about 50 dB/s. and >> 1000 dB/s. The first category covers various types of passing rail, road and air traffic, with the exception of high-speed trains, the categorisation of which was unclear, and low-flying jet fighters, which fall into the second category. All impulse noises fall in the last category. Drawing on Harris' results, the upper limit of the first category can be changed to 30 dB/s. Miedema [1992] gives separate relationships between L_{Aeq}(24h) and annoyance for the first category (miscellaneous traffic) and the last category (impulse noises), with further distinctions within the first category between different types of traffic. He suggests that combinations of impulse noise and non-impulse noise be treated cumulatively (see reference for more details).

6.2 Roughness

Rapid changes in the sound level are perceived as being 'rough'. For a pure tone of 1 kHz, the maximum roughness occurs with an amplitude modulation of about 70 Hz. The roughness of amplitude-modulated pure tones depends on the frequency of the tone, the frequency of the amplitude modulation and the magnitude of the amplitude modulation. Zwicker [1982] indicates how roughness can be calculated for pure tones, and also for other sounds.

Nearby noise sources vary with regard to roughness. It is assumed that for a given loudness, noises with a high roughness cause more annoyance, but this aspect has been little researched. The evidence regarding its influence on annoyance is not consistent. A laboratory experiment by Fastl et al [1990], for example, considered the annoyance experienced as a result of various aircraft flying by. It transpired that the annoyance could be clearly explained by loudness, and that differences in roughness had no impact on annoyance.

7. SPATIAL FACTORS

The L_{Aeq}(24h) and other measures are usually determined for incident sound measured at the external wall at which the noise is greatest. In order to form an accurate picture of the acoustic situation to which someone is being subjected, it is important to know the sound levels at other locations. As well as the L_{Aeq}(24h) at the external wall experiencing the greatest loading, the difference with the external wall with the lowest noise loading (stion 7.1), and with interior noise levels (stion 7.2), may be important.

7.1 Difference between sides of dwelling with highest and lowest loading

If a house has a relatively quiet side, then (some) residents can sleep on that side of the house, ventilation can be provided by opening the windows on that side of the house, and if there is a garden on that side, it is possible to sit outside in relative quiet. Miedema [1992] has confirmed that people living in houses with a relatively quiet side report less noise annoyance. Kryter [1985] has suggested that the lack of a relatively quiet side in houses lying under flight paths partly explains the observed difference in annoyance between air traffic and local road traffic. With buildings parallel to the road, the difference in noise levels between the back and the front of the houses can be as high as 15 dB(A), while under flight paths there will be little significant difference.

7.2 Difference between outdoor and indoor noise

The assumption underlying soundproofing programmes in areas exposed to high noise levels is that this reduces annoyance. A reduction in the indoor level will have a less favourable result than an equal reduction in the outdoor level. Part of the annoyance is caused by exposure directly outside the house. People also need to be able to open windows: it will be an annoyance if windows have to remain shut, but when they are opened the effect of the soundproofing will largely disappear.

The effect of soundproofing houses will not be reflected in noise measures when they are determined in the traditional way by reference to sound levels incident on the external wall with the greatest loading. A noise measure reflecting personal exposure might give the best indication of the effect of soundproofing on annoyance.

Soundproofing reduces the contribution to personal exposure made by time spent indoors, but does not affect the time spent outdoors, for example in the garden. Soundproofing will therefore reduce the value of a measure of personal exposure, and this reduction will lie between the reduction achieved in the measure for the outdoor noise level (no reduction) and the reduction achieved for the indoor noise level. The reduction in the measure of personal exposure could provide a good indication of the extent to which house soundproofing will reduce annoyance. See also stion 4.4.

8. RATING ANNOYANCE FROM ENVIRONMENTAL NOISE

In this chapter a noise measure for a 24-hour period is formulated, based on the considerations set out in chapters 2 to 4, which is expected to correlate strongly with annoyance. The formulation is not complete. For some parameters, only a range can be given, and not exact values. An indication is also given, on the basis of the discussion in stion 4.4, of how the measure could be extended to cover a year. The effect on annoyance of acoustic factors other than the strength of the noise is also considered, based on the matters discussed in chapters 5 to 7.

Noise strength

The base data for the measure are the excitation levels in the various critical bands (see stion 3.2 for an explanation) for each moment during the 24-hour period. The weighting factors and exponents for the formulae by which the base data are aggregated together in each of four successive steps are specified below (see stion 2.2 for an explanation of the notation used).

frequency	time within event	time of event	period of the day
1; 0.23 *)	1; 1 < and < ∞	1; 0 < and < 1	day < night; 1

*) The actual formula in Zwicker's model is more complex, and approaches the formula specified here asymptotically for higher levels.

Extension to measure for the year

In order to consolidate the values for the different days into a value for the whole year, a fifth step is needed. There is little information about the parameters to be used in this step. The simplest possibility is to average the one-day values over the year, as is in principle done when L_{Aeq} or B is calculated for a year. It seems possible to allow for differences in behaviour patterns which cause house occupants to undergo greater personal exposure at weekends and on fine days than at other times. In order to achieve this the noise measure should be based on personal exposure rather than the exposure of the dwelling. It would be possible to achieve this by combining data on the noise loading of the house with an assumed standard behaviour pattern.

Factors other than noise intensity

In many situations, greater annoyance will be associated with a larger value of a noise measure as described above. The following factors complicate the relationship between this measure and the annoyance.

- Tonal components and sharp noise can cause additional annoyance (see chapter 5). These aspects seem to be interrelated. Particularly tonal components with a high frequency, which therefore cause a sharp noise, cause additional annoyance. Although there is no accepted measure for tonality, there is for sharpness. Sharp noise may occur for example where rail traffic must take tight bends, as a metallic noise for some types of helicopters, and in the form of a hissing noise in some industrial installations.
- Extra annoyance is attributable to low-frequency noise because of the vibration it can cause in the house or its contents (see chapter 6). Vibration due to nearby heavy surface transportation (lorries, goods trains) or aircraft near airports can cause annoyance. Vibration can be quantified.
- High rates of increase can cause extra annoyance (see chapter 7). The greatest rates of increase occur for impulsive noises, such as is caused by hammering, piledriving or shooting. High rates of increase in noise due to low-flying jet fighters also cause annoyance. The rate of increase in noise for high speed trains is not sufficiently different from that of other surface transportation to make extra annoyance likely.

- The exposure of a house from several directions results in extra annoyance (see stion 8). Multidirectional exposure occurs, for example, when there are roads on various sides of a house, and it also occurs where houses lie under approach or departure flight paths. It is possible that an overhead noise source has extra annoyance value because of the associated hazard.
- In the case of industrial noise, there will be incidental or special noises which contribute disproportionately to annoyance [Miedema, 1992].

As far as systematic differences between sources in relation to the above factors are concerned, annoyance is not related to the noise measure described in the same manner for all sources. It is possible to allow for some of these differences by adopting different relationships between the measure of noise intensity and annoyance for ground traffic (highway traffic, other road traffic and rail), air traffic and impulse noise.

Multidirectional exposure could be allowed for by determining the (estimated) personal exposure as described above, rather than the exposure on the external wall with the greatest noise loading.

The sharpness of noise, vibration and rates of increase in noise as occur for low-flying jet fighters can in principle be readily quantified. In general, increasing sharpness, noise surges and vibration will be associated with increasing annoyance, but the precise relationship with annoyance remains unclear for the moment.

9. CONCLUSIONS RELATING TO NOISE MEASURES AND ANNOYANCE

The measures used for environmental noise in the Netherlands and most of those used in other countries are based on the A-weighted sound level. This applies, for example, to the measures L_{Aeq}(24h), L_{dn}, BKL, L_{etm} and B. The use of the A-weighted sound level rather than Zwicker's measure for instantaneous loudness is likely to lead to a weaker relationship with annoyance. Since, for example, at higher levels the effect of low-frequency noise is underestimated, the annoyance caused by heavy vehicles can be underestimated.

This applies to all five of the measures, however, so this does not constitute grounds for preferring one of the measures over the others.

None of the measures L_{Aeq}(24h), L_{dn}, BKL, L_{etm} and B fully accords with the specification presented in chapter 8 in relation to the three steps for aggregating over time. A comparison of the specification in the preceding chapter with those in stion 2.2 for these measures shows that each of L_{Aeq}(24h), L_{dn}, L_{etm} and B departs from the chapter 8 specification in a different respect. The manner in which L_{dn} and BKL depart from it is essentially the same. The discrepancies are briefly described below (see chapter 4).

