DEPARTMENT OF TRANSPORT

REPORT OF A FIELD STUDY OF AIRCRAFT NOISE AND SLEEP DISTURBANCE

A study commissioned by the
Civil Aviation Policy Directorate
of the Department of Transport

from the
Department of Safety, Environment and Engineering
Civil Aviation Authority

J B Ollerhead, C J Jones, R E Cadoux, A Woodley, B J Atkinson
Civil Aviation Authority

J A Horne, F Pankhurst, L Reyner
Loughborough University of Technology

K I Hume, F Van, A Watson
Manchester Metropolitan University

I D Diamond, P Egger, D Holmes, J McKean
University of Southampton

December 1992

CONTENTS

TABLES

FIGURES

APPENDICES:

A Simple statistical test of the differences between measured disturbance rates
B Logistic regression analysis
C The Wilkinson-Diamond Analysis
D Prediction intervals for the proportion disturbed by aircraft noise events

PREFACE

This initial report on the Aircraft Noise and Sleep Disturbance Study, which was funded by the Department of Transport, has been prepared to assist them in developing proposals for future restrictions on nighttime aircraft movements at the London Airports of Heathrow, Gatwick and Stansted. It summarises the available results from a major research project initiated in the summer of 1990. Much analysis of the data has been completed and this has allowed a number of important conclusions to be drawn about the effects of aircraft noise on sleep disturbance. However, the subject is an extremely complex one and further analysis of the factors underlying sleep disturbance, and its effects, will continue for some time. This will not affect the conclusions about aircraft noise presented here but additional, more detailed, results will be described in a number of future reports.

The authors wish to express their appreciation to the many people who contributed to the study - through the social survey fieldwork, the noise measurements, the sleep monitoring and the data analysis. Of great importance to the study was the advice freely given by a number of eminent experts on sleep, most of whom attended a three day seminar in the spring of 1992 to discuss the problems in detail. These are:

Prof Torbjorn Åkerstedt ← Lund.
Dr David Dinges ←
Prof Barbara Griefahn
Prof Gerd Jansen
Prof Peretz Lavie ← Haifa Un.
Dr N C Mason
Dr Alain Muzet ← Strasbourg. Centre
Dr Paul Naitoh ×
Prof David Parkes ×
Mrs Peta Pascoe
Dr Mark Rosekind
Prof Ragnar Rylander ←
Dr Barbara Stone
Dr Michel Vallet ←
Prof W B Webb
Dr Robert Wilkinson ×

We also wish to thank BAA and Manchester Airport for providing facilities at the airports for the measurement teams and for hosting the special seminar in London, and British Airways for providing air travel to the visiting experts.

Last but not least, we are grateful to the members of the special Steering Group set up chaired by the Department of Transport to oversee the design and conduct of the project. Its members, including representatives from the airports, IATA, and the airport consultative committees, attended many meetings, analysed many proposals and papers and contributed many helpful ideas to the design and conduct of the study.

It has to be stated, in conclusion, that this report is the work of its authors, and the views expressed are not necessarily those of the above contributors.

J B Ollerhead
Civil Aviation Authority
December 1992

GLOSSARY AND DEFINITION OF TERMS

Frequently used terms, abbreviations and symbols are defined below: others which are used only locally in the text are defined where they first occur.

a	Arousal rate in all epochs, ie with and without aircraft noise events.
α	Value of a, for an individual subject, for the time interval between two successive ANEs.
A-blip	Term used to describe any epoch in an individual's actigram in which an arousal from sleep is identified. So called because arousals appear as 'blips' on a simple graph of disturbance against time.
A-filter	Filter applied to actigrams; see 'filter'.
Actigram	Graphical record of an individual's wrist movements measured by actimeter. A 'raw' actigram gives the actual number of wrist movements per epoch; a transformed or 'filtered' actigram identifies movement onsets only.
Actimeter	Instrument for measuring wrist movements, worn like a wrist watch.
Aircraft Leq	An Leq value which includes all identified aircraft noise energy above 60dBA.
Ambient Leq	An Leq value ecompassing all noise energy except that which comprises the Aircraft Leq.
ANE	Aircraft noise event; the noise experienced when a single aircraft passes by.
ANE-epoch	Epoch encompassing Lmax of an aircraft noise event. Also refered to as noise epoch.
ANIS	UK Aircraft Noise Index Study; a major study of aircraft noise indices (Ref 18).
Arousal	Specifically used in this report to describe the onset of sleep disturbance as measured by an actimeter. Used more generally in the scientific literature to mean various perturbations or disturbances to sleep.
Arousal rate	Incidence of actimetrically determined arousals (A-blips): number of disturbed epochs expressed as a percentage of total epochs.
Artefact	Burst of activity in the EEG record of greater than normal EEG intensity associated with increased muscle activity on the scalp and/or movement of electrodes or electrode leads.
Awaken(ing)	The process of changing from a state of sleep to wakefulness; defined in this study as the start of at least 15 seconds of 'wakefulness' or 10 seconds of 'movement time' in the EEG record.

CAA	Civil Aviation Authority.
Confounded	A relationship between two variables, deduced by analysis of measured data, which is inaccurate or misleading because of hidden effects of other factors, not accounted for in the analysis.
dB	Decibels, units of sound level, or relative sound level, calculated as 10 times the log (base 10) of a sound energy ratio. Used here to define differences between levels measured on the dBA scale.
dBA	Levels on a decibel scale of noise measured using a frequency dependent weighting which approximates the characteristics of human hearing. These are referred to as A-weighted sound levels; these are very widely used for noise assessment purposes.
Designated airports	Airports designated for the purposes of Section 78 of the Civil Aviation Act, 1982. These are the London Airports of Heathrow, Gatwick and Stansted.
Disturbance	Sleep disturbance can be defined in a variety of ways. In this report the expression is used generally to cover both awakenings and actimetrically determined arousals; however, it is also used in a more specific sense to describe events of particular significance such as EEG-awakenings which, if experienced often enough, could have longer-term consequences.
Disturbance rate	Incidence of disturbance; number of disturbed epochs expressed as a percentage of total epochs.
DORA	Directorate of Operational Research and Analysis; CAA directorate formerly responsible for aircraft noise studies.
DOT	Department of Transport.
EDG	Study site at Edgeley near Manchester Airport.
EEG	Electroencephalography: the measurement of very small electrical signals generated within the brain using small electrodes attached to the head - used to determine sleep stage. (Also electroencephalogram or electroencephalograph - the physical record generated by the electroencephalography process)
EMG	Electromyogram; a record of facial muscle tone obtained in a similar way to the EEG.
EOG	Electro-occulogram: a record of eye movements obtained in a similar way to the EEG.
EPNdB	Aircraft noise event level measured on the scale of Effective Perceived Noise Level used internationally for the noise certification of aircraft. Its measurement involves analyses of the frequency spectra of ANEs; thus these units are more complex than dBA. Typically, EPNL values are 3-5 dB greater than those of SEL in dBA.
Epoch	The basic time interval used in the measurement of sleep state; in this study both EEG and actimeter epochs were set to 30 seconds.

Filter	The mathematical procedure used to transform a raw sleep record (ie a hypnogram or actigram) to a record of disturbance onsets (ie A-blips or H-blips). In general, a filter incorporate a 'buffer' - the minimum sequence of undisturbed epochs which must occur before a 'disturbance onset' is defined.
H(b)	Symbol denoting a record of awakenings (H-blips) determined using a filter with buffer b (for all results presented in this report, b=1)
H-blip	Term used to describe any epoch in an individual's hypnogram in which an awakening is identified. So called because awakenings appear as 'blips' on a simple graph against time.
HAT	Study site at Hatfield Heath and Hatfield Broad Oak near Stansted Airport.
HGN	Study site at Heald Green near Manchester Airport.
HLW	Study site at Hounslow near Heathrow Airport.
Hypnogram	An epoch-by-epoch record of sleep stage determined from the Sleep-EEG.
L_{90}	The sound level exceeded for 90% of the measurement period.
Leq	A measure of long-term average noise exposure; for aircraft noise it is the level of a steady sound which, if heard continuously over the same period of time, would contain the same total sound energy as all the ANEs.
LFD	Study site at Lingfield near Gatwick Airport.
LGN	Study site at Langley Green near Gatwick Airport
Lmax	The highest instantaneous sound level recorded during an ANE, in dBA (measured using a standard 'slow' meter setting).
LRA	Logistic regression analysis. A statistical procedure used to distinguish between the effects of different factors which can affect sleep disturbance at the same time - described in Appendix B.
M (MT)	Abbreviation for Movement Time.
Movement Time	Defines EEG output which contains large electrical disturbances called 'artefacts' (see above). They are large enough to mask underlying brain signals and may therefore be considered to indicate significant disturbance of sleep.
N	Number of observations.
n	Arousal rate in ANE epochs.
Noise epoch	See ANE-epoch.
p	The probability that the result of a particular analysis arose purely by chance (ie it is unlikely to be repeated if further measurements were made and analysed).

p	The proportion of a set of observations having a specified characteristic (eg the proportion of subjects being disturbed).
Perturbation	A minor event in a sleep record (EEG, actigram), such as a transient lightening of sleep or a minor body movement, considered to be of less importance than a 'disturbance'.
PNdB	Aircraft noise level measured on the decibel scale of Perceived Noise Level, approximately equal to Lmax (dBA) + 13 numerically . Now little used.
Polygram	Combination of EEG, EMG and EOG records used to determine sleep stage more loosely described collectively as a 'sleep-EEG'.
q	Arousal rate in non-ANE epochs (ie epochs which do not encompass ANEs also referred to as 'quiet' epochs).
Random effects	A term having a special meaning when applied to LRA; a modification which permits the results to be controlled for the effects of serial correlation.
Regression	The statistical procedure of fitting a descriptive mathematical relationship to a set of measurements.
REM	Rapid Eye Movement; a stage of sleep usually accompanied by dreaming; its position in the natural sleep stage hierarchy is uncertain.
SEL, *SEL*	ANE level measured on the decibel Sound Exposure Level scale, in dBA (*SEL* is the energy average of a set of SEL values used in the calculation of aircraft Leq); like EPNL, this scale accounts for both the duration and the intensity of the noise event.
Sensitivity	Also termed arousability; the susceptibilty of an individual to sleep disturbance - subjects of high sensitivity have high arousal rates.
Serial correlation	The non-independence of observations obtained from a single individual which may confound any analysis based on normal probability statistics (see Appendix B).
Sleep diary	The daily record made by each subject of his/her daytime activity and sleepiness.
Sleep-EEG	See 'polygram'.
Sleep log	The daily record made by each subject of sleeping times about bedtime, lights out, estimated sleep onset, number of night awakenings and reasons for them, morning awakening, rising time and sleep quality.
Sleep onset	The time of first falling asleep: in this study defined (a) subjectively in the sleep log and/or (b) from the hypnogram or actigram. The two are not necessarily the same.
Sleep stage	State of sleep as measured by sleep-EEG. Sleep stages include wakefulness, MT, and REM as well as stages 1 to 4, the latter indicating depth of sleep (from 'shallow' to 'deep').

Sound level	The magnitude of noise measured on a decibel scale.
Statistically significant	Describes a result for which *p* is less than a specified value called the level of significance, set at 5% for all results in this report.
SWM	Study site at Stanwell Moor near Heathrow Airport
Time	All clock times are local times.
Twitch	Small sudden involuntary movement often accompanying REM sleep, for example.
W	Abbreviation for state of wakefulness.
W-D analysis	See Appendix C.
Wakefulness	The state of not being asleep; positively identifiable as an EEG stage.
Waken(ing)	Being awakened by an external stimulus (such as an aircraft noise event)
WSB	Study site at West Sawbridgeworth near Stansted Airport.
y/n	Yes or no (characteristic present or absent).

SUMMARY

OBJECTIVES

1 Current night restrictions at Heathrow and Gatwick Airports are based, in part, on the results of studies of the effects of noise on sleep carried out more than ten years ago. As these policies were due to be reviewed, the Department of Transport asked the Civil Aviation Authority to undertake further studies of aircraft noise and sleep disturbance, with emphasis on objective measurements. The study has been conducted by the CAA in conjunction with research teams from the Universities of Loughborough, Manchester Metropolitan and Southampton.

2 The objectives of the study were to determine:

(a) the relationships between outdoor aircraft noise levels* and the probability of sleep disturbance,

(b) the variation of these relationships with time of night.

To meet these, it was also necessary to investigate the influence of non-acoustical factors upon disturbance of people's sleep including their age, sex and personal characteristics, their general views about the neighbourhood, their perceptions about sleep quality and the ways in which this might be affected by aircraft noise.

3 It may be postulated that sleep disturbance involves three different kinds of effects, (1) interference with the sleep process itself, (2) short-term after-effects which include, for example, daytime sleepiness and annoyance, and (3) possible long-term health effects. As the latter effects are consequent upon the first, a major aim of this study was to observe the sleeping patterns of people in homes which are affected by aircraft noise.

BACKGROUND

4 The traditional method for monitoring sleep is electroencephalography or 'sleep-EEG' in which brainwaves are measured by electrodes attached to the scalp. A *hypnogram* is a record of sleep stage changes during the night obtained from EEG data. Sleep stages in the hypnogram include light, deep and REM (rapid eye movement - indicative of dreaming) as well as wakefulness. However, the method is complex and expensive and, partly for these reasons, most EEG work has been performed in laboratory situations using relatively small numbers of subjects. In order to avoid the statistical constraints of such limited studies and because of a strong possibility that laboratory results are not representative of the way people react in their homes, this study made use of actimeters to gather a large quantity of field data. Actimeters are used to measure fine limb movements, usually of the wrist, which are indicative of sleep disturbance. Actimeters are small, relatively inexpensive devices, worn like a wrist watch, and easily used in the home without supervision. They log and store data for many nights which is subsequently transferred to a computer for conversion to *actigrams*, the graphical records of limb movements.

5 Actimetry is widely used in sleep research, but an important part of the study was to validate its use for measuring the effects of aircraft noise on sleep. This was done by

---

* Although people in bed hear aircraft noise as attenuated by the walls, windows and furnishings of their bedrooms, indoor noise levels naturally vary very widely from room to room and from ear to ear. These variations cannot be accounted for in planning or policymaking; only outdoor levels are known or can be estimated with any degree of confidence. Noise level measurement inside all subjects' bedrooms was not practicable. The unknown variability is viewed simply as one of the many uncontrollable factors affecting sleep disturbance.

direct comparison of EEG and actimeter measured disturbance, both in the main study itself and in a preliminary pilot investigation.

DEFINITION OF SLEEP DISTURBANCE

6 In order to establish a working definition of 'sleep disturbance' within the context of this study, views were sought from a number of eminent sleep experts both at the outset and when the initial experimental results became available. Opinions differed on precise definitions, particularly with regard to effects which might in any way be regarded as injurious to health. However, there was broad agreement on three points:

(a) Any identified period of EEG-measured wakefulness is definitely indicative of sleep disturbance.

(b) Lesser EEG responses, such as sleep stage changes, may be considered as minor perturbations.

(c) Brief awakenings, of less than about 30 seconds, are most unlikely to result in daytime sleepiness or otherwise impair health unless, in sum, they occur more than about six times an hour through the night. Longer awakenings, depending on their duration and number, can be increasingly more harmful. Awakenings are not usually remembered the next day unless they last beyond 1 to 2 minutes. A high proportion of awakenings are very brief with durations measured in seconds rather than minutes.

7 Accordingly, for the purposes of this work, an EEG-disturbance was defined as an episode of wakefulness lasting 15 seconds or more, or 'movement time' (a distorted EEG response usually related to wakefulness) lasting 10 seconds or more. Onsets of such disturbances, identified from EEG records or hypnograms, were defined as *awakenings.*

8 Disturbances identified from actigrams, ie any onsets of wrist movement following still periods, were termed *arousals*. These arousals often coincide with EEG-awakenings or movement time (nearly 90% of these are detected) but they also include minor perturbations such as twitches of the kind that commonly occur during dreaming (REM) sleep.

MEASUREMENTS

9 In the main study, volunteer subjects were recruited from homes in 8 study areas, two near to each of four major UK airports - London-Heathrow, London-Gatwick, London-Stansted and Manchester. The sites were chosen (a) to cover a wide range of nighttime aircraft noise exposures (Leq) and widely different combinations of event noise levels and numbers, (b) to be large enough to provide statistically adequate samples of residents but small enough to limit the variation in outdoor noise exposure, ideally to within 3dB, and (c) to be free of excessive noise from non-aircraft sources.

10 At each site, at least 200 people were interviewed in a preliminary social survey. Each sample was chosen to match the wider local population with respect to sex and age distribution. As well as providing a pool of potential subjects, the social survey was designed to yield information on factors other than noise which affect sleep patterns. These included personal characteristics, general views about the neighbourhood, perceptions of sleep quality and the ways in which that might be affected by aircraft noise.

11 From the survey respondents at each site, 50 participants were selected who met various sampling and test criteria. People who said they were deaf, that they suffered from serious sleep-disturbing ailments, that they were taking medications that affect sleep or that they were shift workers were excluded. At each site, all 50 subjects wore actimeters for a fifteen night monitoring period; 6 of them also underwent simultaneous EEG monitoring on four sequential nights.

disturbed over 60% more than average. There appear to be no strong personal factors contributing to this sensitivity; a large number of possible variables have been specifically ruled out , although further analysis is being undertaken.

Aircraft noise

17 The results indicate that, below outdoor event levels of 90 dBA SEL (80 dBA Lmax), aircraft noise events (ANEs) are most unlikely to cause any measurable increase in the overall rates of sleep disturbance experienced during normal sleep. For outdoor event levels in the range 90-100 dBA SEL (80-95 dBA Lmax) the chance of the average person being wakened is about 1 in 75. Again, individual deviations from the average are substantial. It is possible that, for aircraft noise related disturbance, the variability is even greater; compared with the average, the 2-3% most sensitive people could be over twice as likely to be disturbed and the 2-3% least sensitive less than half as likely.

Sex and age

18 The results indicate that, overall, men are disturbed from sleep about 15% more frequently than women and that this is true for all causes of disturbance, not especially aircraft noise. No statistically significant effects of age were found.

Time of night

19 Statistically, time of night and time from sleep onset are significant factors. When the data are broken down by time of night, people appear to be most resistant to disturbance, from any cause, after first falling asleep. Then, starting with a pronounced fluctuation having a cycle time of about 90 minutes, the overall disturbance rate increases steadily, from the equivalent of about two awakenings an hour at the beginning of the night to about three per hour at the end of the night.

20 Arousals related to aircraft noise seem to follow a stronger cyclic pattern. After the first 45 minutes of sleep, which appears to be insensitive to the noise, noise-related disturbances repeatly rise and fall in a way that cannot be explained by the rates at which aircraft noise events occur. Although difficult to verify statistically, natural biological rhythms of sleep may be the reason. The possibility that people are most sensitive to disturbance by noise when sleep lightens, and less vulnerable when sleep deepens, is the subject of continuing analysis.

21 As well as being minimal during the first hour of sleep, sensitivity to aircraft noise seems to diminish at the end of the night's sleep. However, this may be due to greater overall rates of awakening from all causes and, consequently, a diminishing proportion of people asleep from 0600 onwards. Further analysis is continuing in an attempt to shed more light upon this important but difficult question.

NON SIGNIFICANT EFFECTS

Site

22 There were no statistically significant differences between the average arousal rates over the night at the different study sites.

Window state

23 The reported 'window state' each night, ie open, single glazing shut or double glazing shut, was included in the analysis but, although increased noise insulation was accompanied by reduced arousal rates, this has not been found statistically significant.

12 The fieldwork was conducted during the summer of 1991. In all, 400 subjects were monitored for a total of 5742 subject-nights. Sleep-EEG were obtained from 46 subjects for 178 subject-nights (the 'EEG sample' - 3% of the total; data from two subjects were lost). In total, some 40,000 subject-hours of sleep data were analysed, broken down into more than 4.5 million 30-second *epochs*. Outdoor aircraft noise levels (Lmax and SEL) were measured at up to three positions at each site using noise monitors set to record all levels in excess of 60dBA (use of a lower threshold would have increased the difficulty of identifying and measuring aircraft noise events due to interference from non-aircraft noise sources). Aircraft movements causing noise events were identified from airport runway logs; the events were accurately timed for synchronisation with the sleep measurements. A total of 4823 aircraft noise events were logged during the 120 measurement nights at outdoor noise levels from 60 dBA to more than 100 dBA Lmax. Accompanying data from pre-test and debrief interviews, sleep logs and diaries comprised another 100,000 items of data.

13 The data were analysed to determine the relationships between sleep disturbance and aircraft noise taking into account the effects of other relevant factors including time of night and the age and sex of the subjects. Because the main results, such as overall disturbance rates, are based on analyses of large data samples, there is a high level of statistical confidence that they are reliable estimates of true 'population' values. However, when the data were divided into subsamples to determine the effects of the other factors, confidence intervals inevitably widened and considerable care was necessary to ensure that the conclusions are statistically valid. Wherever possible, a procedure known as *random effects logistic regression analysis* was used to take proper account of the combined effects of the various factors of importance. This technique also overcomes a limitation inherent in measured data of this kind: that although each individual subject provides an independent set of disturbance data, the many measurements from one individual are not statistically independent of each other.

VALIDITY OF ACTIMETRY

14 Actimetry was shown to be a convenient and valid technique for investigating sleep disturbance in the home. For the EEG-sample, the agreement between actimetrically determined *arousals* and EEG-measured *awakenings* was very good: 88% of all awakenings coincide with actimetric arousals. For the noisiest site, the agreement was 92%. The agreement in the case of undisturbed epochs is even higher, 97% overall . This is important support for the actimetry method, given that undisturbed epochs were 95% of the total.

OVERALL DISTURBANCE

15 The mean arousal rate (ie the proportion of epochs with movement arousals) for all subjects, all causes, all nights, all epochs, was 5.3%. For the average sleeping period of 7.25 hours, this is equivalent to about 45 arousals per night. Of these, some 40%, ie about 18 per night, are likely to be awakenings of 10-15 seconds or more, the remainder being minor perturbations* .

FACTORS AFFECTING SLEEP DISTURBANCE

Individual sensitivity

16 Individual rates of sleep disturbance varied markedly; after statistically controlling for the effects of aircraft noise, sex and time of night, the 2-3% most sensitive individuals were

---

* The awakening-to-arousal ratio of 40% is an example of a statistic which is subject to a sampling uncertainty, in this case of perhaps ±10%. Thus, although the average number of nightly awakenings (all causes) is likely to be *about* 18, it would be more accurate to state that it probably lies in the range 18±4.

Aircraft type

24 Allowing for noise level, ie comparing their effects at the same event noise levels, no significant differences were found between the average noise-related arousal rates for large jets, small Chapter 2 jets, small Chapter 3 jets* and propeller aircraft types.

Length of residence

25 No subjects were selected who had lived locally for less than one month. With this proviso, there is no significant effect of length of residence on arousal rates, ie there appear to be no adaptation effects after the first month of residence.

Other noise variables

26 Because of the predominance of approach noise in this study (which rightly reflects the high proportion of arrivals in nighttime aircraft movements) as well as the generally weak effect of aircraft noise level, it is impossible to distinguish between the performace of Lmax and SEL as indicators of sleep disturbance.

RECOLLECTIONS OF SLEEP DISTURBANCE

27 The secondary or after effects of sleep disturbance include subjects' recollections of being wakened and adverse perceptions of their sleep quality. For 57% of subject-nights, no awakenings were reported the next day. On the remaining 43% of occasions, at least one awakening was reported (all causes), the average number being three per night. In 26% of reported awakenings, the reason was given as 'not known'. For the remainder, the most frequently reported cause was 'toilet' (16%). The next most common was 'children' (13%) mainly among women in the lower age groups. 'Illness' was also mentioned frequently (>9%), again mostly by women. 'Aircraft' was a relatively minor cause (<4%); about one quarter of all actimetry subjects specifically reported being disturbed by aircraft noise during the study - on average by these subjects, once every five nights.

28 The agreement between individuals' measured arousal rates and their general self-ratings of sleep quality (recorded during the prior social survey interview) is poor. However, there is better agreement between the measured arousal rates and next-day reports of sleep quality obtained from the daily sleep logs. This suggests that when social survey methods are used for investigating sleep disturbance, emphasis should be placed on collecting data about disturbance experienced during the previous night.

29 The measurements of sleep disturbance, which were the main subject of this study, are quite distinct from those of annoyance, which must be counted among secondary effects. The relationship between measured disturbance and annoyance reports as well as the question of daytime sleepiness, are the subject of continuing study.

CONCLUSIONS

30 All subjective reactions to noise vary greatly from person to person and from time to time and sleep disturbance is no exception; deviations from the average can be very large. Even so, this study indicates that, once asleep, very few people living near airports are at risk of any substantial sleep disturbance due to aircraft noise, even at the highest event noise levels.

31 At outdoor event levels below 90 dBA SEL (80 dBA Lmax), average sleep disturbance rates are unlikely to be affected by aircraft noise. At higher levels, and most of the events upon which these conclusions are based were in the range 90 to 100 dBA SEL (80 to 95 dBA Lmax), the chance of the average person being wakened is about 1 in 75. Compared

---

* These 'Chapters' refer to international aircraft noise certification standards; at comparable weights, Chapter 3 aircraft are quieter than (earlier) Chapter 2 aircraft.

with the overall average of about 18 nightly awakenings, this probability indicates that even large numbers of noisy nighttime aircraft movements will cause very little increase in the average person's nightly awakenings. Therefore, based on expert opinion on the consequences of sleep disturbance, the results of this study provide no evidence to suggest that aircraft noise is likely to cause harmful after effects.

32 At the same time, it must be emphasised that these are estimates of *average* effects; clearly more susceptible people exist. At one extreme, 2-3% of people are over 60% more sensitive than average; some may be twice as sensitive to noise disturbance. There may also be particular times of the night, perhaps during periods of sleep lightening, when individuals could be more sensitive to noise. Although the relationship cannot be verified statistically, the data do indicate that aircraft events with noise levels greater than 100 dBA SEL (95 dBA Lmax) out of doors, will have a greater chance of disturbing sleep. The most sensitive people may also react to aircraft noise events with levels below 90 dBA SEL (80 dBA Lmax) (approximating to 95 EPNdB on the noise scale used internationally for the noise certification of aircraft).

33 These conclusions are based on actimetric measurements of arousals from sleep supported by EEG data.

34 Work is continuing on a number of detailed points to supplement the findings in this report including further analysis of the possible effects of noise in preventing sleep onset at the beginning of the night, or delaying return to sleep after awakening during the night or in the early morning. This will not change the conclusions about aircraft noise presented here but additional results will be published subsequently.

1 INTRODUCTION

1.1 Existing night noise criteria

Existing restrictions on nighttime movements of aircraft at Heathrow and Gatwick Airports are based, in part, on the results of previous scientific studies of the effects of aircraft noise on sleep. These effects can generally be divided into two kinds, primary and secondary. The primary effects are direct disturbances of the sleep process itself. These include awakenings, changes of sleep state and other physiological reactions. The secondary or after effects are the consequences of the disturbance such as daytime sleepiness and perceptions of other adverse effects including annoyance. Theoretically, given such disturbances, there might also be longer term detriment to health and well being; this could be a tertiary effect.

From a review of available research data on primary effects, it was previously concluded (Ref 1) that little sleep disturbance would be caused by aircraft noise events (ANE) whose maximum level did not exceed an outdoor aircraft noise level around 95 PNdB (approximately 82 dBA Lmax). This was used as supporting evidence for the 1980 noise insulation grant scheme boundaries and the 1989 extension schemes. Subsequent social survey studies of secondary effects (Refs 2, 3) indicated that the total night noise exposure, as quantified by Leq, was an appropriate index of total perceived night noise disturbance. Current night restrictions at Heathrow and Gatwick Airports are aimed at limiting night noise exposure.

It was recognised that the subjective data on which the conclusions of the previous studies were based could have limitations. People tend to be poor at estimating their sleep quality and quantity and how often they are disturbed. Therefore, it was recognised that any future changes to the policy needed to be supported by more detailed and reliable evidence as to the likely effects of aircraft noise at night, preferably based on *objective* measurements of sleep disturbance. Accordingly, the DOT asked the CAA to undertake a new study of aircraft noise and sleep disturbance with the aim of providing scientific evidence to assist future policymaking with respect to night traffic at the designated airports.

1.2 Previous studies of sleep disturbance

There is a substantial body of research information on the primary effects of noise on sleep. Most of this research has been carried out in laboratory situations. A much smaller number of studies has been performed in 'field' (ie at home) conditions.

In both types of studies, the immediate responses of the subjects have been determined from sleep electroencephalograms; less frequently there have been measurements of cardiovascular reactions, body movements or by reported awakenings. Such studies have shown that the disturbance of sleep by noise can depend on both the magnitude of the noise and the 'state' of the individual. The latter includes personal variables such as age and sex, and sleep variables such as the individual's stage of sleep, accumulated sleep time and when the noise occurs.

In a small number of studies, the effects of a noise-disturbed night on the individual's performance the following day has been determined using, for example, reaction time tests. Little research appears to have been carried out into the relationships between marked sleep disturbance and any chronic health effects.

1.3 The nature of sleep disturbance

There is no absolute definition of what constitutes sleep disturbance. At best it is a relative term which takes an operational definition depending on the nature of the problem being investigated and the measurement procedures in use. The traditional 'gold standard' of sleep assessment involves the interpretation of all-night recordings of the sleep polygram, often referred to as the 'sleep-EEG'. This is a combination of the electroencephalogram or EEG records of very small electrical signals generated within the brain; the electro-occulogram or EOG records of eye movements; and the electromyogram or EMG records of facial muscle tone - all of which are detected by a number of small electrodes attached to the head. Various measures are derived from these records to gauge the degree of sleep disturbance. Most research divides disturbance into two categories; *specific*, where a brief event in the sleeper's environment is investigated for a relatively short period of time, eg disturbance caused by aircraft noise events, and *general*, where the whole night's sleep is thought to be affected, eg by a raised bedroom temperature.

Short term disturbances can be very brief, often lasting just a few seconds. However, people are generally quite unaware and unaffected by these unless they happen very frequently when the consequences could include insidious effects such as increased daytime sleepiness.

1.4 Laboratory or field studies?

Measurements of sleep disturbance, like those of other subjective reactions to noise, show a large amount of variation between individuals. For this reason, large data samples are needed to distinguish between the effects of different sources of disturbance. There is also evidence of a substantial lack of agreement between field and laboratory data. A recent reanalysis of available data (Ref 4) illustrated in Figure 1 suggested that, for the same levels of noise, subjects are much less likely to be awakened at home than in the laboratory. This raised concerns that, although further laboratory studies could be designed to provide more controlled test data on the effects of aircraft noise on sleep, uncertainties would remain about their relevance to 'real life' situations. From this point of view, a large scale field study would be preferable. However, because of its high cost, EEG monitoring is not a practical option for such studies.

1.5 Actimetry v EEG

Another indication of sleep state is given by nighttime limb movements; these only cease to any marked extent during sleep. Poor or disturbed sleep is reflected by increased movements and there is good evidence that these usually accompany noise disturbed sleep (eg Refs 5,6).

These movements can be accurately measured using activity monitors, or actimeters - small, unobtrusive strap-on devices which can record all significant arm or leg movements over long periods of time. The costs of actimetry (alternatively described as actigraphy) are only a fraction of those of EEG monitoring. Generally good correlations between actimetry and EEG have been demonstrated (eg Refs 7-11). However, the use of actimetry for studying sleep disturbance due to aircraft noise has not previously been demonstrated (except in a small-scale experiment in Israel - Ref 12). Use of this technique needed verification by comparison with alternative and accepted procedures; for this purpose it was logical to rely on the 'gold standard' EEG.

1.6 The Study

The study, which involved extensive measurements of aircraft noise and sleep disturbance, was carried out by a team of research staff from Loughborough University of Technology, Manchester Metropolitan University, the University of Southampton, and the CAA. It was funded by the Department of Transport and guided by a special Steering Group with members drawn from the air transport industry and airport consultative committees. Technical advice was obtained from a panel of internationally recognised sleep experts, both before the study commenced and during its progress, prior to the main data analysis phase. Social survey fieldwork was carried out by Public Attitude Surveys Ltd.

Full technical details of the work will be given in a number of separate specialist reports to be published subsequently (Refs 13-17). The present report, which was prepared by the whole research team, summarises the study and its results. The experimental work was carried out between September 1990 and October 1991: analysis of the experimental data commenced immediately after the measurements had been made. The bulk of the analysis has been completed and this has enabled important conclusions to be published in this report. However the subject of sleep disturbance is an extremely complex one; much new methodology, both experimental and analytical, had to be developed to investigate it adequately and progress has necessarily been cautious. The data unquestionably contain more information on aircraft noise and sleep disturbance than it has been possible to extract within the scope of the present study. At the time of publication of this report, analysis of a number of aspects is continuing. Further work will not affect the conclusions about aircraft noise presented here but additional results may be added in References 13-17.

2 DESIGN OF THE STUDY

2.1 Approach

The objectives of the study were to determine:

(a) the relationships between outdoor aircraft noise levels* and the probability of sleep disturbance,

(b) the variation of these relationships with time of night and with other factors.

To meet these, it was also necessary to investigate the influence of non-acoustical factors upon disturbance of people's sleep including their age, sex and personal characteristics, their general views about the neighbourhood, their perceptions about sleep quality and the ways in which this might be affected by aircraft noise.

A two phase programme was undertaken, commencing in July 1990. It comprised:-

(i) July - December, 1990: A pilot study at a single site near Manchester Airport to develop and validate the experimental procedures.

(ii) 1991-2: A main study involving eight sites, two each near Heathrow, Gatwick, Stansted and Manchester Airports.

The study sites contained sufficient homes to provide the statistically adequate samples of subjects yet were small enough to be considered as 'constant noise areas'; ie areas over each of which outdoor aircraft noise exposures are relatively uniform (ideally to within 3dB).

2.2 Pilot study

The specific aims of the pilot study were:

- to develop the necessary experimental procedures,
- to evaluate actimetry by comparing it with EEG measurements, and
- to provide statistical data on which to base the design of the main study.

The conclusions were:

a) actimetry is a suitable method; however, additional sleep EEGs to calibrate the results would be essential;

b) the link between noise exposure and sleep disturbance is weak; other factors (personal/psychological) were identified as playing an important role and would need to be examined as closely as possible if the results were to be adequately explained;

c) to yield results of statistical significance, about 50 subjects would be required at each main study site;

---

* Although people in bed hear aircraft noise as attenuated by the walls, windows and furnishings of their bedrooms, indoor noise levels naturally vary very widely from room to room and from ear to ear. These variations cannot be accounted for in planning or policymaking; only outdoor levels are known or can be estimated with any degree of confidence. Noise level measurement inside all subjects' bedrooms was not practicable. The unknown variability is viewed simply as one of the many uncontrollable factors affecting sleep disturbance.

d) each subject should be monitored for a period of two weeks.

The effects of aircraft noise upon people at night are likely to be greatest during the summer when they may have their windows open and when air traffic is usually at a peak due to summer holiday flights. Allowing for limits on data gathering rates, the need to complete the main study fieldwork between Spring and early Autumn 1991 meant that 8 sites could be envisaged, yielding up to 6000 subject nights of actimetry in total. In addition, up to 200 subject-nights of Sleep-EEG data would be needed to back this up.

The pilot work suggested that about 25% of potential subjects contacted would meet the criteria for participation (available, good hearing, not shift workers etc) and would be willing to participate. Therefore to recruit 50 subjects at each site, about 200 residents would need to be approached at each of the main study sites. In view of the number of questions that had to be asked, this search needed to be organised along the lines of a restricted-scope social survey.

2.3 Main study

The main experiment involved, at each site:

a) A social survey with interviews of 200 respondents. The aims were:-

  - to provide a pool of subjects for the sleep measurements,
  - to provide data on the personal and socio-psychological factors that would be likely to affect sleep disturbance, and
  - to allow comparison with previous social survey studies of sleep disturbance.

b) Selection of 50 subjects to take part in the actimetry monitoring.

c) 15 nights of actimetry on all 50 subjects. During the monitoring period all actimetry subjects should also complete sleep logs and daytime sleepiness reports.

d) EEG monitoring on 6 of the 50 subjects for 4 nights each (simultaneously with actimetry).

e) Outside noise measurements at locations around the study area.

In order to make final checks on the experimental arrangements and procedures, a second small pilot study was also performed, immediately prior to the start of the main study.

2.4 Site selection

Forty seven possible study sites were identified from available night noise data from the areas surrounding the four airports. The nighttime aircraft noise conditions (Leq, *SEL* and N) at these sites are illustrated in Figure 2. However, site visits revealed that many of them were actually unsuitable because of insufficient quantity of housing or the presence of noise from other sources such as roads and railways which would confuse the data analysis. The eight sites eventually selected for the main study (shown in bold in Figure 2) were chosen to cover the four airports and a wide range of *SEL*, N combinations. Their locations relative to the airports and main flight paths are shown in Figure 3 (the second pilot test was carried out at Mogden near Heathrow).

The study sites and the predominant mode of operation of overflying aircraft were:

Airport	Location	Abbreviation	Predominant mode
Heathrow	Hounslow	HLW	Arrivals
Gatwick	Langley Green	LGN	Departures
Heathrow	Stanwell Moor	SWM	Departures
Gatwick	Lingfield	LFD	Arrivals
Manchester	Heald Green	HGN	Arrivals
Manchester	Edgeley	EDG	Arrivals
Stansted	Hatfield Heath/Hatfield Broad Oak	HAT	Departures
Stansted	West Sawbridgeworth	WSB	Arrivals

2.5 Need for control sites

The need for comparable 'control' sites with no aircraft noise was considered in detail. It was recognised that, at best, control sites which adequately matched the (unknown) relevant characteristics of the test sites (other than aircraft noise) would be very difficult to identify. To reflect the distributions of these characteristics, several control sites would in fact be required, possibly one for each of the eight test sites. The conclusion, endorsed by the Steering Group, was that the inclusion of control sites would not be the best use of study resources. This is not expected to place serious limitations on the statistical results, in particular because the nighttime aircraft noise exposures at the eight sites were intended to cover a very wide range; indeed the quietest sites - at Stansted - experienced little nighttime aircraft noise on average, on some nights none at all. The position was reviewed later in the study; as will be seen (Section 7.3) the results supported the view that control sites were not necessary.

2.6 Fieldwork

This was carried out between 3rd February and 31st October 1991. The programme of work at each site spanned thirteen weeks; thus, at any time, work was in progress at several sites simultaneously. At each site, the programme involved a site survey and listing of available dwellings (weeks 1-4), social survey interviews (weeks 5-7), selection of test subjects (weeks 8-10) and measurement of noise and sleep disturbance (weeks 11-13). Table 1 lists the main study sites and indicates the schedule of work at each of them.

3 DATA GATHERED

3.1 Study sites

The general characteristics of each of the study sites shown in Figure 3(a) to 3(d) are summarised below.

Heathrow: Hounslow (HLW): Suburban area of mainly 1920s semi-detached houses, with rather narrow roads. The area, just over 3 km from the end of runway 27L, is affected predominantly by westerly landings but also by some easterly departure noise from Heathrow, and was within the Heathrow Noise Insulation Grants Scheme area.

Gatwick: Langley Green (LGN): Estate of late 1960s terraces, semis and a few bungalows with little road traffic. The area is about 2 km due south of the centre of the Gatwick runway, and is affected by both take-off and landing noise from movements in both runway directions.

Heathrow: Stanwell Moor (SWM): Mostly 1960/70s semis and bungalows bounded by gravel pits, farms, a nursery and a reservoir. The area, just over 1 km from the end of runway 27L and about 500m to the side of the extended runway centreline, is affected by westerly take-offs and by both landing and take-off noise during easterly operations. It was within the Heathrow Noise Insulation Grants Scheme area.

Gatwick: Lingfield (LFD): A mixture of very varied housing of most types and ages, with some narrow roads, 10 km east of Gatwick. The area is affected predominantly by westerly approaches but some departures (in both directions) are also heard.

Manchester: Heald Green (HGN): Just over 1 km from the threshold of Manchester runway 24 and under the flight path, consisting of bungalows and semi-detached houses. The area is affected by both westerly landings and easterly departures; some westerly take-offs are also audible. Within the Manchester Noise Insulation Grants Scheme area.

Manchester: Edgeley (EDG): Turn-of-the-century terraced housing with only small back yards, making a very compact area 5 km from the end of Manchester runway 24, and under the approach flight path. The area is affected by westerly approaches and by some easterly departures.

Stansted: Hatfield (HAT): As Figure 3(d) shows, because no second single site was available at Stansted, the Hatfield site was made up of two separate areas; at Hatfield Heath and Hatfield Broad Oak. Both are affected by westerly departures from the airport; but because the flight tracks tend to lie between them, their aircraft noise exposures are very similar. Altogether, a very varied mixture of housing - very old cottages including listed buildings, old semis and detached houses, some modern bungalows and houses, some early 1900s and 1950s terraces and 1970s detached houses.

Stansted: West Sawbridgeworth (WSB): Mostly modern semi-detached houses plus some detached houses and bungalows, but also some early 1900s houses. Some of the gardens are quite small. The area is about 9 km from the end of Stansted runway 05, and is affected by easterly approaches and by a few westerly departures.

3.2 Social survey and subject selection

Subjects for the sleep monitoring phase of the investigation were chosen from social survey respondents who expressed a willingness to participate, so as to give a sample representative of the local area. The primary purpose of the social survey was to provide a pool of potential subjects. Thus much of the questionnaire was concerned with specific nighttime factors; sleeping habits, sleep quality assessments, the incidence and perceived causes of disturbance. Other parts were concerned with the respondent's availability and suitability for subsequent participation in the sleep experiment. From their questionnaire responses, potential subjects were identified who said they were (a) willing to participate,

(b) available during the test period, (c) not deaf, (d) not suffering from nighttime pain that seriously disrupts sleep (eg severe arthritis and rheumatism), (e) aged between 20 and 70, (f) not currently taking sleeping tablets or other medications that affect sleep, and (g) not a shift worker.

A secondary aim was to collect data on factors which might help to explain observed sleep patterns. These include personal characteristics, general views about the neighbourhood, perceptions about sleep quality and the ways in which this might be affected by aircraft noise. To this end, the survey incorporated a number of questions to probe attitudes and reactions to aircraft noise, several of which had been used in previous studies of aircraft noise impact, including the CAA sleep disturbance surveys (Refs 2,3) and the UK Aircraft Noise Index Study (ANIS, Ref 18).

The social survey, which is fully described in Reference 14, was not intended to be a definitive study of subjective reactions to aircraft noise of the kind reported in References 2, 3 and 18. A comprehensive social survey study of perceived sleep disturbance, and the factors which influence it, would have been much more elaborate and would have been designed and administered rather differently.

In selecting respondents, the objective was to draw representative samples of people who lived in the local areas (which would *not* necessarily be representative of *all* people affected by noise around each of the airports). Quota samples for each site were set on age and sex using available census data for the area. Age was categorised into three groups, 20-34, 35-49 and 50-70, which were chosen to reflect possible differences in lifestyles and sleeping habits.

The survey fieldwork for the main study, including pilot work to develop the questionnaire, was conducted by professional market research interviewers. The interviewers were asked to exclude shift workers and people who had lived in the area for less than 1 month. The average interview duration was 25 minutes. Respondents were not told of the reasons for the study in advance.

Altogether, nearly 4000 addresses in the eight study sites were targeted and a total of 1636 initial interviews were conducted, with more than 200 at each site. These potential subjects underwent a subsequent structured interview centred on a sleep habit questionnaire (Ref 16). The questions covered such topics as: anxiety, illnesses, worry about health, medicines taken, smoking, tea and coffee intake, evening exercise, medical and other reasons for not getting good sleep, difficulty getting to sleep, alertness on arising, level of alertness at bedtime, and sleeping arrangements. During this interview, subjects also completed (a) a questionnaire discriminating between 'morning-types' (larks) and 'evening types' (owls) (Ref 19), (b) the Bortner Type A/Type B personality questionnaire (Ref 20) and (c) The State-Trait Anxiety Inventory (Ref 21).

Table 2 lists the subject recruitment figures and includes the distributions of age and sex of the 400 participants selected. At the first social survey contacts, 971 of the 1636 respondents volunteered for the sleep experiments. Of these, 524 were rejected and 47 were not required. The subjects selected were paid £5 for each night's participation, with additional payments of £15 per night for those who also underwent EEG monitoring.

Full complements of subjects were recruited at all sites. Table 3 compares the essential characteristics of the 1636 social survey interviewees with those of the 400 actimetry subjects and the 50 EEG-subjects.

3.3 Noise exposure

The noise exposures during the test-periods were determined from data gathered by remotely operated noise monitoring equipment using Brüel & Kjær Type SBK 1323 noise monitoring terminals. This equipment stores the measured noise data in its internal memory, then transfers it (usually at 24-hour intervals) to a central computer via cellular telephone links. Noise monitors were positioned to determine the range of outdoor aircraft noise event (ANE) levels at each site; the aim was to limit this to about 3dB. Figure 4 shows, for example, the monitor positions at Heathrow/SWM.

Usually, ANEs were readily identified as those sounds which triggered all noise monitors simultaneously (within the passage time of the aircraft). Other noises, eg from individual road vehicles, influenced much smaller areas and therefore tended to be picked up by a single monitor only. The ANEs were subsequently related to specific aircraft flights by analysis of the airport runway controller's logs from which the aircraft type and operating mode (arrival or departure) could be identified. The time of the event, specifically the time when maximum noise level was reached, was recorded to the nearest second for subsequent correlation with the sleep data. The sound levels Lmax and SEL of all events exceeding a threshold of 60 dBA were recorded whether they were due to aircraft or not (use of a lower threshold would have caused a large increase in the incidence of non-aircraft sounds) . In addition, hourly values of aircraft Leq, ambient Leq and background L_{90} levels were recorded at all sites. Reference 13 describes the noise measurement programme in detail.

The noise monitoring system operated for more than 1300 hours during the experimental programme. A total of 4823 individual ANEs were logged during the 120 measurement nights. The night-average aircraft noise variables are compared with the target values in Figure 5. In this graph, average SEL (in dBA) is plotted against average hourly number of events - for the 8-hour period 2300-0700 - at each of the sites. Except at the Manchester sites, where the numbers of events were a little higher, the traffic was lower than expected, especially at Gatwick where the shortfall was more than 30%. This is attributed to the effects of a major airline failure during 1991 as well as possible after-effects of the Gulf War.

The nighttime site noise exposures during the measurement period 2200-0800 are summarised in Table 4. These are indicated by the hourly values of (a) the aircraft noise Leq(1-hr) - encompassing all aircraft noise energy above 60dBA, (b) the 'ambient' Leq(1-hr) - calculated by removing the aircraft component (a) from the total Leq(1-hr) (this thus includes any aircraft noise below the 60dBA threshold), and (c) the background noise levels L_{90}.

Figures 6(a) to 6(h) show plots of Lmax against time of night for each of the sites. All ANEs recorded during the 15-day test periods are shown. These clearly indicate the wide range of nighttime aircraft noise exposure conditions existing at the chosen test sites. Figure 7 shows the distribution of the ANE Lmax and SEL values from all sites divided into 5dB wide bands.

To ensure that the study covered the highest possible indoor noise exposure conditions, subjects at the noisiest site (Manchester/HGN) were asked to sleep with their bedroom windows open for one of the two study weeks.

3.4 Sleep data: Actimetry and EEG

Sleep EEGs

Sleep-EEGs were obtained using Oxford Instruments Ltd Medilog 9200 recorders . These record the EEG signals on cassette tapes, and allow freedom of movement for the subject and actimetry measurements. The EEG instrumentation did not affect movement. This was subsequently confirmed by the fact that there was no significant difference between the EEG subjects' average actimetrically measured disturbance rates on nights with and without EEG.

At each site, a total of 24 subject nights of data were collected in two sessions of four nights each. Each volunteer wore the associated electrodes on four successive nights; on all occasions, these were fitted by a skilled EEG technician before 2130 hours and checked for integrity using an independent portable signal monitor. The tape recordings were collected and checked the following day prior to analysis. The analysis, including the generation of 30-second epoch hypnograms (Section 5.1) based on normal sleep stage scoring methods (Ref 22), was performed using an Oxford Instruments Ltd computerised sleep analysis system. All automatic scoring was checked visually (by the EEG technician) and corrected in clear cases of incorrect staging. Reference 15 describes the EEG work in detail.

Actimetry

Fine wrist movements were measured using Gähwiler actimeters (of Swiss manufacture). This instrument incorporates a programmable microprocessor with 32K of memory and a quartz clock. It is small (51 x 36 x 21 mm, 68 g [2 x 1.4 x 0.8 in, 2.4oz]), and no more uncomfortable than a wrist watch. It measures single-axis accelerations eight times per second and records the number which exceed 0.1g in sequential epochs of time (presettable by the researcher), in this case of 30-second duration. Although, at this setting, they could continue to accumulate readings for 11 days, data from each actimeter was downloaded to a portable computer after 7 or 8 days for transfer to longer term storage and subsequent analysis. The first step in the latter was to use Gähwiler software to generate actigrams, the basic epoch by epoch records of wrist movement (Section 5.3).

All recording instrumentation, noise, EEG and actimetry, were synchronised at intervals to a master clock controlled by time transmissions from the National Physical Laboratory. The test design aim was to ensure that no instrument ever had a time drift exceeding 15 seconds (ie half an epoch).

Recording period

On test nights, the Medilog units and noise instruments were set to record data continuously for the 10 hours between 2200 and 0800. Once switched on, the actimeters operated continuously; however, relevant data was only collected whilst they were worn, normally between the times of each individual wearer going to bed and getting up.

During the 15-day measurement period, subjects completed sleep logs each morning to record information about bedtime, lights out, estimated sleep onset, number of night awakenings and reasons for them, morning awakening, rising time and sleep quality. They also kept diaries recording sleepiness state (Ref 23) and activity every two hours during the day. On completion of the study, subjects were again interviewed using a debrief questionnaire designed to probe their perceptions of the study. It was at this point that the purpose of the study was explained prior to further questions about the incidence of, and reasons for, disturbances to their sleep and their attitudes to aviation. All the actimetry work and associated tasks are reported in full in Reference 16.

Data gathered

Overall, nearly 50,000 subject-hours of sleep data were recorded, although not all of this has been used in the analysis (which was mostly concentrated on the period 2300-0700). Of possible totals of 6000 subject nights of actimetry and 192 subject nights of sleep-EEG, actual samples achieved were 5742 and 178 subject-nights respectively. In total, 4,603,298 'valid' epochs (38,358 subject-hours) of actimeter data were collected, subdivided by:-

- site (8)
- noise (encompassing an ANE) or quiet (not encompassing an ANE); of the total number of epochs, 1.9% (87,729) were ANE epochs
- noise level (60-100plus dBA, Lmax)
- subject (400)
- time of night
- subject age (3 groups) and sex
- bedroom window state (eg open, single glazing closed, double glazing closed)

'Valid epochs' are those estimated to occur between falling asleep and final awakening. Table 5 shows the distributions of the times of sleep onset and getting up and the sleep periods estimated from the 5742 actigrams. The overall average sleep period was 7h 15m (7.25 hours, 870 epochs) with a standard deviation of 1h 15m; the average times of sleep onset and getting up were 2349 and 0704 hrs respectively.

Sleep was also monitored by actimetry in 46 bed partners (i.e. both people concurrently monitored) for 8 consecutive nights; total 368 nights. The aim of this latter pilot investigation was to assess the extent to which disturbance was common to both partners.

Altogether, the accompanying data from the pre-test and debrief questionnaires, the sleep logs and diaries comprised another 100,000 items of data.

4 SOCIAL SURVEY RESULTS

The social survey is described fully in Reference 14. This section summarises the main observations and compares them, where possible, with the results of the earlier CAA studies.

For the purposes of this analysis, the average site noise exposures are expressed in dBA, Leq for daytime (0700-2300 local) or nighttime (2300-0700). It is usual to conduct social surveys of this kind during late summer/ early autumn and to relate the responses to the summer noise exposures which are fresh in the respondents' minds. However, the interviews in this study had to be finished well before the sleep measurements began, and covered the period 16 March - 28 July. It was not possible to determine for each site long-term average noise levels matched to the specific interview period, nor would it have been appropriate to do so in cases where the averaging periods covered months of winter and early spring. Thus, the values adopted, purely for comparative purposes, are those for the conventional summer period, mid-June to mid-September, of the preceding year, 1990 - on the assumption that responses to general questions about the effects of aircraft noise are likely to be most strongly influenced by experience of that period (being the busiest in the year). Again it is stressed that the social survey was not designed to yield definitive relationships between aircraft noise exposure and reactions to it .

4.1 Distributions of main responses

The percentage response distributions are summarised in Figures 8 to 31. In each figure, the study areas are ordered, from left to right, by nighttime aircraft noise exposure (Leq 8-hr), with the noisiest to the right.

Age

The age distributions of respondents by survey site are shown in Figure 8. Although, overall, ages were divided fairly evenly between the three categories, the distribution for Manchester/EDG is noticeably different, reflecting a larger proportion of younger residents. Figure 9 shows that EDG also had a high percentage of people in the manual occupational group C2DE; it was only exceeded at Heathrow/HLW. EDG was also different with respect to length of residence; Figure 10 shows this to be rather shorter on average.

Local environment

Figure 11 shows that more than 75% of the 1636 respondents rated their area as good or excellent, ranging from ~50% at Heathrow/HLW to ~90% at Gatwick/LFD and Stansted/WSB. Possible reasons for these opinions can be gleaned from the spontaneously mentioned 'likes' and 'dislikes' shown in Figures 12 and 13. Figure 12 shows how the main reasons for liking the local area vary from site to site. Half the respondents quoted essential facilities. These include, for example, workplaces, shops, schools. General environmental aspects are the next most common (eg clean and pleasant), followed by local amenities (including community aspects, family and friends etc) and good transport links. Nearly 30% of all respondents described their areas as 'quiet'; locally the figure varied from 18% at Manchester/EDG to 47% at Stansted/WSB. Among dislikes, (Fig 13) aircraft noise was the most prevalent; particularly at Manchester/HGN and the two Heathrow sites HLW and SWM (Heathrow/SWM has the greatest daytime aircraft noise exposure of all the sites.)

Noise

Aircraft and, to a lesser extent, road traffic, were reported to be the most noticeable sources of noise in all eight areas (Fig 14). Figure 15 shows how quiet or noisy people considered their areas to be. Again the two Heathrow sites were prominent - over 50% of respondents of both areas considered them to be 'noisy' or 'very noisy' by contrast with Stansted/HAT, Gatwick/LGN and Stansted/WSB where the proportions were less than 20%.

Figure 16 gives the distribution of annoyance caused by aircraft noise. Figure 17 compares, for all respondents and sites, the distributions of annoyance caused by aircraft noise and noise from other sources mentioned spontaneously. Of the 14% of subjects who did not mention aircraft noise spontaneously, all but 2% did so afterwards when prompted.

Sleep

More than 80% of respondents said that they went to bed before midnight on weekdays, with those at Manchester (EDG and HGN) being latest and most likely to go to bed after 11pm (Fig 18). Very few described themselves as bad sleepers (Fig 19) and, according to their answers, they were roughly evenly divided between deep and light sleepers (Fig 20). Once in bed, between 30% and 45% of respondents reported difficulty getting to sleep, typically on two or three nights a week (Fig 21).

Most people reported being woken from sleep, but this occured 'regularly' in under 20% of cases (Fig 22). Figure 23 suggests that 'regularly' means every night or perhaps every other night; otherwise, reported awakening rates are fairly evenly distributed across the intervals between 'almost every night' and 'less than one night a month'. Typically, respondents said they were only awakened once per night (Fig 24). Most found it easy to get back to sleep although a significant minority (~25%) found it rather harder (Fig 25). Few were woken up at any particular times of the night although, of those who were, most mentioned midnight to 4am (Fig 26). Aircraft noise was given as a common reason for waking up (Fig 27). However, the main cause cited was being disturbed by partners or own children. Other reasons were noise from traffic and other outside sources and using the toilet.

Most respondents got up between 6am and 8am except at Heathrow/ HLW where many got up earlier (Fig 28). In total, slightly less than 50% of respondents felt refreshed or very refreshed after waking up and 25% feel tired or very tired (Fig 29). A majority slept with windows open, except at Manchester/HGN where aircraft noise exposure is highest (Fig 30). Stansted (HAT and WSB) excepted, there was a clear tendency for there to be a higher fraction of windows reported shut as the site noise exposure increases. Three of the sites (SWM, HLW and HGN) are within Noise Insulation Grant Scheme areas and this appears to be reflected by higher proportions with double glazing (Fig 31). Only at Manchester/EDG was the incidence of double glazing low, perhaps due to the high proportion of younger families.

4.2 Comparison with previous CAA studies

Several of the questions used in this survey are very similar to ones incorporated in the 1979 and 1984 CAA sleep survey studies (Refs 2, 3) and in the UK Aircraft Noise Index Study, ANIS (Ref 18). Therefore, it is possible to compare some of the present noise-response relationships with those observed previously.

Spontaneous identification of aircraft noise

Figure 32 compares the percentage of respondents spontaneously mentioning aircraft noise as a reason for disliking the area with those obtained in the ANIS. This indicates that although the results from the two studies for Heathrow, Gatwick and Manchester are in broad agreement, residents of Hatfield and West Sawbridgeworth, near to the expanding airport at Stansted, report more awareness of aircraft noise than people with similar daytime noise exposure levels at the other, more established airports. A similar effect may be seen when comparing percentages of people 'very much annoyed' by aircraft noise with those of ANIS respondents (Fig 33); again the reactions of the Stansted (WSB and HAT) residents are rather greater than the general trend.

Reported sleep disturbance

Comparisons with the 1980 and 1984 CAA sleep survey results from References 2 and 3 are made in Figures 34 to 37. In each of these, the results from four surveys are plotted:

(a) the present survey, 8 sites, 1636 respondents, (b) the 1980 interview survey (Ref 2), 8 sites, 964 respondents, (c) the 1980 postal survey (Ref 2), 22 sites, 3188 respondents, and (d) the 1984 postal survey, 5 sites, 1000 respondents (Ref 3). Figure 34 shows the percentage of respondents at each site giving aircraft noise as the *main reason for sleeping with the windows closed*. The wording of the questions in the three questionnaires was very similar and the responses clearly exhibit similar trends, albeit with the large scatter typical of social survey data. The percentages giving aircraft noise as the *main reason for having difficulty getting to sleep*, for being *awakened once asleep* and for *having difficulty getting back to sleep*, once awakened, are plotted in Figures 35 to 37. There are similar degrees of agreement in all three cases, suggesting that, in relation to night noise exposure in Leq, general perceptions of nighttime aircraft noise effects have changed little since 1980.

4.3 Factors contributing to sleep disturbance

Returning now to the results of this study alone, caution has to be exercised when interpreting 'raw' results of the kinds presented in Sections 4.1 and 4.2 because of the possibility of *confounding* effects. These could arise if factors other than aircraft noise which influence sleep and sleep disturbance are not distributed randomly across the different study sites. However, it has been possible to identify, from the present social survey, some factors which appear to affect reported sleep disturbance and other responses to aircraft noise. This analysis is described in Reference 14.

The analysis indicates that reported reactions to aircraft noise, both annoyance and sleep disturbance, are influenced by numerous intervening factors, notable amongst which are age, sex, marital status, and whether subjects describe themselves as light or deep sleepers. However, no clear relationships emerge between aircraft noise exposure and reported disturbances, whether known intervening factors (confounding effects) are controlled or not. A trend which appears to nullify any systematic noise effect, and which perhaps holds the key to better understanding, is that reactions from the two Manchester sites tend to be the reverse of what might be expected from their relative noise exposures. Residents in Heald Green, in the main, reported less disturbance and annoyance than those in Edgeley, sometimes markedly less. This is despite the fact that nighttime aircraft noise exposure at Heald Green is the highest of all the sites and considerably more than at Edgeley. Further study of this finding could throw important new light on factors which contribute to aircraft noise annoyance.

5 MEASUREMENT OF SLEEP DISTURBANCE USING ACTIMETRY

5.1 EEG definitions of sleep

As already noted, EEG is generally recognised as the 'gold standard' method for assessing sleep state. For this reason, a sub-sample of EEG records was obtained to provide baseline measurements of sleep disturbance to which the actimetric measurements could be compared.

The EEG records define the predominant sleep stage in each 30-second 'epoch' of the night. The epoch by epoch record of sleep stage during the night is known as a *hypnogram*. The stages are:

W	Wakefulness
M (or MT)	Movement Time
Stage 1	"Shallow" sleep
Stage 2	"Light" sleep
Stages 3 and 4	"Deep" sleep
REM	Rapid Eye Movement (dreaming sleep)

It should be noted that REM sleep does not represent a distinct level in the natural sleep stage hierarchy; its position in the list vis-a-vis 'depth' is unclear. Movement Time reflects the presence of large electrical disturbances called *'artefacts'*. Such events are bursts of non-EEG activity of greater than normal EEG intensity associated with increased muscle activity on the scalp and/or movement of electrodes or electrode leads. They are large enough to mask underlying brain signals and may therefore be considered to indicate significant disturbances of sleep. Any single epoch may encompass waves of more than one type; the standard scoring method (Ref 22) records the predominant stage present during the epoch.

5.2 Definition of sleep disturbance

The main question to be addressed is, "does aircraft noise cause sleep disturbance within sleep itself ?" This is distinct from the questions of:

a) whether such noise at bed-time interferes with the process of getting to sleep, or

b) whether such noise causes premature awakening at the end of sleep.

Data gathered during this study may well throw light on these latter questions, which are the subject of continuing analysis. The results presented here are mainly concerned with the primary line of enquiry.

In pursuing the main question, the initial problem was to establish an acceptable and feasible definition of "sleep disturbance", based on EEG criteria, and applicable within the context of aircraft noise. Such a definition was central to the assessment of actimeter sensitivity. The scientific literature is unhelpful in defining sleep disturbance as there are so many interpretations, ranging for example, from a transient sleep stage shift to an awakening lasting for several minutes. To help resolve this, several internationally-known sleep experts were asked for their definitions. Whilst there were differences of opinion, it was generally agreed that any measurable period of wakefulness is definitely indicative of sleep disturbance. A sleep stage change, particularly a lightening of sleep from, say, stage 2 to 1, was considered as a minor 'perturbation'. Any adverse effects of awakenings on daytime sleepiness and performance would probably not be evident until these occurred over six times per hour during the night (Ref 24). Such daytime effects may be apparent after fewer but more lengthy 'wakefulness' episodes; the longer and more frequent the episodes, the greater the possibility of secondary effects. The decision was therefore to define as a 'sleep disturbance' any EEG-determined arousal to wakefulness or movement time.

Although the EEG is seen as the standard, it still has certain limitations; in particular, there can be considerable difficulty in identifying transitory wakefulness in the EEG signal particularly when the EEG signal is overlaid by movement artefact from the electrical activity in muscles around the scalp. In general, the shorter the episode of W-type waveform (the pattern in the EEG record which signifies wakefulness), the less is the certainty that this is indicative of any real disturbance. It is common sleep monitoring practice (Ref 22) to present EEG data in 30-second epochs, recording the predominant sleep stage for each one. Usually, this means that the recorded sleep stage occupies more than 15 seconds of the epoch. However, for movement time, the Oxford Instruments system (Section 3.4) provides the facility to set the time limit as low as 30% of the epoch. Advantage was taken of this to adopt a slightly more cautious EEG definition of 'awakening' for the purpose of calibrating the actimetry method. For the purposes of this study, *an EEG-awakening was defined as any period of wakefulness of 15 seconds or more, or any movement time lasting 10 seconds or more*. Subsequently, finer grained responses were considered in a separate examination of EEG records for direct evidence of aircraft noise induced disturbance (Section 8).

5.3 Determination of arousals: 'filtering' of sleep records

A 'sleep record' is defined as the epoch-by-epoch trace of the measured quantity, ie sleep state (hypnogram) or wrist movements (actigram). The requirement is to transform these into simpler records of 'disturbance onsets'. Figure 38 shows examples, for one subject-night, of the basic actigram (a) and hypnogram (d), together with the transformed sequences of 'disturbance onsets' (b) and (c). The appearance of the latter graphs in Figure 38(b) and 38(c) prompted these onsets to be described as 'blips'; *H-blips* from the hypnogram and *A-blips* from the actigram.

The process by which raw EEG and actimetry sleep records are transformed into disturbance blips is termed *filtering*. The filtering process received a great deal of attention during the study.

Hypnogram filter

The transformation of standard hypnograms was relatively straightforward. The hypnogram record assigns a particular sleep stage, W, M, 1, 2, 3, 4 or REM, to each sequential 30-second epoch, epochs labelled W or M being assessed according to the definitions in Section 5.2. An 'awakening' or H-blip was marked in any W or M epoch which was not preceded by another W or M within a 'buffer' interval of a specified number of epochs. This buffer defines the minimum allowable separation between successive awakenings. The filtered record is defined as H(b) where b is the length of the buffer in epochs. The minimum buffer is 1, so that H(1) leads to the maximum possible number of arousals in the record.

Actigram filters

The choice of method for transforming actigrams was less obvious because there is no unique or systematic pattern of disturbance within the movement count histories. The search for a suitable actigram filter (A-filter) was based on maximising the coincidence between A-blips and H-blips in the records of those subjects who took part in the EEG monitoring. The extent to which A-blips coincide with H-blips may be quantified by the 'hit rate'. This is illustrated in Figure 39 which compares two hypothetical blip sequences. The first and second rows contain 57 A-blips and 50 H-blips respectively. The bottom row contains 28 'hit-blips' at those positions where the A- and H-blips coincide. Thus, in this example, the hit rate of A on H is 28 out of 50 or 56%. The reverse hit rate, of H on A, is 28 out of 57 or 49%.

It was observed that H-blips tend to be associated with the initiation of bursts of wrist movement following periods of relative immobility. Thus a filter algorithm of the form *a,b,c* was adopted where a blip was assigned to any epoch in which the actimeter count is not less than *a* <u>and</u> none of the counts in the preceding *b* epochs reaches *c*. Detailed comparison of many paired actigrams and hypnograms yielded a number of filters (ie particular combinations of *a*, *b* and *c*) which performed reasonably well.

Extensive analysis showed that a very large number of different filters yielded similar results. It was also apparent that the better performing filters were those which matched the buffers *b* of the H- and A-filters. The simplest actigram filter, upon which most attention was focussed, had the form *m,1,m*. This assigned a blip to any epoch in which the count was *m* or more and which was not immediately preceded by a similar count, ie consecutive blips were inadmissable.

A factor which had to be taken into account is that the epochs of different records might not be exactly synchronised due to differences in the clock settings of different actigraphs and medilog recorders. Although small, these differences might be sufficient to cause overlap, ie an apparent mis-match of blips which were in fact concurrent. To allow for possible overlap, hits were scored when blips coincided within a time lag of ±1 epoch.

A major task of the study was an attempt to determine optimum A-filter characteristics. This involved a very large number of computations. Many of these involved independent variation of the three filter parameters *a*, *b* and *c* as well as the corresponding H-buffer. In some, numerous other EEG-events were admitted in addition to the basic W/M arousals. These included artefact events, shifts to Stage 1, and movement time episodes of less than 10 seconds duration.

A fundamental aspect of A-filtering is that 'coarser' filters, ie those which yield more A-blips, naturally lead to higher hit rates of A on H. However an *increase* of this hit rate tends to be accompanied by a greater *decrease* in the reverse hit rate of H on A, ie the fraction of A-blips which are coincident with H-blips. In most cases, the reverse hit rate is markedly lower than the forward rate and it is arguable that the best filter is that which gives the highest 'average' hit rate (this average can be defined in numerous ways and various forms of correlation coefficients were considered as filter performance indicators). However, a conclusion from the expert seminar was that the most important requirement is not to underestimate the incidence of awakenings, as indicated by the number of EEG H-blips. This is achieved by maximising the hit rate of A on H, accepting that this will result in a relatively low reverse hit rate - in other words, a significant fraction of the resulting actimetric A-blips will indicate minor perturbations of less significance than awakenings.

The A-filter that best fitted EEG arousals was one that registered a blip in any epoch where actimetrically recorded movement, no matter how little, followed one epoch of nil movement. This filter is referred to as "1,1,1" (i.e. assigning a blip to any epoch with a movement count of 1 or more, following at least 1 epoch with a count of less than 1, ie zero). Again, *this filter detects the maximum possible number of movement onsets in any record*. It is the simplest possible filter and was used to generate all the main results of the study presented in Section 7.

5.4 Estimation of Sleep Onset

It also had to be gauged from the actigrams when sleep actually began, that is, the time of sleep onset each night. Again, the EEG data subset were used to determine a 'best fit' between EEG and actimetric criteria. The EEG definition for sleep onset (Ref 22) was the start of the first ten minute period of continuous sleep consisting of stage 2 or deeper. It was found by matching the corresponding actigrams that movements tend to cease for at least 14 epochs (7 minutes) after sleep is established following 'lights out' (as reported in the sleep logs). However, the above stage 2 condition was not met, on average, until ten epochs (five minutes) into this 7-minute period. A sleep onset algorithm ('14,10') based on this result identified 72% of EEG-determined sleep onsets from the actigrams, to within ±10 minutes. This algorithm was used to define sleep onsets for all subjects, although it is possible that the accuracy might be further improved by applying different expressions to different age groups.

5.5 Comparison of EEG and actimetrically measured disturbance

Table 6 shows the agreement between EEG and actimeter measured sleep disturbance, overall and for each each of the study sites separately. These results were calculated from the 178 paired records obtained from those subjects whose sleep was monitored simultaneously by both EEG and actimeter. Including only those periods between initial sleep onset and final awakening, this comprised a set of 135,643 matched 30-second epochs.

Of these epochs, 2530 contained awakenings (H-blips), of which 2226 (88.0%) were matched (within ± 1 epoch) by actimetric arousals (A-blips). Broken down by site, the detection rates varied from 83.4% to 92.1%. In addition, of the 133,113 epochs in which there were no H-blip awakenings, A-blip arousals were also absent in 129,184 of them; ie actimetry accurately identified 97.0% of non-awakenings. Given the uncertainties associated with all sleep measurements (eg, manual identification of EEG sleep stages, considered to be the most reliable technique, is only repeatable to about 95%) these figures, of 88% and 97%, confirm that actimetry provides a very satisfactory method for detecting awakenings.

For all sites, the EEG sample awakening rate of 1.86% translates to approximately 16 awakenings per 7.25-hour night (Section 3.4), for the average individual. By comparison, the average A-blip arousal rate is 4.74% or 41 arousals per night. This suggests that about 40% of A-blips represent awakenings, the remaining 60% being minor perturbations. These include small movements and natural twitches, some of which are associated with shifts to Stage 1 sleep and short duration arousals. This 'awakening-to-arousal' ratio of 40% is an estimate based on a limited data sample of 178 subject-nights. It is thus subject to a sampling error, which, on the basis of normal probability theory would be expressed by a 95% confidence interval of ±7%. However, to allow for additional uncertainties discussed later (Section 6.3), it is more realistic to accept an error of perhaps ±10%; ie the true average ratio which would be determined from a very much larger set of measurements probably lies in the range 30-50%.

6 METHODS OF ANALYSIS

6.1 Actimetry data

Figure 40 shows one of the 120 sets of 50 actigrams obtained in this study (one night at Stansted/HAT). Figure 41 shows the corresponding A-blip records obtained by '1,1,1' filtering (Section 5.3). In both of these figures, the diamond-shaped markers at the beginning of each record denote the estimated times of sleep onset.

Figure 42 shows the sum of the 50 individual A-blip records from Figure 41; ie the total number of subjects aroused from sleep in each epoch. The lower trace shows the times of occurrence of ANEs; in each case the height of the line is indicative of the sound level Lmax of the event. It was expected, initially, that such aggregations would provide a simple means of identifying the incidence of aircraft noise induced sleep disturbances. However, in Figure 42, which is no different in its general features from any of the 120 such records available, there is no obvious correlation between the ANEs and the incidence of arousals. Indeed, most of the arousal peaks occured in the 'quiet' periods, a clear illustration that factors other than aircraft noise controlled sleep-wake patterns.

To determine the specific rôle of aircraft noise, it has been necessary to resort to more elaborate techniques for analysing the A-blip records. These are:

- a simple, but statistically limited, comparison of disturbance rates in 'noisy' and 'quiet' epochs (Section 6.2)
- a more soundly based statistical analysis of these *n*- and *q*- disturbance rates, allowing for the 'confounding' effects of underlying factors of importance such as age and sex (Section 6.3), and
- an alternative method, which also allows for the effects of various non-noise factors but, additionally, takes specific account of any disturbance experienced in the quiet periods immediately before the ANEs (Section 6.4).

6.2 Simple estimates of aircraft noise induced arousal

These are based on a direct epoch by epoch analysis of the actimeter data. Each epoch is described by the following variables:-

Site, night, subject, time, noise/quiet, noise level, arousal (y/n)

Although the noise of an individual aircraft movement may span more than one epoch, the 'noise epochs' are defined as those which contain maximum sound levels Lmax occurring during ANEs; 'quiet' epochs are those which do not . Noise epochs are also described by the *sound levels* of the ANE in dBA, both Lmax and SEL. Subjects are categorised by *sex* and *age group* but each subject can also be described by a large number of ancillary variables determined from the questionnaires, sleep logs and diaries. Arousals are the actimetrically determined disturbance onsets described in Section 6.3.

For any data sample, three sleep disturbance variables are defined:

a = overall arousal rate, the ratio of aroused epochs to total epochs expressed as a percentage
q = arousal rate in quiet epochs only
n = arousal rate in noise epochs only

The effects of aircraft noise on sleep disturbance may be expressed in terms of the difference n-q between the arousal rates in noise and quiet epochs. The total rate of arousal in noise epochs is n but, of this, q would have occurred in quiet anyway. The difference n-q is therefore an estimate of the *rate* at which arousals may be *caused* by noise, ie the *aircraft noise related arousal rate.*

Because of the random nature of sleep disturbance, small values of n-q may arise purely as 'sampling fluctuations', the chance results of particular measurements; ie they may not arise regularly if the measurements were repeated many times. Such occurrences would not be statistically significant. Provided the sample sizes are not too small, the probability that any particular value of n-q is not significant can be estimated using standard tests based on normal probability theory such as that described in Appendix A. However, there are serious difficulties with this simple approach which are discussed below.

6.3 Multivariate analysis: logistic regression with random effects

It was noted in Section 6.1 that there is unlikely to be any simple relationship between aircraft noise and an individual's sleep disturbance. This is because aircraft noise is only one of many factors which can affect sleep disturbance.

In an ideal study (ie a hypothetical one with unlimited resources) the measurements would be fully representative of the entire population of interest and would be made in such a way that the effects of the influencing factors are independent of each other. Then to determine the true effects of noise, for example, it would only be necessary to plot a simple graph of measured disturbance against noise level. The effects of non-acoustical factors such as age, sex, time of night, etc could be determined by similar analyses.

Such graphs are presented here but, in evaluating the results, it is most important to recognise that the underlying factors will not, in general, be statistically independent. For example, disturbance may be dependent upon age, sex and noise level, but these factors may not be distributed randomly with respect to each other. For example, subjects who are more disturbed by virtue of their age and sex might be concentrated at the low noise sites and *vice-versa*. In this case the noise-disturbance relationship inferred from a simple graph would be *confounded* by these hidden age and sex effects.

In this example, the true relationship between noise and disturbance could only be determined by an analysis in which the effects of age and sex are properly *controlled*. To do this, the effects of all three factors, noise, age and sex, must be determined concurrently using *multivariate* analysis methods. For reasonably straightforward problems, ie those involving a few known factors whose effects are simple (eg linear) the necessary statistical tools are readily available. One of these is *multiple linear regression* which is commonly used to generate a linear mathematical model of the relationship between a number of independent 'explanatory' variables (eg noise level, age, sex, etc) and a single dependent 'outcome' variable (eg level of annoyance). An essential assumption underlying the use of linear regression is that all variables are measured on (or can be converted to) continuous, equal interval scales and that the effect of any independent variable is linear (although non-linear terms can sometimes be handled via a transformation of variables). Dichotomous or binary variables (eg whether the subject is male or female) can be accommodated as independent variables.