- For L_{Aeq}(24h), the weights applied for night and day in step 4 are equal. The exponent used in the aggregation scheme for step 2, i.e. 1, is too low. This means that the non-maximum noise levels during an event weigh too heavily. The exponent in step 3 is also 1, which is too high for this step. This means that the favourable effect of relatively quiet periods (lulls) is underestimated.
- The first discrepancy mentioned above does not apply to L_{dn} or BKL, but the other two do.
- L_{etm} is also subject to these two discrepancies, as well as the following. The exponent in the aggregation formula in step 4, namely ∞, is too high. This means that the influence of periodes of the day when the value is below the maximum is underestimated.
- For B there is just one discrepancy. The exponent in step 1, namely ∞, is too high. This means that the influence of the non-maximum levels during an event on the annoyance is underestimated. In other respects, B corresponds with the chapter 8 specification for a noise measure for predicting annoyance.

Because each of L_{Aeq}(24h), BKL, L_{etm} and B differs in a different way from the chapter 8 specification, it is difficult to express a preference for one of the measures.

The purpose of the present study is to indicate which noise measure should be used in setting standards for environmental noise, in particular for air traffic. The utility of noise measures depends in the first place on how well they correlate with various adverse impacts, in this case specifically on their relationship with annoyance.

We have already noted, in stion 2.1, that it is not possible to make a choice between the candidate noise measures on the basis of calculated correlations between them and annoyance. We have therefore broken the measures down into the constituent steps in their calculation, and examined each of these steps to see to what extent we can determine the best way of carrying it out. The specification arrived at in this manner is described in chapter 8. The four measures now under consideration all deviate from this specification.

The main respect in which B does not conform to the chapter 8 specification is that the maximum level is evaluated for a flyover, and that the non-maximum levels make no contribution to the value of B. This is one of the factors which led to the introduction of BKL for light aviation. This drawback associated with B is less important for residential districts around major civil airports because there is a closer relationship between, for example, L_{Amax} and L_{AX}. The exact relationship between L_{Amax} and L_{AX}

could be studied around Schiphol, for example.

Based on the foregoing, we conclude that at the moment there is no evidence to suggest that one of the other noise measures considered is more closely related to noise annoyance due to large-scale civil aviation than B. There are no appropriate correlation data based on dose-response studies which would suggest that another measure is to be preferred. For this reason we have concentrated on an analysis of the steps which underlie the various measures. This analysis shows that the only clear drawback of B is that it is based on the maximum noise level during a flyover. For large-scale civil aviation this disadvantage is of little importance. Furthermore, other measures are subject to other drawbacks.

This conclusion refers to B determined for a 24-hour period, without a lower limit for the peak values which are included in the calculation. Annual values of B are determined by averaging the day values. In view of the lack of data for assessing the aggregation formula used in this step, there appears to be no reason for departing from this simple procedure. It could be useful, however, to base B on personal exposure assuming a standard behaviour pattern, rather than the external exposure of the house.

The results of this study do not provide sufficient basis for a single measure to be proposed which could replace the various measures currently used in the Netherlands. A specification was given in chapter 8, however, of a noise measure which would in principle be the most suitable for this. But this specification is incomplete. If this specification were to be completed it would also be possible to make a quantitative comparison of different noise measures. Of the noise measures which would qualify, having regard for example to practical considerations, one measure could be chosen on the basis of this comparison if it is able to predict annoyance from different sources as well as or better than the various measures presently used. This would allow the multiplicity of noise measures now in use to be reduced.

It would be wise to first complete the specification and to perform the comparisons mentioned before going for changes in current practice. Two alternative approaches to completing the specification are possible.

The first approach involves further investigating the correlation between various noise measures and annoyance. It is important that these correlations be based on samples in which a very wide range of contrasting situations are equally represented. In order, for example, to study the need for and the magnitude of a night-time penalty factor, equal numbers of respondents could be selected, for the following four types of situation, from the large database of exposure and annoyance data available to NIPG-TNO: high night-time, high day-time; high night-time, low day-time; low night-time, high day-time; low night-time, low day-time. In these situations the correlation between, for example, L_{Aeq}(24h) (no night-time penalty) and L_{etm} (10 dB(A) night-time penalty) will be weaker than for the database as a whole, and there will be greater discrimination between the different predictors of annoyance. The idea would be to select sub-samples for which the correlations between the different measures would be as low as possible.

The second approach would involve describing, in sufficient detail, typical examples of different types of noise situation. These would be contrasting situations for which the differences in annoyance are known. In each of the following pairs of situations, for example, situation a causes more annoyance than situation b for the same L_{Aeq}(24h):

1a highway traffic
1b typical local road traffic situation
2a local road traffic with a high proportion of goods traffic
2b local road traffic with a low proportion of goods traffic
3a local traffic around traffic lights
3b local traffic, flowing

4a local traffic, accelerating
4b rail traffic

On the basis of situations of this nature it is possible to identify what the further specifications of the chapter 8 measure need to be to predict differences of this kind in annoyance.

In aggregating 24-hour values together to produce a value for the year it is also important to consider whether measures of personal exposure to noise could be estimated using standard behaviour patterns, and whether these are more closely related to annoyance than measures of the maximum noise loading on the exterior of the house. This would also provide insight into the reduction in personal exposure to be achieved by soundproofing, and the resulting reduction in annoyance.

PART II: SLEEP DISTURBANCE

10. AWAKENINGS DUE TO NOISE: GRIEFAHN'S METHOD

Sleep disturbance can be considered in somatic, in functional and in disturbance terms. A somatic factor which has been studied extensively is its effect on EEG patterns. In functional terms it may lead to reduced ability to perform tasks. And the nuisance effects relate to difficulty in getting to sleep, being awakened and the feeling of being 'fit' on awakening.

In this and the following chapter we consider awakenings as reflected by the EEG. A shortcoming of the approach is that, even where noise does not awaken, it can impair the quality of sleep as assessed by other factors. We focus here on awakenings because data are relatively abundant, and because Griefahn's method, which is currently the subject of considerable attention, relates to awakenings. Chapter 13 discusses another indicator of sleep disturbance - self-reported sleep quality - and this is related to night-time noise exposure.

Section 10.1 describes the 10% awakening curve proposed by Griefahn to assess the likelihood that people are awakened at night by noise events. Griefahn's curve is based on two components, namely an analysis of data reported in the literature and a derivation procedure. These two elements will be discussed separately in sections 10.2 and 10.3 respectively.

10.1 Griefahn's 10% awakening curve

Griefahn has presented a curve which claims to indicate the combinations of peak noise levels and numbers of events which awaken 10% of those exposed. For brevity this will be referred to as Griefahn's curve.

The peaks involved here are peaks in the A-weighted sound level, measured at the location where a person is sleeping. It therefore deals in practice with sound levels in the bedroom. It should be noted that this distinguishes it from the annoyance studies, in which it is usually the exterior sound level which is involved.

The curve covers a range from 2 to 32 events. The level is also given in the text for which 1 event during the night awakens 10% of people.

A safety margin is built in to the curve by basing it not on the 'mean' sleeping state of an 'average' person, but on the REM phase during the sleep of 71-year-olds. Griefahn assumed that the waking threshold in the REM phase is lower than in other phases of sleep, and that 71-year-olds are woken more easily by noise than younger people.

10.2 The study data

The base data for determining Griefahn's curve were taken from laboratory studies in which subjects are exposed to noise while asleep. The results were collected from a literature search by Griefahn, Jansen and Klosterkötter [1976].

On the basis of data from various studies, it was ascertained that there was a linear relationship between the peak level and the percentage of those exposed who were awakened. This percentage is then used as an estimate of the likelihood that a given peak level will lead to awakening. No account is taken of the context in which a given event occurs. Determining the percentage in this way implicitly makes the simplifying assumption that the probability of being awakened by each event is independent of other separate events. Hofman [1991] observed of the data used in the analysis that the noise events in the various studies drawn upon were very diverse in nature, and that the spread around the straight line derived by Griefahn is very wide. Hofman excluded the data which did not relate to passing aircraft or cars, and supplemented Griefahn's data, which dated from prior to 1976, with more recent data. These adjusted data relating the percentage awakened to the peak level gave a much better fit to a straight line. This new straight line indicated a higher percentage of subjects being awakened by a given peak level than Griefahn's line. This may be partially explained by the fact that Hofman included the data from the first six nights of exposure, while these were omitted by Griefahn because of the higher sensitivity during the first nights in a new situation.

Hofman [1991: pg 100] concluded that it is difficult to base a procedure of the sort proposed by Griefahn on these research results, because of the considerable residual spread in the results.

10.3 The derivation procedure

Griefahn [1990] stated that the linear relationship which she had established between the peak level and the percentage awakened is only reliable for percentages between 10 and 90%. This latter limit seems too high, but because such a level corresponds, according to the linear relationship, to very high peak levels, this is in practice less important. Although Griefahn stated that the straight line does not give reliable results below 10%, it was in fact extrapolated to 0%. This value corresponds to a peak value of approximately 60 dB(A). The suggestion is here made that, given the uncertainty, this might be a reasonable extrapolation. In itself an extrapolation of this kind may be the best or simplest description. Griefahn used this extrapolation between 0 and 10% in order to construct her 10% awakening curve.

Jansen [1992: 42], however, correctly points out that the resulting 10% awakening curve implies that the linear extrapolation to 0% is incorrect. Griefahn's curve indicates that events *below* 60 dB(A) can lead to the awakening of 71-year-olds during REM sleep. This, despite the fact that according to the linear extrapolation, at about 60 dB(A) and below, 0% of subjects should be awakened, and therefore also 0% of 71-year-olds during their REM sleep (see also point ii below). The result of Griefahn's derivation therefore contradicts an important assumption underlying this derivation. This indicates that the derivation must contain errors or inaccuracies. The following points, for example, should be noted:

i. The influence of age, phase of sleep and number of events on the percentage of subjects awakened was analysed without controlling for the peak level. Since peak level is one of the most important determinants of the likelihood of waking the analysis must either control for this variable or else ensure that the distribution of the peak levels does not differ significantly for the different ages, for example. Neither of these was done, however.

If, for example, higher peak levels were used in the experiments with older subjects, then an age effect might be caused fully or partly by this factor. The same applies for numbers of events. If high peak levels were used in the experiments with low numbers of events then it might be spuriously concluded that a night with few events results in more subjects being awakened than a night with many events.

Two consequences of this approach, for example, are:

- Figure 1 in Griefahn [1990] implicitly assumes that the likelihood of awakening is independent for each of the different events. This means that the curve in figure 4 of the same paper should be a straight line. This is not even approximately the case, however.
- Suppose that 100 20-year-olds are exposed to the noise from one flyover with a peak level of 85 dB(A). Figure 1 in Griefahn [1990] implies that 33 of these people are expected to be awakened, whereas according to figure 2 the number is 12, and according to figure 3, the number is 6.

ii. The manner of correcting for age and stage of sleep seems to assume that the relationship between the peak level and the probability of being awakened for various ages and for different stages of sleep can be described by various parallel straight lines. For a given *difference* in the percentage of 71-year-olds in REM sleep awakened, a *difference* in peak level is determined by making use of the linear relationship in figure 1. This figure 1 applies to a group of people of mixed ages, exposed during different stages of sleep. The assumption of different parallel straight lines for subgroups is inconsistent with the straight line presented in figure 1 to describe the relationship for the entire group. For the total group the awakening probability is asserted to be nil at a level where for subgroups with a higher probability of being awakened (71-year-olds or people in REM sleep) it is not nil.

iii. It is a simple matter to indicate how Griefahn's curve should be determined if we know the relationship between peak level and the probability of being awakened and if we assume, just as Griefahn must have done, that the probabilities of being awakened for various events are independent of one another (in fact an implausible assumption). Although this is not spelt out explicitly, the purpose of Griefahn's curve seems to be to indicate the combinations of numbers and (constant) peak levels for which the expected number of awakenings is equal to 10% of the subjects exposed. Because the same person may be awakened more than once, this is not the same as the curve for which 10% of those exposed are expected to be awakened, that is, that 90% are not awakened. Other possible formulae, such as a curve for the combinations for which 10% of the population are expected to be awakened are also simple to derive.

The Q% awakening curve is determined by the following equation:

$$Q = 100 \sum_{i=1}^{N} p(x_i),$$

where N is the number of events, with peak levels x_i for i = 1, 2, ..., N, and p is the function which gives the (estimated) probability of awakening for a given peak level.

Getting down to concrete figures, we will use Griefahn's function p for the probability of being awakened for 71-year-olds during the REM sleep for peak level x:

$$p(x) = (1.32x - 72.47)/100$$

Like Griefahn we consider the special case where Q = 10 and where the events have an equal peak level X. We then obtain the following formula for Griefahn's curve:

$$X = [10/(1.32N)] + 54.9.$$

Griefahn presents a quite different equation for the 10% awakening curve, however. In addition to the equation for the function p, she also uses an equation for the awakening probability for various numbers of events. Why she does this, and what the thought underlying the way in which these two equations are combined, is unclear. We can only conclude that the equation presented is incorrect.

11. COMPARISON OF GRIEFAHN'S METHOD WITH L_{etm} AND B

Despite the weak theoretical basis for Griefahn's curve, the possibility that the curve might be broadly correct cannot be excluded. Based on this possibility, this section examines the relationship between a criterion based on Griefahn's curve and one based on L_{etm} or B. For this purpose various hypothetical situations are considered.

As a first example we take a road traffic situation. For this purpose we consider only goods vehicles in the 8-hour period 23.00 - 7.00 hrs. Noise measures and numbers of vehicles only relate to goods vehicles. Assume that all goods vehicles have a speed of 50 km/h. and that they emit 103 dB(A). The situation is made more concrete by specifying the situation in some other respects, such as a 'soft' ground surface between source and dwelling and a transmission loss of 25 dB(A) in the bedroom. For any given distance from the house to the road, the peak level and the L_{AX} in the bedroom can then be calculated for each passing goods vehicle. Using this L_{AX}, the L_{Aeq}(23-7h) in the bedroom can be calculated for any number of passing goods vehicles. In this way the number of passing vehicles required, for the various distances (and corresponding L_{Amax} in the bedroom), to make L_{Aeq}(23-7h) equal to a given value can be calculated. The results of these calculations are shown in figure 11.1. Curves are drawn for the combinations of numbers and peak levels for which the L_{Aeq}(23-7h) has the values 20, 25, 30 and 35 dB(A) respectively. Griefahn's curve is also shown.

The diagram shows that a limit value for L_{Aeq}(23-7h) of 20 dB(A) is more stringent than the limit defined by Griefahn's curve when there are at least 6 vehicles passing. If less vehicles pass then Griefahn's curve is more stringent than a limit of L_{Aeq}(23-7h) = 20 dB(A).

If L_{Aeq}(night) = 20 dB(A), the L_{etm} is equal to at least 30 dB(A). This means that a limit of L_{etm} = 30 dB(A) is more stringent than Griefahn's curve where there are more than 6 passing vehicles, while otherwise Griefahn's curve is somewhat more stringent. It can similarly be seen that the higher limit value L_{etm} = 35 dB(A) is substantially less stringent, when few vehicles pass, than Griefahn's curve.

A number of points should be noted in relation to this example. The limits referred to above for indoor values of L_{etm} for road traffic have a special status in Dutch law. In principle, L_{etm} = 50 dB(A) is the legal maximum for incident noise at the exterior of a house. Furthermore houses are required to provide soundproofing of at least 20 dB(A). This therefore means that a limit of 30 dB(A) is built in under existing legislation. This suggests that there are already more stringent implicit standards applying, for more than 6 passing vehicles, than would be required by Griefahn's curve, at least for the situation described. For less than 6 passing vehicles Griefahn's curve imposes somewhat more stringent standards.

In fact the law allows greater noise loadings on exterior walls in many cases, but states that indoors the L_{etm} must remain below 35 dB(A). As we saw above, this represents a further relaxation in relation to Griefahn's curve.

The force of the example given depends on how representative it is of situations which occur in practice. A number of simplifying assumptions are made in the example, such as an equal transmission loss for the peak level and L_{AX}. The example also represents a simplified situation. No allowance is made, for example, for the possibility that two goods vehicles might pass at the same time, thus producing a higher peak without affecting L_{Aeq}(23-7h). Probably more important is that no allowance is made for the contribution from private cars. These can make a substantial contribution to the L_{Aeq}(23-7h). But Griefahn's curve is no longer applicable for different peak values or where the number of passing goods vehicles exceeds 32, so that it is no longer possible to compare standards based on the curve and on L_{etm}.
The representativeness of the example can also be explored by examining situations with other base

data. This has not been done in any detail. The results from one other example are shown in figure 11.2. The difference compared with the first example is that it is now assumed that the goods vehicles all pass with a speed of 80 km/h. This diagram can be used in the same way to compare standards based on L_{etm} with a limit based on Griefahn's curve. For a relatively small number of passing vehicles the 'severity' of standards based on Griefahn's curve exceeds those based on L_{etm}.

The following example deals with air traffic, and examines the relationship between the measure B currently used in the Netherlands for air traffic, and Griefahn's curve. To keep matters simple we consider situations in which the peak level for all flyovers is the same. The relationship between Griefahn's curve and B is simple to determine because both are based on the peak value during an event. Figure 11.3 shows the *outdoor* value of B for two different situations as a function of the combinations of numbers and peak values which lie on Griefahn's curve. In one case we have again assumed that the transmission loss in the bedroom in relation to outdoor noise is 25 dB(A), and in the other, 20 dB(A). Of the numbers/peak values data pairs, only the numbers are shown, along the horizontal axis.

The zoning limit for aircraft noise in Dutch law is set at 35 Ke. The figure shows that, for a transmission loss of 25 dB(A), below 35 Ke a limit based on Griefahn's curve would not be exceeded for more than 5 flyovers. For 5 or less flyovers, exceedance might occur, however. With the same transmission loss, a limit of B = 25 Ke and (for more than one flyover) also B = 30 Ke is more stringent than Griefahn's curve.