The above is only a simple example of the kind of complex effects of underlying variables which could confound the results of this study. Some factors are and may remain unknown. Some may have unknown non-linear effects. A special feature of the actimetry data is that the measured dependent or 'outcome' variable, probability of disturbance, is a *proportion* rather than a continuous linear variable. This means that linear regression is inappropriate; it cannot generate a mathematical model of a variable which has upper and lower limits (0 and 100%).

In Section 7, the effects of various factors on sleep disturbance are examined separately. In most cases uncontrolled (ie unadjusted) results are presented first to indicate possible trends. In some of these cases, especially when apparently large effects emerge from very large data samples, simple statistical tests are adequate to confirm that the effect is not simply a sampling fluctuation. In others, where the data samples are small, simply estimated confidence intervals may be very unreliable. In such cases, they are described as approximate estimates. Where possible, multivariate methods have been used to control for the effects of other confounding variables in order estimate how sleep disturbance depends on the primary factor of interest alone. For this purpose, an elaborate statistical procedure known as *logistic regression analysis* (LRA) was used. Unlike linear regression, this can make allowance for the fact that 'disturbance' is not a quantity measured on a continuous linear scale but a proportion lying between 0 and 100%. The principles of LRA are summarised in Appendix B.

A difficulty with conventional statistical methods such as tests for the differences between proportions (Section 6.2), as well as the more common multivariate procedures, is that they generally rely on the assumption that test observations are independent of each other. In other words, they assume that the data represent a sample drawn randomly from a large population, with any particular observation having the same chance of being selected as any other. Whilst it is obviously reasonable to suppose that all 400 test subjects behave independently, the same cannot be assumed of different observations (repeated measures) from the same subject. Because a subject's sleep state at any particular time must depend to some extent on their previous sleep state, the observations (epochs) are said to be *serially correlated.* Simple statistical tests are not valid for serially correlated data.

It has been possible here to control for serial correlation effects upon *n* and *q* values by using a modified *random effects* version of logistic regression analysis (LRA) discussed in Appendix B. Unfortunately, because of computational limitations, it has not been possible to apply this technique to all 4.6 million epochs simultaneously. The computations have instead been made using data subsets and this inevitably entails some loss of statistical power (ie the confidence in the results is less). This increases the risk of not detecting small effects which may nevertheless be real. However, this limitation has been largely overcome in the third method of analysis described below.

6.4 The Wilkinson-Diamond method

This addresses directly the possibility that whether or not a person will be aroused from sleep depends upon his or her immediate past history of arousal. A suitable method was suggested independently by Dr Wilkinson and Prof Diamond and is described therefore as the W-D method (see Appendix C). This involves comparing arousals that coincide with aircraft noise events (ie ANE-epoch arousals), not with a long term average arousal rate in quiet, but with arousals in specified quiet periods preceding the noise events themselves. In this case, the data set is restricted to ANE epochs alone and the extent to which an individual is disturbed in the preceding quiet period has been represented in two ways:

(i) according to whether the individual is disturbed in an epoch chosen at random from within the quiet period - the answer (y/n) is represented in the analysis by a 'dummy' variable.

(ii) according to the rate of arousal in the quiet period. This is the ratio of the number of disturbed epochs (A-blips) to the total number of epochs in the preceding quiet interval (effectively the time since the previous ANE).

The W-D analysis was also performed using LRA with random effects. Because it was possible to manipulate all of this reduced data set simultaneously, this approach yields firmer conclusions than the *n,q* analysis about the effects of explanatory variables upon sleep disturbance. The method and the analyses performed are described in Appendix C. The results are summarised at the appropriate points in Section 7.

7 MAIN RESULTS: FACTORS AFFECTING SLEEP DISTURBANCE

7.1 Actimetrically measured arousals: definition

The results in this section are based on analyses of A-blips. *For present purposes, an A-blip is defined as a sleep arousal.* This point is important because the word *arousal* is ill-defined in the scientific literature and is often applied to different kinds of sleep events. Correlations of actigrams and hypnograms (Section 5.5) from the EEG subsample have shown that about 40% of A-blips (movement onsets) coincide with H-blips, onsets of EEG measured wakefulness or movement time which are here collectively termed *awakenings*. The remaining, unmatched A-blips indicate lesser degrees of arousal, including sleep stage changes, large REM-twitches and other minor perturbations. When interpreting actimetric data in terms of likely awakening rates it is assumed in what follows that the 'awakening-to-arousal' ratio is the same for all sleep disturbances, whatever their causes. That is, statistically, about 40% of arousals represent actual awakenings although *which* 40% is not known. Also, it must be remembered that the 40% proportion is based on the relatively small EEG-subsample of subject nights and it is therefore subject to a separate sampling error, of perhaps ±10% which affects all actimetrically based estimates of awakening rates.

7.2 Timing of aircraft noise induced arousals

An initial question concerned the possibility of time lag in responses to aircraft noise: do noise induced arousals occur in epochs following those containing the ANEs? A comparison of arousal rates a in the noise epochs with those occurring in the 1st, 2nd, 3rd following epochs, etc. indicated a significant difference between the noise epoch and the 1st following epoch, but not between the 1st following epoch and any remaining following epochs. This indicates that any significant noise effects are confined to the noise epochs themselves; responses to ANEs occur within a very short time and arousal lag is not significant. Thus the noise arousal rates n in Table 7, which lists the actimetric results by site, relate to the noise epochs only.

7.3 Sleep disturbance rates and site differences

Of the 4.6 million valid epochs, A-blip arousals occurred in 243,602 or 5.29% of them. Again, expressed as a fraction of the average 7.25-hour night, this translates to 46 arousals per night, also on average. This rate is higher than the 4.76%, or 41 per night average for the EEG subjects during Medilog nights (Section 5.5). In fact, the EEG subjects were, over all their actimetry nights, around 10% less disturbed than the overall average subject (this is an example of a 'sampling fluctuation'). An inference from this is that the average awakening rate for all 400 subjects is 10% greater than the figure of 16 per 7.25-hour night derived from the EEG data sample, ie around 18 per night. Remembering the sampling error of perhaps ±10% associated with the estimated awakening-to-arousal ratio (Section 5.5), this is more properly stated as a number in the range 18±4 awakenings per night. It is emphasised again that most of these awakenings are very short; Figure 43, which shows the distribution, in epochs, of EEG-measured episodes of wakefulness and movement time, indicates that more than 80% of awakenings last less than 1 minute and two thirds are less than 30 seconds.

There are differences between the overall arousal rates for different sites, which range from 4.80% at Heathrow/SWM to 5.60% at Manchester/EDG. However, the W-D analysis (Appendix C) shows that when confounding factors are fully controlled, the differences between the arousal rates at the eight sites are not statistically significant. As these sites covered a very wide range of aircraft noise exposure, this finding supports the decision not to include control sites (see Section 2.5).

7.4 Distributions of disturbance and noise sensitivity: individual differences

Figure 44 shows how 'arousability', expressed by the average A-blip rate a, varies across all 400 subjects. The overall average rate (Table 7) is 5.29% but individual variation about this is large; 3% of subjects have rates outside the range 3% to 9%. It must also be remembered that the study sample excluded people who reported taking sleeping tablets (4.4% of the social survey sample). Thus the possibility exists that some especially sensitive people may have been excluded because they were taking sleeping tablets. This would, of course, have blocked arousals anyway so that inclusion of these individuals could have led to an underestimation of sleep disturbance.

Because the actimetry data comprise many repeated observations from single individuals, it is possible for a small number of subjects with very large or very small arousal rates to bias the results quite markedly. A major focus of the analysis has therefore been the need to control for the powerful confounding effects of individual variations in arousability. Analysis of part of the actimetry data (in three separate time periods) using LRA (Appendix B), confirmed that, even when this and other confounding factors are controlled, variations in arousal rates due to all causes remain large, the most sensitive subjects being aroused over 2.5 times more often, on average, than the least sensitive (together, these most and least sensitive subjects comprise 5% of the total).

Concentrating attention upon noise epochs only suggets that the variability in noise-related arousal rates may be higher. The W-D analysis of all ANE epochs between 2300 and 0530, described in Appendix C, showed that subjects of high arousability were disturbed 78% more than average; those of low arousability, 44% less, a ratio of just over 3 to 1 (Table C3). (Again, subjects of 'high' and 'low' arousability together comprise 5% of the total.) And these all-site results may be masking a trend for noise related arousal rates to be even more variable; a separate analysis restricted to the Manchester data, which contributes most of the high noise level information, gave a ratio of 4 to 1; subjects of high arousability having twice the average chance of being disturbed and those of low arousability, half the chance (Table C2).

Figure 45 shows the distribution of individual subjects' 'aircraft noise related arousal rates', *n-q*, calculated for all ANE epochs (*q* here is the subject average - in all non-ANE epochs). Although a majority of subjects (57%) register positive *n-q* values, the remainder have negative ones. At first sight, this seems to point to the unlikely possibility that aircraft noise actually suppresses sleep disturbance. However, this paradox can be explained as a consequence of two confounding factors.

The first is that, as will be seen, the arousal rate in noise epochs, *n*, is dependent upon the ANE sound levels; much of the variation of *n-q* in Figure 45 is simply a natural variation of general arousability in epochs where low level ANEs have little or no effect upon *n*. The second is that the number of ANEs differs greatly from site to site, from night to night and from hour to hour; many subjects experience so few events that their measured *n*-rates are statistically unreliable estimates of the true rates.

Figure 46 shows the results of removing parts of the data to reduce the effects of these factors. Here, the *n-q* values have been calculated only for ANE epochs with Lmax ≥ 80dBA. Two distributions are shown: (i) for all subjects and (ii) only for those subjects who experienced more than 100 ANEs during their 15-day measurement periods. There are relatively few of the latter subjects, but it is clear from Figure 46 that they include a higher proportion with positive *n-q* values. The two curves have been normalised in Figure 47 by plotting the percentage of subjects, rather than their actual numbers, against *n-q*. Of the subjects with over 100 ANEs, more than 75% have positive *n-q* rates.

Accepting that aircraft noise is unlikely to *reduce* sleep disturbance, it must be concluded that the remaining negative *n-q* values in Figure 47 also reflect sampling errors - although these particular individuals may well be among the least noise sensitive. To understand this, a 'dummy' disturbance variable, *q'-q*, may be imagined where *q* is still the arousal rate in the large number of non-ANE epochs, but *q'* is the arousal rate in a random sample of non-ANE epochs, equal in number to the ANE epochs themselves. Because of sampling errors, the *q'-q* values will also vary from subject to subject, the amount of the variation increasing as the sample sizes decrease. It is reasonable to assume that a similar sampling phenomenon is responsible for much of the variation of *n-q* in Figure 45.

7.5 Noise level

Counting all ANEs with Lmax > 60 dBA, the overall arousal rate in the presence of aircraft noise (ie in ANE epochs) was found to be 6.18% (Table 7). Thus, subtracting the overall non-ANE arousal rate *q* of 5.27%, an estimate of the overall average aircraft noise related arousal rate, *n-q*, is 0.91%. However, this encompasses noise events of all levels, from 60 dBA to more than 100 dBA, Lmax.

Relationships between the 'unadjusted' ANE-epoch arousal rate *n* (ie without controlling for confounding effects) and aircraft noise event level, measured in (a) Lmax and (b) SEL, are shown in Figure 48, together with the overall non-ANE arousal rate, *q* (the incidence of actimetric arousals in all epochs that do not coincide with aircraft noise). These graphs have been generated by grouping the ANE epochs into 5dB wide bands.

There is really no material difference between these two graphs; the two are shown for information as both scales are widely used for measuring ANEs. Figure 49 shows the relationship between Lmax and SEL determined from the ANEs measured in this study. The points are arithmetic averages of the SELs for the ANEs grouped into 3dB bands of Lmax; the error bars denote ±1 standard deviation of the data. The regression line fitted to these averaged points is

$$\text{SEL} = 23.9 + 0.810\ \text{Lmax}$$

This gives the following approximate equivalencies:

Lmax:	70	75	80	85	90	95	100
SEL:	81	85	89	93	97	101	105

Use of this transformation between Lmax and SEL will reveal that the two graphs in Figure 48 are equivalent. The fact that SEL and Lmax are so highly correlated is partly attributable to the predominance of approach noise in this study (which rightly reflects the high proportion of arrivals in nighttime aircraft movements) which means there is not very much variation in the duration of the ANEs. Because of this, and the generally weak effect of aircraft noise level, it is impossible to distinguish between the performance of Lmax and SEL as indicators of sleep disturbance. For the purpose of interpreting the results in this report, the two scales may be regarded as completely interchangeable (using the above conversion).

Comparison of the rates of ANE-arousals and non-ANE arousals in Figure 48 suggests that aircraft noise begins to cause additional sleep disturbance when ANE levels exceed 80 dBA SEL (70 dBA Lmax). However the results shown in Figure 48 have not been adjusted to remove the statistically confounding effects of non-acoustical factors. When these effects are controlled in an LRA analysis (see Appendix B), the effect of aircraft noise is not statistically significant for ANE levels below about 90 dBA SEL (80 dBA Lmax*). This conclusion is supported by the separate W-D analyses, described in Appendix C (Table C3), of a large samples of ANE-epochs. Although in this case, unlike that of Appendix B, quiet epoch categories were excluded, no significant differences were found between the chances of being disturbed by ANEs in the four bands <75, 75-79, 80-84 and 85-89 dBA SEL. Only at levels of 90 dBA SEL and above was the chance of disturbance significantly greater.

The LRA analysis in Appendix B gives an estimated 3.3% increase in the average arousal rate (equivalent to *n-q*), for all ANEs above 90 dBA SEL (80 dBA Lmax). That is, the additional chance of being aroused during a higher level aircraft noise event is about 1 in

---

* Here, the SEL-Lmax difference is rounded to 10 dB as the statistical estimates of the threshold of disturbance are not as precise as the physical measurements of the noise levels.

30. Assuming that approximately 40% of actimetric arousals are awakenings (Section 5.5), about 1 in 75 ANEs above 90 dBA SEL (80 dBA Lmax) would waken the average person†. Events 'above 90 dBA SEL' essentially means events between about 90 and 100 dBA SEL (between about 80 and 95 dBA, Lmax) because most higher-level events measured in this study lay in this range (Fig 7(b)). However, the likely variation of average arousal rates within this sound level range have been estimated using a procedure described in Appendix D.

This makes use of the results of the W-D analysis described in Appendix C to give estimates of the ANE-related arousal rates with the particularly strong confounding effects of individual arousability controlled (other factors of lesser importance are disregarded). These are shown in Figure 50 together with 95% prediction intervals (these are similar to confidence intervals; see Appendix D). The events above 95 dBA SEL are grouped into a single category; no separate estimates could be made for smaller bands because of the small amount of data (Fig 7(b)). Included for comparison is the overall q-rate (5.1%) for the period 2330-0530 used in the W-D analysis. Figure 50, which shows the ***estimated** disturbance rate for an individual of average arousability*, is a graphical illustration of the finding (Appendix B) that no statistically significant increase in sleep arousal rates can be associated with ANE levels below 90 dBA, SEL (80 dBA, Lmax). Only at higher ANE levels do the prediction intervals exclude the non-ANE arousal rate.

The simplest interpretation of Figures 48 and 50 is that the incidence of disturbance increases by about 1% with each 5dB increase in aircraft noise event level. However, ignoring the sampling errors, the relationships in Figure 48 exhibit some curvature which might be better represented by non-linear sigmoid-shaped curves. Because of the statistical constraints of limited data samples it is not possible to be definitive on this point, but it does seem important to recognise that disturbance rates could increase relatively rapidly at ANE levels above 100 dBA SEL (95 dBA Lmax).

These observations relate to the average person. It may be seen from the results in Appendix B that the most arousable subjects are more disturbed than these average rates suggest. However, to some extent, people who are more sensitive to aircraft noise are more likely to be aroused for other reasons, thus lessening the effect which should be attributed to aircraft noise.

7.6 Time interval between events

An important practical question, which is the subject of continuing analysis, concerns the possible effects of the time interval between successive ANEs; ie that shorter intervals might increase the probability of the second event causing arousal. The W-D analysis (Appendix C) revealed that the probability of arousal by an aircraft noise event (ie in an ANE epoch) increases with the arousal rate in the 'quiet' interval since the last event; ie a person who has been recently disturbed is more likely to respond to an aircraft noise. This was corroborated by the independent analysis of the EEG data (Section 8). Thus it may be inferred that, if one ANE causes disturbance, this will increase to some extent the probability of disturbance by an immediately following ANE. However, as the independent probabilities of either noise causing disturbance are low, any additional disturbance attributable to repeated events is likely to be very small. This appeared to be confirmed by the W-D analysis (Appendix C) in which 'time since the last ANE' was not found to be a factor of significance. However, in this analysis, ANEs which occurred within 5 minutes of a preceding event were omitted; further examination of this effect is continuing.

---

† Remembering the ±10% sampling uncertainty associated with the 40% proportion, it is more accurate to say that the wakening rate probably lies somewhere between 1 in 60 and 1 in 100)

Again, it has to be stressed that these observations relate to arousals from sleep. No conclusion can yet be stated about the possibility that a second ANE might impede return to sleep after an awakening. Whether or not this has an important bearing upon end-of-night sleep disturbance is a question still being examined.

7.7 Age and sex

Figure 51, which shows unadjusted sleep arousal rates for subjects divided by age and sex, suggests that males are more susceptible to disturbance than females and younger people more than older people. However, the difference between *n* and *q* does not change uniformly with age. Figure 51 indicates that in females it varies little and that males are more noise-sensitive than females.

Again, however, the conclusions have to be modified when the effects of individual variability and other variables are taken into account. Two analyses have been undertaken. The first, using LRA, considered all Manchester site epochs within three time periods: 0100-0130, 0300-0330 and 0500-0530 (Appendix B). Here the outcome was whether or not the individual is disturbed in a particular epoch and variables examined included the noise level and the individual's age and sex. This analysis indicated that the sex related difference was small; on the basis of the Manchester data, men are 10% more likely to be disturbed from any cause than women.

The second analysis, using the W-D approach, considered all sites but only ANE epochs. This indicated that although, in general, men's sleep is about 15% more likely to be disturbed than women's, a statistically significant difference, aircraft noise does not affect them differently, ie men are no more susceptible than women to aircraft noise than the general differences would suggest.

7.8 Time of night

Figure 52 shows the percentage of subjects asleep and the incidence of aircraft noise events, also expressed as a percentage, during the course of the 'average night'. These figures have been calculated from the full set of epoch data. The sleep onset times have been estimated from the actimeter data (Section 5.4); the end of sleep is the time of getting up reported in the daily sleep logs. The two curves naturally show opposite trends; aircraft traffic diminishes as more people fall asleep and it increases as people are waking up. (although no causal relationship can be inferred).

Figure 53 shows the 'in-quiet' arousal rate *q*, averaged over all subjects, in 15-minute intervals from sleep onset, a time scale referred to subsequently as 'time of sleep'. The upper and lower lines give approximate 95% confidence intervals for the population proportions (Appendix A). It can be seen that there is a clear trend for sleep to become more disturbed as time progresses. Sleep is deepest during the first hour of sleep, and it is here where disturbance is least. The underlying disturbance rate increases steadily, from about 4% or 5 arousals per hour (which would include about two awakenings an hour) at the beginning of the night to about 6.5% (more than 7 arousals or 3 awakenings per hour) at the end of the night.

Controlling for individual differences in arousability shows no lessening of this time of night effect. LRA showed there to be an increase of around 25% in the probability of being disturbed in the period 0500-0530 as compared with 0100-0130 (Appendix B). In the W-D analysis (Appendix C), all noise events for the period 2330-0530 were included. Relative to the first part of the night, 2330-0100, aircraft noise events between 0400 and 0530 are about 37% more likely to cause a disturbance. This effect is independent of the event noise level.

In Figure 53, an approximately 90 minute undulation in *q* is evident during the first three hours of sleep. This is the well known "ultradian" rhythm of sleep which is purely a biological phenomenon (Ref 25). Sleep becomes naturally and transiently more disturbed when it periodically lightens during the night, and is less disturbed when it deepens periodically and naturally. This periodicity is superimposed on the general trend for sleep to lighten progressively over the night. The fact that it shows so clearly in Figure 53 is further evidence of the validity of actimetry for measuring sleep disturbance.

The same curve is compared in Figure 54 with the ANE arousal rate *n*, again plotted against time of sleep. The approximate 95% confidence limits for *n* are also given, but because of the very much smaller sample sizes (ANE epochs are only 2% of the total - during the middle hours of the night the percentage is even less), the confidence intervals are much greater than in the *q* case. There are two particularly notable features of the *n* curve. First, it shows a pronounced rhythmicity over the sleep period, with a cycle time of about 90 minutes; i.e. there are certain times of the night when subjects appear to be more sensitive to ANEs than at others. This rhythmicity is not related to patterns in the occurrence over the night of the ANEs themselves, which may be seen by comparing Figures 52 and 54. It is likely to be a biological phenomenon, which is elaborated upon below. Second, for the first 45 minutes of sleep, *n* is indistinguishable from *q*. That is, most subjects are unaffected by ANEs, despite there being many ANEs at this time of night.

The same data were used to generate Figure 55, but here time of night (rather than time of sleep) is used, again in 15 minute periods, starting at 2200 and going on to 0800 the next morning. Because people have varying bed-times, their cyclic sleep rhythms are not so well synchronised in this representation of the data and they are therefore less apparent in both records. Also, as many subjects had not yet fallen asleep (Fig 52), the first few intervals of Figure 55 are distorted by the smaller subject numbers. The same applies to the last few intervals as increasing numbers of subjects get up from 0600 onwards. Nevertheless, the general trend for all arousals to increase during the night remains clear, as does the low average noise sensitivity during the first hour of the night. Disregarding the fluctuations of the *n*-arousal rates (which are similar in magnitude to the approximate confidence intervals), it appears that the underlying trend of noise sensitivity is for it to increase until around 0300-0400 and to decrease thereafter. However, the apparently low rates of response to ANEs from 0630 onwards, which coincides with a general increase in the number of ANEs, requires further analysis. It may, at least in part, be a consequence of some people waking up at this time and not going back to sleep. Either they continue to move frequently, not generating the movement *onsets* that are registered as arousals, or they get up and remove their actimeters. More detailed analysis of the actimetry data including comparison with morning sleep log data may help to resolve the issue but this has not yet been undertaken.

A possible reason for a rhythmicity in response to noise is that sleep is most sensitive to disturbance by noise and other extrinsic factors when sleep lightens, and less vulnerable when sleep deepens. This is suggested by comparisons of the variation of *n* and *q* with variations in the the incidences of deep sleep and REM ('dreaming') sleep determined from the sleep-EEG records (Ref 16). The roughly 90-minute peaks of the *n*-cycle seem to correspond to the initial rises in REM sleep which tend to follow periods of sleep lightening. Conversely, the troughs seem to match the peaks of REM sleep and increases in deep sleep. During deep sleep, auditory input to the higher centres of the brain is blocked, and arousal by noise, here, depends mostly on the sound being of high amplitude. In REM sleep people can be equally unresponsive to noise, although, if the noise is of personal significance to the subject (due to a strong antipathy to aircraft for example) even low amplitudes can arouse (Ref 25).

7.9 Window state

Subjects reported whether their bedroom windows were open or closed during each measurement night. Figure 56 shows that the average arousal rates in noise epochs are lower, but only slightly so, when windows are reported closed: open - 7.9%; single glazing closed - 7.6%; double glazing closed - 6.6%. However, these differences are small and, when confounding factors are controlled, they turn out to be statistically insignificant (Appendix C).

The reasons why the effects of 'window state' are so small cannot be isolated from the data. It is likely that among contributing factors are (a) that the reports of window state are unreliable (certainly the amounts of noise insulation obtained would inevitably vary widely) and (b), probably of greater importance, is a likely interaction between subjects' arousability, noise sensitivity and likelihood of opening windows at night. It is unlikely that any further analysis will shed more light on this question, although the effects of the special request for the Manchester/HGN subjects to sleep with their windows open during part of the measurements (Section 3.4) are still being examined.

7.10 Aircraft type

ANEs were classified by noise level band and one of four aircraft categories: large jets (eg 747, MD11/DC10, L1011), medium and small Chapter 2 jets, (eg 727, DC9, 737-200), medium and small Chapter 3 jets (eg 767, 757, 737-300, A320, BAe146) and propellers (eg F27, ATP) on the grounds that these reflect fairly basic differences in the character of the noise heard on the ground. No marked differences were found between the average *n-q* arousal rates for the various aircraft categories within each of the different noise level bands. A minor exception was a small difference between Chapter 3 jets and the other jets in the 80-90dBA Lmax range. However, although the analysis has yet to be performed, it is almost certain that this difference will disappear when confounding factors are controlled.

7.11 Other sources of disturbance

Interview responses showed that seventy three percent of subjects regularly shared their bed with a partner and, a further 6% sometimes. As part of a small subsidiary study to investigate whether sleep was disturbed by partners, the partners of 46 subjects, at four sites, agreed to wear actimeters on 8 of the 15 test nights.

Statistical comparisons were made between each subject's actigrams and those of his or her partner and those of a 'pseudo partner', matched for control purposes by site, night, sex and age group. This revealed a strong relationship between the sleep patterns of bed-partners. Further comparisons, between subjects who shared their bed and those who slept alone, confirmed that the movements of a partner are a significant source of sleep disturbance (Ref 16).

7.12 Temporary cessation of night flights

A few weeks after the September 1990 pilot trials in the Heald Green area of Manchester, the single runway at Manchester Airport was closed for repairs on weekday nights (2230-0600) of two successive weeks of November 1990. Although it was appreciated that the cooler weather of November would increase window closures, sixteen of the original 20 subjects took part in a further period of actimetry for 16 nights, encompassing 10 weekday and 6 weekend nights. During the measurement periods, between sleep onset and getting up, there were an average of 28 ANEs per subject night at weekends, and 7 per night (outside the period 2230-0600) on weekdays.

The overall arousal rates *a* measured during the September and November weekday nights were identical, 6.24%. The rates for the September and November weekend nights were 6.30% and 6.07% respectively. This difference was not significant according to a simple test of proportions.

On completion of the second phase, subjects were asked if they had noticed anything about the number of ANEs at night. Five out of the sixteen reported fewer aircraft and three guessed that the airport may have been shut, at least during the late evening and around bed-time.

Whilst these simple comparisons provide no evidence of a significant effect of nighttime cessation of aircraft noise, it must be appreciated that these particular noise free nights were a temporary phenomenon in the area and that subjects may not have adapted fully to the changed situation. On the other hand, given the relatively small influence of aircraft noise on sleep disturbance described in Section 7.5, it now seems doubtful that any changes caused by the cessation of night flying could have been reliably detected from an analysis of such a limited data sample.

7.13 Length of residence

No subjects were selected who had lived locally for less than one month. With this proviso, there is no significant effect of length of residence on arousal rates, ie there appear to be no adaptation effects after the first month of residence.

8 EEG RESPONSES TO AIRCRAFT NOISE EVENTS

The main aim of the EEG measurements was to provide a 'gold standard' against which to compare the actimetry method. However, the EEG records themselves, although limited in quantity by comparison with the actimetry data (3%) have also been searched for evidence of direct EEG responses to aircraft noise events (ANE). This involved visual scrutiny of the original EEG traces, a process which is labour-intensive and lengthy, but which may detect fine sleep responses that are missed by actimetry or sleep-stage analysis.

EEG responses were classified into major or minor types. *Major* responses were episodes of wakefulness of 15 seconds or more or movement time lasting 10 seconds or more. *Minor* responses included shifts to stage 1 sleep, simultaneous movement artefacts in all channels, episodes of wakefulness lasting less than 10 seconds or abrupt increases in EMG associated with two or more K-complexes within 3 seconds. (K-complexes are minor EEG responses, often to external stimuli, not usually regarded as arousals, particularly when they occur singly.)

A response was associated with a particular ANE if it occurred within a window of 64 seconds starting 16 seconds before the start of the ANE (ie before its level exceeded 60dBA). The window was extended to 16 seconds after the time of Lmax if this did not occur during the 64 seconds. The initial 16 seconds allowed for variations in the event times at different parts of the site as well as the possibility that the event might be audible before its outdoor level reached 60dBA.

A response which coincided with an ANE in this way may have occurred by chance; ie it may not have been caused by the noise. To ascertain the probability that such a response was caused by the ANE, two matched 'background responses' were recorded in every case. These were whether or not any EEG response occurred within two other 64-second windows randomly chosen as follows:-

(a) within the period 2 to 5 minutes before the ANE ('immediate background'), and

(b) within the period 2 to 5 minutes before the nearest ANE to which there was no response (general background). The average interval between these two ANEs was 28 minutes.

A total of 3189 individual subject-ANEs were examined covering the periods between 2200 to 0800 when subjects were asleep. Associated responses were detected in 459 cases; ie 14.4% of all noise events. Immediate background responses occurred before 93 of these (20.3%). General background responses occurred in 47 cases (10.2%).

If it is assumed that the general background response rate of 10.2% applies throughout, ie that this is the probability that a 'residual' EEG response will occur in *any* 64-second window, then it follows that the probability of any ANE causing such a response is 14.4 - 10.2 = 4.2%. The higher response rate in the immediate background supports the W-D analysis conclusion (Section 7.6) that people are more susceptible to noise disturbance immediately following a prior disturbance. (The immediate background rate is *not* an appropriate estimate of the true residual response rate because of its serial correlation with the ANE linked responses.)

The response rate of 4.2% is considerably higher than the actimetrically based estimate of the noise-induced arousal rate of 0.91% (for all noise events > 60dBA; Section 7.5). However, the EEG responses include many of a very minor nature. Splitting the responses into major and minor ones reveals marked differences between the separate response rates:-

EEG response classification	ANE window	General background window
Major	4.2%	4.1%
Minor	10.2%	6.1%

This analysis indicates that the EEG responses to aircraft noise are almost entirely of a minor nature; there is little difference between the major disturbance rates (W or M) measured in the ANE windows and the general background windows. However, the data samples in this analysis are small and the above proportions are not necessarily reliable estimates of true values. The errors are considered more fully in Reference 15.

9 RELATIONS BETWEEN REPORTED AND MEASURED SLEEP DISTURBANCE

The secondary or after effects of sleep disturbance include subjects' recollections of awakenings and perceptions of their sleep quality. In 57% of subject-nights, no awakenings were reported the next day. On the remaining 43% of occasions, at least one awakening was reported (all causes), the average number being three per night. The causes given for these are summarised in Figure 57. In 26% of cases, the reason for awakening was given as 'not known'. For the remainder, the most frequently reported cause was 'toilet' (16%). The next most common was 'children' (13%) mainly among women in the lower age groups. 'Illness' was also mentioned frequently (>9%), again mostly by women. 'Aircraft' was a relatively minor cause (<4%); about one quarter of all actimetry subjects specifically reported being disturbed by aircraft noise during the study - on average, once every five nights.

In Figure 58, subject-nights have been grouped by reported sleep quality, ie how well or badly the subjects reported sleeping in their sleep logs the next day. For each of these groups, the overall measured arousal rate *a* is plotted, together with the estimated 95% confidence intervals (some are too small to show). This graph shows a good correspondence between these 'objective' and subjective measures (notwithstanding the displaced zero in the vertical axis) and bears out the validity of the experimental approach. Also shown in Figure 58 for comparison are similar responses from the social survey carried out a few weeks before the sleep measurements, in this case one per subject corresponding to 'general' sleep quality. It is clear that the association between the categories of reply and the values of *a* bear little relationship to each other, and that there is no clear increment of *a* over the values 1 to 5 for the questionnaire, as is the case for the sleep log reports. One conclusion that might be drawn is that daily sleep logs provide more reliable data about sleep quality than 'one shot' social surveys. This suggests that when social survey methods are used for investigating sleep disturbance, emphasis should be placed on collecting data about disturbance experienced during the previous night.

10 CONCLUSIONS

The two main conclusions which emerge from the study are *(a) that actimetry provides a very cost-effective way of measuring arousals from sleep in people at home* and *(b) that aircraft noise is a relatively minor cause of these arousals.* In the latter regard it is important to note that the study sites covered a very wide range of nighttime aircraft noise exposures.

It is also very clear that people vary greatly in their susceptibility to sleep disturbance in general and in response to aircraft noise events (ANEs). This personal part of the variability is the major determinant of sleep disturbance; among subjects participating in this study, the most susceptible people experienced two or three times as much general sleep disturbance as the least sensitive, with around 95% of people between these two extremes.

It is mainly because of this individual variability that proper statistical controls were essential to eliminate several confounding effects. Direct inspection of 'unadjusted' results point to the possible importance of numerous influencing factors; including people's age and sex and where they live - there appeared to be differences in disturbance rates at some of the study sites. However, when all these influences were controlled in a multivariate analysis of the data, some of them turned out to be statistically insignificant, ie they probably arose purely as a chance result of the particular combinations of measurements made. As to the individual variability itself, no statistically significant relationship has been found between this and subjects' personal and psychological characteristics as determined from questionnaire responses.

It was recognised at the outset that the statistical constraints upon the study would be severe; it is for this reason that an unprecedented amount of data was gathered. But even with nearly 6000 'subject-nights' of data, over 4.5 million measurements, the difficulties of unravelling the relationships between sleep disturbance and the various factors of influence remain considerable. The inevitable statistical limitations on the results must be borne clearly in mind.

Considerable efforts have been made to isolate the most important factors and to quantify their effects. This has involved the use of 'random effects logistic regression analysis'. This allows a large number of variables to be handled simultaneously, to take account of serial correlation effects (the fact that at any particular time, an individual's measured sleep disturbance is not independent of previous measurements) and the fact that the disturbance itself is expressed as a percentage, not an unbounded linear variable. The method is powerful, but it also consumes a lot of time and computing resources. Its use so far has therefore been limited but it is hoped that more can be accomplished in the future.

Actimetry detects around 90% of awakenings of 10-15 seconds or more. It also picks up a large number of minor arousals which include very brief awakenings, some sleep stage changes and minor body movements. Nearly all of these 'sleep events' are quite natural; they occur frequently during normal sleep. The average subject experienced about 46 arousals per night although individual rates varied greatly - from 26 to 54 per night (3% were outside this range). Of this average of 46, about 40% are likely to be significant awakenings, more than 10 seconds or so, although nearly 80% of awakenings last less than 1 minute and two thirds are less than 30 seconds. Allowing for statistical uncertainty in the 40% factor, the average number of nightly awakenings probably lies in the range 14 to 23. Most of these are not remembered; no awakenings were reported by subjects on days following 57% of the measurement nights. In the remaining 43% of cases, subjects recalled an average of three awakenings during the previous night.

This rate of awakening is normal; only if there were many more, probably in excess of 6 per hour throughout the night, would any after-effects of sleep disturbance be noticed (eg daytime sleepiness, deterioration of performance etc.). In terms of total arousals, as measured by actimetry, this probably corresponds to more than 100 per night (which may be compared with the average rate in this study of about 45 per night).

Subjects who reported awakenings the next day often did not state a cause (26%). Of reported causes, the most frequent were toilet use, children and illness. Aircraft noise was among the minor reported causes; less than one quarter of all subjects mentioned aircraft, on average about once every 5 nights.

The results indicate that, below outdoor event levels of 90 dBA SEL (about 80 dBA Lmax, 95 EPNdB*), aircraft noise events (ANEs) are most unlikely to cause any measurable increase in the overall rates of sleep disturbance experienced during normal sleep. For all ANE events above this level, the average sleep arousal rate was about 1 in 30. This corresponds to a wakening rate of about 1 in 75 (somewhere between 1 in 60 and 1 in 100).

Analysis of the records in 15-minute periods has shown that arousals from sleep steadily increase during the night. During the first part of the night, there is a small but clear 90-minute rhythmicity in the 'non-ANE' arousal rate (ie measured in 'quiet' between aircraft noise events) which mirrors the well known cyclic pattern of sleep. This kind of pattern is more pronounced in the ANE arousals (ie the rate measured in ANE epochs themselves), and the fluctuations continue throughout the night. Such cycles in sensitivity to aircraft noise can be explained in biological terms by relating them to the natural periodic changes in sleep stage. This suggests that people are indeed more sensitive to noise disturbance at particular times during their night's sleep. However, such conclusions have to be interpreted with caution. The data samples for each 15-minute period are relatively small and therefore subject to larger sampling errors (which cannot reliably be estimated by standard statistical tests) than are the comparable 'in quiet' rates. In particular, they are too small for meaningful multivariate analysis. As a consequence, it has not yet been possible to ascertain how much of this apparent rhythmicity in ANE-related arousals is 'real' and how much is statistical fluctuation. However, it is hoped that continuing work will resolve this.

If the fluctuations are ignored, it can be concluded that sensitivity to arousal by aircraft noise is low during the first part of sleep, increases until 0300-0400, and then falls to a low level again at the end of the night. Here it must be remembered that these arousals have been measured by actimeter. Although reliable for detecting arousals and awakenings from sleep, actimetry reveals less about sleep onset, especially as to whether or not this is prevented or delayed by the presence of noise. The insensitivity of subjects to aircraft noise at the beginning of sleep appears to be a real effect (it is consistent with the fact that people tend to descend into deep sleep fairly rapidly at the beginning of the night). However, the same cannot be said of the period in which people end their night's sleep. Arousals are only measured when subjects stir; if they are already awake and moving fairly continuously, perhaps because aircraft noise is bothering them, no movement onsets are registered. Equally, subjects may simply get up and remove their actimeters. More detailed analysis of the actimetric records, in conjunction with the personal sleep log data, is required to shed further light on this question, but this has not yet been done.

A related question, which has been addressed, is whether the time interval between ANEs affects the probability of arousal. The 'Wilkinson-Diamond' analysis of ANE epochs showed that the likelihood of ANE-related disturbance increases with the incidence of arousal in the interval since the previous ANE, but not upon the duration of the interval. In other words, there is no evidence that increasing or decreasing the frequency of flights is likely to affect the probability of being disturbed by any particular event. However this analysis excluded ANEs which occurred within 5 minutes of preceding ANEs and more detailed analysis is still required.

It is clear from this analysis that, in general, aircraft noise has a negligible effect upon overall patterns of arousal from sleep. Even at locations near to airports with higher levels of night aircraft traffic, the additional disturbance caused by the aircraft noise, both wakenings and lesser arousals, is likely to be very small by comparison with that occurring

---

* 95 EPNdB, on the noise scale used internationally for the noise certification of aircraft, is roughly the equivalent of 90 dBA SEL.

'naturally' due to all other causes. Aircraft noise itself is most unlikely to increase sleep disturbance rates to the point at which after-effects upon health or performance would be noticeable.

At the same time, it must be emphasised that these are estimates of *average* effects; clearly more susceptible people exist. At one extreme, 2-3% of people are 60% more sensitive than average; some may be twice as sensitive to noise disturbance. There may also be particular times of the night, perhaps during periods of sleep lightening, when individuals could be more sensitive to noise. Although the relationship cannot be verified statistically, the data do indicate that aircraft events with noise levels, greater than 100 dBA SEL (105 EPNdB, 95 dBA Lmax) out of doors, will have a greater chance of disturbing sleep. Finally, the most sensitive people may also react to aircraft noise events with levels below 90 dBA SEL (80 dBA Lmax).

The relationships between sleep disturbance and aircraft noise depend on numerous other factors, important among which are time of night, the individual's sex (men are about 15% more susceptible than women, with or without aircraft noise) and the incidence of disturbance in the period preceding the ANE.

Of many factors which it was thought might have affected sleep disturbance, most have been ruled out statistically in the analysis. Among these are location (site), subjects' age, length of residence, window state (open or closed, single or double glazing) and aircraft type (controlling for noise level).

A period of nighttime runway maintenance at Manchester airport during the pilot stage of the study provided an opportunity to examine the effect of limited cessations of night flying upon the sleep of subjects living in a high aircraft noise exposure area. No statistically significant differences were found between the overall arousal rates on nights with or without high levels of aircraft noise exposure. In the light of the subsequent main study results, this finding is not surprising; it appears to reinforce the conclusion that even relatively high levels of aircraft noise are unlikely to add significantly to overall sleep disturbance.

Finally, it must be made clear that most of the findings presented here are concerned with the primary effects of aircraft noise, ie in causing sleep arousals or wakenings. Possible relationships between sleep disturbance and annoyance have yet to be investigated more fully.

REFERENCES

1 Brooker, P, *Noise and Sleep*, Civil Aviation Authority Paper 78011, June 1978

2 DORA staff, *Aircraft noise and sleep disturbance*: Final Report, Civil Aviation Authority, DORA Report 8008, 1980

3 Brooker, P, *Noise disturbance at night near Heathrow and Gatwick Airports: 1984 Check Study*, Civil Aviation Authority, DR Report 8513, 1985

4 Pearsons K S, et al., *Analyses of the predictability of of noise-induced sleep disturbance*, US Air Force Report HSD-TR-89-029, 1990

5 Griefahn, B, *Sleeping in noisy environments: effects, limits and preventative measures*, In: Sleep '90, (ed) J A Horne, Pontenagel Press, Bochum, 391-393, 1990

6 Libert, J P, et al., *Relative and combined effects of heat and noise exposure on sleep in humans*, Sleep, 14, 24-31, 1991

7 Mullany, D J, et al., *Wrist-actigraphic estimation of sleep time*, Sleep, 3, 83-92, 1980

8 Webster, J B, et al., *An activity-based sleep monitor system for ambulatory use*, Sleep, 5, 389-399, 1982

9 Cole, R J and Kripke, D F, *Progress in automatic sleep/wake scoring by wrist actigraph*, Sleep Research, 13, 331, 1988

10 Sadeh, A et al., *Actigraphically determined bedtime sleep-wake scoring: validity and clinical applications*, Journal of Ambulatory Monitoring, 2, 209-216, 1989

11 Dunham, D W, et al., *Wrist activity and sleep/wake estimation revisited*, Sleep Research, 20A, 491, 1991

12 Lavie P, Personal communication to K I Hume, 1990

13 Cadoux, R E and Atkinson, B J, *Noise disturbance at night: 1991 Study noise measurements*, Civil Aviation Authority, Report to be published

14 Diamond, I D, et al., *Sleep disturbance due to aircraft noise: social survey report*, University of Southampton, Department of Social Statistics, Report to be published

15 Hume, K I, et al., *Effects of aircraft noise on sleep; EEG-based measurements*, Manchester Metropolitan University, Department of Biological Sciences, Report to be published

16 Horne, J A et al., *Aircraft noise and sleep disturbance: findings with actimeter, sleep log and questionnaire data*, Loughborough University Sleep Research Laboratory, Report to be published

17 Diamond, I D et al., *Aircraft noise and sleep disturbance: statistical analyses of actimetry and questionnaire data*, University of Southampton, Department of Social Statistics, Report to be published

18 Brooker, P, et al., *United Kingdom Aircraft Noise Index Study: main Report*, Civil Aviation Authority DR Report 8402, January 1985

19 Horne, J A and Östberg, O, *A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms*, International Journal of Chronobiology, 4, 97-110, 1976

20 Wills, G and Cooper, C L, *Pressure Sensitive*, Sage Publications, London, 1988

21 Spielberger, C D, *State-Trait Anxiety Questionnaire*, Consulting Psychologists Press, California, 1977

22 Rechtschaffen, A and Kales, A, *A manual of standardised terminology, techniques and scoring system of sleep stages in human subjects*, UCLA Brain Information Service, Los Angeles, 1968

23 Åkerstedt, T, *Self assessment sleepiness questionnaire*, (unpublished), Institute for Stress Research, Stockholm, 1989

24 Bonnet, M H, *Infrequent periodic sleep disruption: effects on sleep, performance and mood*, Physiology and Behaviour, 45, 1049-1055, 1989

25 Horne, J A, *Why we sleep*, Oxford University Press, 1988

TABLE 1   FIELDWORK SCHEDULE

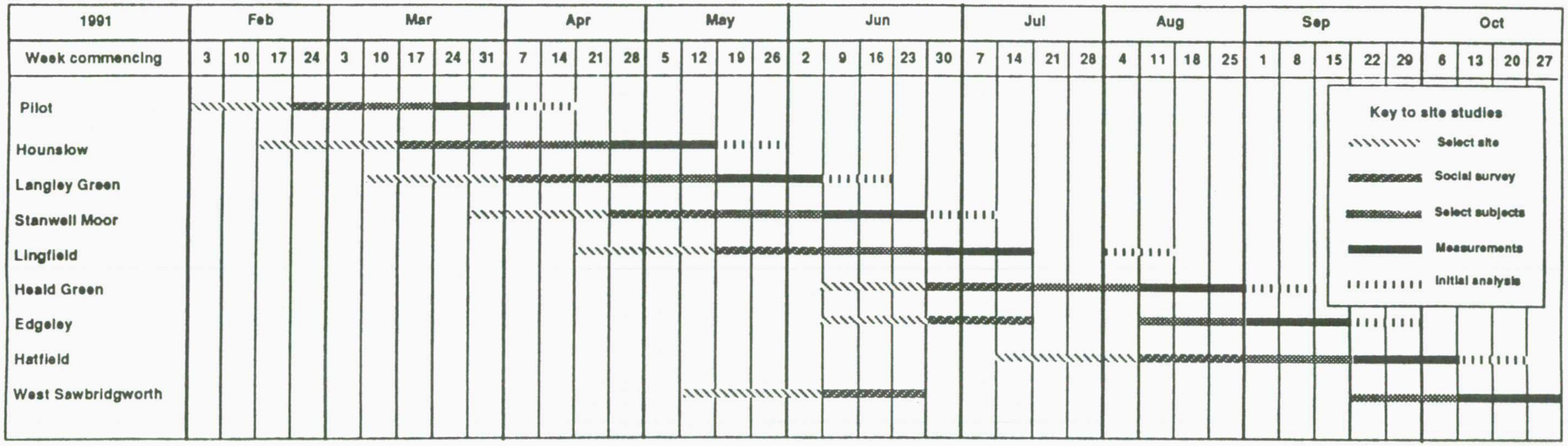


TABLE 2 SUBJECT SELECTION

Site	Addresses Listed	SS I'views	Volunteers			20-34 years		35-49 years		50-70 years		Total	
			Initial	Final	AM only	F	M	F	M	F	M	F	M
HLW	397	203	122	56	36	12	10	7	9	7	5	26	24
LGN	469	203	134	53	21	9	7	11	9	5	9	25	25
SWM	436	207	120	59	31	12	7	11	6	10	4	33	17
LFD	471	208	126	56	25	8	9	8	7	9	9	25	25
HGN	445	203	120	57	33	7	8	9	7	8	11	24	26
EDG	471	204	130	57	20	15	14	6	6	4	5	25	25
HFD	669	204	119	53	32	9	5	13	11	6	6	28	22
SBW	538	204	100	56	29	8	8	8	7	9	10	25	25
Total	3896	1636	971	447	227	80	68	73	62	58	59	211	189
					Age group totals:	148		135		117		400	
					Age group %:	37		33.75		29.25			

Abbreviations: SS I'views = social survey interviews AM = actimetry F, M = Female, Male

TABLE 5 DISTRIBUTION OF SLEEP PERIODS ESTIMATED FROM ACTIMETER DATA (hours.mins)

Sample	Number	Sleep onset		Get-up		Sleep period	
		Mean	standard deviation	Mean	standard deviation	Mean	standard deviation
All	5742	23.49	1.08	7.04	0.58	7.15	1.15
Males							
All	2705	23.56	1.12	6.57	1.03	7.01	1.18
20-34	952	0.02	1.13	7.10	1.01	7.08	1.19
35-49	904	0.00	1.18	6.47	1.05	6.46	1.20
50-70	849	23.45	1.03	6.53	1.01	7.08	1.11
Females							
All	3037	23.43	1.04	7.11	0.52	7.28	1.11
20-34	1157	23.45	1.07	7.13	0.57	7.27	1.18
35-49	1038	23.46	1.08	7.13	0.44	7.27	1.10
50-70	842	23.36	0.54	7.07	0.53	7.31	1.01

TABLE 6 SUMMARY OF RESULTS FROM EEG SUBSAMPLE (178 subject-nights)

Site	A	X	H	E	X/A%	X/H%	H/E%	A/E%	H/A%
Heathrow/HLW	848	322	355	17300	38.0	90.7	2.05	4.90	41.86
Gatwick/LGN	952	327	392	19183	34.3	83.4	2.04	4.96	41.18
Heathrow/SWM	689	294	334	17540	42.7	88.0	1.90	3.93	48.48
Gatwick/LFD	461	146	169	9240	31.7	86.4	1.83	4.99	36.66
Manchester/HGN	761	258	280	17821	33.9	92.1	1.57	4.27	36.79
Manchester/EDG	875	281	312	16833	32.1	90.1	1.85	5.20	35.66
Stansted/HAT	916	342	385	20170	37.3	88.8	1.91	4.54	42.03
Stansted/WSB	957	256	303	17556	26.8	84.5	1.73	5.45	31.66
All	6459	2226	2530	135643	34.5	88.0	1.87	4.76	39.17

Notation:
A = Number of actigraph blips (arousals)
X = number of coincident blips (±1 epoch)
H = number of Hypnogram blips (awakenings)
E = total number of epochs

TABLE 7 SUMMARY OF ACTIMETRY RESULTS BY SITE (5742 subject nights)

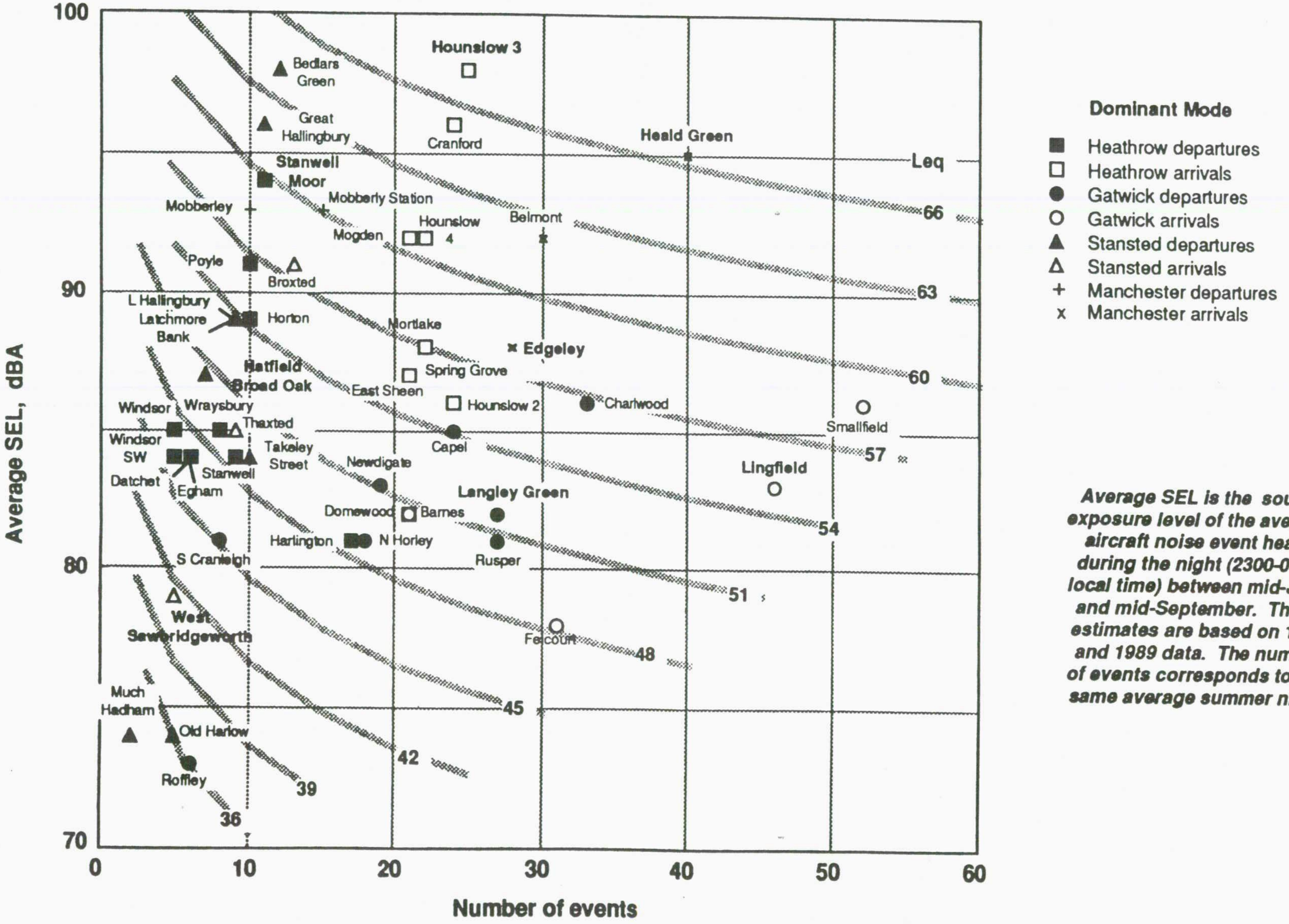
Site	A(N)	N	n	A(Q)	Q	q	A(E)	E	a
Heathrow/HLW	557	8904	6.26	27310	504861	5.41	27867	513765	5.42
Gatwick/LGN	619	10517	5.89	31794	576641	5.51	32413	587158	5.52
Heathrow/SWM	164	3425	4.79	25576	532443	4.80	25740	535868	4.80
Gatwick/LFD	990	17612	5.62	31195	595414	5.24	32185	613026	5.25
Manchester/HGN	1943	27881	6.97	30573	564784	5.41	32516	592665	5.49
Manchester/EDG	939	15314	6.13	28476	510208	5.58	29415	525522	5.60
Stansted/HAT	125	2374	5.27	33512	627435	5.34	33637	629809	5.34
Stansted/WSB	85	1702	4.99	29744	603783	4.93	29829	605485	4.93
All	5422	87729	6.18	238180	4515569	5.27	243602	4603298	5.29

A(N) = Number of arousals in noise epochs (Lmax ≥ 60)
N = Number of noise epochs (Lmax ≥ 60)
n = % of noise epochs with arousals

A(Q) = number of arousals in quiet epochs
Q = number of quiet epochs
q = % of quiet epochs with arousals

A(E) = total arousals
E = Total epochs
a = % of all epochs with arousals

Figure 2 - Night noise exposures (2300-0700) at possible study sites



Average SEL is the sound exposure level of the average aircraft noise event heard during the night (2300-0700 local time) between mid-June and mid-September. These estimates are based on 1988 and 1989 data. The number of events corresponds to this same average summer night.

Figure 1 - Noise induced wakenings: laboratory and field data (from Reference 4)

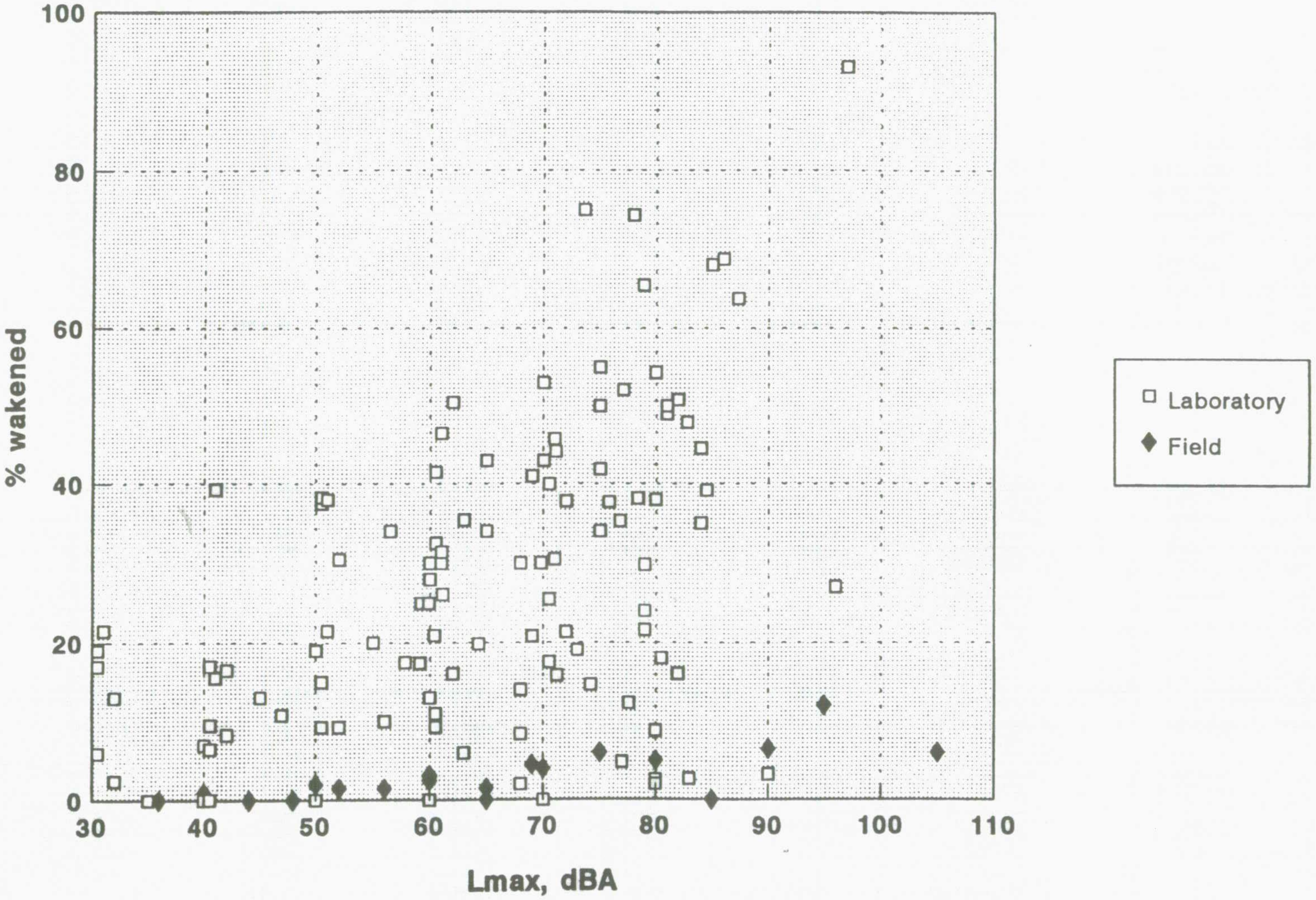
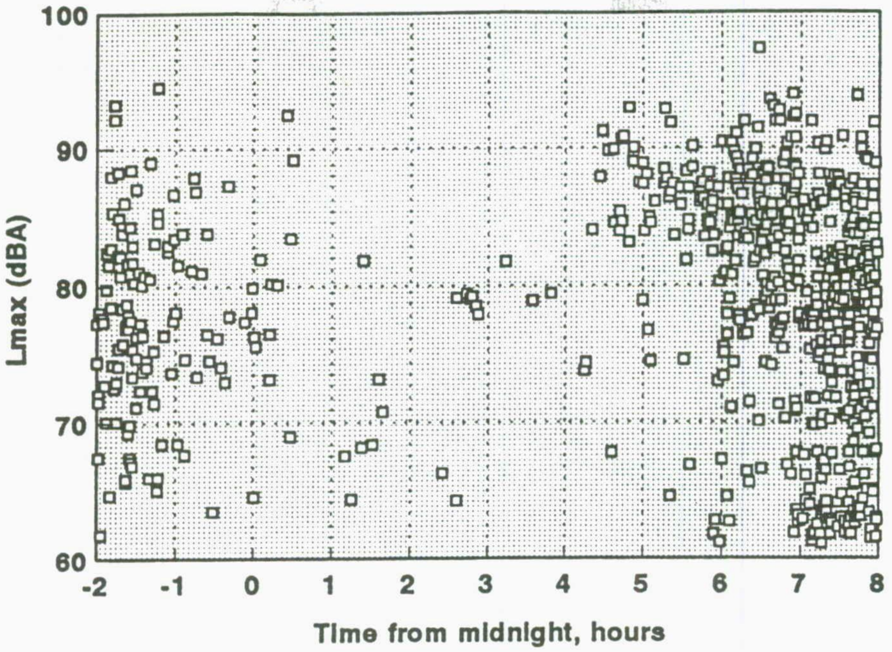
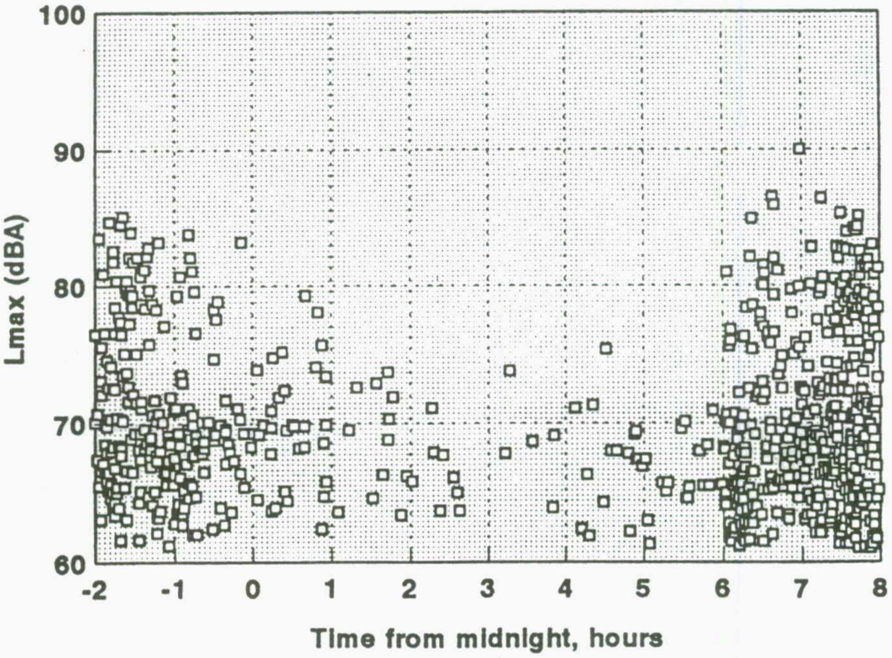


Figure 6(a) - ANEs measured at Heathrow, Hounslow (HLW)



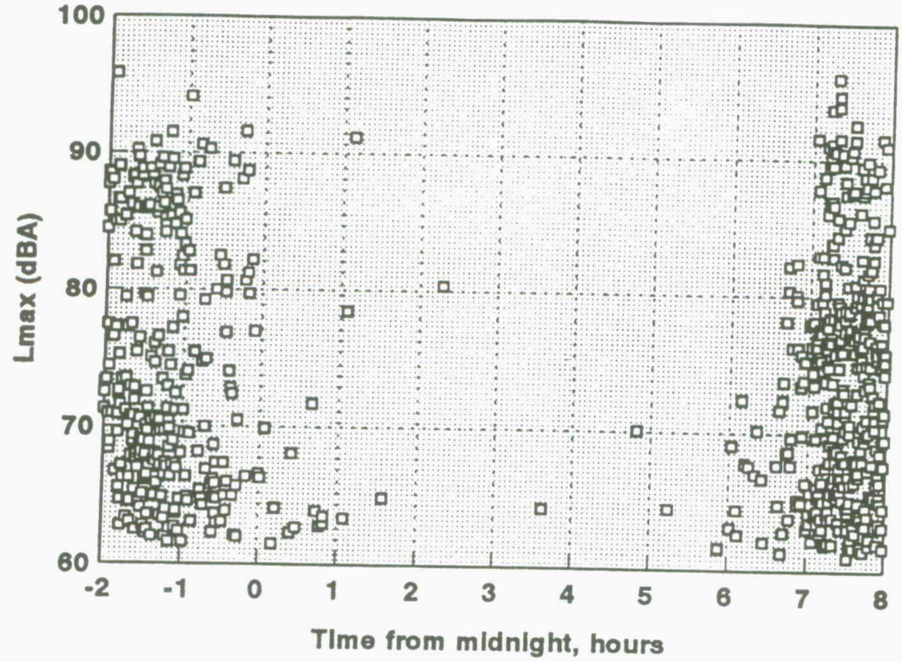
jopfig6a

Figure 6(b) - ANEs measured at Gatwick, Langley Green (LGN)



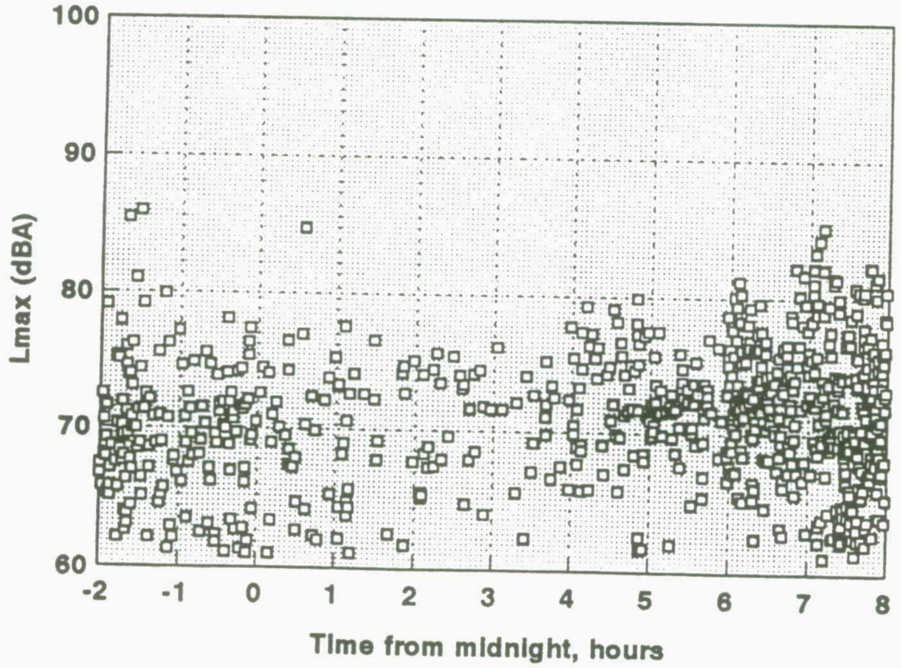
jopfig6b

Figure 6(c) - ANEs measured at Heathrow, Stanwell Moor (SWM)



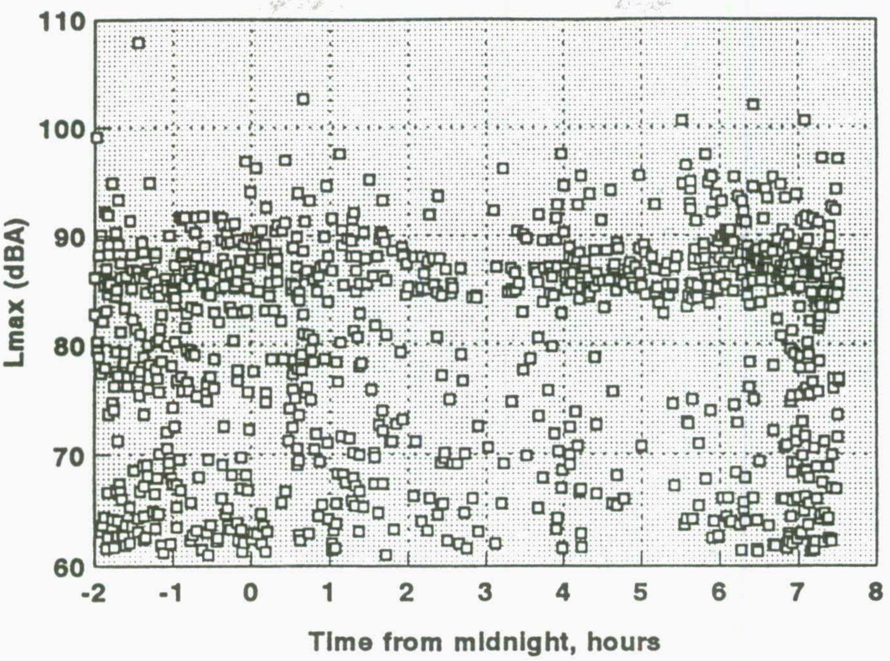
jopfig6c

Figure 6(d) - ANEs measured at Gawick, Lingfield (LFD)



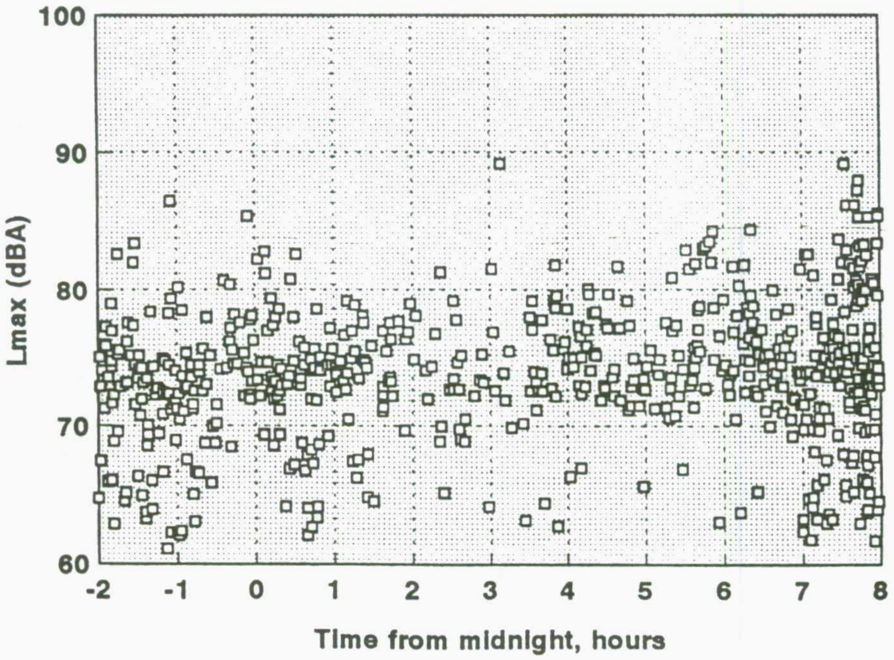
jopfig6d

Figure 6(e) - ANEs measured at Manchester, Heald Green (HGN)



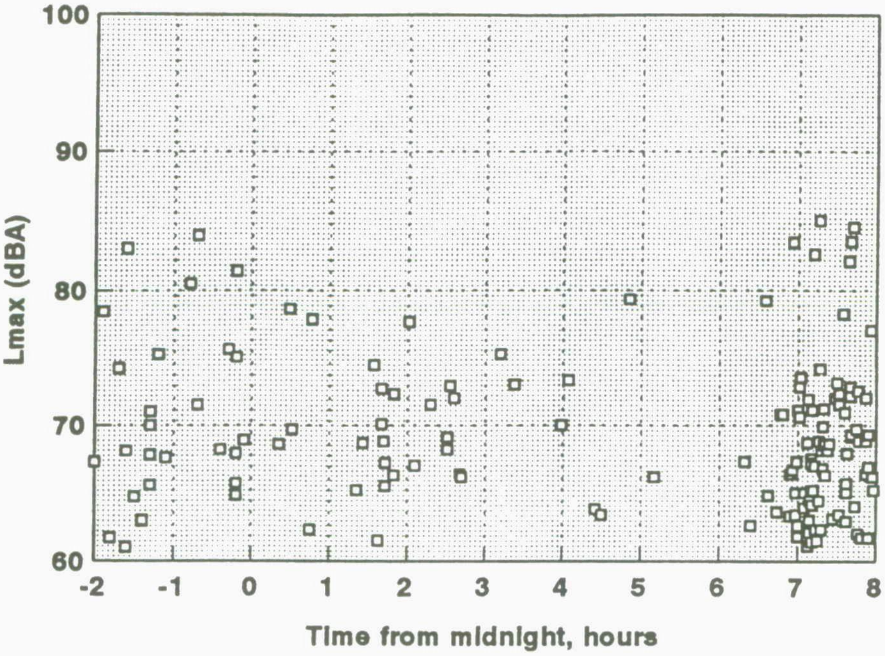
tgfig6e

Figure 6(f) - ANEs measured at Manchester, Edgeley (EDG)