It is known that with a limit value of B = 35 Ke much more annoyance occurs than for L_{etm} = 55 (or 50) dB(A) for road traffic. The limit value for air traffic is therefore relatively high on the basis of annoyance. A comparison with Griefahn's curve shows that this also applies in certain circumstances in relation to the likelihood of being awakened.

The value of B which corresponds to the annoyance caused by road traffic when L_{etm} = 55 dB(A) is significantly lower than B = 35 Ke. A more consistent approach to standard-setting based on health criteria would mean that, based on the standard for road traffic for example, and taking annoyance as the criterion, the standard for air traffic should be lower. It can be seen that a limit of B = 20 Ke would mean that a standard based on Griefahn's curve would be met automatically.

Figure 11.1 The dotted lines indicate the combinations of numbers (N) and peak levels (L_{Amax}) for which L_{Aeq}(23-7h) is equal to 20, 25, 30 and 35 dB(A) respectively. The figure also indicates Griefahn's curve (extrapolated). The figure applies to goods vehicles passing with a speed (v) of 50 km/h.

Figure 11.2 The dotted lines indicate the combinations of numbers (N) and peak levels (L_{Amax}) for which L_{Aeq}(23-7h) is equal to 20, 25, 30 and 35 dB(A) respectively. The figure also indicates Griefahn's curve (extrapolated). The figure applies to goods vehicles passing with a speed (v) of 80 km/h.

Figure 11.3 The value of B *outdoors* which, for equal peaks, lies on Griefahn's curve, as a function of the number of flyovers. The two curves correspond to a transmission loss in the bedroom with respect to outdoor noise of 25 dB(A) and 20 dB(A) respectively.

12. SELF-REPORTED QUALITY OF SLEEP

Various indicators are used to reflect impaired sleep quality, such as increased awakening, sleep latencies, movements during sleep, shifts between stages of sleep, decreased feeling of fitness on waking and decreased performance in performing tasks the following day. As part of noise annoyance surveys, questions may be asked about the extent to which sleep is disturbed by a given source. Subjects may be asked for example how often they are disturbed by a given noise while asleep or going to sleep, choosing from the responses 'never', 'seldom', 'sometimes' or 'often'. This allows a great deal of data to be collected on sleep disturbance as perceived and reported by the subjects themselves.

The discussions on standards for night-time aircraft noise focused mainly on information on the relationship between exposure to noise during sleep and the probability of being awakened. Self-reported sleep disturbance played little role in this.

This chapter considers the relationship between exposure to night-time noise and self-reported sleep disturbance. Because the discussion concentrates exclusively on self-reported sleep disturbance and not any of the other indicators of sleep disturbance mentioned earlier, the qualifier 'self-reported' will often be omitted.

Section 12.1 describes the data used in the analyses, and the analytical methods. Readers seeking a general overview only can omit this section. The analysis and results are presented in section 12.2. This chapter concludes in section 12.3 with some observations on the results.

12.1 Data and methodology

The air traffic data here analyzed are a subset of a database established earlier for a variety of mobile and stationary sources. Miedema [1992] described the data and derived dose-response relationships. These primarily addressed non-specific noise annoyance, whereas the emphasis here is on sleep disturbance. The air traffic data in the database were taken from an EC study carried out in three countries in a largely comparable manner. Data were collected by questionnaire on, *inter alia*, sleep disturbance due to noise from overflying aircraft, from 1758 subjects living around Amsterdam-Schiphol, Paris-Orly and Glasgow airport. Various corresponding noise measures, including the L_{Aeq}(23-7h), were also determined by measurement and calculation.

Details of the determination of the L_{Aeq}(23-7h) are to be found in Miedema [1992] and the study reports there referred to. The questions on sleep disturbance due to aircraft noise related to people being kept awake at nights, being awakened in the morning and, in the Dutch and French studies only, being prevented from falling asleep when they go to bed. If respondents indicated that passing aircraft affected them, they were asked about the annoyance caused. (If they were not affected, this was equated to no annoyance.) The questions distinguished between working days and weekends. There were therefore 4 questions on annoyance in the British study (2 types of sleep disturbance x 2 types of day) and 6 questions in the Dutch and French studies (3 types of sleep disturbance x 2 types of day). For each of these questions respondents could choose from the same 4 responses.

Scores were assigned to the annoyance categories [see Miedema, 1992 for further details and an evaluation of the procedure]. The lower limit for the first category was set at 0 and the upper limit for the last category to 100. If the categories are assumed to be equally broad, the midpoint scores for each category are 12.5, 37.5, 62.5 and 87.5 respectively. Category scores are determined according to the general rule:

score for category i = 100(i - ½)/m,

where m is the number of categories and i = 1,..., m is the rank of the category (1 for the least annoyance, m for the most annoyance).

The four or six scores thus obtained are then aggregated together as follows. First the maximum is taken of the responses to the 2 or 3 questions on the different types of sleep disturbance, for weekdays and weekends separately. The two values thus obtained, for weekdays and weekends, are then averaged. The result was then designated the 'sleep disturbance score'. The great majority of the respondents, i.e. 1629 out of 1758, scored similarly for weekends and weekdays. 53 of the remaining respondents scored more heavily one way and 76 the other.

Scores were also assigned to the boundaries between categories, determined according to the general rule:

score for inter-category boundary i = 100i/m,

where m is the number of categories and i = 1,..., m-1 is the rank of the boundary beginning with the least annoyance.

Two methods are used to summarise the information on the relationship between L_{Aeq}(23-7h) and the responses to the questions. These methods are alternative ways of processing the same data, and can be used side-by-side.

The first method makes use of the sleep disturbance score described above. The data pair consisting of a L_{Aeq}(23-7h) and the score can be represented by plotting a point on a plane. Plotting the points corresponding to all the respondents produces a scatter diagram. A function can be sought which fits these points as closely as possible. A straight line fitted by means of linear regression is an example of a simple function.

The second method uses the scores for the boundaries between the four response categories. The exposures are grouped in classes and for each class the percentage of respondents is calculated for whom the sleep disturbance exceeds a certain limit. Percentages obtained with the limits 28, 50 and 72 are described as the percentages experiencing 'at least some sleep disturbance', 'sleep disturbance' and 'serious sleep disturbance' respectively. The limits of 28 and 72 do not correspond to the scores for the boundaries between the annoyance categories used. The former were obtained by interpolating between the percentages applying for the next lesser and greater category boundary (25/50 and 50/75 respectively). The details of this interpolation are not dealt with here, nor any of the complications arising from the fact that we are determining percentages based on responses to 4 or 6 questions, rather than a single question.

12.2 Results

Figure 12.1 shows the sleep disturbance score plotted as a function of L_{Aeq}(23-7h). The correlation coefficient is 0.28.

The minimum sleep disturbance score for a subject is 12.5. This corresponds to the midpoint value for the lowest category (least disturbance) for each of the questions on which the sleep disturbance score is based. The maximum is 87.5 (see chapter 2). A value around 12.5 for the function in figure 12.1 should therefore point to almost no sleep disturbance in the exposed population, and a value around 87.5 to serious sleep disturbance for all members of the exposed population.
Figure 12.2 shows the percentage experiencing 'at least some sleep disturbance', 'sleep disturbance' and 'serious sleep disturbance' respectively, as a function of the L_{Aeq}(23-7h).

The correlation coefficient of 0.28 for the line in figure 12.1 is not high, particularly when compared with the correlation coefficient for the non-specific annoyance in the same European study. If similar analyses are carried out of the relationship between this annoyance and the L_{etm} using data from the same study, the correlation coefficient is 0.56. It is interesting to note that for rail traffic, which like air traffic consists of noise events with intervening pauses, the correlation with sleep disturbance is also low compared with non-specific annoyance. For road traffic (highway and other road traffic) on the other hand, the correlations with sleep disturbance and with non-specific annoyance are approximately equal. The ratios between correlation coefficients mentioned by Miedema [1992] are:

(correlation coefficient between L_{Aeq}(23-7h) and sleep disturbance) / (correlation coefficient between L_{etm} and non-specific annoyance)

air traffic	0.28 / 0.56	=	**0.50**
rail traffic	0.22 / 0.40	=	**0.55**
road traffic	0.44 / 0.46	=	**0.96**

The correlation between L_{Aeq}(23-7h) and sleep disturbance for air traffic is therefore of the same order of magnitude, both absolutely and relative to the correlation between L_{etm} and annoyance, as that for rail traffic, but low compared with that for road traffic.

It is in theory possible that a measure of exposure other than L_{Aeq}(23-7h) might correlate better with sleep disturbance for both road traffic and air and rail traffic. Further analyses were carried out to investigate this.