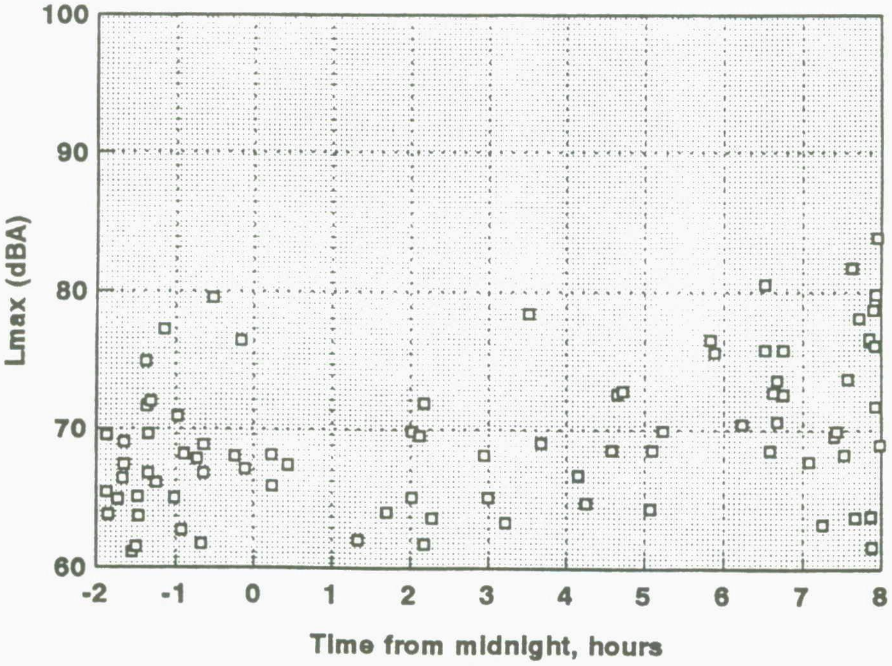
jopfig6f

Figure 6(g) - ANEs measured at Stansted, Hatfield (HAT)



jop6g6g

Figure 6(h) - ANEs measured at Stansted, Sawbridgeworth (WSB)



jop6g6h

Figure 7 - Overall distribution of aircraft noise epochs

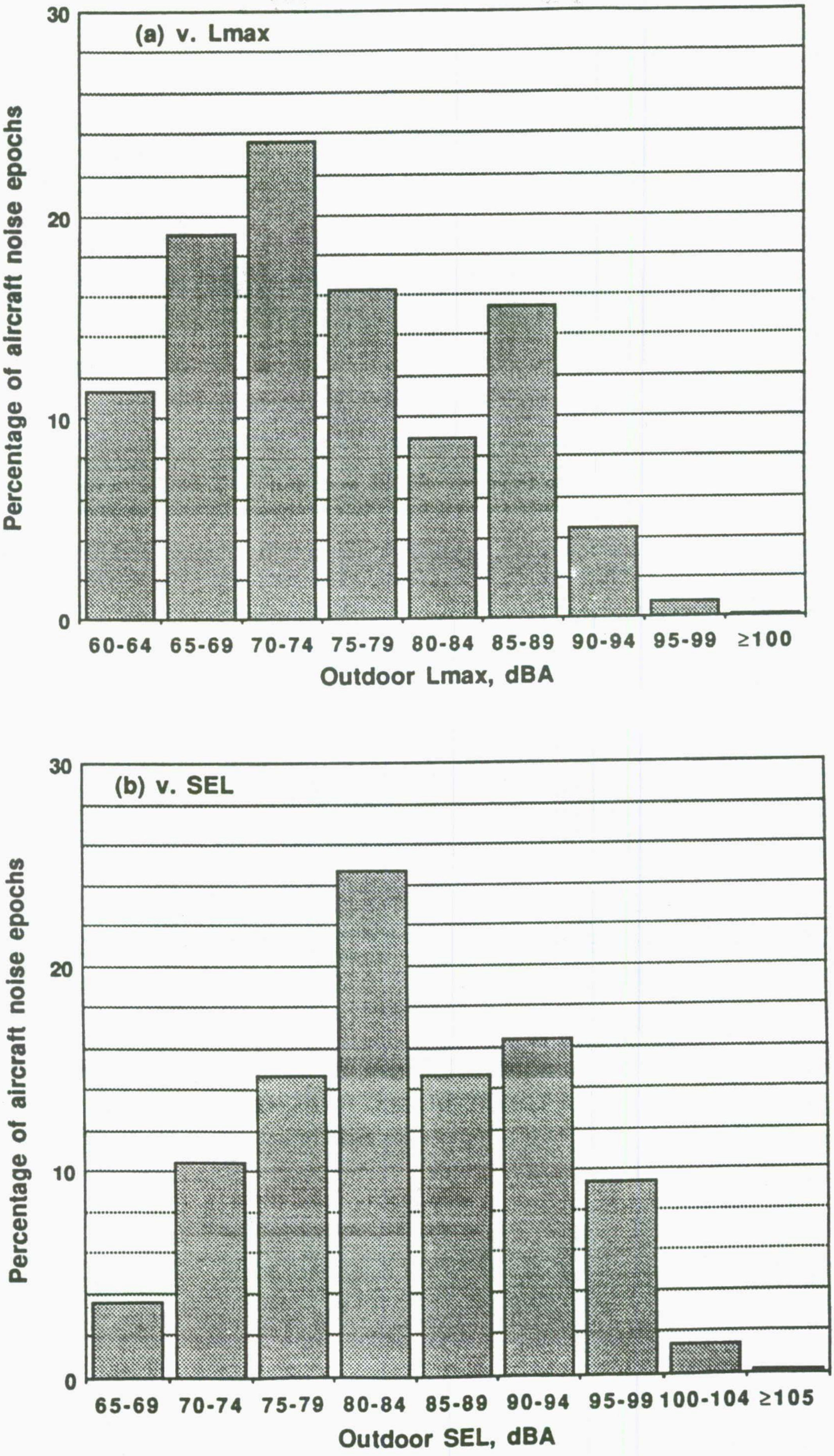


Figure 8 - Age distribution of social survey respondents

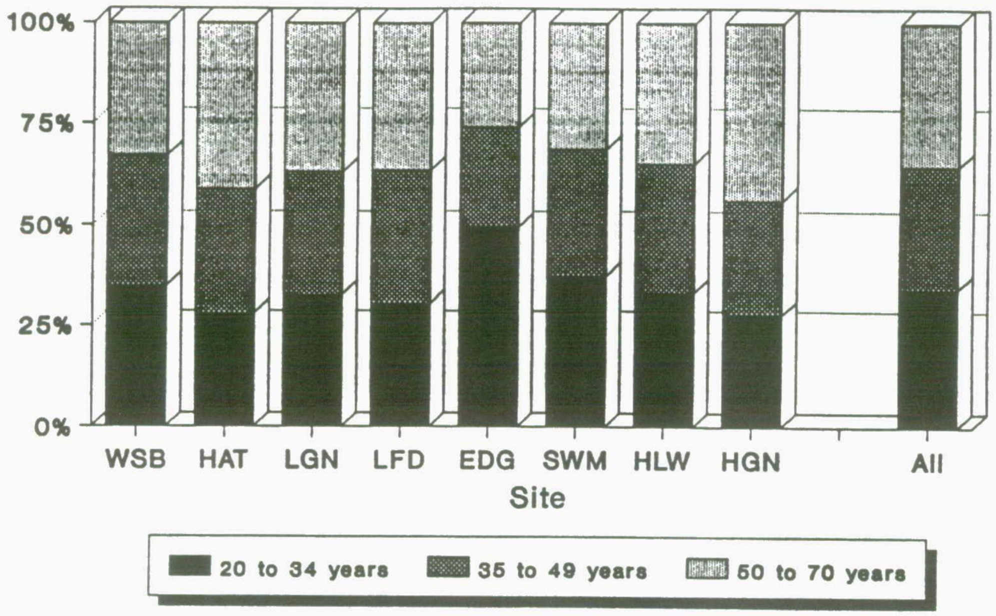


Figure 9 - Occupational group of social survey respondents

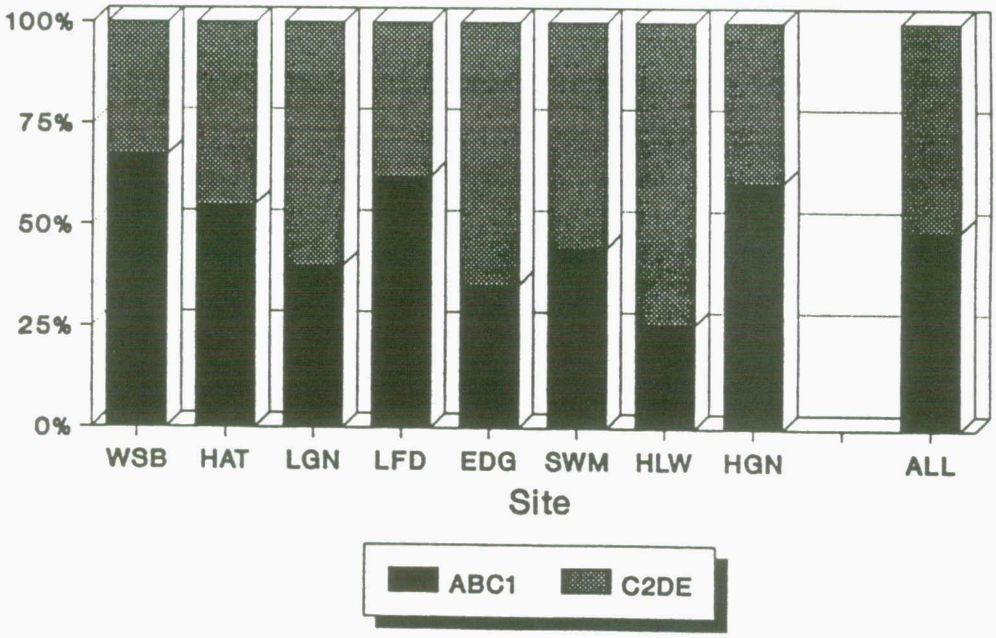
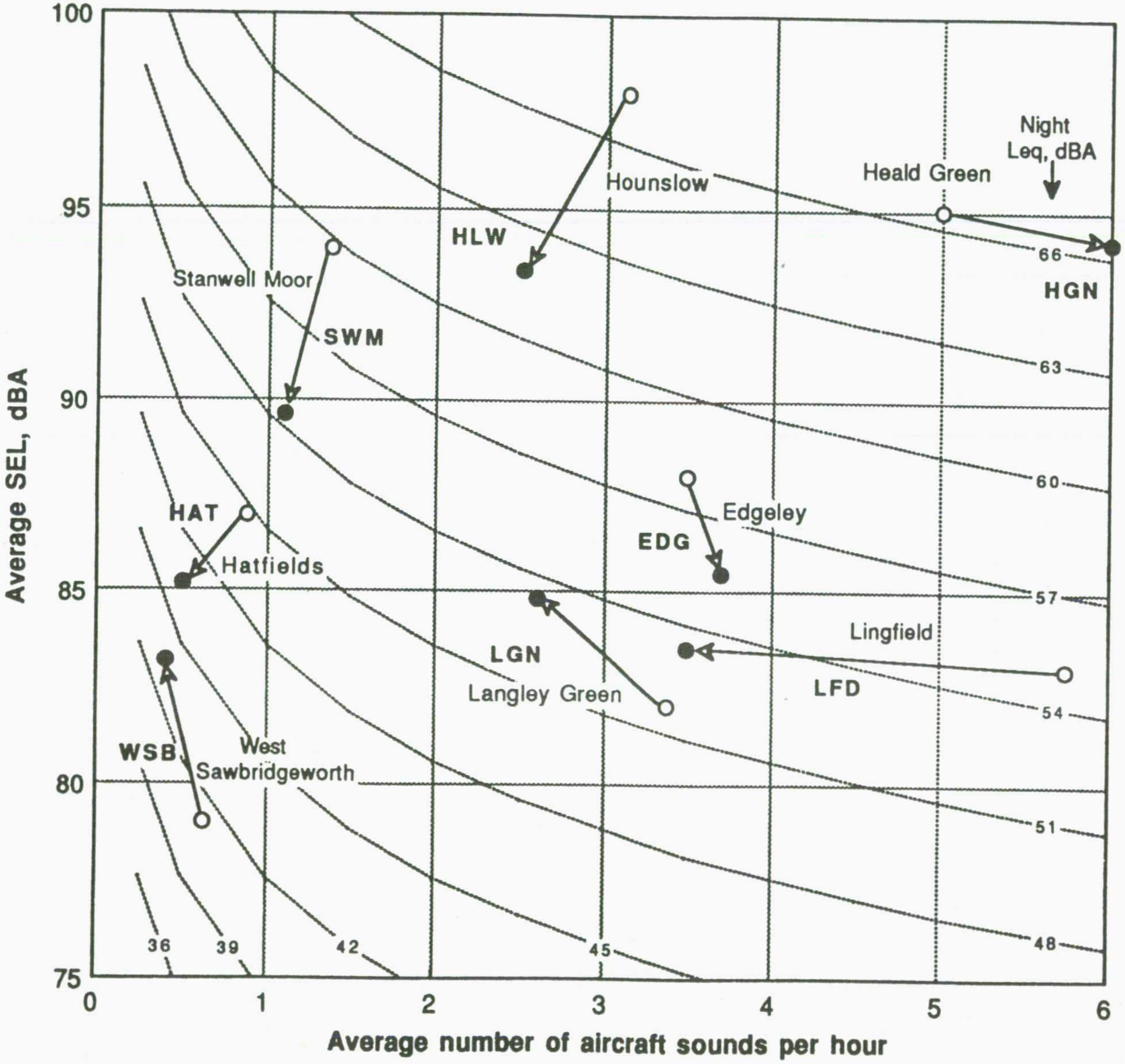


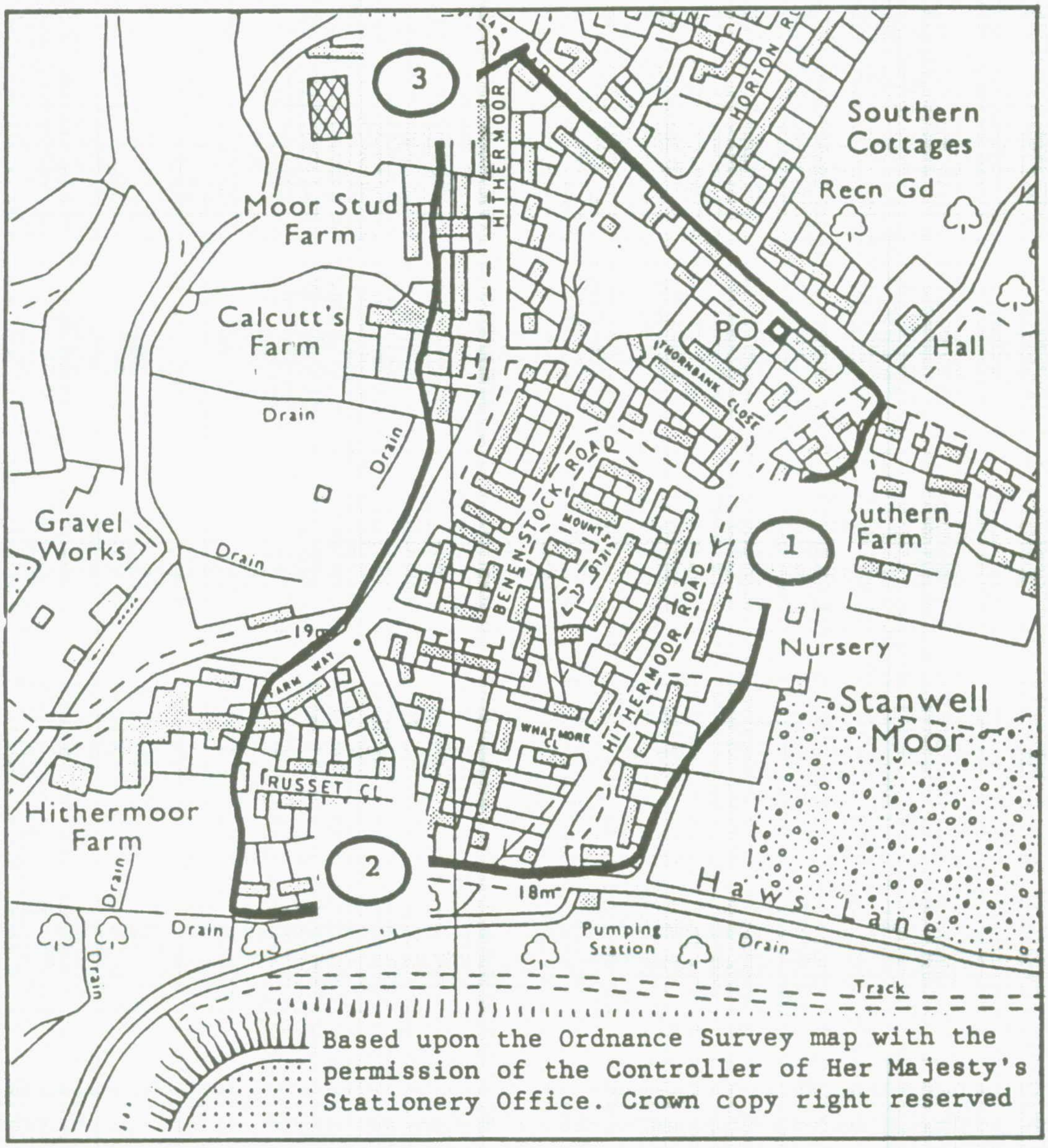
Figure 5 - Measured night noise exposures at study sites



The open symbols are the original 8-hr (2300-0700) estimates

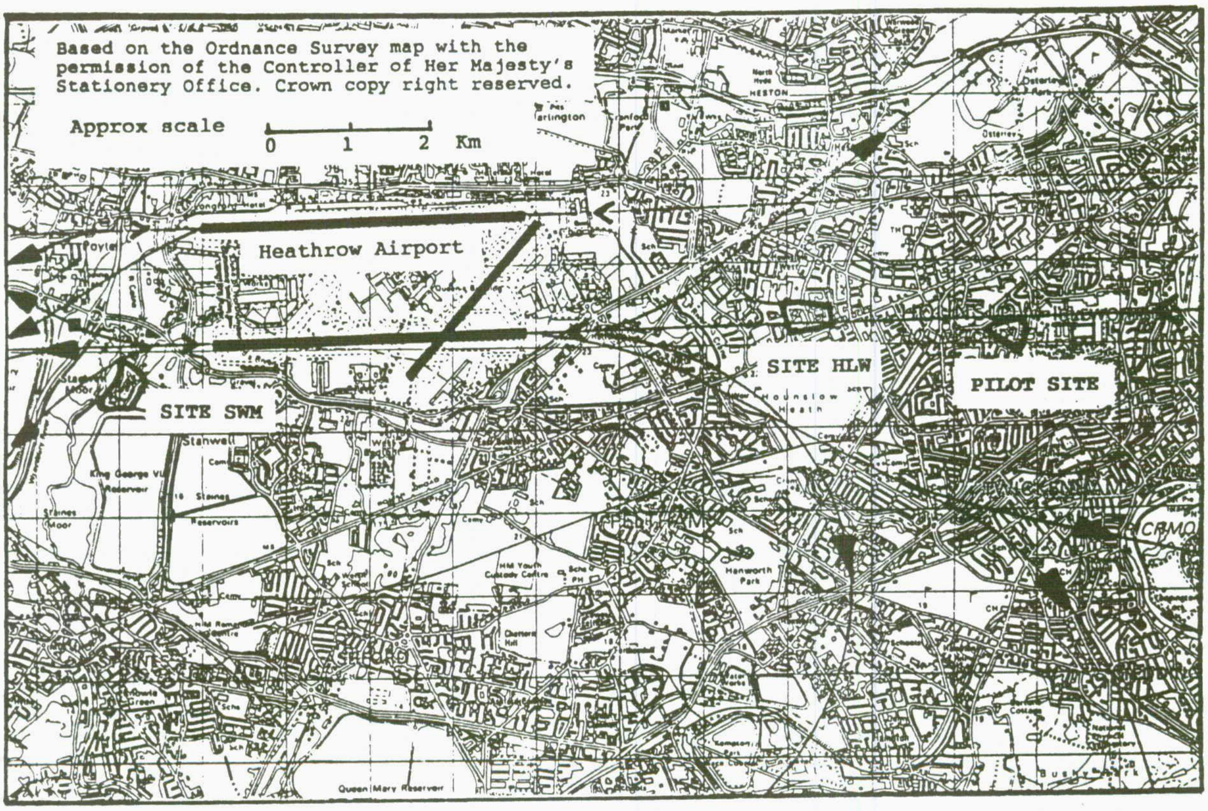
The solid symbols are measured 8-hour (2300-0700) values

Figure 4 - Location of noise monitors at Heathrow, Stanwell Moor (SWM) Site



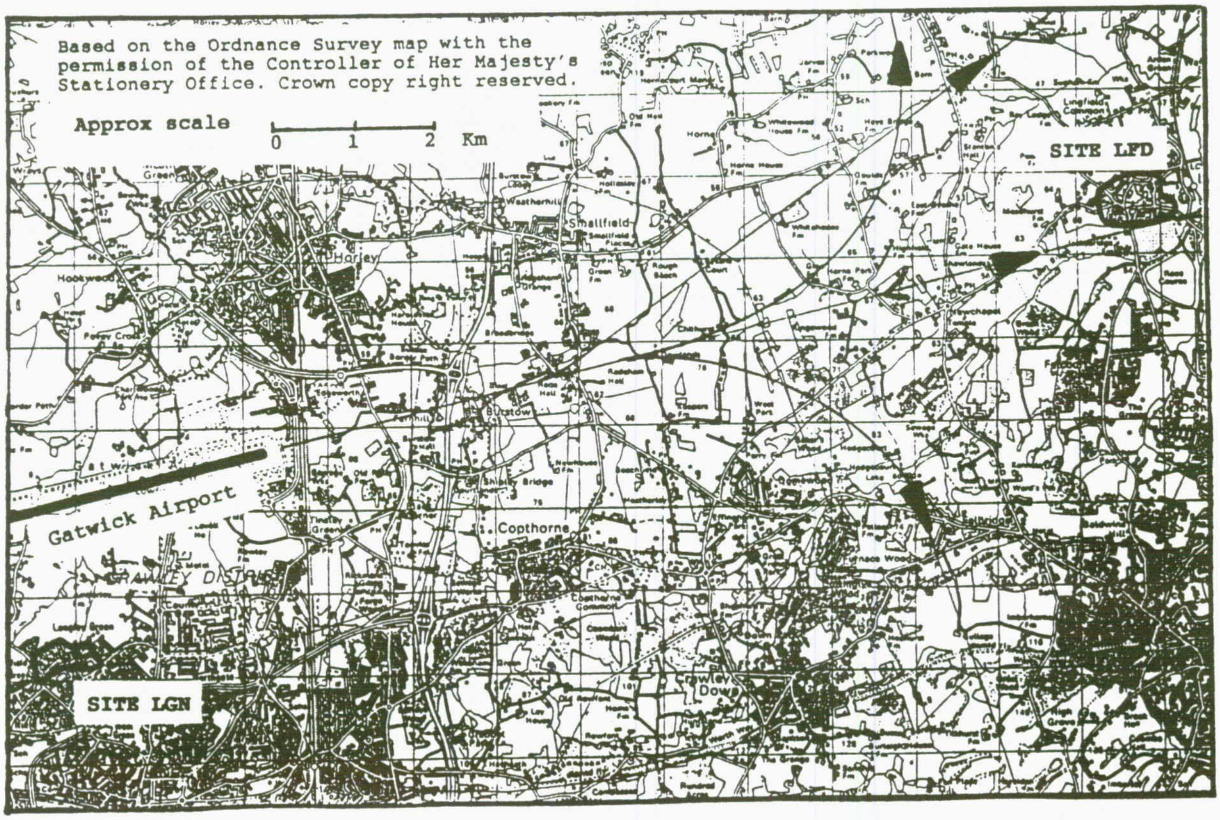
Based upon the Ordnance Survey map with the permission of the Controller of Her Majesty's Stationery Office. Crown copy right reserved

Figure 3(a) - Heathrow study sites



Based on the Ordnance Survey map with the permission of the Controller of Her Majesty's Stationery Office. Crown copy right reserved.

Figure 3(b) - Gatwick study sites



Based on the Ordnance Survey map with the permission of the Controller of Her Majesty's Stationery Office. Crown copy right reserved.

Figure 3(c) - Manchester study sites

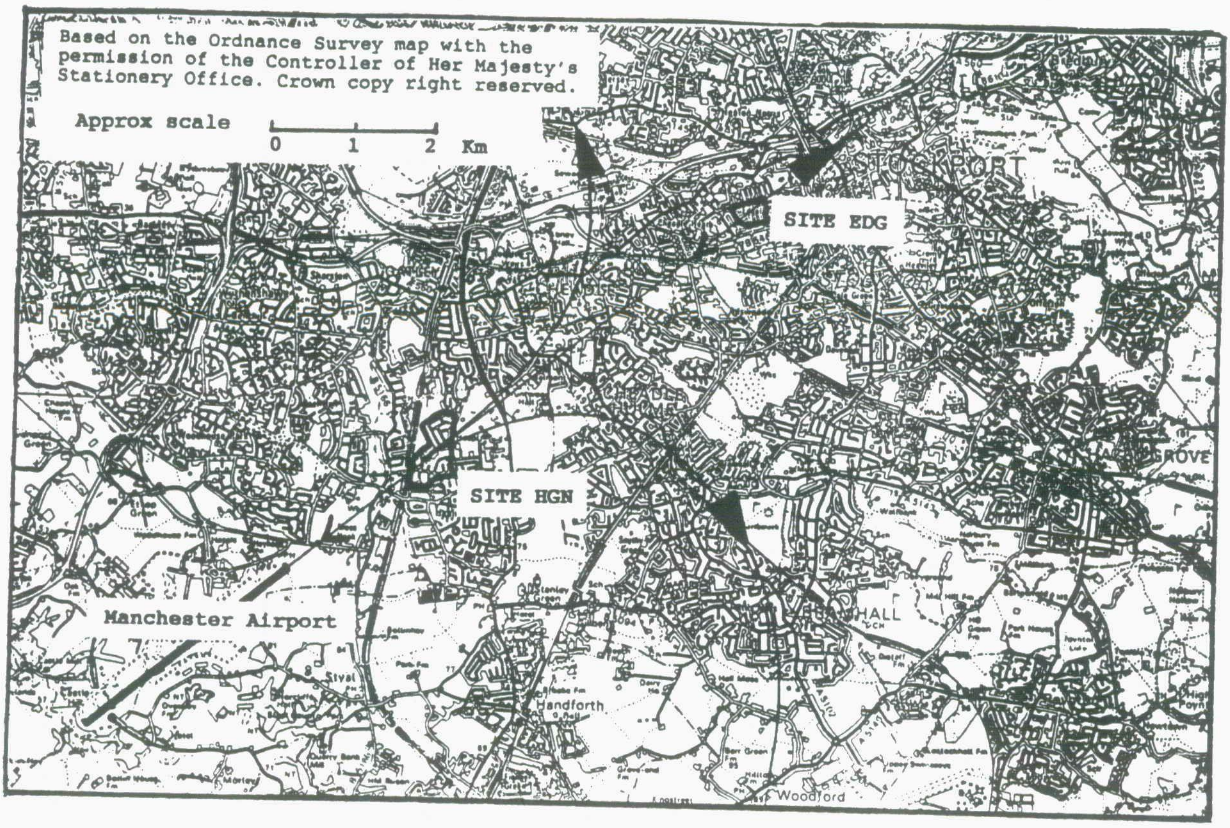


Figure 3(d) - Stansted study sites

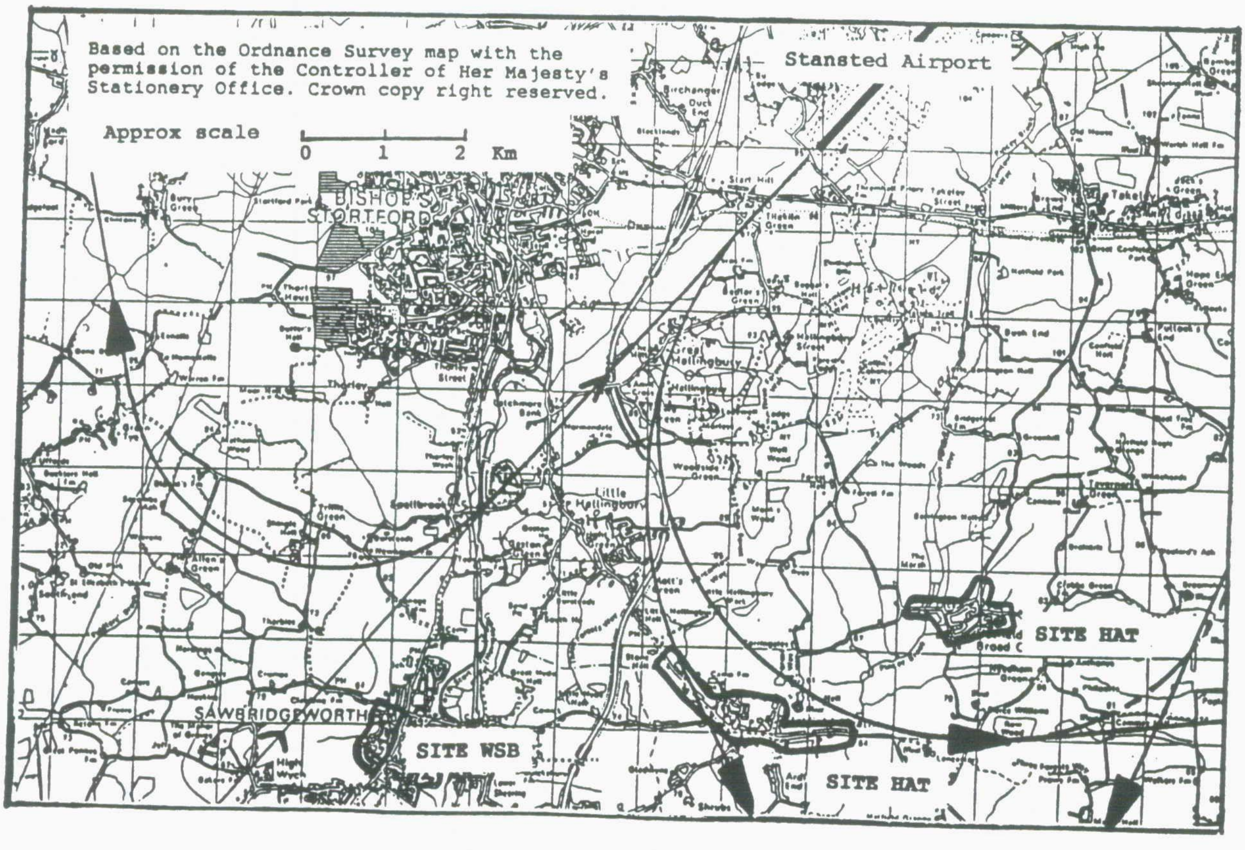


TABLE 3 CHARACTERISTICS OF SOCIAL SURVEY RESPONDENTS, ACTIMETRY SUBJECTS AND EEG SUBJECTS: % BY CATEGORY

Variable	Category		Subject group	
Number		Social survey 1636	Actimetry 400	EEG 50
Age	20-24	33	37	35
	35-49	29	33	33
	50-70	37	29	33
Sex	Male	50	48	47
	Female	50	52	53
Marital Status	Married	71	75	78
	Single	16	14	16
	Separated etc	12	11	6
Occupational group	ABC1	48	50	47
	C2DE	51	50	53
Years in residence	<5	31	40	33
	≥5	69	60	67
Children	None	63	55	55
	1 or more	37	45	45
Windows	Single	64	64	65
	Double	36	36	35
Windows at night	Shut	79	80	78
	Open	20	19	22
Sleeper	Light	42	44	45
	Deep	58	56	55
Bed-partner	None	29	25	20
	never disturbs	48	46	51
	disturbs	23	29	29
ANGEN	yes	71	71	69
	no	29	29	31
ANWK	yes	80	77	78
	no	20	22	22
HEARNT	yes	83	81	84
	no	17	19	16
COMPLAIN	yes	89	91	84
	no	11	9	16
DIFFGET	yes	60	63	55
	no	40	37	45
WOKENREG	yes	84	81	84
	no	16	19	16
GETBACK	yes	68	72	69
	no	32	28	31

ANGEN	Very much annoyed by aircraft noise
HEARNT	Very much annoyed by aircraft noise at night
ANWK	Awakened at night by aircraft noise
COMPLAIN	Has made a formal complaint about noise
DIFFGET	Has difficulty getting to sleep
WOKENREG	Regularly woken up once asleep
GETBACK	Has difficulty getting back to sleep once woken

TABLE 4 SITE HOURLY NOISE EXPOSURES, dBA

(a) Aircraft Leq

4,5km +9km 7km

Hour	HLW	LGN	SWM	LFD	HGN	EDG	HAT	WSB
2200-2259	59.5	58.3	63.1	52.6	64.5	56.4	47.9	44.6
2300-2359	52.2	52.9	56.7	50.7	65.3	53.1	46.9	44.7
0000-0059	43.3	47.9	42.8	45.0	65.0	55.9	42.6	38.2
0100-0159	39.6	42.8	42.2	43.0	62.3	53.7	42.1	36.4
0200-0259	43.5	41.4	40.1	42.4	55.7	49.5	40.6	38.1
0300-0359	39.7	41.7	40.9	41.6	62.0	51.3	39.0	36.9
0400-0459	57.1	48.7	45.4	49.1	63.4	53.8	39.1	36.3
0500-0559	61.3	47.7	46.7	48.3	64.6	52.9	40.9	39.5
0600-0659	64.2	58.4	52.8	52.7	67.4	57.2	46.7	44.4
0700-0759	61.2	62.4	65.5	56.4	71.6	63.1	53.8	52.6

≈ 68,5 66,5 56,7 0

(b) Ambient Leq

Hour	HLW	LGN	SWM	LFD	HGN	EDG	HAT	WSB
2200-2259	44.0	47.5	48.7	43.8	48.8	45.1	45.0	41.7
2300-2359	40.9	45.5	45.2	43.7	48.3	44.2	43.5	40.7
0000-0059	39.0	42.8	42.0	40.6	46.5	43.5	40.8	37.7
0100-0159	37.3	40.1	40.0	39.4	45.4	42.0	39.1	36.1
0200-0259	36.9	40.4	38.9	38.5	43.3	40.9	37.1	35.5
0300-0359	37.2	39.8	40.9	37.9	43.7	41.2	37.1	35.1
0400-0459	47.0	47.7	45.2	43.1	44.4	43.0	37.5	34.7
0500-0559	47.9	46.7	46.6	44.3	45.8	44.4	40.7	36.2
0600-0659	48.9	49.8	48.1	47.0	48.7	48.0	45.1	38.8
0700-0759	50.8	51.7	51.4	49.0	52.7	49.2	49.4	46.8

(c) Background L90

Hour	HLW	LGN	SWM	LFD	HGN	EDG	HAT	WSB
2200-2259	37.6	38.8	42.3	34.2	43.3	39.1	35.3	34.4
2300-2359	36.6	36.0	40.6	33.2	42.1	38.6	34.1	33.9
0000-0059	34.8	34.0	38.6	31.5	39.1	39.5	32.6	31.9
0100-0159	33.4	33.3	36.8	31.2	36.4	34.5	31.6	30.5
0200-0259	32.8	32.9	35.8	30.3	34.6	33.7	30.8	30.3
0300-0359	33.0	33.5	37.1	30.2	34.9	33.2	30.7	30.0
0400-0459	37.0	38.4	41.3	33.8	36.4	33.6	31.2	30.0
0500-0559	43.1	40.0	43.5	35.0	38.7	35.3	33.2	31.7
0600-0659	43.0	42.9	45.2	38.4	42.8	39.5	35.8	33.8
0700-0759	44.6	44.3	46.8	40.7	47.9	43.6	42.1	37.0

Figure 10 - Length of residence in local area

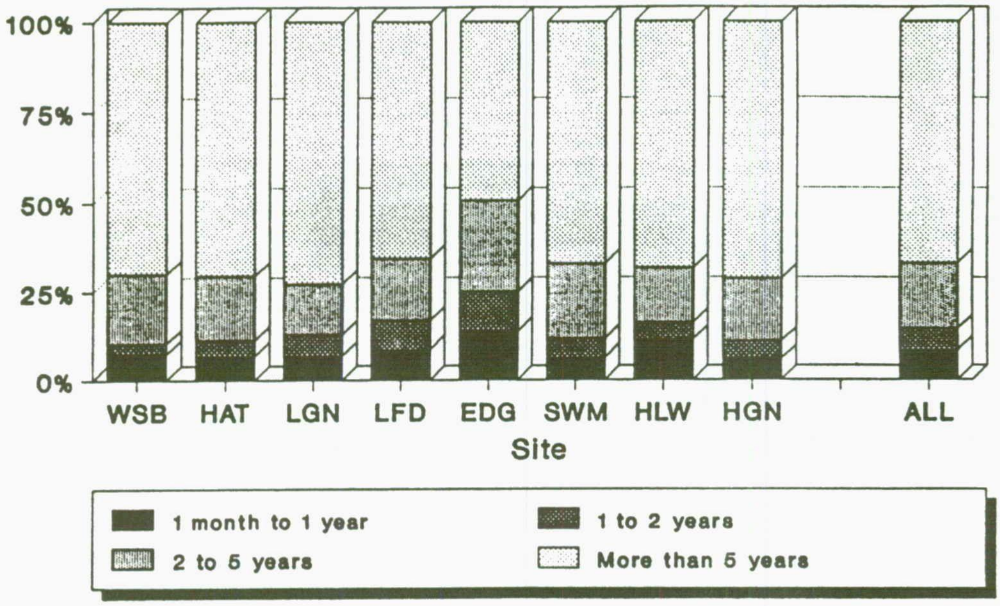


Figure 11 - On the whole, how do you rate living in this area?

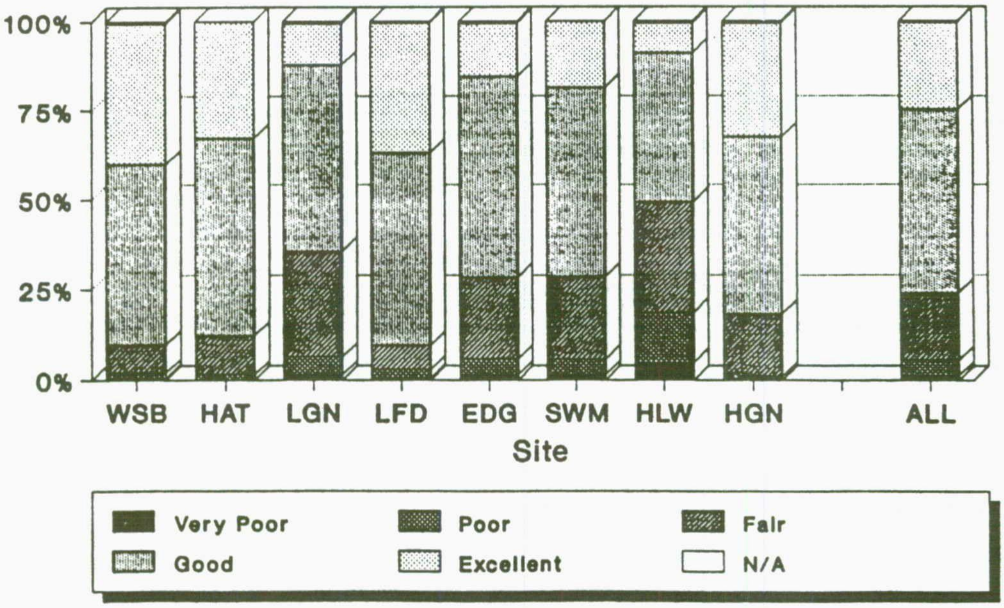


Figure 12 - What are the things you <u>like</u> about living around here?

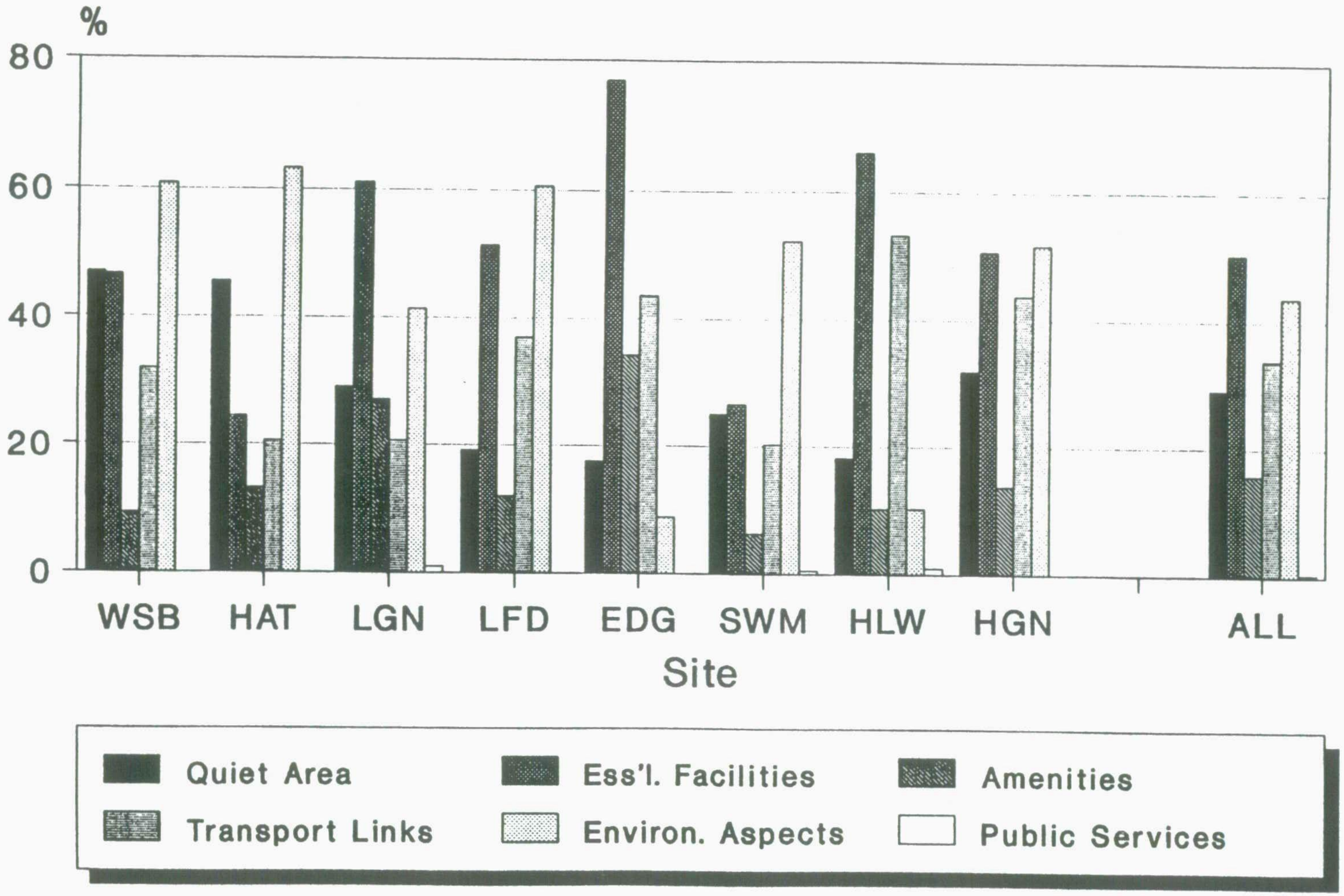


Figure 13 - What are the things you <u>don't like</u> about living around here?