The technique used to produce the lines in figures 12.1 and 12.2 tends at least in part to smooth any irregularities in the relationship. This is not the case in figures 12.3 and 12.4. The study areas around the airports at Amsterdam, Paris and Glasgow were divided into a total of 13 sub-areas within which the noise loading due to aircraft was assumed for the study to be similar. The L_{Aeq}(23-7h) in the 13 sub-areas differed to a greater or lesser extent. Figure 12.3 was obtained by calculating the mean sleep disturbance score for each sub-area, plotting this against L_{Aeq}(23-7h), and joining up the points thus obtained by straight line segments. This produces a correlation coefficient of 0.32.

This correlation coefficient represents an upper bound on the correlations at the individual level between the sleep disturbance score and a measure of the noise dose or a combination of such measures. Assuming that the sub-areas were indeed homogeneous in terms of their noise loadings from aircraft, this applies to all measures of the dose, including those not determined in the study. It might be possible to find a measure of the dose which is monotonically related to the sleep disturbance score and with a correlation coefficient higher than the 0.28 obtained with L_{Aeq}(23-7h), but it could not exceed the value of 0.32 for these data. The ratio to the correlation for the non-specific annoyance will therefore remain below 0.32/0.56 = 0.57, which is low compared with the ratio of 0.96 mentioned above for road traffic.

The reason for this is not clear. The table below figure 12.3 contains a row with data for each of the zones in the study. As well as L_{Aeq}(23-7h), data are given on the number of flyovers during the various periods, and the number of respondents in the zone. Numbers of flyovers are averaged over a month or longer. The 'mean' L_{AX} per overflight between 23.00 and 7.00 hrs. can be simply derived from the first two columns. It can be seen that the mean number of flyovers between 23.00 and 7.00 hrs. in no case exceeds 12.

In figure 12.4, the percentage reporting sleep disturbance in each zone is plotted against the L_{Aeq}(23-7h), and these percentages are joined up by straight line segments. In order to facilitate legibility only two, rather than three, lines are drawn.

12.3 Some observations

The results show that, starting from L_{Aeq}(23-7h) = 40 dB(A) outside sleep disturbance, and also serious sleep disturbance, occur. At present it is not possible to indicate exactly what the relationship between sleep disturbance and the L_{Aeq}(23-7h) looks like. This is particularly difficult to indicate for the lower exposure levels. Further analysis will be required. A point requiring attention in this regard is also the different ratios found of the correlation between disturbance and L_{Aeq}(23-7h) and the correlation between non-specific annoyance and the L_{etm}: for air and train traffic somewhat over 0.5, and for road traffic almost 1. An explanation for this may be important in order to set standards for night-time noise related to sleep disturbance.

Figure 12.1 Air traffic: sleep disturbance score as a function of L_{Aeq}(23-7h)

Figure 12.2 Air traffic: the percentage of those exposed reporting sleep disturbance as a function of L_{Aeq}(23-7h). The upper, middle and lower lines correspond to the categories 'at least some sleep disturbance', 'sleep disturbance' and 'serious sleep disturbance' respectively.

Figure 12.3 Air traffic: the sleep disturbance score as a function of L_{Aeq}(23-7h).
The table below gives, for each value of L_{Aeq}(23-7h) in the figure, additional data on the numbers of flyovers per time period, and the numbers of respondents in the zones with the corresponding aircraft noise level.

L_{Aeq}(23-7hrs)	N(23-7hrs)	N (23-24 hrs)	N (0-6 hrs)	N (6-7 hrs)	No. respondents
19.5	1	0	0	1	202
40.2	4.1	0.6	1.1	2.4	100
42.9	5	0	3	2	100
46.1	3	0.6	1.4	1	209
46.6	1.4	0.4	0.5	0.5	101
47.9	5	1	2	2	69
50.3	1.4	0.4	0.5	0.5	123
51.5	4.1	0.6	1.1	2.4	97
56	4.1	0.6	1.1	2.4	160
57.1	8	1	3	4	35
57.2	4.8	1	2.9	0.9	203
59.4	11.8	2	9	0.8	195
60.7	9	1	3	5	164

Figure 12.4 Air traffic: the percentage of those exposed reporting sleep disturbance as a function of L_{Aeq}(23-7h). The upper and lower lines correspond to the categories 'at least some sleep disturbance' and 'serious sleep disturbance' respectively.The table below gives (as for figure 12.3), for each value of L_{Aeq}(23-7h) in the figure, additional data on the numbers of flyovers per time period, and the numbers of respondents in the zones with the corresponding aircraft noise level.

L_{Aeq}(23-7hrs)	N(23-7hrs)	N (23-24 hrs)	N (0-6 hrs)	N (6-7 hrs)	No. respondents
19.5	1	0	0	1	202
40.2	4.1	0.6	1.1	2.4	100
42.9	5	0	3	2	100
46.1	3	0.6	1.4	1	209
46.6	1.4	0.4	0.5	0.5	101
47.9	5	1	2	2	69
50.3	1.4	0.4	0.5	0.5	123
51.5	4.1	0.6	1.1	2.4	97
56	4.1	0.6	1.1	2.4	160
57.1	8	1	3	4	35
57.2	4.8	1	2.9	0.9	203
59.4	11.8	2	9	0.8	195

60.7	9	1	3	5	164

Figure 12.5 Rail traffic: sleep disturbance score as a function of L_{Aeq}(23-7h)

Figure 12.6 Rail traffic: the percentage of those exposed reporting sleep disturbance as a function of L_{Aeq}(23-7h). The upper, middle and lower lines correspond to the categories 'at least some sleep disturbance', 'sleep disturbance' and 'serious sleep disturbance' respectively.

Figure 12.7 Highway and other road traffic: the sleep disturbance score as a function of L_{Aeq}(23-7h).

Highway **Other road traffic**

Figure 12.8 Highway and other road traffic: the percentage of those exposed reporting sleep disturbance as a function of L_{Aeq}(23-7h). The upper, middle and lower lines correspond to the categories 'at least some sleep disturbance', 'sleep disturbance' and 'serious sleep disturbance' respectively.

Highway **Other road traffic**

13. CONCLUSIONS WITH REGARD TO NOISE MEASURES AND SLEEP DISTURBANCE

Griefahn's method

Using data from various studies, Griefahn has constructed a curve representing the relationship between exposure to noise and awakenings. Because such work is of great practical importance, it can expect to be subjected to considerable scrutiny and critical attention. This applies particularly when, as in this case, it is one of the first attempts to propound a synthesis in a given area. Whether or not the original proposal manages to hold its own, a synthesis of this nature serves to initiate the process of organising many separate facts into a practical result.

Griefahn's proposal in fact appears not to withstand scrutiny. Two elements in the construction of her 10% awakening curve need to be distinguished: the base data set and their analysis, and the consequent derivation of the curve. It was indicated in section 10.2 that the base data used were very heterogeneous, and were only very partially represented by the straight line constructed by Griefahn. In section 10.3 it was also pointed out that the procedure followed to derive the curve is incorrect. These points together mean that the theoretical basis for Griefahn's curve is weak. We shall briefly discuss the scope for developing a more theoretically sound Q% awakening curve.

If the corrections for age and phase of sleep are left out of the analysis, it is a simple matter to indicate the correct procedure for deriving a Q% awakening curve (see section 10.3). The effect of the simplifying assumption made that the probabilities of being awakened by the various events in one night are independent of one another could be studied further. In addition, the correctness of the resulting Q% awakening curve will depend greatly on the description of the relationship between the peak level and the probability of awakening. Hofman has derived a relationship for relatively homogeneous noises, which in her opinion however, is not sufficiently accurate to serve as the basis for a Q% awakening curve.

For the moment it remains an open question whether it will be possible to establish a sufficiently accurate relationship between peak levels and the probability of being awakened. The following two related points in particular need to be addressed. The first is whether or not the results of the first nights in the laboratory should be included in determining the relationship. One reason why Griefahn omitted them was that people wake more quickly in a new situation than in one with which they are familiar. Hofman, on the other hand, included these nights in her analysis. The second point relates to how representative the laboratory results are, even after some nights of habituation, of the normal situation in the subject's own bedroom.

Hofman and De Jong [1993], although not establishing a complete relationship between the peak level and the probability of being awakened, succeed in characterising this relationship in two respects. Based on laboratory results including the first nights, they suggest that peaks of less than 35 dB(A), independent of the context, do not lead to awakening, and that starting from 55 dB(A) the probability of being awakened rises clearly with the peak value.

There are therefore few grounds for using Griefahn's curve to set limit values for night-time noise. If B is being used as the measure of annoyance, the simplest approach would be to use the part of B relating to the night period as a measure of sleep disturbance. Where annoyance is being measured by L_{etm} or L_{dn} then it would be easiest to use the L_{Aeq} for the night period as the measure of sleep disturbance. Simplest of all would be if a single limit value in terms of B, L_{etm} or L_{dn} were to suffice for both annoyance and sleep disturbance. This is considered further below.

B and sleep disturbance

In principle, a value of B corresponds to a certain degree of annoyance. The question is whether there is a clear-cut relationship between the part of B relating to the night period and the disturbance of sleep during the night. It can be seen in figure 11.3 that, relative to Griefahn's procedure, B is rather less determined by the loudness of the events and rather more by the number of events. In view of the considerations discussed by Hofman and De Jong [1993] at the end of their report on the effect of the number of night flights on sleep, it appears that the trade-off between peak level and numbers of flights in relation to awakening characteristics is better in B than in Griefahn's curve. In any case there is no reason to think that B, calculated on the basis of night-time noise only, is less closely related to awakenings than Griefahn's curve.

We now consider the relative stringency of limit values based on B and on Griefahn's curve respectively. The relationship between B and annoyance has been studied in situations with different distributions between day and night-time flights. At the zoning limit for aircraft noise, where B = 35 Ke, the annoyance is considerable. With less than 5 night flyovers with a relatively high noise level the standard based on Griefahn's curve can be exceeded. With more than 5 night-time flyovers and a transmission loss of 25 dB(A) this standard cannot be exceeded without also exceeding the limit B = 35 Ke, irrespective of day-time noise. If there are also day-time flights then the room for exceeding Griefahn's curve is even more limited. It should be stated that the great majority of people in the Netherlands prefer to sleep with the window open, particularly in summer, in which case a transmission loss of 25 dB(A) will not be achieved.

The value of B for air traffic which in nuisance terms is comparable to the zoning limit L_{etm} = 50 dB(A) for road traffic lies well below 35 Ke. If B = 20 Ke, annoyance is very slight. Figure 11.3 shows that, for a transmission loss of 20 dB(A) and with B at this level, a standard based on Griefahn's curve would be comfortably complied with. This applies even with zero day-time exposure, and therefore *a fortiori* if there is also day-time aircraft noise. Where there is an open window in a bedroom then, depending on the precise circumstances, a transmission loss of about 15 dB(A) can be assumed. In this situation, Griefahn's curve would not be exceeded if there are more than three night-time flyovers, and assuming a limit of B = 20 Ke applied. With less flyovers an exceedance could occur.

To sum up briefly, for B = 20 Ke annoyance would be minimal, and allowing for a transmission loss of at least 20 dB(A), a standard based on Griefahn's curve would always be met, even if there were no day-time flights. Even with a partly open window compliance with Griefahn's standard would be achieved if there were more than 3 night-time flyovers. With a value B = 35 Ke, annoyance is considerable, but the standard will still be complied with if the transmission loss is 25 dB(A) and there are more than 5 flyovers. One problem is that it will be necessary to sleep with the windows shut to achieve this level of sound reduction.

L_{dn} and sleep disturbance

The L_{dn} value for air traffic which is comparable in annoyance terms with the zoning limit for road traffic of L_{etm} = 50 dB(A) is less than 50 dB(A), at about 45 dB(A). For L_{dn} = 45 dB(A), annoyance due to aircraft noise is slight. If L_{dn} = 45 dB(A) then the L_{Aeq}(23-7h) is equal to 40 dB(A) at most. Figures 12.1 to 12.4 show that for night-time loadings at this level, self-reported sleep disturbance would be slight.

For higher values of L_{dn}, annoyance will be greater. When L_{dn} is greater than 45 dB(A), the value of L_{Aeq}(23-7h) can be greater than 40 dB(A). Figures 12.1 to 12.4 show that the self-reported sleep disturbance increases as L_{Aeq}(23-7h) increases. When the L_{dn} is at a level at which annoyance corresponds with the zoning limit B = 35 Ke, the L_{Aeq}(23-7h) can be high enough to cause considerable self-reported sleep disturbance.

14. CONCLUSIONS

There is no evidence that an *existing* measure other than B would be a better indicator of annoyance due to aircraft take-off and landing activities around major civil airports. There are two aspects of B which are probably not ideal on technical grounds. Firstly, B uses only the maximum noise level during a flyover, whereas annoyance is also influenced by the duration of the exposure. And secondly, it is questionable whether the penalty factors for the different times of day adequately reflect the differences in sensitivity to noise during the different periods.

As we have already seen, if a 24-hour limit value in terms of B were set at a level at which the annoyance is slight, than allowing for 20 dB(A) sound transmission loss indoors, this would always be more stringent than a limit value based on Griefahn's curve.

As far as a limit value is concerned, a target value for B can be set which corresponds to a similar, existing value for road traffic, for example. The value of B which corresponds to L_{etm} = 50 dB(A) for city traffic is well under 35 Ke. At 20 Ke annoyance is only slight and, assuming noise transmission loss of at least 20 dB(A), a standard based on Griefahn's curve would be comfortably met.

Noise measures such as the L_{dn} (US), L_{den} (Denmark) and BKL (Netherlands), which are very closely related one another are, like B, based on the A-weighted noise level, but use the L_{AX} per event rather than the L_{Amax}. For the moment there is not sufficient evidence to choose between these measures or B. While for B the non-maximum values during a flyover have too little (i.e. no) influence, they probably have too much influence for measures based on L_{AX} because they are weighted as strongly as the peak level. Furthermore, a change in the L_{AX} probably has too much effect on measures such as L_{dn}, L_{den} and BKL as compared with the number of flights. It seems likely that a doubling of the number of flights will increase annoyance more than an increase of 3 dB(A) in the L_{AX}.

If a measure such as L_{dn} were used, a target value for L_{dn} could be set which corresponds to a similar, existing value for road traffic, for example. The L_{dn} value for air traffic which corresponds in terms of annoyance with L_{etm} = 50 dB(A) for city traffic is 45 dB(A). At this level, the annoyance caused by air traffic is therefore slight. A limit of L_{dn} = 45 dB(A) implies a limit of L_{Aeq}(23-7) = 40 dB(A) for night-time noise. Figures 12.1 to 12.4 show that self-reported sleep disturbance is limited at that level.

This means that in formulating a target value, both annoyance and sleep disturbance can be covered by a single measure, and that a single value will suffice. If additional insight is needed into disturbance during specific periods of the day, such as night-time, for the purpose of taking appropriate measures for example, then the contribution of that period to the measure can be determined. For the region between a target value for air traffic of, for example, B = 20 Ke or L_{dn} = 45 dB(A), and a maximum permissible value for B or L_{dn}, trade-offs can be made between health-based and other, for example economic, factors, as they are for other sources. If a choice is made, because of these other factors, to permit higher noise levels, then some annoyance and sleep disturbance will occur in consequence. This is then more a consequence of the decision to accept higher loadings than of the choice of measure.

In order to help determine which measure should replace the presently used measures in the context of a rationalisation, a number of steps are necessary, related to:

- **the harmonisation of measures for the different sources**
  At the end of Part I, an indication is given of the research needed to further reduce the uncertainties in specifying an optimum measure. There tends to be a resurgence of interest in this type of research whenever the most appropriate measure for a new source is being discussed. This is the case for helicopters at present, and this question could be addressed further in the relevant study planned. Another matter which needs to be addressed in further

research is the way in which 24-hour values should be aggregated together to give a value for the year.

- **the harmonisation within the EC of noise measures for air traffic**
  In chapter 2 the relationships between noise measures in use in the Netherlands are indicated by means of a tree structure. It would be a relatively simple matter to extend this tree to include the measures used in the EC. This tree would then provide insight into the interrelationships between the measures used in the EC. By progressing through the tree from top to bottom it is possible to decompose the decision about a common measure into a series of rational sub-decisions. This will provide a clearer and more systematic framework for the discussion than if the various measures are simply placed side by side in their totality. The various properties of an ideal measure and pragmatic considerations discussed above can be applied to the various sub-decisions.

- **the harmonisation of measures for noise annoyance and sleep disturbance**
  The relationship between noise measures and annoyance was studied in Part I. Part II considered the relationship with sleep disturbance. There are various indicators of sleep disturbance, and two of these, the probability of being awakened and self-reported sleep quality were considered here. The question as to whether a single measure for a 24-hour period could be adopted for both annoyance and sleep disturbance by aircraft noise has been answered, on the basis of *present* knowledge, in the affirmative. It has also been concluded that a single target value would suffice. There are various gaps in present knowledge, however, and further research into sleep disturbance is desirable. This research would explore the relationships between various noise measures and various indicators of sleep disturbance. This research could be appropriately accommodated within the research programme which is expected to be established in connection with the expansion of Schiphol.

REFERENCES

ADVIESCOMMISSIE GELUIDHINDER DOOR VLIEGTUIGEN (ADVISORY COMMITTEE ON AIRCRAFT NOISE). Geluidhinder door vliegtuigen (Aircraft noise). Delft: 1967.

BAUCH H. Die Bedeutung der Frequenzgruppe für die Lautheit von Klangen. Acustica 1956; 6: 40-5.

BERANEK, MARSHALL, CUDWORTH, PETERSON. J Acoust Soc Am 1951; 23: 261-9.

BERGLUND B, BERGLUND U, LINDVALL T. Scaling loudness, noisiness and annoyance of aircraft noise. J Acoust Soc Am 1975; 57: 930-4.