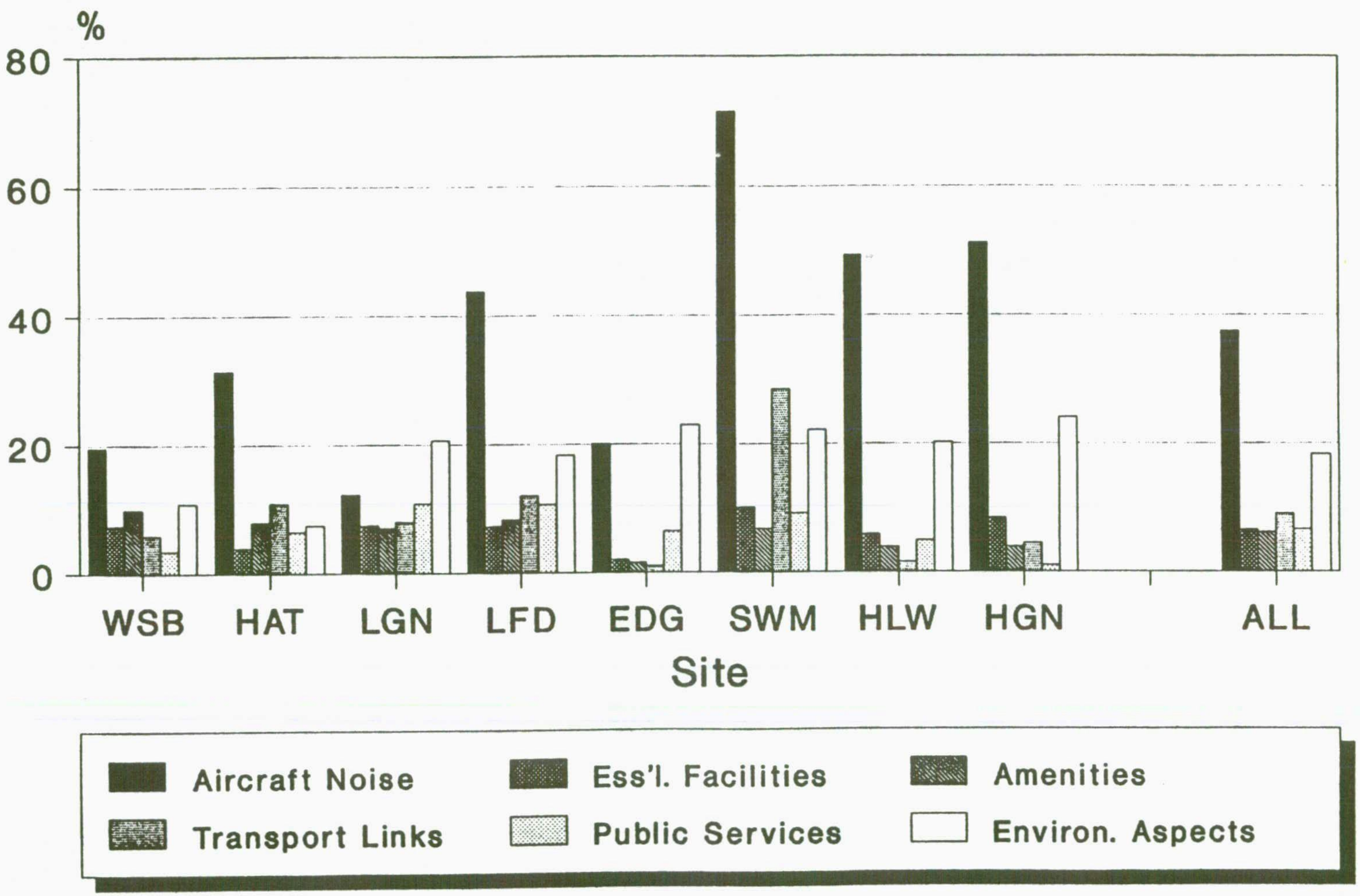


Figure 14 - What different kinds of noise do you hear around here?

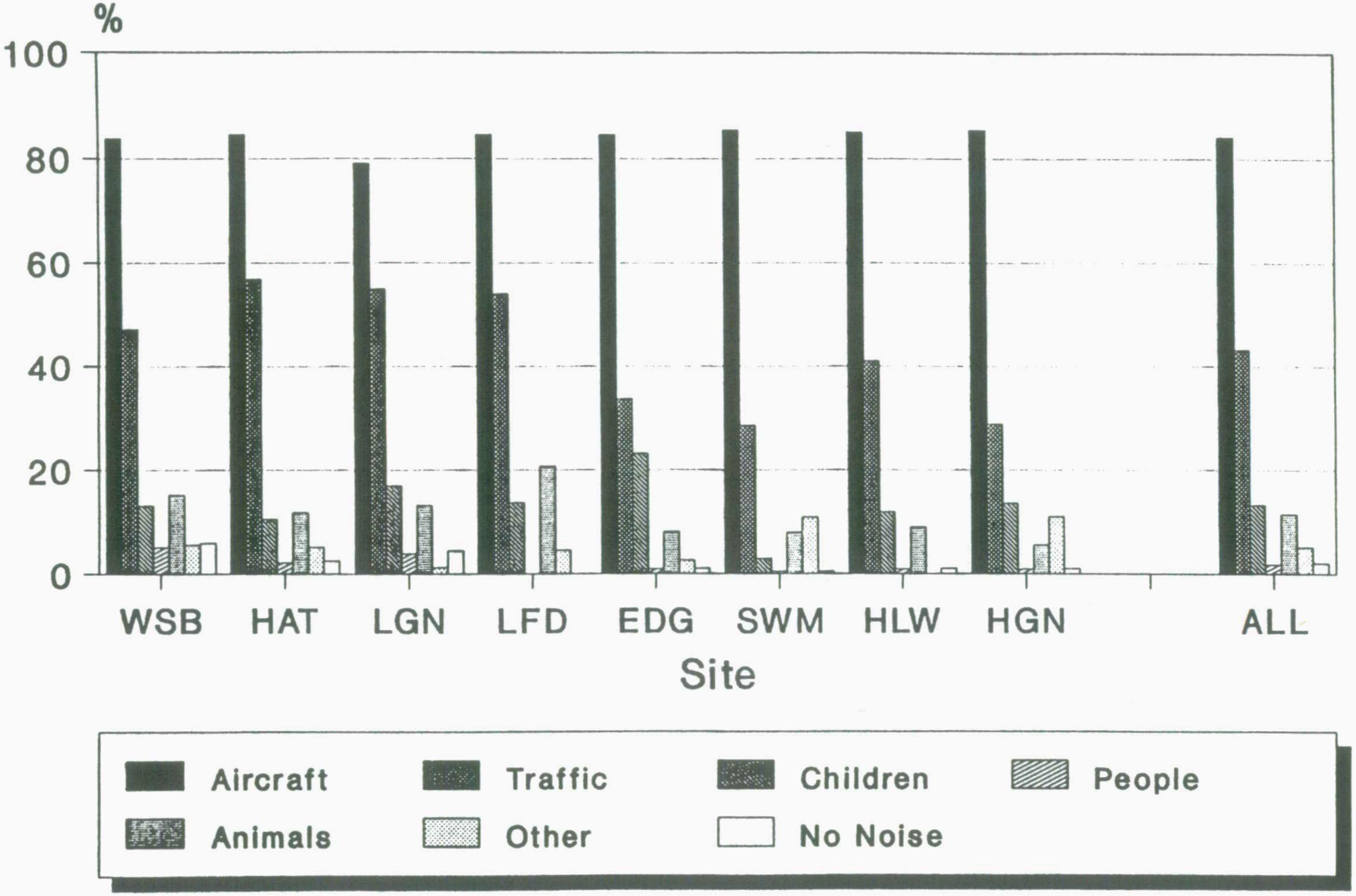


Figure 15 - On the whole, is this a quiet or noisy area?

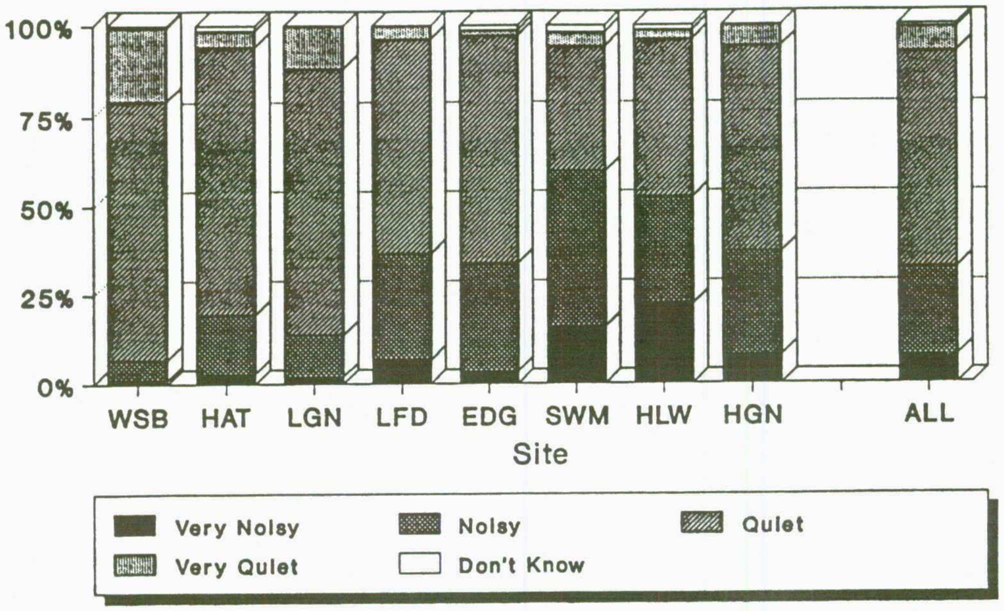


Figure 16 - How much does aircraft noise around here bother or annoy you?

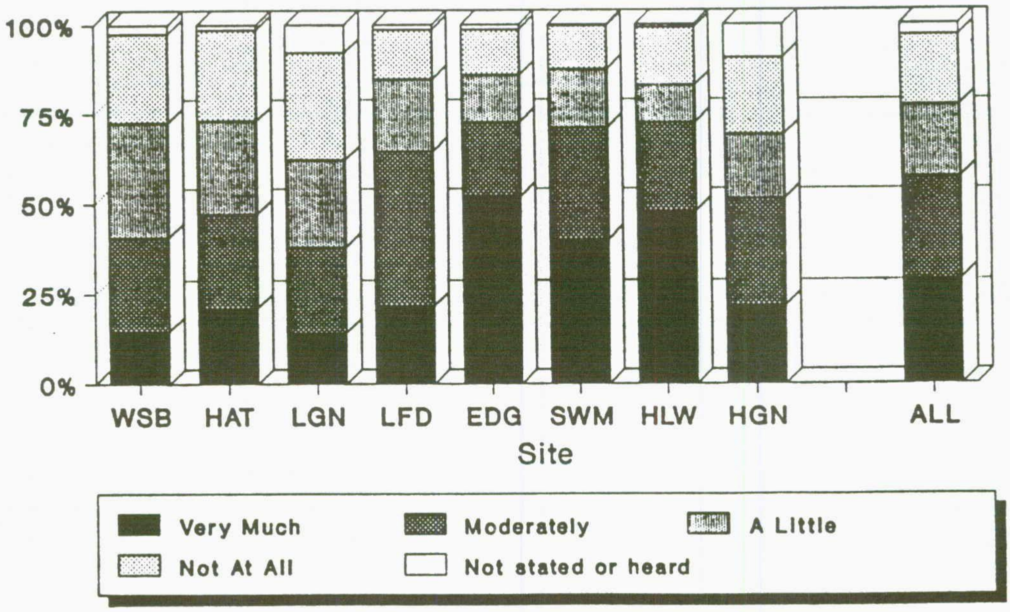


Figure 17 - How much does the noise of .... here bother or annoy you?

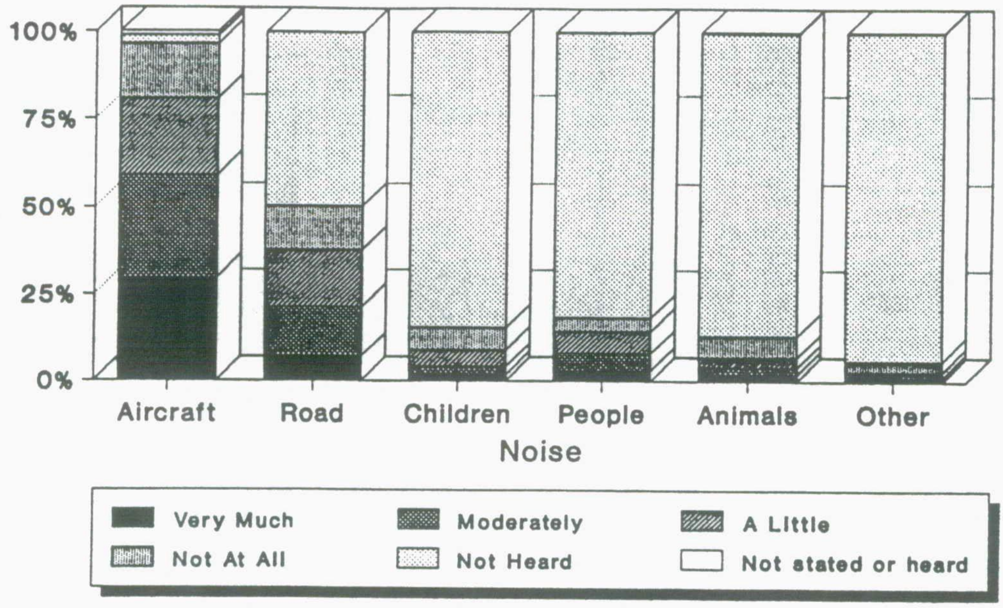


Figure 18 - At what time do you normally go to bed on weekdays?

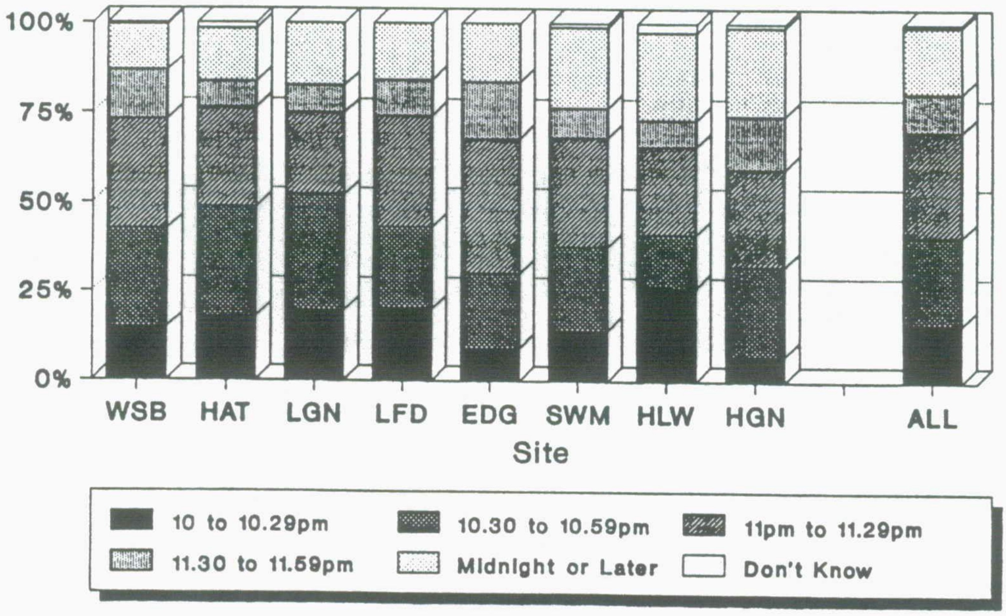


Figure 19 - How well or badly do you normally sleep at night?

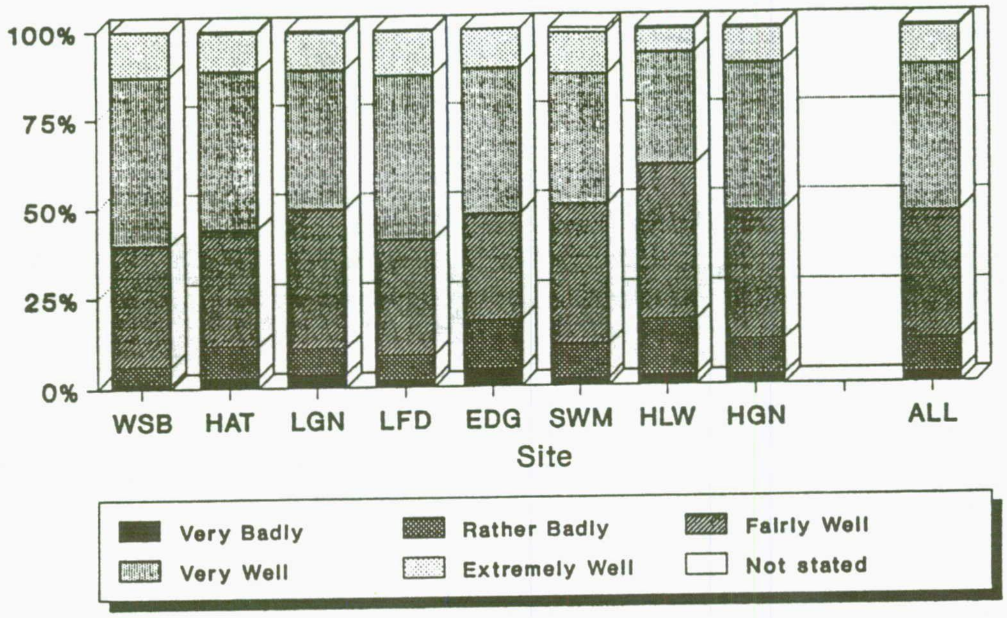


Figure 20 - Would you describe yourself as a light or deep sleeper?

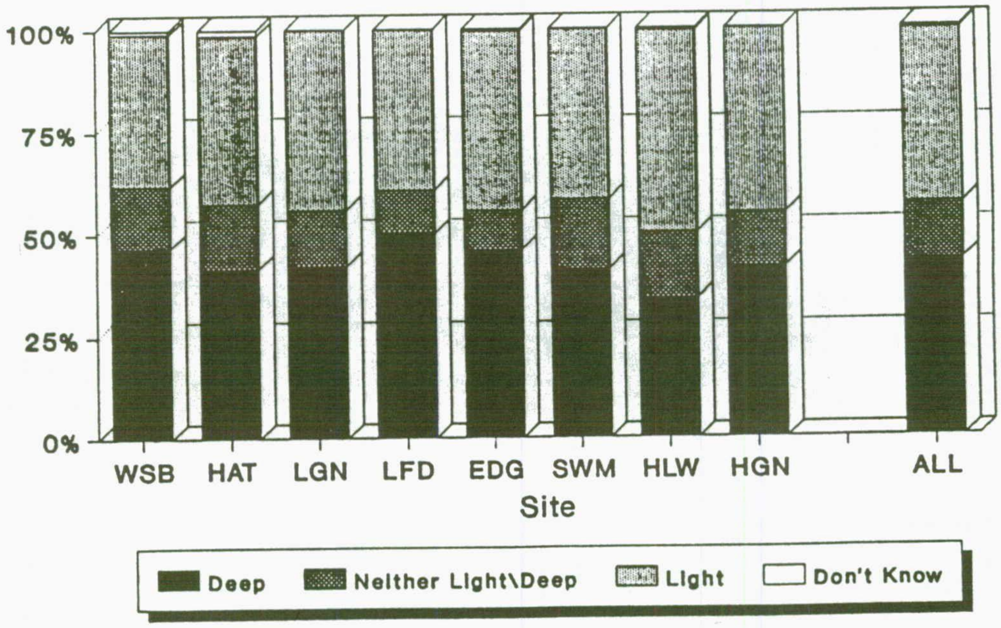


Figure 21 - How often do you have difficulty getting to sleep?

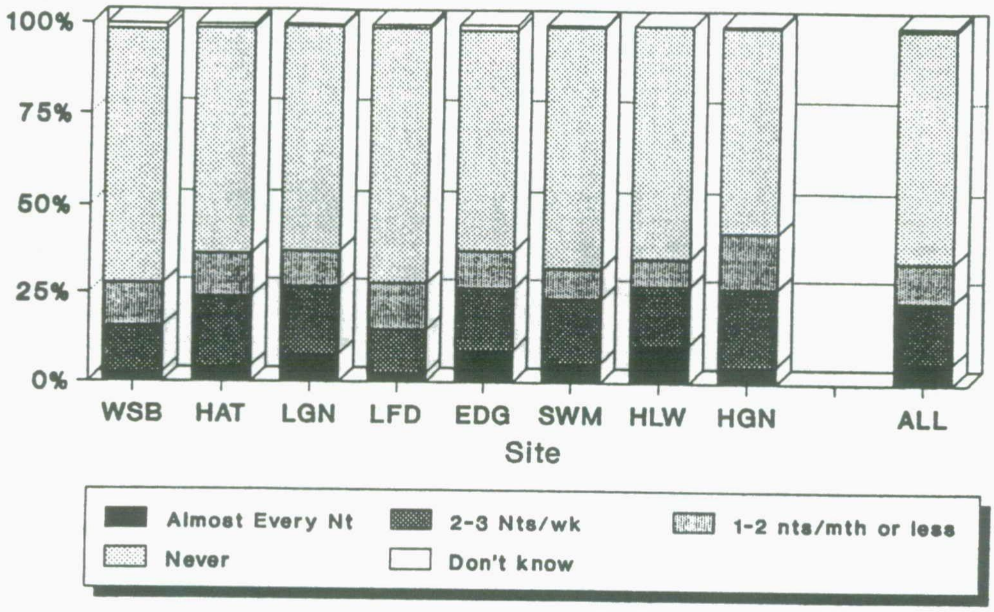


Figure 22 - Are you ever woken up once asleep, is that regularly or sometimes?

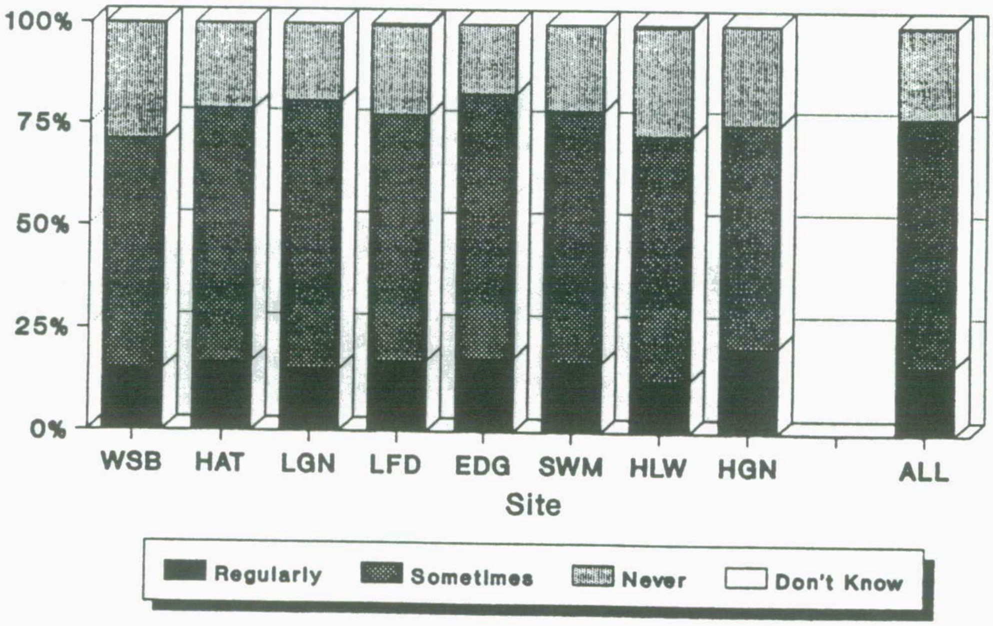


Figure 23 - How often are you woken up once you are asleep?

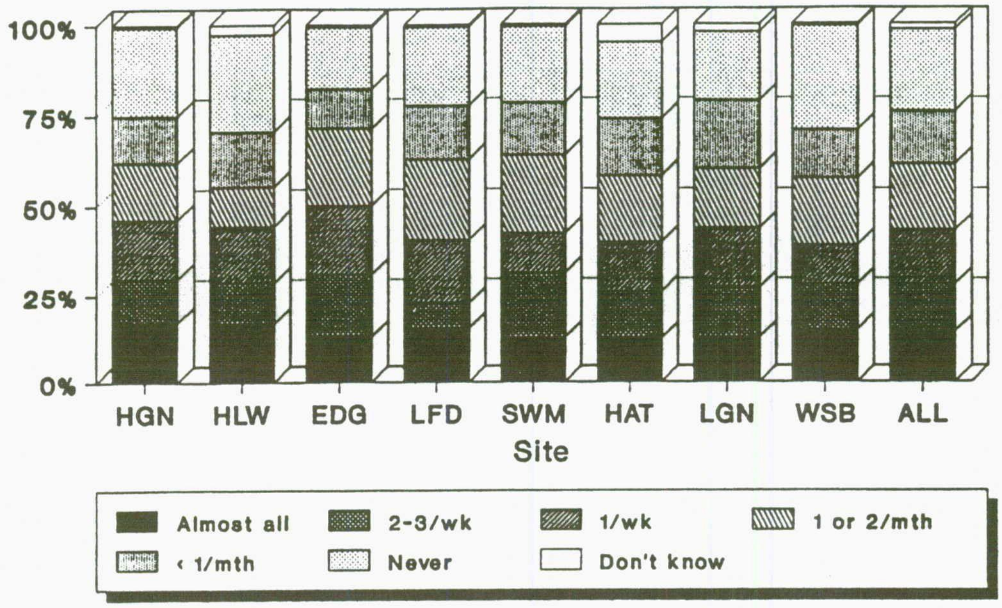


Figure 24 - Typically, how many times a night are you woken up?

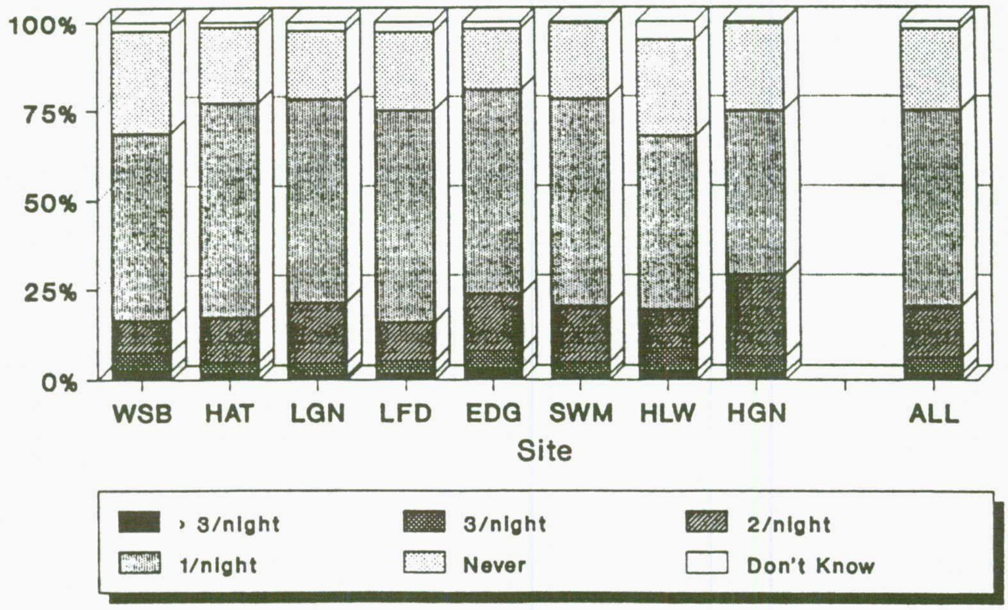


Figure 25 - If woken at night, how difficult is it to get back to sleep?

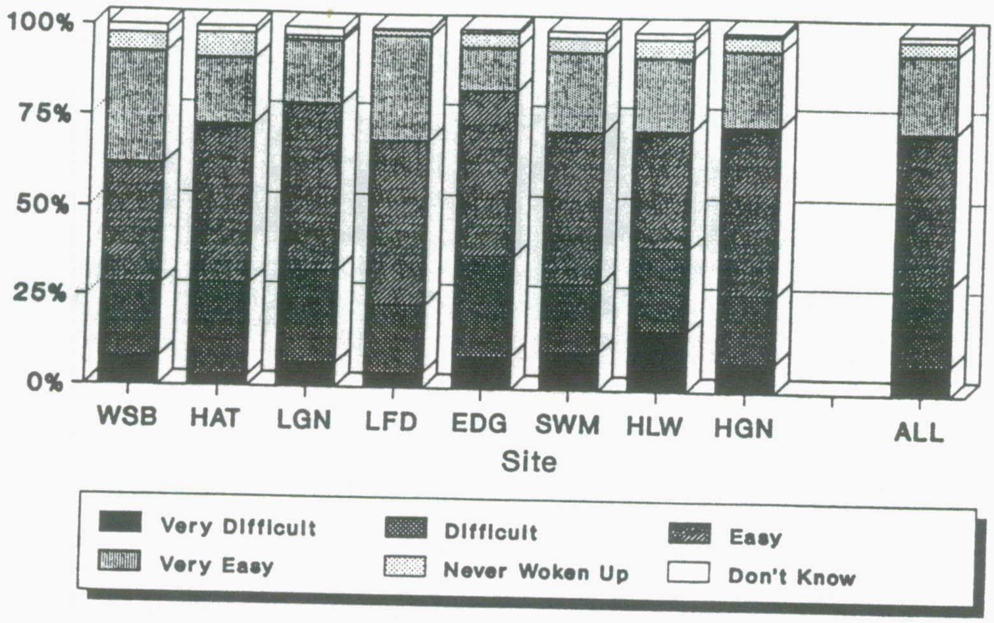


Figure 26 - Are there any particular times of the night when you wake up?

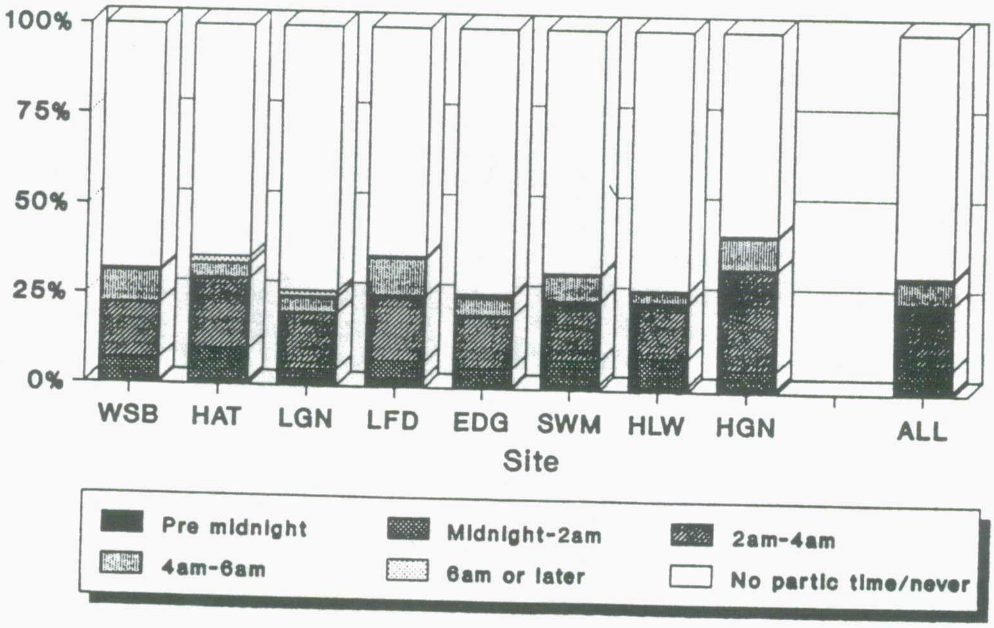


Figure 27 - What causes you to wake up?

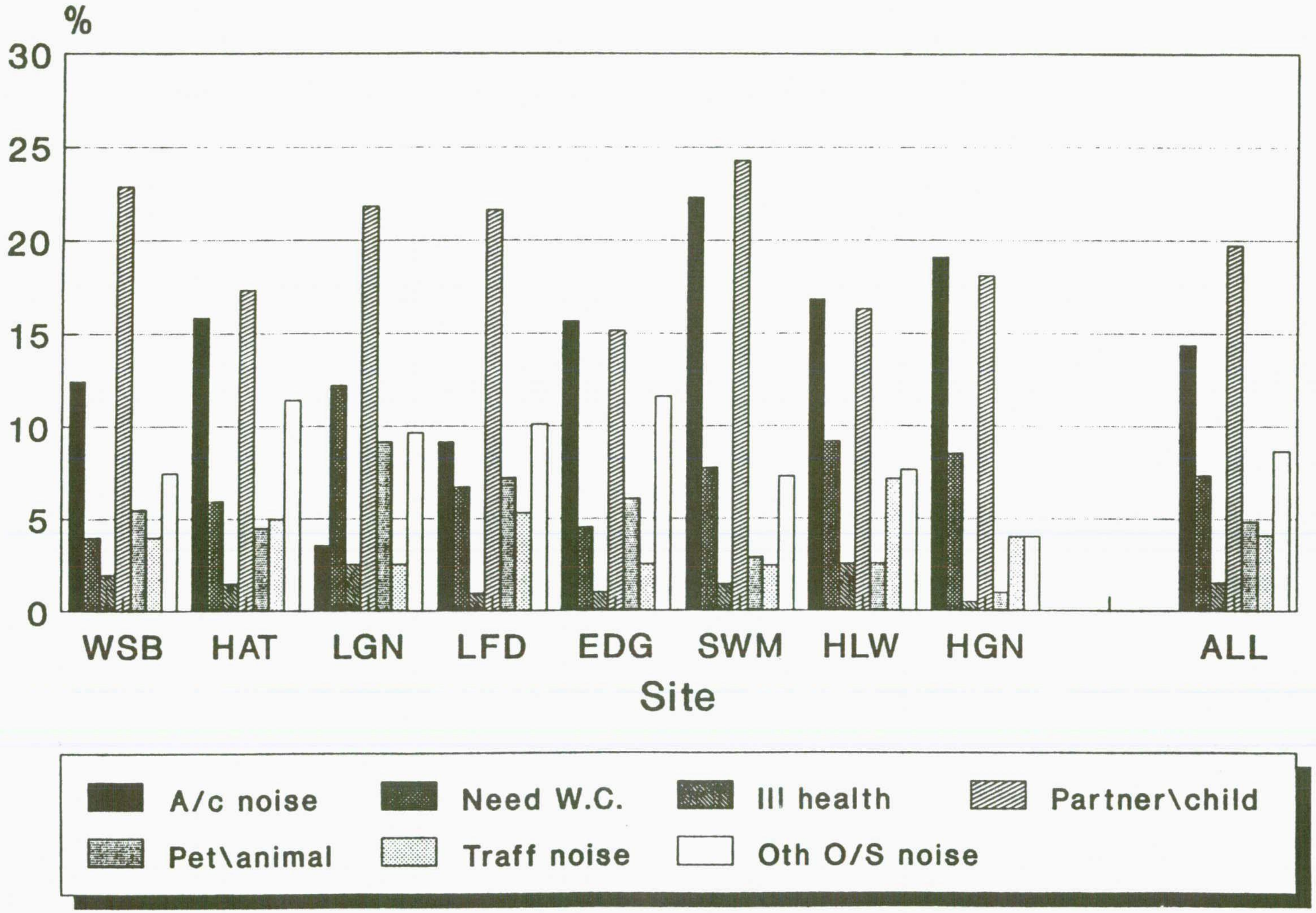


Figure 28 - At about what time do you normally get up on weekdays?

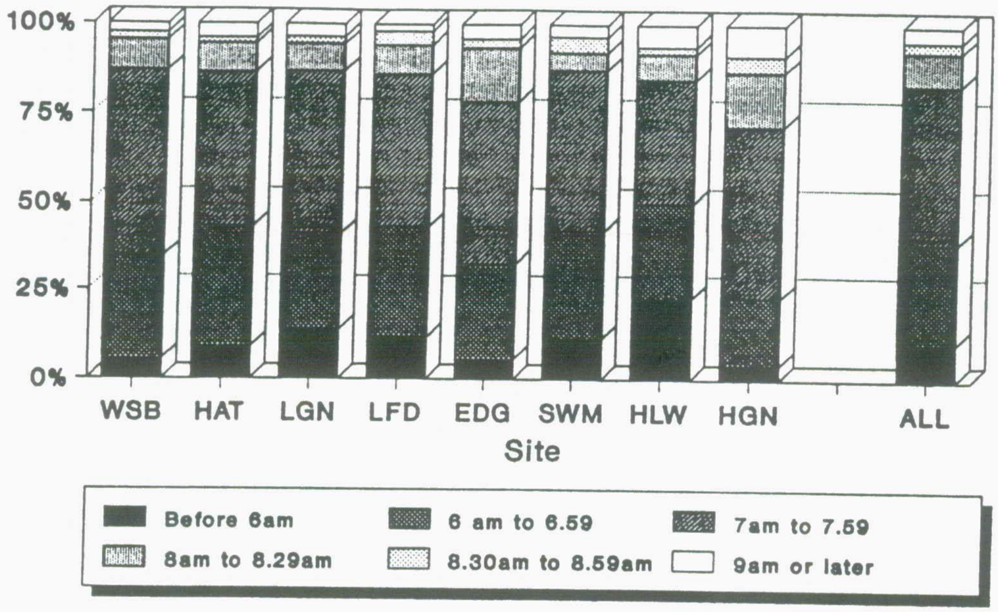


Figure 29 - How do you feel when you wake up after a typical night's sleep?

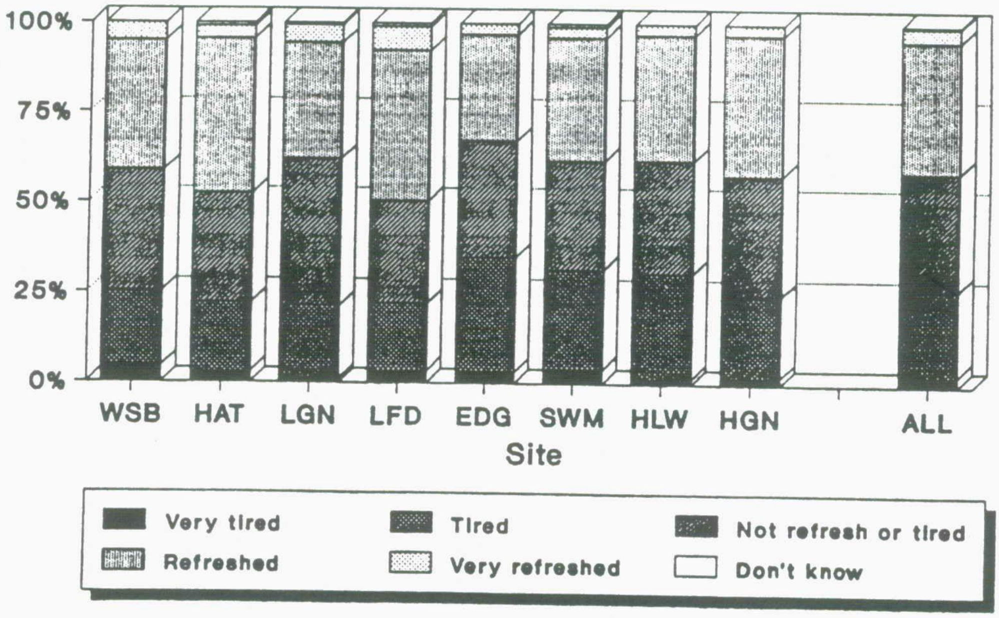


Figure 30 - In good weather do you sleep with your windows... ?

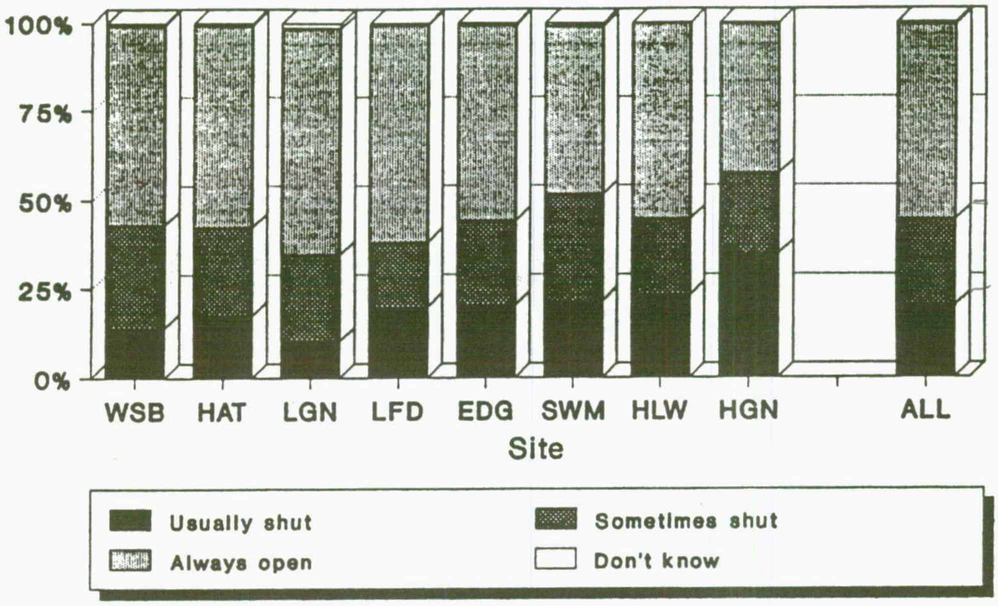


Figure 31 - Does your bedroom have double or secondary double glazing?

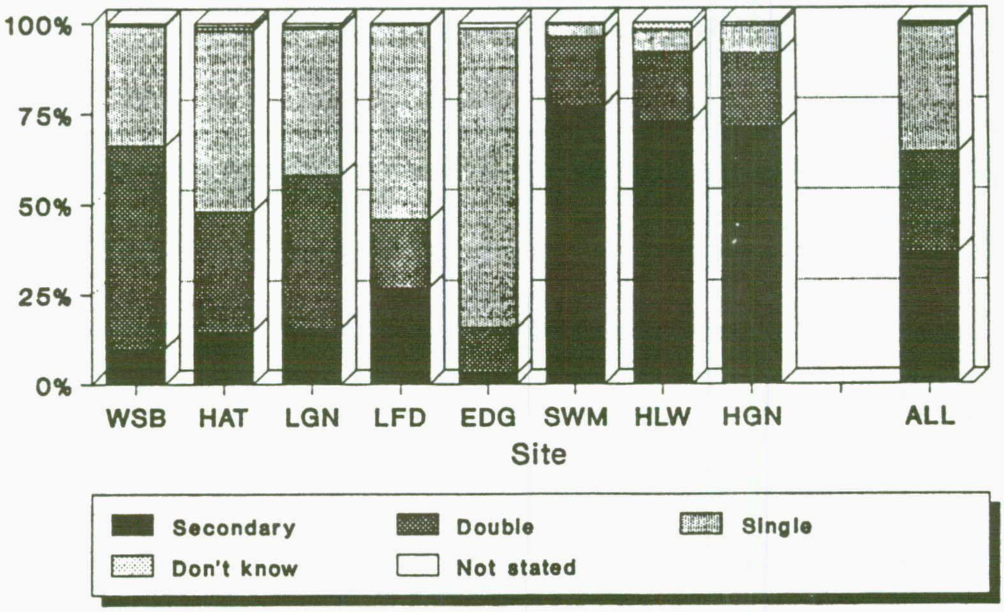


Figure 32 - Percentage of respondents spontaneously mentioning aircraft noise as a reason for disliking area

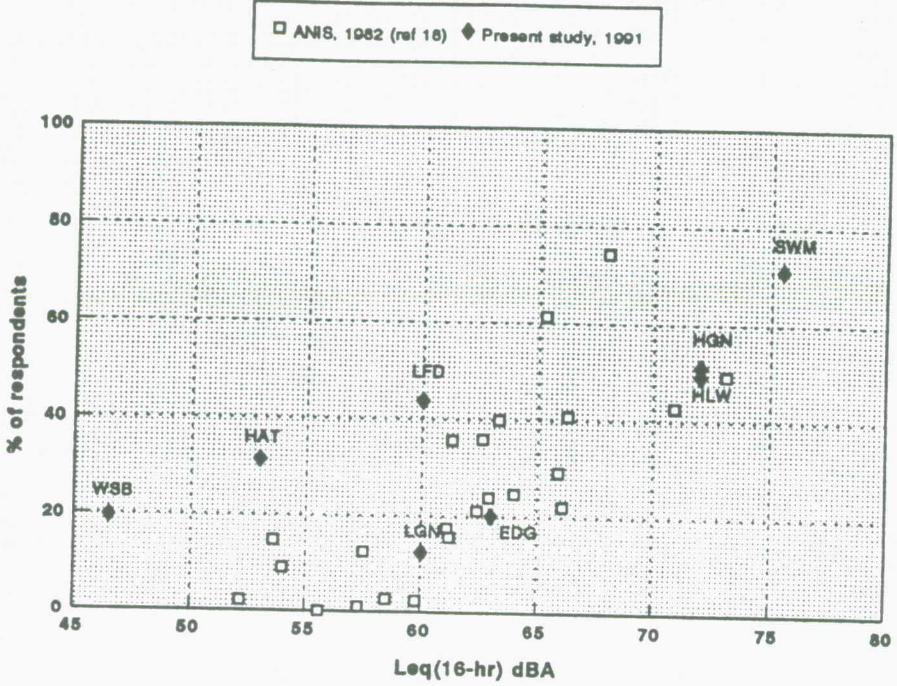


Figure 33 - Percentage of respondents very much annoyed by aircraft noise

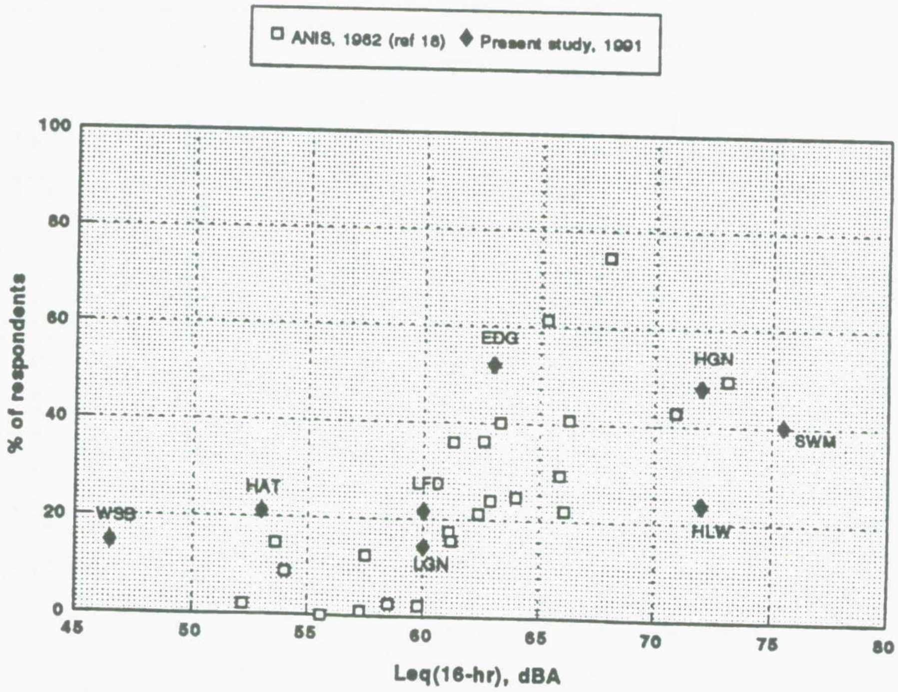
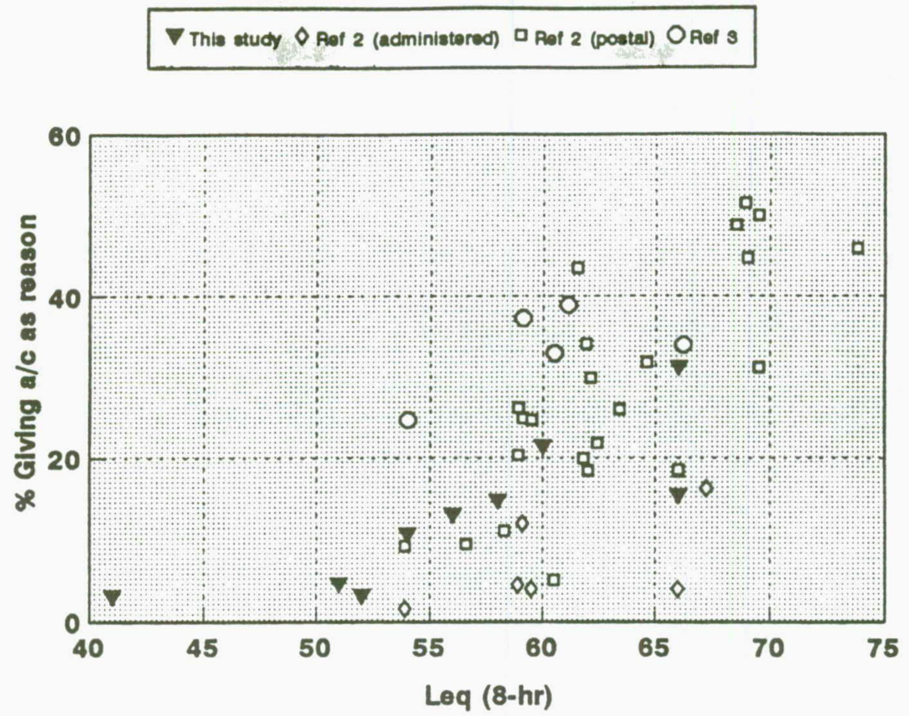
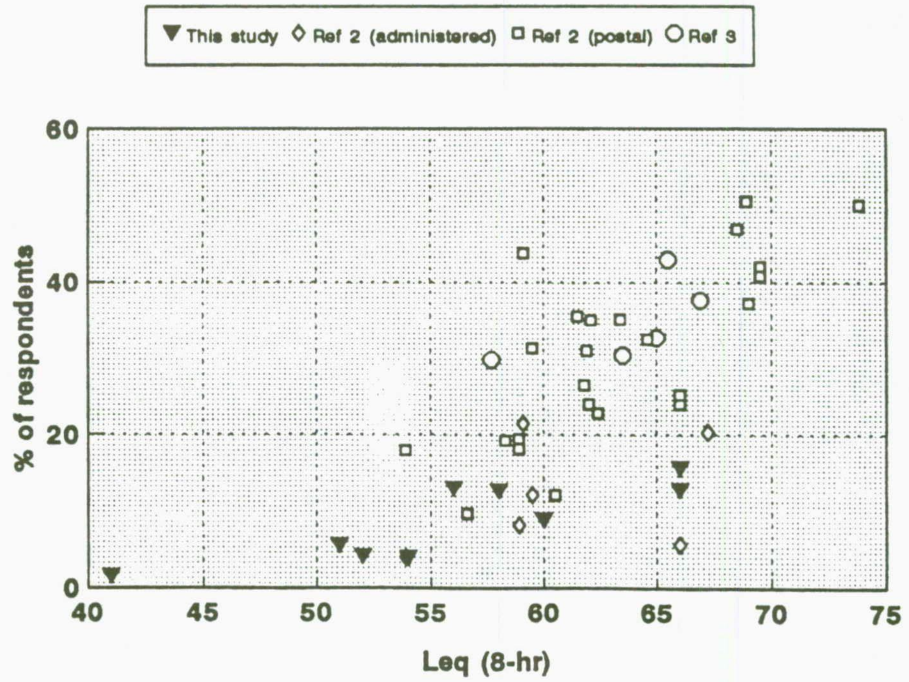


Figure 34 - Aircraft given as reason for having windows closed



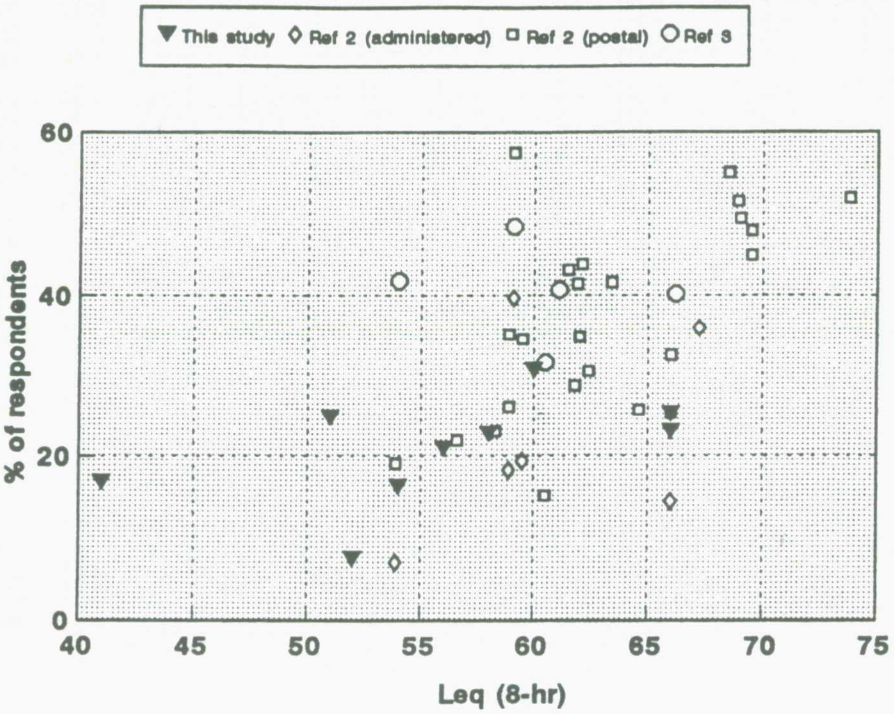
jop6g34

Figure 35 - Aircraft given as reason for having difficulty getting to sleep



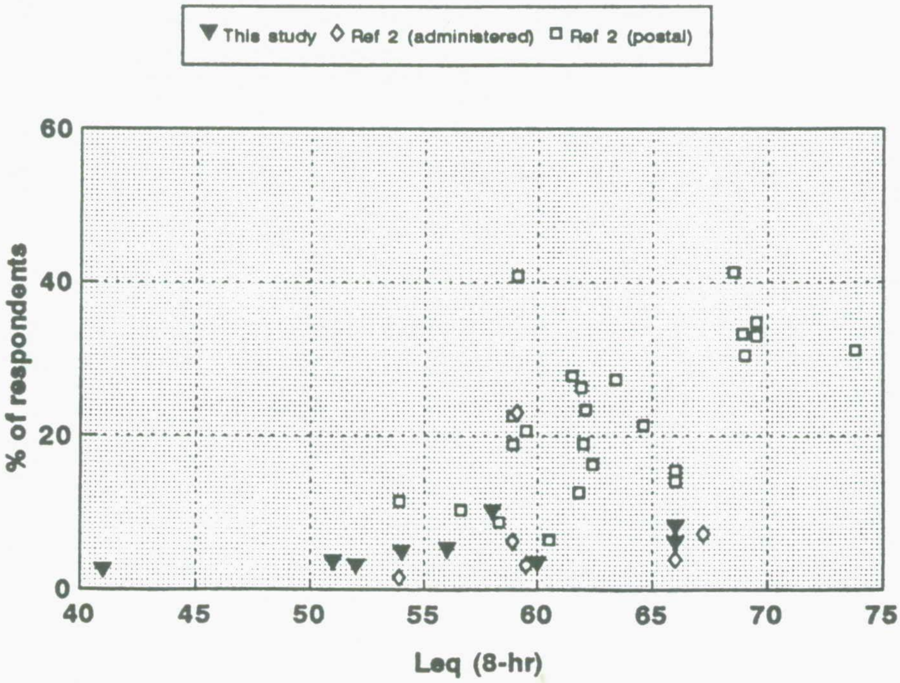
jop6g35

Figure 36 - Aircraft given as reason for awakening



jopfig36

Figure 37 - Aircraft given as reason for having difficulty getting back to sleep



jopfig37

Figure 38 - Sleep disturbance records

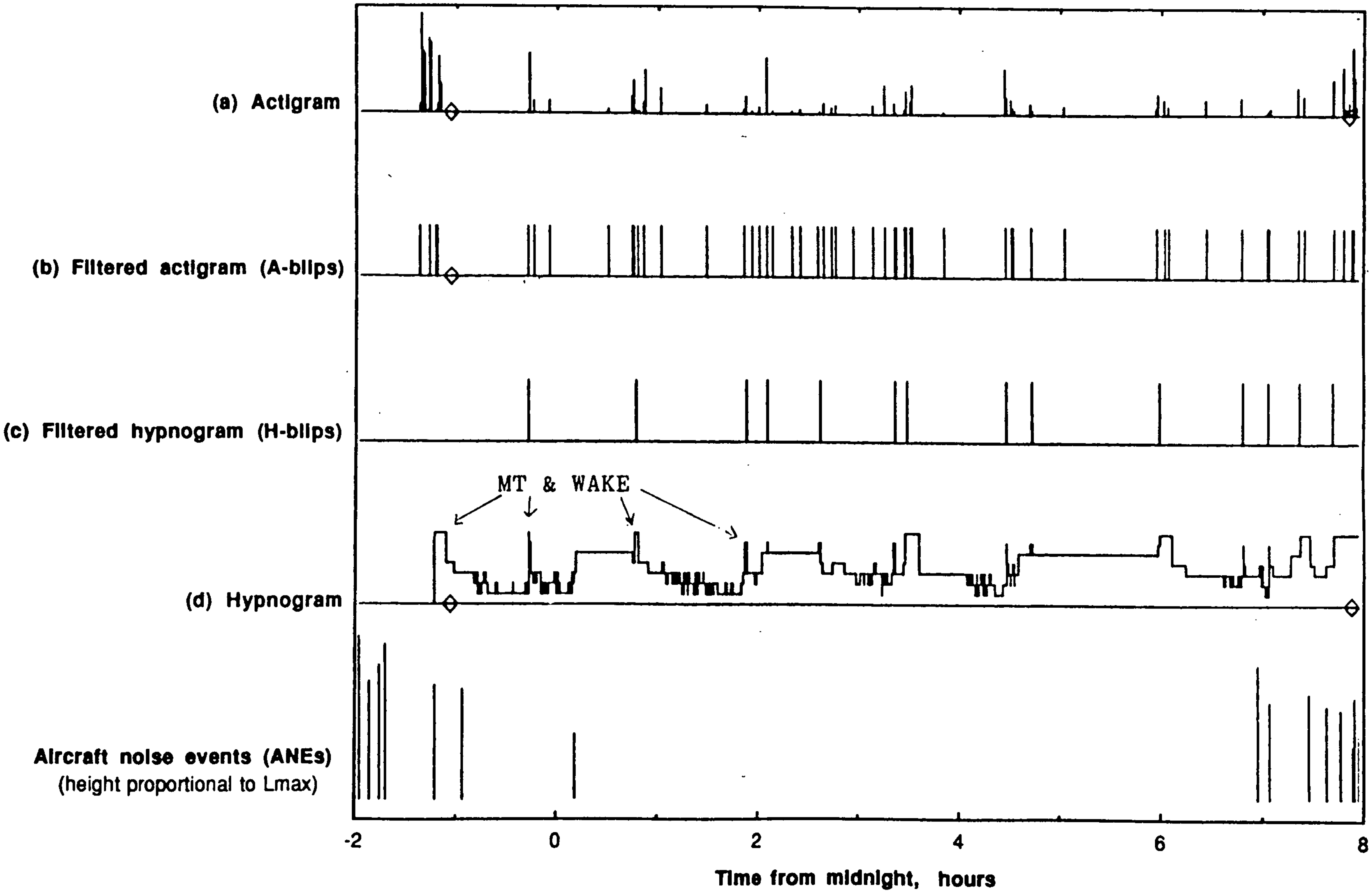


Figure 39 - Comparison of actigrams and hypnograms: calculation of hit rates

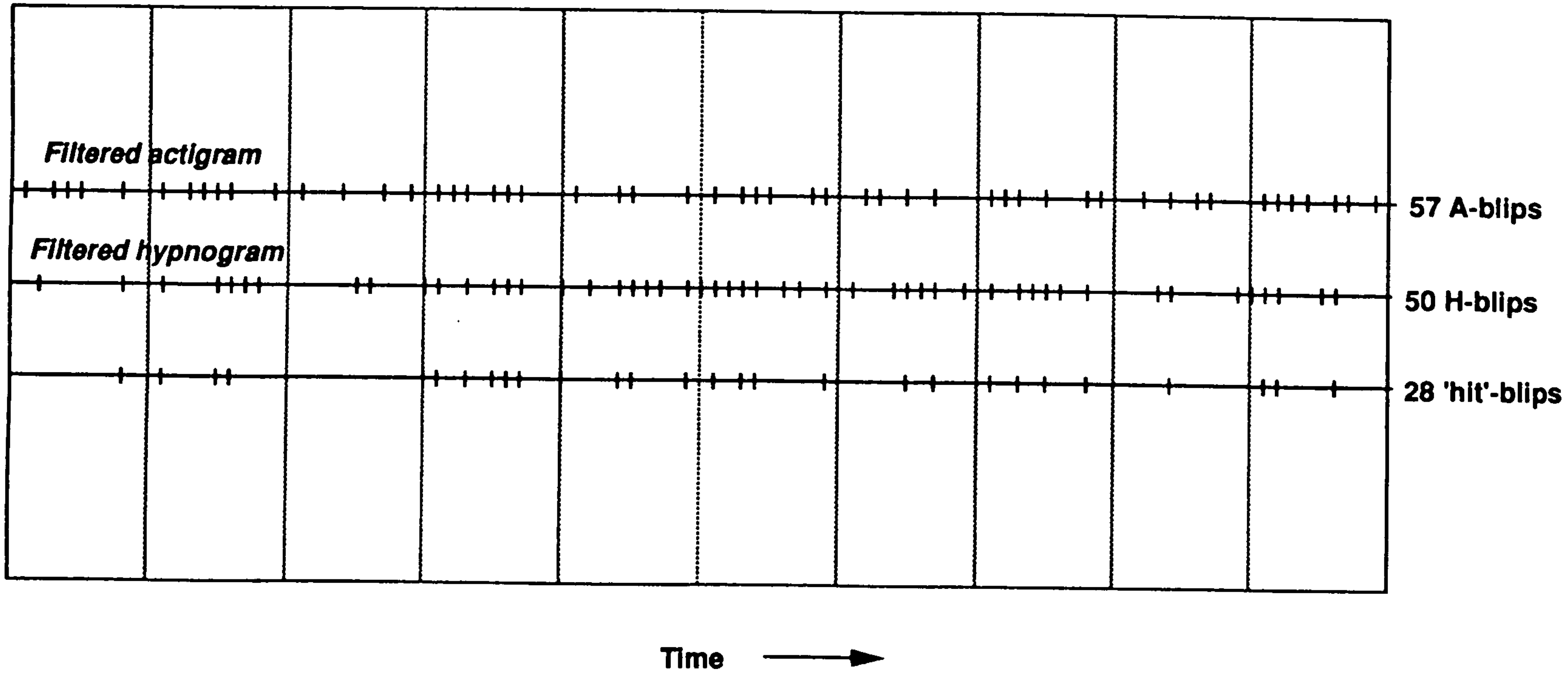


Figure 40 - Set of actigrams from 50 subjects on one night

Each trace is a single actigram from one subject. The vertical displacement of the trace at any point is proportional to the acceleration count in a single epoch.

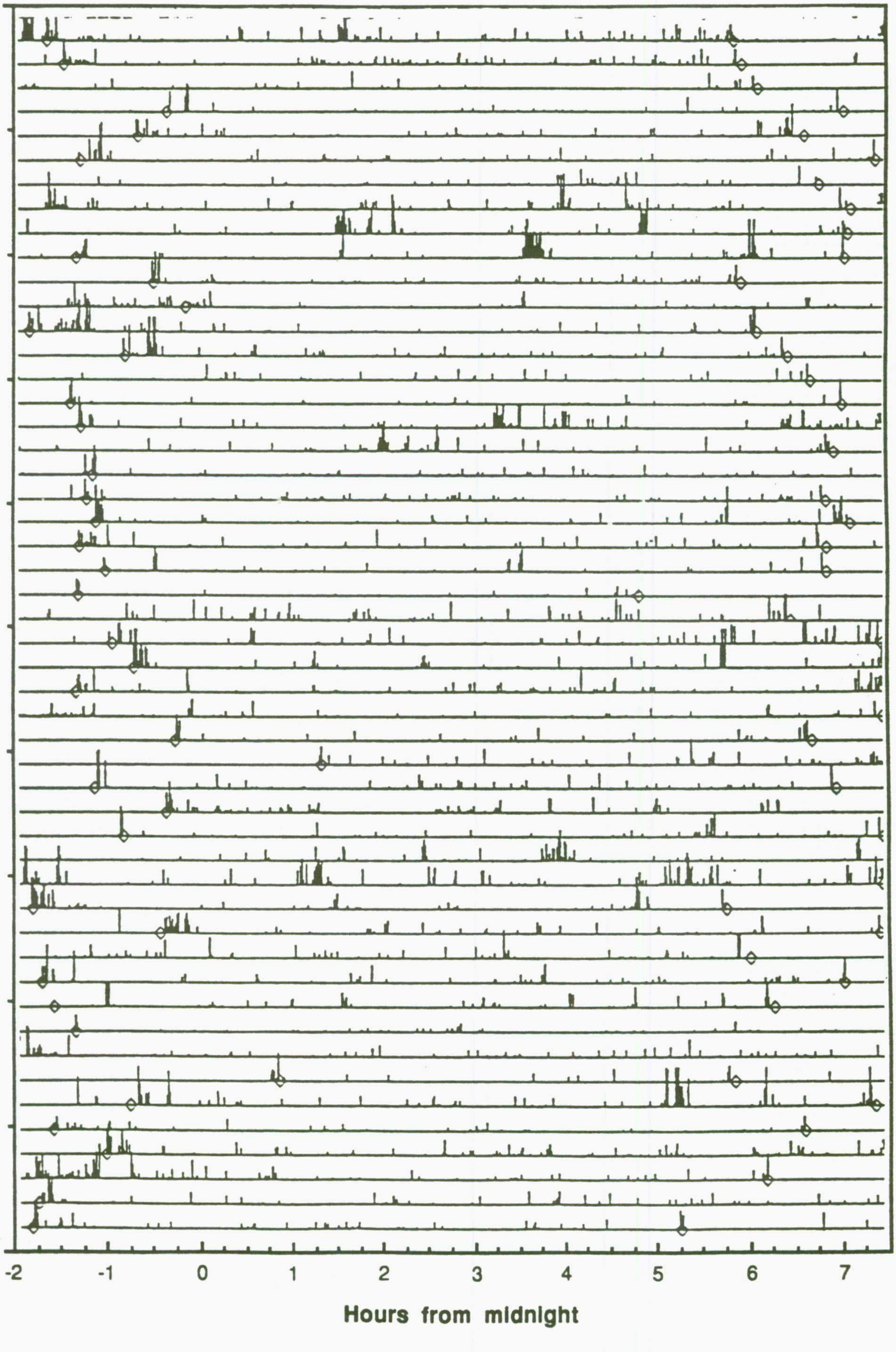


Figure 41 - Set of filtered actigrams from 50 subjects on one night

Each trace is a single actigram from one subject. Each vertical line on the trace (blip) indicates an arousal (disturbance onset)

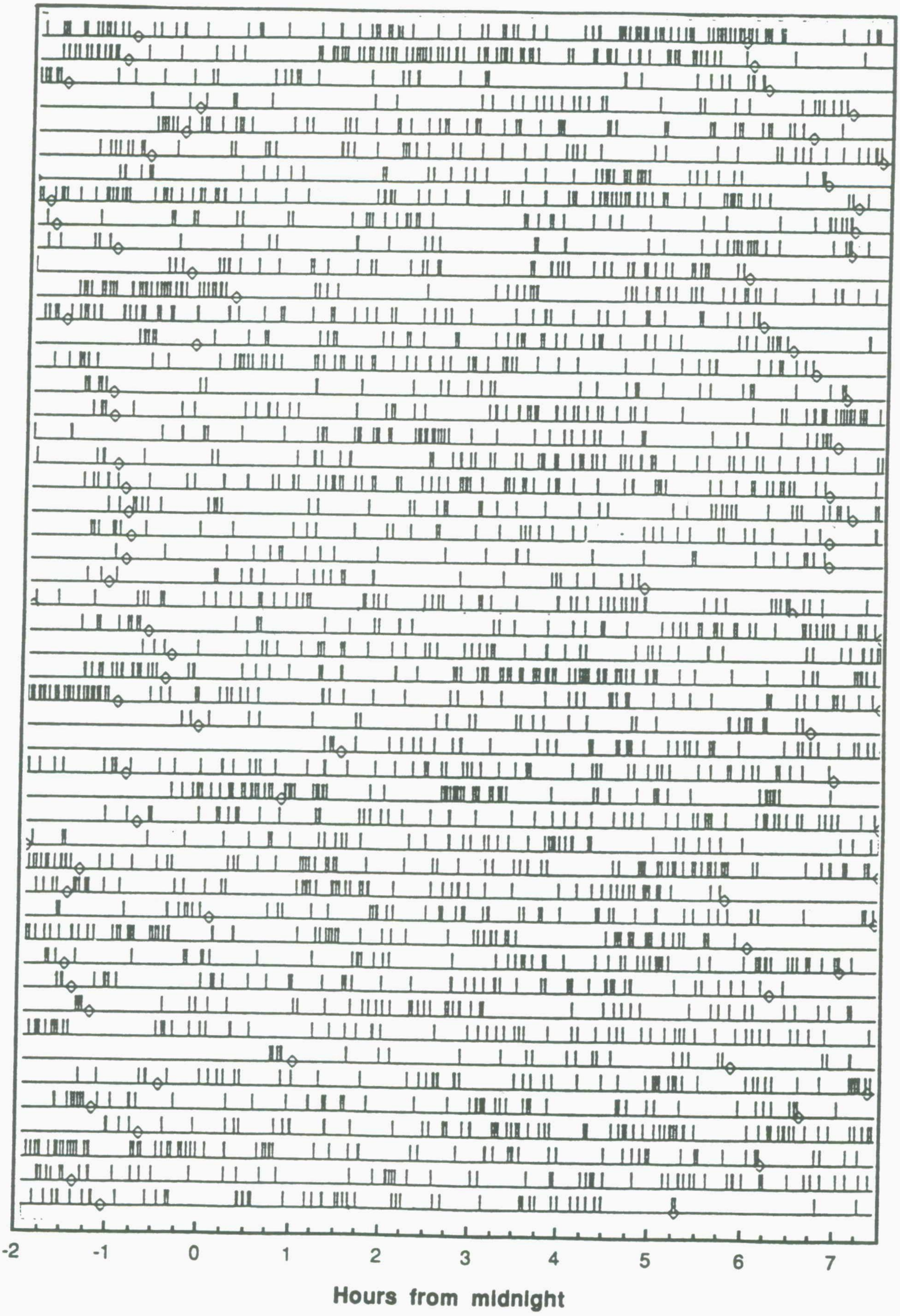


Figure 42 - Total disturbance (one night) represented by sum of actimetric disturbances (A-blips)

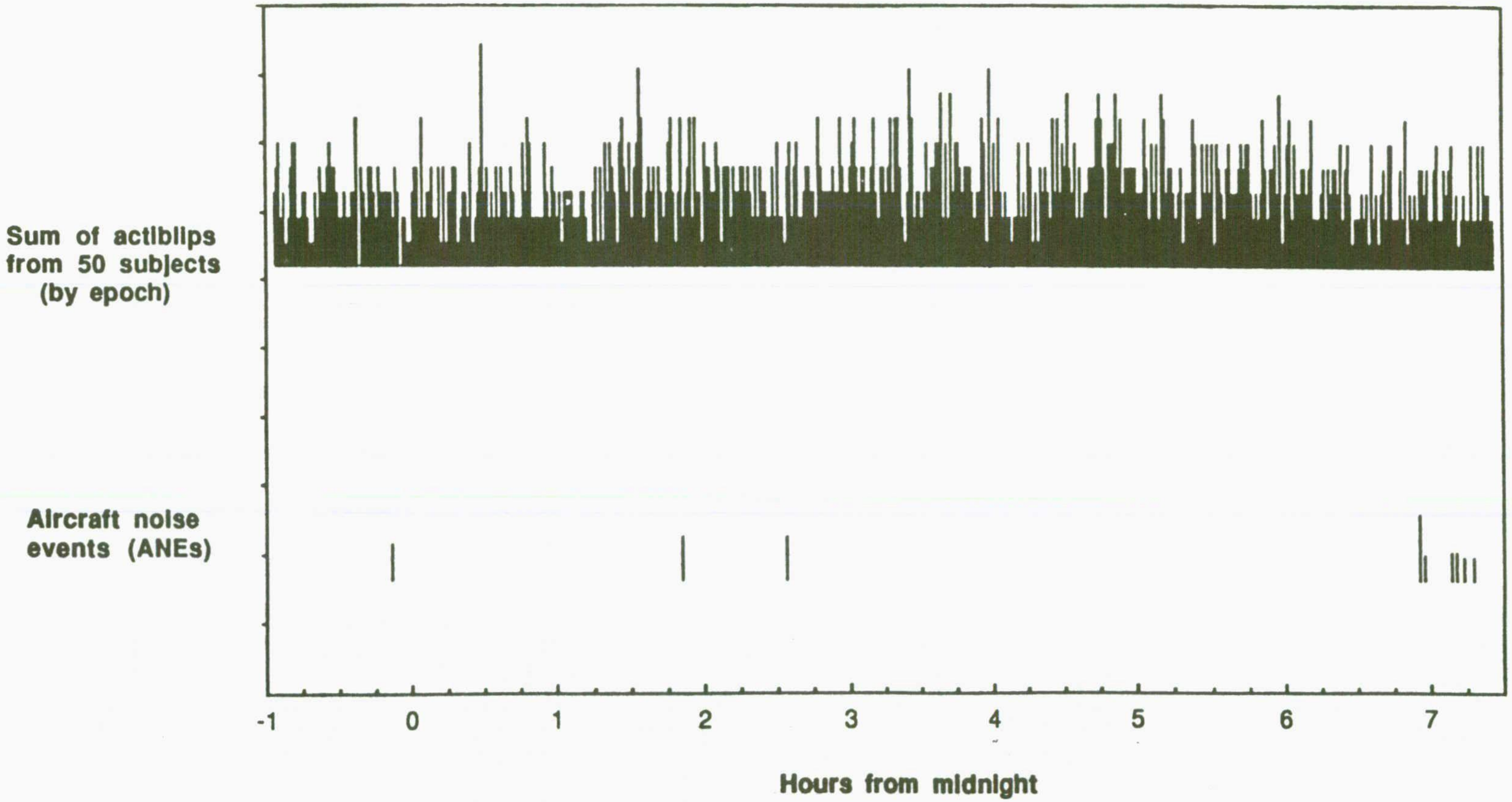


Figure 43 - EEG sample: durations of wakefulness or movement time

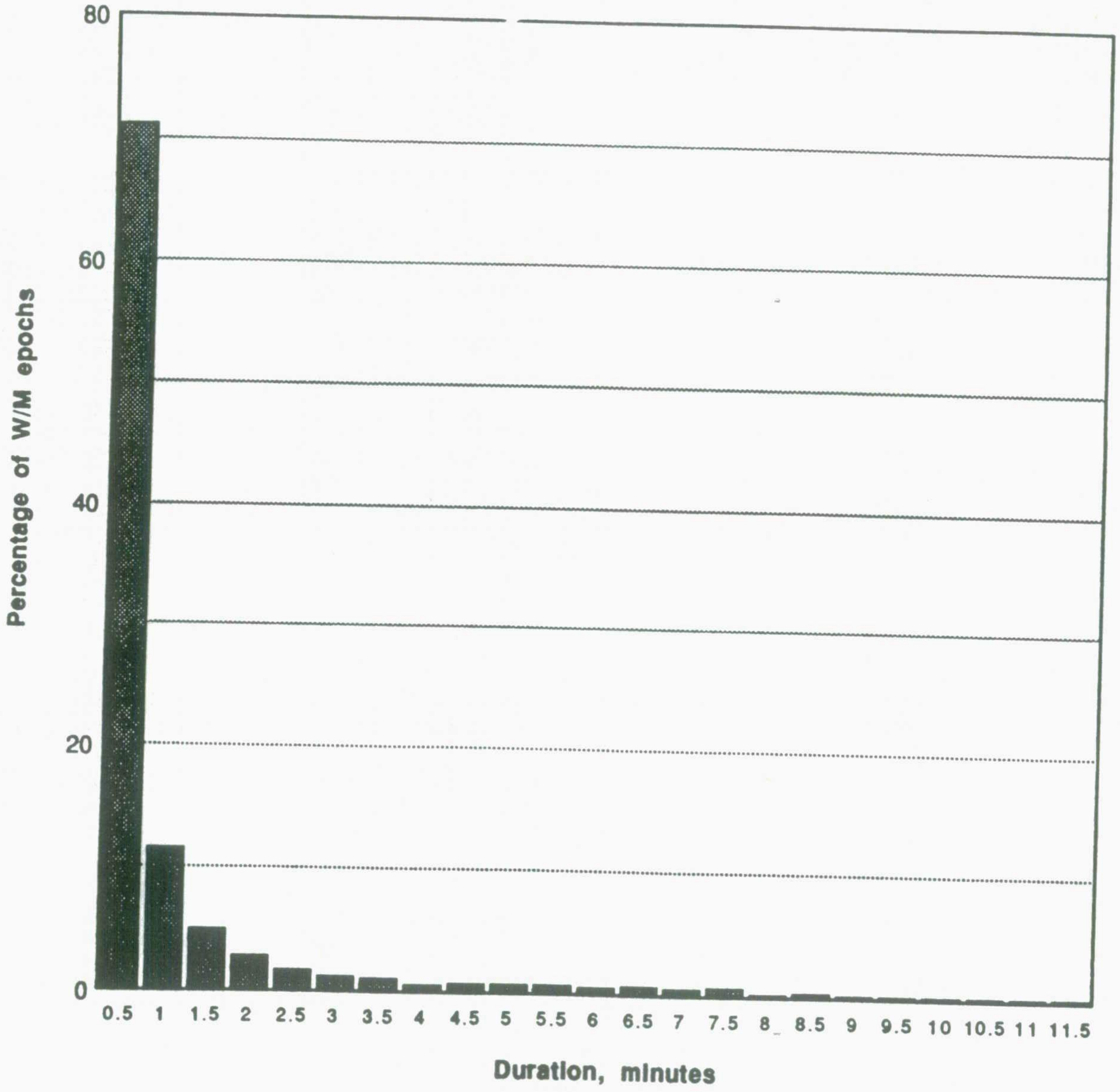


Figure 44 - Distribution of sleep arousal rate, all causes, all subjects
(disturbed epochs as a percentage of total epochs)