BERGLUND B, BERGLUND U, PREIS A, RANKIN K. Does equal loudness mean equal annoyance? In: Berglund B, Berglund U, Karlson J, Lindvall T. (eds) Noise as a Public Health Problem. Stockholm: Swedish Council for Building Research, 1988; Volume 3.

BITTER C. Geluidhinder door vliegtuigen (Aircraft noise nuisance). In: Lammers CJ et al (eds) 'Menswetenschappen Vandaag' (Social Sciences Today). Meppel: Boom, 1972: 251-68.

BITTER C. Beleving van geluidwerende voorzieningen tegen vliegtuiglawaai in de woonsituatie (Subjective evaluation of sound-proofing against aircraft noise in the home). The Hague: Ministry of Housing, Spatial Planning and the Environment, 1980. ICG report LL-HR-14-03.

ERICZ WJ, NOORDAM A, SCHOONDERBEEK W. Trollificering van buslijn 9 in Arnhem: onderzoek naar effecten van geluidhinder (Converting bus route 9 to trolley service in Arnhem: study of the effects of noise nuisance). The Hague: Ministry of Housing, Spatial Planning and the Environment , 1986. Noise series GA-HR-12-1.

FASTL H. Methodenvergleich zur Lautheitsbeurteilung. In: Schick A (ed). Akustik zwischen Physik und Psychologie, Stuttgart: Klett-Cotta, 1981.

FASTL H, MARKUS D, NITSCHE V. Zur lautheit und Lästigheit von Fluglärm. In: Proceedings of DAGA '85 Bad Honnef: DPG-Verlag: 227-230.

FASTL H, WIDMANN U. Technical note: subjective and physical evaluation of aircraft noise. Noise Control Engineering J 1990; 35: 61-4.

FELDTKELLER R, ZWICKER E. Das Ohr als Nachrichtenempfanger, Stuttgart: Hirzel, 1956.

FIELDS JM. The relative effect of noise at different times of day: an analysis of existing survey data. Langley (Virg.): NASA Contractor Report 3965, 1986.

FLETCHER H. Auditory patterns. Rev Mod Physics 1940; 12: 47-65.

FLETCHER H. Speech and hearing in communication. New York: Van Nostrand, 1953.

FLETCHER H, MUNSON WA. Loudness, its definition, measurement and calculation. J Acoust Soc Am 1933; V: 82-106.

FLETCHER H, MUNSON WA. Relation between loudness and masking. J Acoust Soc Am 1937; 9 no. 1: 1-11.

GALLOWAY WJ. Critical issue - military aircraft. In: Berglund B, Lindvall T: Noise as a Public Health Problem. Stockholm: Swedish Council for Building Research, 1990; Volume 5.

GATES BG. Discussion appended to paper by Churcher and King. J Inst Elec Engineers 1037; 81: 85-6.

GEZONDHEIDSRAAD: Committee on Sleep Disturbance. Vliegtuiglawaai en slaap (Aircraft noise and sleep). Health Council, 1991. Publication 1991/05.

GRIEFAHN B. Präventivmedizinische Vorschläge für den nächtlichen Schallschutz. Zeitschrift für Lärmbekämpfung 1990; 37: 7-14.

GRIEFAHN B, JANSEN G, KLOSTERKÖTTER. Zur Problematik lärmbedingter Schlafstörungen: eine Auswertung von Schlaf-Literatur. Berlin: Umweltbundesamt, 1976.

GRIFFITHS ID, LANGDON FJ, SWAN MA. Subjective effects of traffic noise exposure: Reliability and seasonal effects. J Sound Vibr 1980; 71(2): 227-40.

HOFMAN WF. Vliegtuiglawaai, slaap en gezondheid (Aircraft noise, sleep and health). The Hague: Health Council 1991 report A91/01.

HOFMAN WF, JONG, RG de. The rol van de aantallen passages bij gezondheidseffecten van nachtvluchten (The role of the number of flyovers in the health effects of night flights). Leiden: NIPG-TNO, 1993. Publication no. 93.012.

JANSEN G. Ärtzliche stellungname zu Fluglärmbedingten Störwirkungen des natürlichen Nachtschlafes. Essen, 1991.

JONG RG de. De bruikbaarheid van kosten-eenheid en bitter-index voor het bepalen van dosis-effectrelaties bij militaire luchtvaart (The practicality of the Kosten unit and the Bitter index in determining dose-effect relationships for military aviation). Delft: IMG-TNO, 1983, report D72.

JONG RG de. Geluidhinder onder laagvliegroutes in Overijssel (Noise nuisance under low flight paths in Overijssel). Leiden: NIPG-TNO 1986. Publication no. 86024.

JONG RG de, GROENEVELD Y. Geluidhinder rond militaire vliegvelden: een nieuwe benadering (A new approach to noise around military airfields). Delft: IMG-TNO 1983, report D77.

KRYTER KD. The meaning and measurement of perceived noise levels. Noise Control 1960; 6(5): 12-27. Addendum: Noise Control 1961; 7(2): 48.

KRYTER KD. Scaling human reactions to the sound from aircraft. J Acoust Soc Am 1964; 14(1): 1415-29.

KRYTER KD. Community annoyance from aircraft and ground vehicle noise. J Acoust Soc Am 1982; 72: 1212-42.

KRYTER KD. Response of KD Kryter to modified comments by Th J Schultz on KD Kryter's paper 'Community annoyance from aircraft and ground vehicle noise' J Acoust Soc Am 1983; 773: 1066-8.
KRYTER KD. The effect of noise on man. New York (etc.): Academic Press, 1985.

KRYTER KD, PEARSON KS. Judged noisiness of a band of random noise containing an audible pure tone. J Acoust Soc Am 1965; 38: 106-12.

LITTLE JW. Human response to jet engine noises. Noise Control Shock Vib 1961; 7(3): 11-3.

LÜBCKE E, MITTAG G, PORT E. Subjektive und objektive Bewertung von Maschinengeräuschen. Acustica, 1964; 14: 105-14.

MIEDEMA HME. Lawaai op arbeidsplaats (Noise in the workplace). The Hague: Ministry of Housing, Spatial Planning and the Environment, 1984. ICG report LA-DR-08-01.

MIEDEMA HME. Response functions for environmental noise in residential areas. Leiden: NIPG-TNO, 1992. Publication no. 92021.

MIEDEMA HME. Quantification of environmental quality: foundation and application with respect to sound and odour. Leiden: NIPG-TNO, 1993. To be released as Publication no. 92022.

MIEDEMA HME, BERG R van den. Hinder door geluid van tram- en wegverkeer (Annoyance by tram and road traffic noise). The Hague: Ministry of Housing, Spatial Planning and the Environment, 1985. Noise Series GA-HR-08-04.

MINTZ F, TYZZER FG. A loudness chart for octave band data on complex sounds. J Acoust Soc Am 1952; 24: 80-2.

MOORE BCJ. An introduction to the psychology of hearing. London: Academic Press, 1982.

PAULUS E, ZWICKER E.Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln. Acustica 1972; 27: 254-66.

PEARSONS KS. Assessment of the validity of pure tone corrections to perceived noise level: progress of NASA research relating to noise alleviation of large subsonic jet aircraft. 1968; NASA SP-189: 573-86.

PEETERS AL, DEVENTER FWJ van. Geluidhinder rondom vliegvelden voor de kleine luchtvaart (Noise nuisance around general aviation aerodromes). The Hague: Ministry of Public Health and Environmental Protection, 1981. ICG report LL-HR-16-03.

PLOTKEN KJ, SUTHERLAND L, MOLINO J. Environmental noise assessment for military training routes: Recommended noise metric. Ohio: Wright-Patterson Air Force Base, 1987. Report no. AAMRL-TR-001.

ROBINSON DW. Annoyance due to discrete tones in broadband background noise. Southampton: ISVR 1992. Technical Report part II. No 202.

ROBINSON DW, WRIGHT O. Annoyance due to discrete tones in broadband background noise. Southampton, ISVR 1991. Technical report Part I no. 201.

SCHARF B, HELLMAN R. How best to predict human response to noise on the basis of acoustic variables. In: Tobias JV, Jansen G, Dixon Ward W (eds.). Proceedings of the Third International Congress on Noise as a Public Health Problem. Rockville: The American Speech-Language Hearing Association, 1980. ASHA Reports 10.

SCHULTZ THJ. Synthesis of social surveys on noise annoyance. J Acoust Soc Am 1978; 64: 377-405.

SCHULTZ THJ. Community noise rating. 2nd ed. London (etc.): Applied Science Publishers, 1982.

SCHULTZ THJ. Comments on KD Kryter's 'Community annoyance from aircraft and ground vehicle noise'. J Acoust Soc Am 1982; 72: 1243-52.

SMOORENBURG GF. Voorlopige evaluatie van de geluidhinder van schietinrichtingen (Provisional evaluation of noise annoyance due to firing ranges). Leidschendam: Ministry of Public Health and Environmental Protection, 1979. ICG report BG-HR-10-01.

SPERRY WC. Aircraft noise evaluation. Washington DC: US Federal Aviation Administration. 1968; Report FAA-NO-68-34.

STEVENS SS. The measurement of loudness. J Acoust Soc Am 1955; 27: 815-29.

STEVENS SS. Calculation of the loudness of complex noise. J Acoust Soc Am 1956; 28: 807-32.

STEVENS SS. Calculating loudness. J Acoust Soc Am 1957: 11-22.

STEVENS SS. Procedure for calculating loudness: Mark VI. J Acoust Soc Am 1961; 33(11): 1577-85.

STEVENS SS. Power-group transformations under glare, masking and recruitment. J Acoust Soc Am 1965; 4: 726-35.

STEVENS SS. Perceived level of noise by Mark VII and Decibels (E). J Acoust Soc Am 1971; 51: 575-601.

TERHARDT E. On the perception of spectral information in speech. In: Creutzfeld O, Scheidch H, Schreiner C. Hearing Mechanisms and Speech. Berlin: Springer: 281-91. EBBS Workshop, Göttingen, 1979: 26-8.

US FEDERAL AVIATION ADMINISTRATION. Noise standards: aircraft type certification. Federal Aviation Regulations, 1969; Federal Aviation Regulation, Part 36.

VOS J, GEURTSEN FWM. Geluidhinder rond onregelmatig gebruikte schietbanen: een studie naar het verwachte effect van verschillende tijdschema's (Noise nuisance around irregularly used firing ranges: a study of the expected effects of different timetables). Soesterberg: IZF-TNO, 1992. Report no. IZF 1992 A-13.

WYK AJ van. A comparison of measurement methods for assessing human perception of loudness: an international survey. Acustica 1981; 49: 33-46.

ZWICKER E. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J Acoust Soc Am 1961; 33(2): 248.

ZWICKER E. Der Einfluß der zeitlichen Struktur von Tönen auf die Addition von Teillautheiten. Acustica 1969; 21: 16-25.

ZWICKER E. Loudness and excitation patterns of strongly frequency modulated tones. In: Moskowitz HR, Scharf B, Stevens JC (eds.) Sensation and Measurement: Papers in honor of SS Stevens. SL, S.N. 1974: 325-35.

ZWICKER E. Procedure for calculating loudness of temporally variable sounds. J Acoust Soc Am 1977; 62(3): 675-82.

ZWICKER E. Weniger L_a = Größere Lautstärke? In: Fortschritte der Akustik. Berlin: VDE, 1980: 159-62.

ZWICKER E. Psychophysics of hearing. Noise Pollution. Scope 24 1986: 147-67.

ZWICKER E, FASTL H. A portable loudness meter based on ISO 532B. In: Proceedings of 11th ICA, Paris, 1983.

ZWICKER E, FASTL H, DALLMAYZ C. BASIC Program for calculating the loudness of sounds from their _-octave band spectra according to ISO 532B. Acustica 1984; 55: 63-7.

ZWICKER E, FELDTKELLER R. Über die Lautstärke von gleichformigen Geräuschen. Acustica 1955; 5: 306-16.

ZWICKER E, STEVENS SS, FLOTTORP G. Critical bandwidth in loudness summation. J Acoust Soc Am 1957; 29(5): 548-57.

ZWICKER E, TERHARDT E. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. J Acoust Soc Am 1980; 68(5): 1523-5.

- page left blank -

ANNEX : DEFINITIONS OF NOISE MEASURES

The following are the definitions customarily used in the literature. An excellent summary of most of the measures defined here, and many others, can be found in Schultz [1982].

(instantaneous) sound level
The sound level, L, is expressed in dB and defined as:

$$L = 10 \log \Sigma_i I_i,$$

where I_i is the noise intensity in frequency band i, expressed in 10^{-12} W/m^2, integrated over a short time period. Most noise meters permit integration over 100 ms ('fast') and 1 s ('slow').

(instantaneous) A-weighted sound level
The A-weighted sound level, L_A, is expressed in dB(A) and defined by:

$$L_A = 10 \log \Sigma_i A_i I_i,$$

where I_i is the noise intensity in a _-octave or octave band, expressed in 10^{-12} W/m^2, integrated over a short time period. Most noise meters permit integration over 100 ms ('fast') and 1 s ('slow'). The A_i are the standard A-weights.

total A-weighted sound level for an event
The A-weighted sound level for an event (sound exposure level), L_{AX} or SEL, is expressed in dB(A) and defined by:

$$L_{AX} = SEL = 10 \log \Sigma_t \text{antilog}[L_A(t)/10],$$

where $L_A(t)$ is the A-weighted sound level at time t during a single event.

maximum A-weighted sound level for an event
The maximum A-weighted sound level, L_{Amax}, is expressed in dB(A). It is defined, for an event, as the maximum of the A-weighted sound levels during the event. Note that L_{Amax} determined with the meter in the 'fast' position is at least equal to, but often higher than, the maximum found when the meter is set to 'slow'.

A-weighted equivalent sound level for the period T
The A-weighted equivalent level for a period of duration T seconds, $L_{Aeq}(T)$, is expressed in dB(A) and defined by:

$$L_{Aeq}(T) = 10 \log \{ \Sigma_t \text{antilog}[L_A(t)/10]/T\},$$

where $L_A(t)$ is the A-weighted sound level at time t within the period of duration T. Note that:

$$L_{Aeq}(T) = 10 \log \{ \Sigma_i \text{antilog}[L_{AX,i}/10]/T\},$$

where $L_{AX,i}$ is the A-weighted 'sound exposure level' for event i.

A-weighted equivalent sound level for 24 hours and for a year
The A-weighted equivalent sound level for 24 hours, $L_{Aeq}(24h)$, is obtained from the above definition by setting T equal to 24 x 60 x 60 = 86,400 secs. For $L_{Aeq}(year)$, T is equal to 365 x 86,400 secs.

A-weighted day-(evening)-night average level and related measures for a year

The A-weighted day-evening-night average level, L_{den}, or day-night average level, L_{dn}, is expressed in dB(A), and is defined by:

$$L_{d(e)n} = 10 \log \{ \Sigma_i w_i \times (T_i/T) \times \text{antilog}[L_{Aeq}(T_i)/10] \},$$

where $L_{Aeq}(T_i)$ is the A-weighted equivalent sound level for period i of the day, duration T_i expressed in seconds, and T = 86,400 secs. The w_i are weighting factors depending on the period of the day. For L_{dn}, a distinction is made between day-time: 7 to 22 hrs. ($w_i = 1$) and night-time: 22 - 7 hrs. ($w_i = 10$). For L_{den}, day-time (7 to 19 hrs, $w_i = 1$), evening (19 to 22 hrs, $w_i = 3.16$) and night-time (22 to 7 hrs, $w_i = 10$) are distinguished.

Both of these 24-hour measures can be readily modified to give a measure for the year by setting $L_{Aeq}(T_i)$ to the A-weighted equivalent sound level for day i, and summing over the days in the year. Note also that:

$$L_{d(e)n}(\text{year}) = 10 \log \{ \Sigma_i \text{antilog}[L_{d(e)n,i}/10]/T \},$$

where $L_{d(e)n,i}$ is the $L_{d(e)n}$ for day i in the year.

The measure 'B for light aviation' (*B voor de Kleine Luchtvaart* - BKL), is expressed in dB(A), and involves aggregating the L_{den} values which occur during a year in a different manner.

24-hour value and a related measure for a year

The 24-hour value, L_{etm}, is expressed in dB(A) and is defined as the maximum of:

$L_{Aeq}(7\text{-}19h)$,
$L_{Aeq}(19\text{-}23h) + 5$,
$L_{Aeq}(23\text{-}7h) + 10$.

For road traffic L_{etm} is also defined as the maximum of the first and last value, but the first definition is assumed here. Regulations for example make use of the twelfth highest 24-hour value as a measure for the year.

noise load in Ke for 24 hours and for a year

The noise load measure B is expressed in Ke (named after Kosten, chairman of the committee which proposed this measure), and is defined by:

$$B = 20 \log \{ \Sigma_i w_i \text{antilog}[L_{Amax,i}/15] \} - C,$$

where $L_{Amax,i}$ is the maximum of the A-weighted sound levels during flyover i. When B is being used for regulatory purposes, A-weighted sound levels are used, the meter is set to 'slow' and B is calculated for the year. For a year, C = 157. Schultz [1982] also defines B for a 24-hour period, and then C = 106. In practice, only maxima of at least 65 dB(A) are included in the calculation of B. The w_i are weights which depend on the time of day when flyover i occurs:

period	w_i
0 - 6	10
6 - 7	8
7 - 8	4
8 - 18	1
18 - 19	2

19 - 20	3
20 - 21	4
21 - 22	6
22 - 23	8
23 - 24	10

-page left blank-