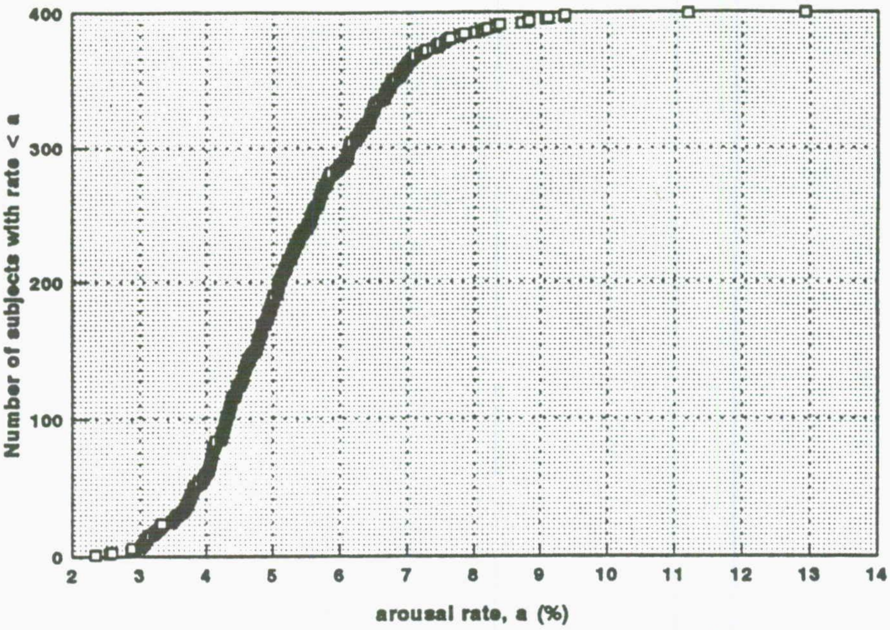
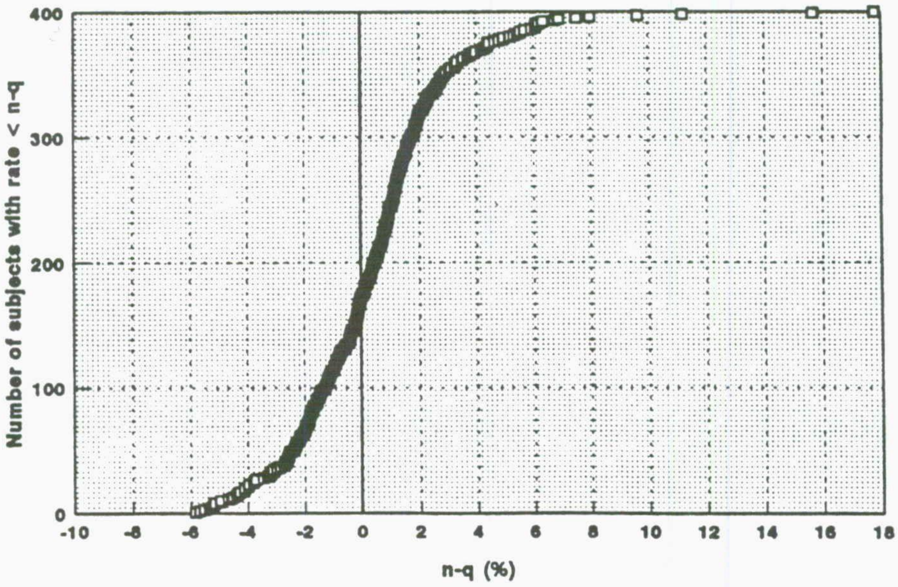


Figure 45 - Distribution of n-q: all subjects, all ANEs
n = disturbed ANE epochs as percentage of all ANE epochs
q = disturbed quiet epochs as a percentage of all quiet epochs



n-q is a crude estimate of aircraft noise induced arousal rate

Figure 46 - Distributions of n-q: ANEs with Lmax greater than or equal to 80 dBA
(a) for N>0, (b) N>100; N = total number of ANEs experienced

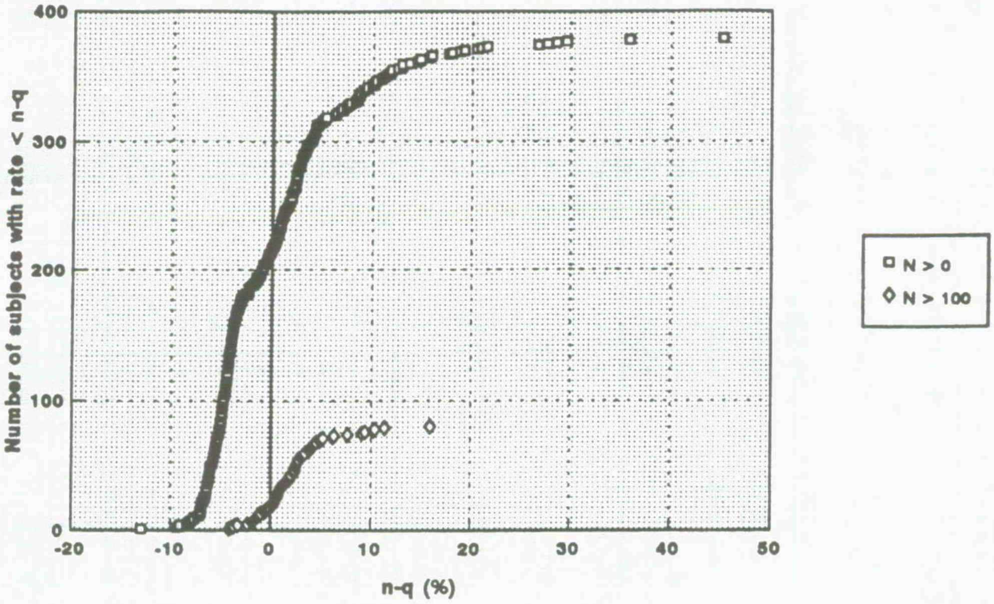


Figure 47 - Percentage distribution of n-q: ANEs with Lmax greater than or equal to 80 dBA
(a) for N>0, (b) N>100; N = total number of ANEs experienced

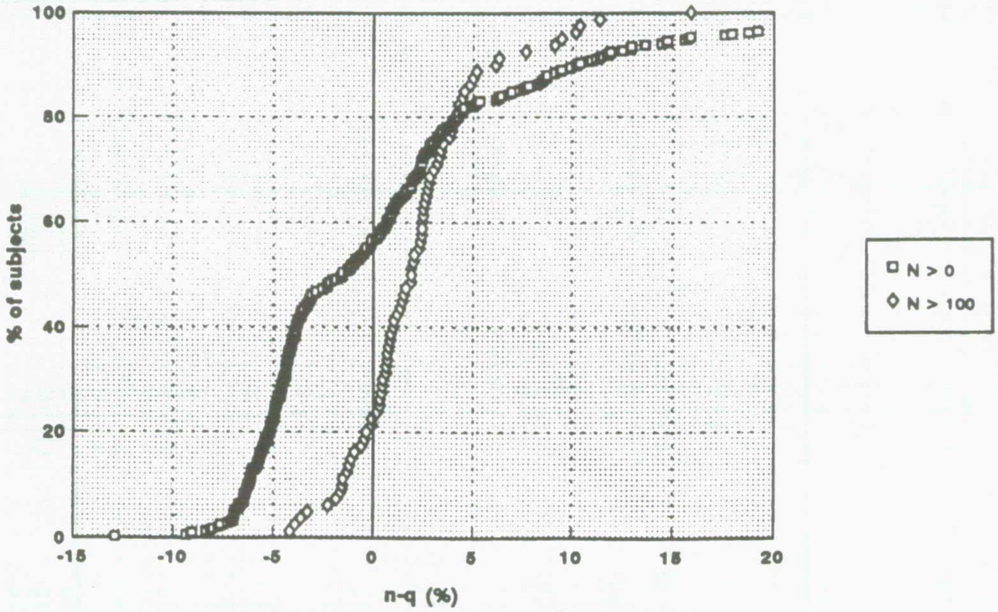
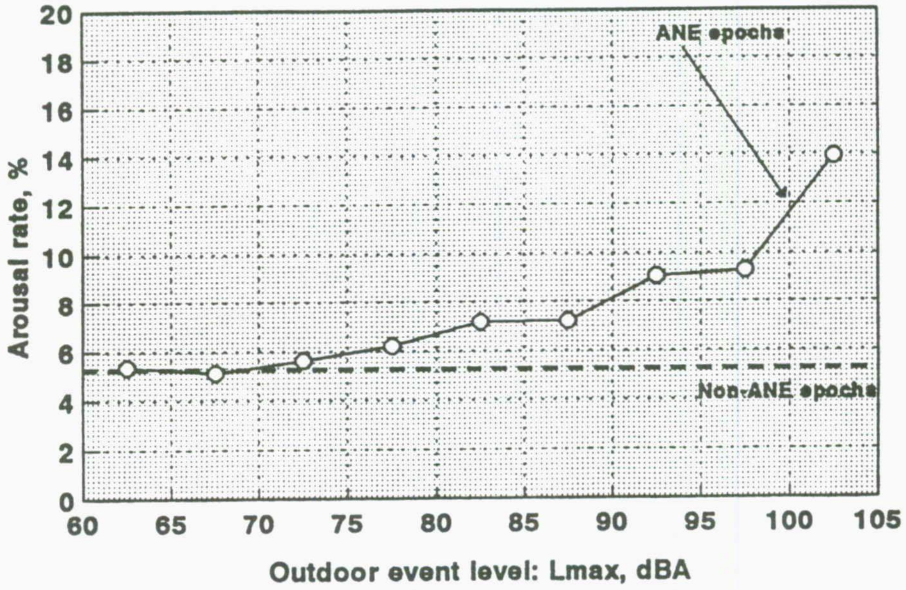


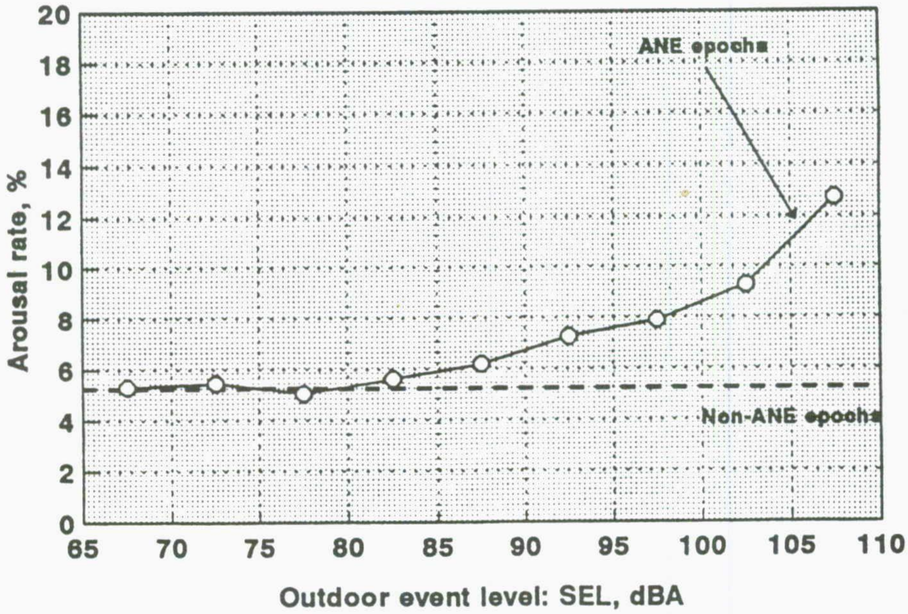
Figure 48 - Variation of unadjusted arousal rate with ANE level

(a) Lmax



jopfig48

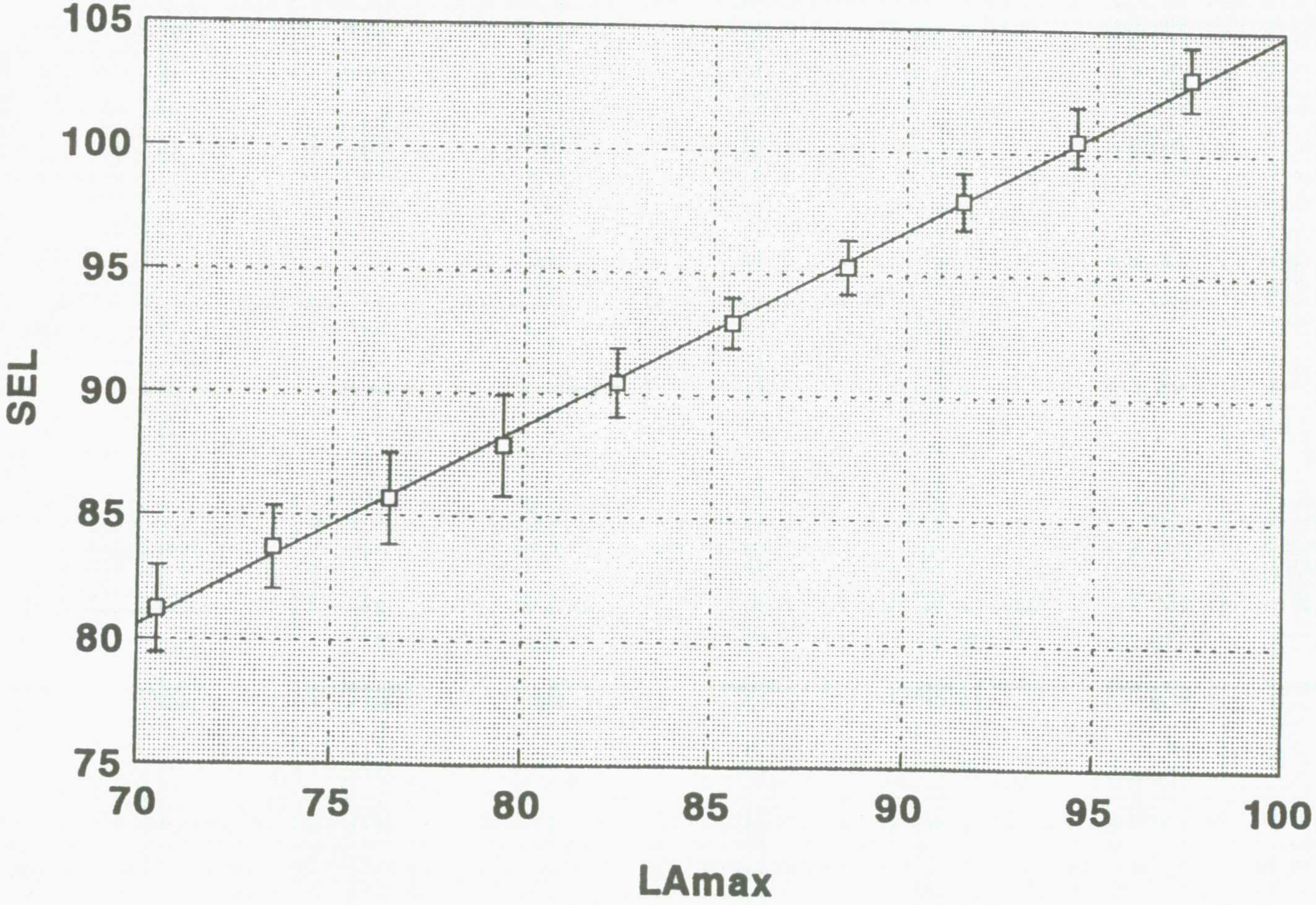
(b) SEL



jopfig48b

Figure 49 - Relationship between measured SEL and Lmax values: all ANE events

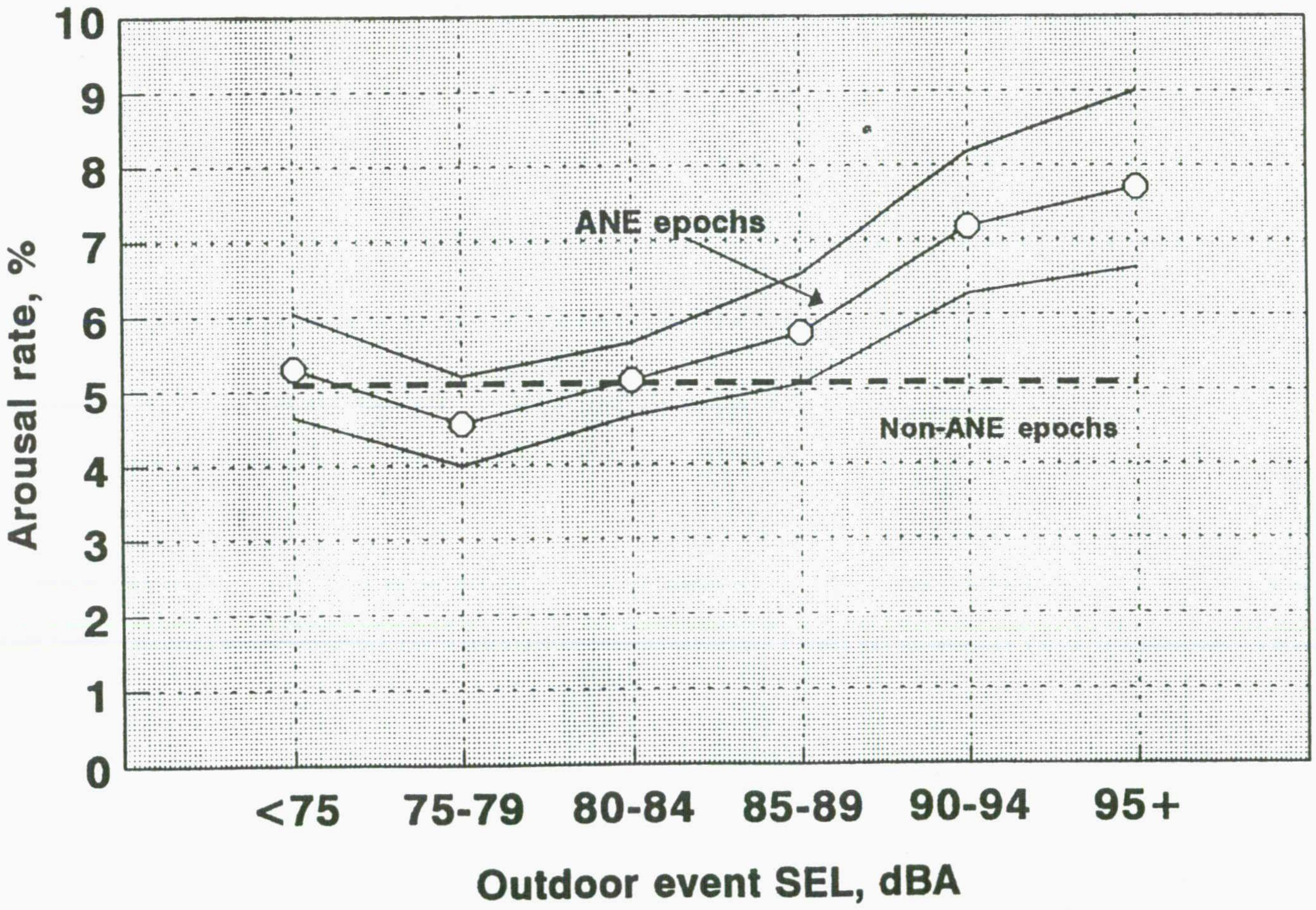
(confidence interval is + or - 1 standard deviation)



jopfig48

Figure 50 - Relationship between average sleep disturbance and aircraft noise level (showing 95% prediction interval)

Estimates controlled for the effects of individual arousability



jopfig50

Figure 51 - Variation of sleep arousal rates with subjects' age and sex

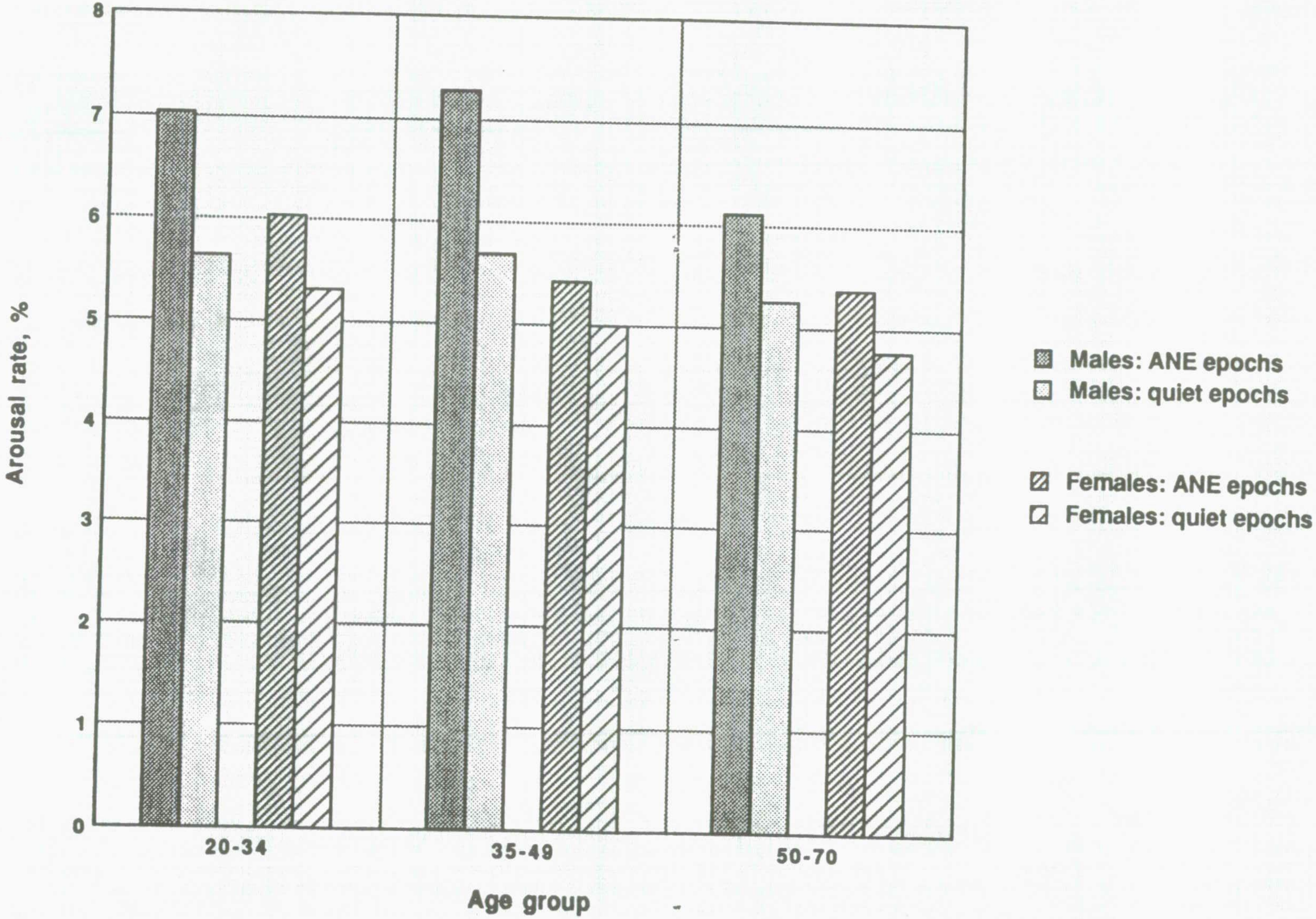
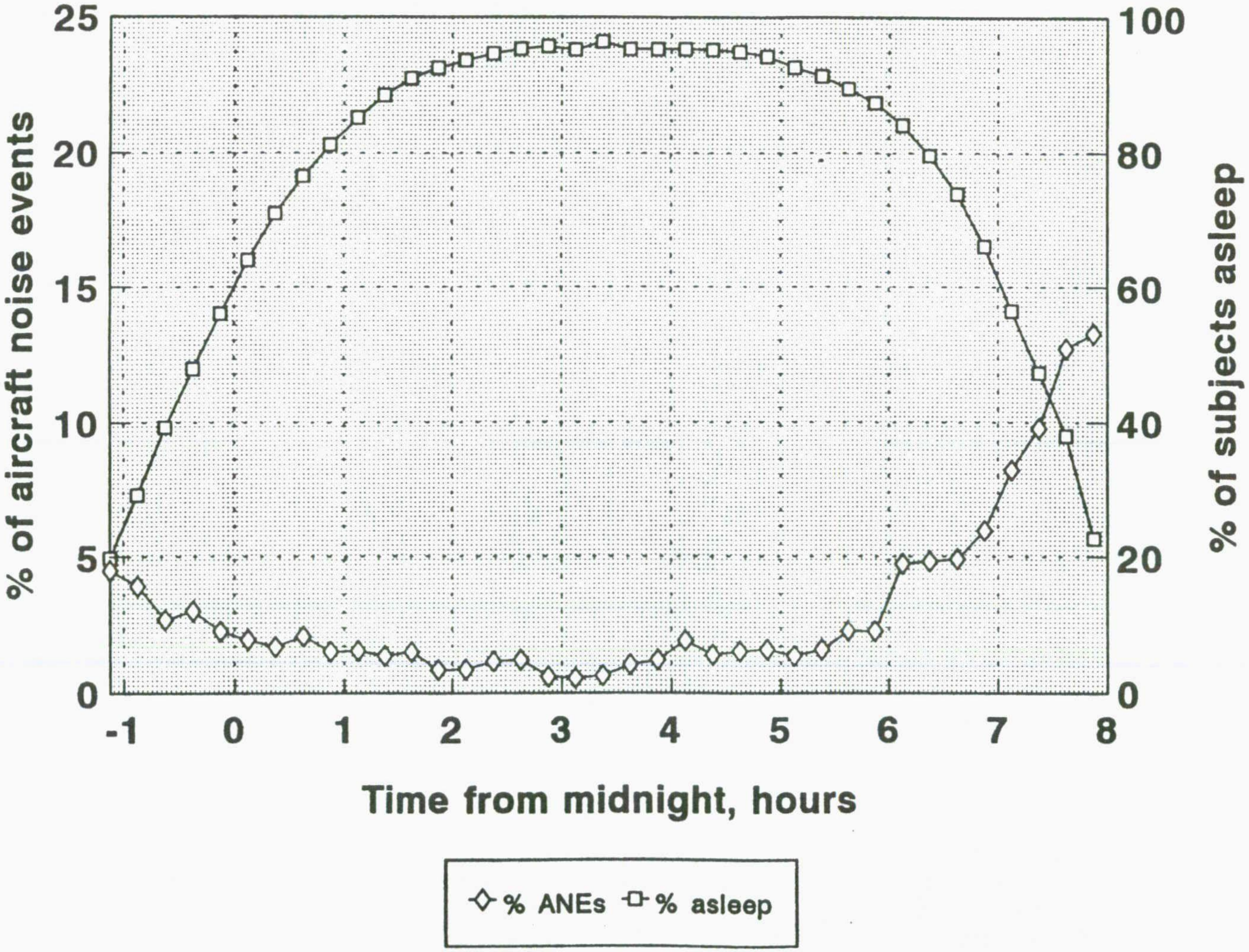
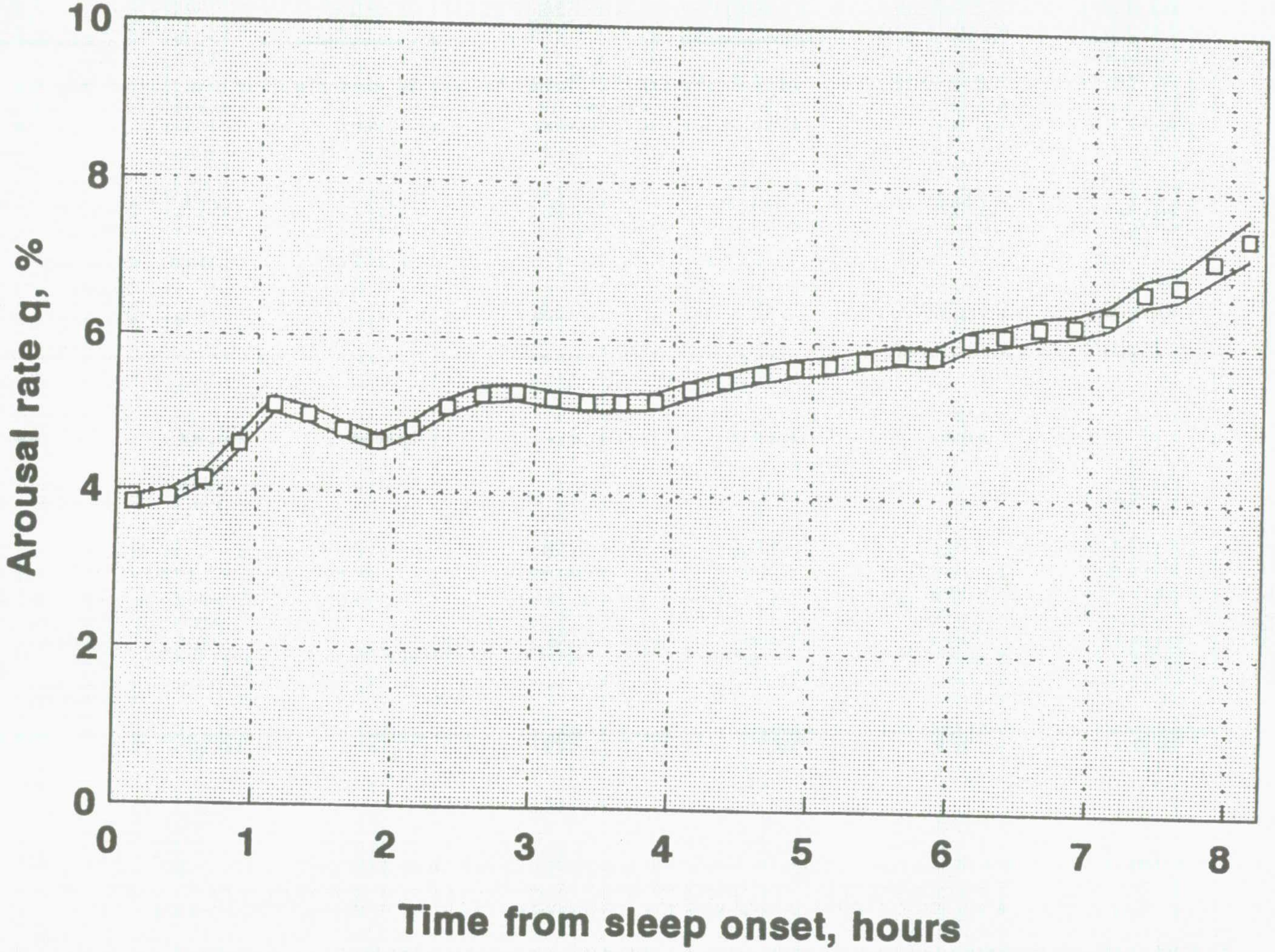


Figure 52 - Nighttime distribution of sleeping subjects and aircraft noise (by 15 minute intervals)



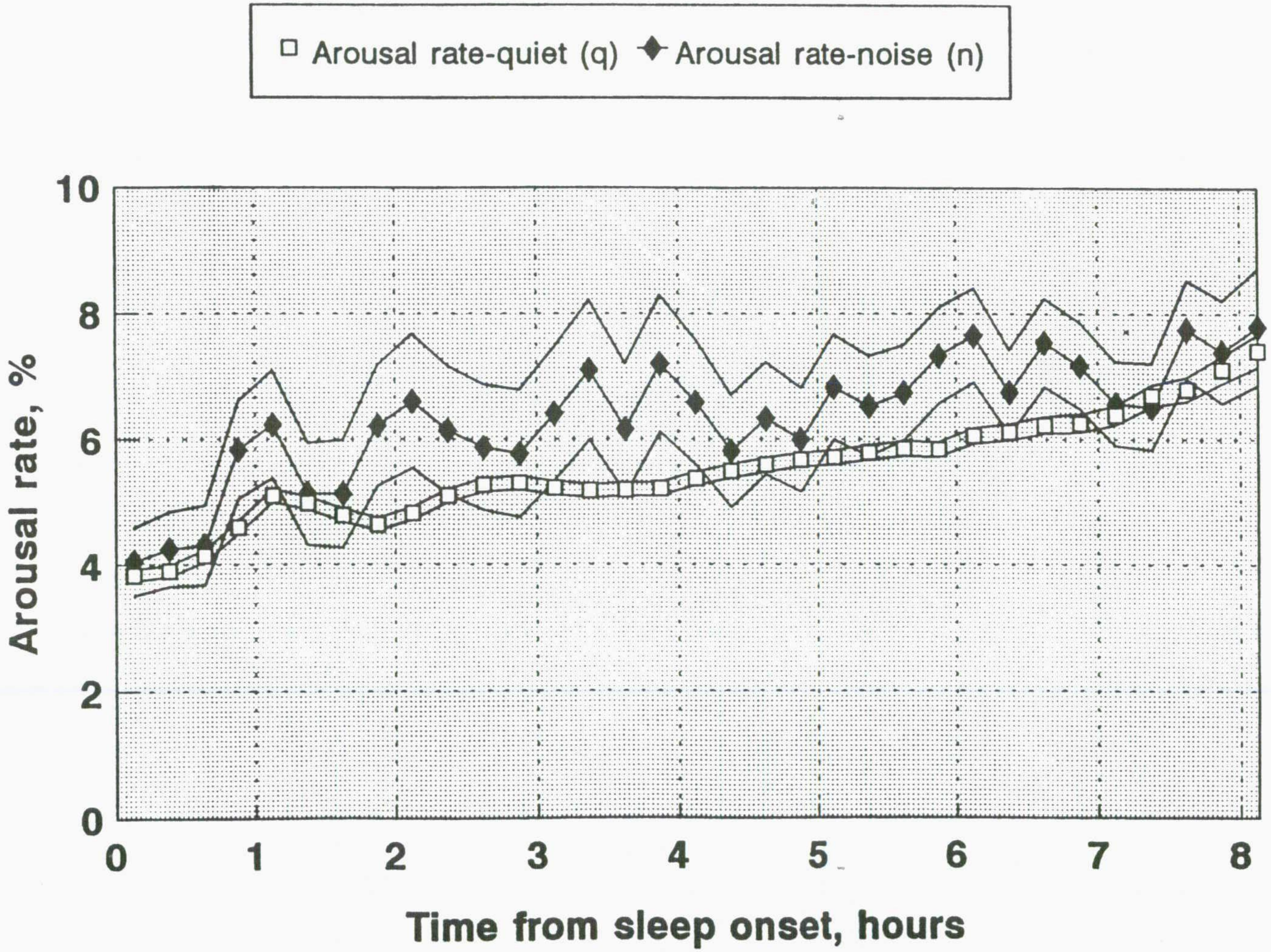
jopfig52

Figure 53 - Arousal rate in 'quiet' epochs: variation with time of sleep (showing very approximate 95% confidence intervals)



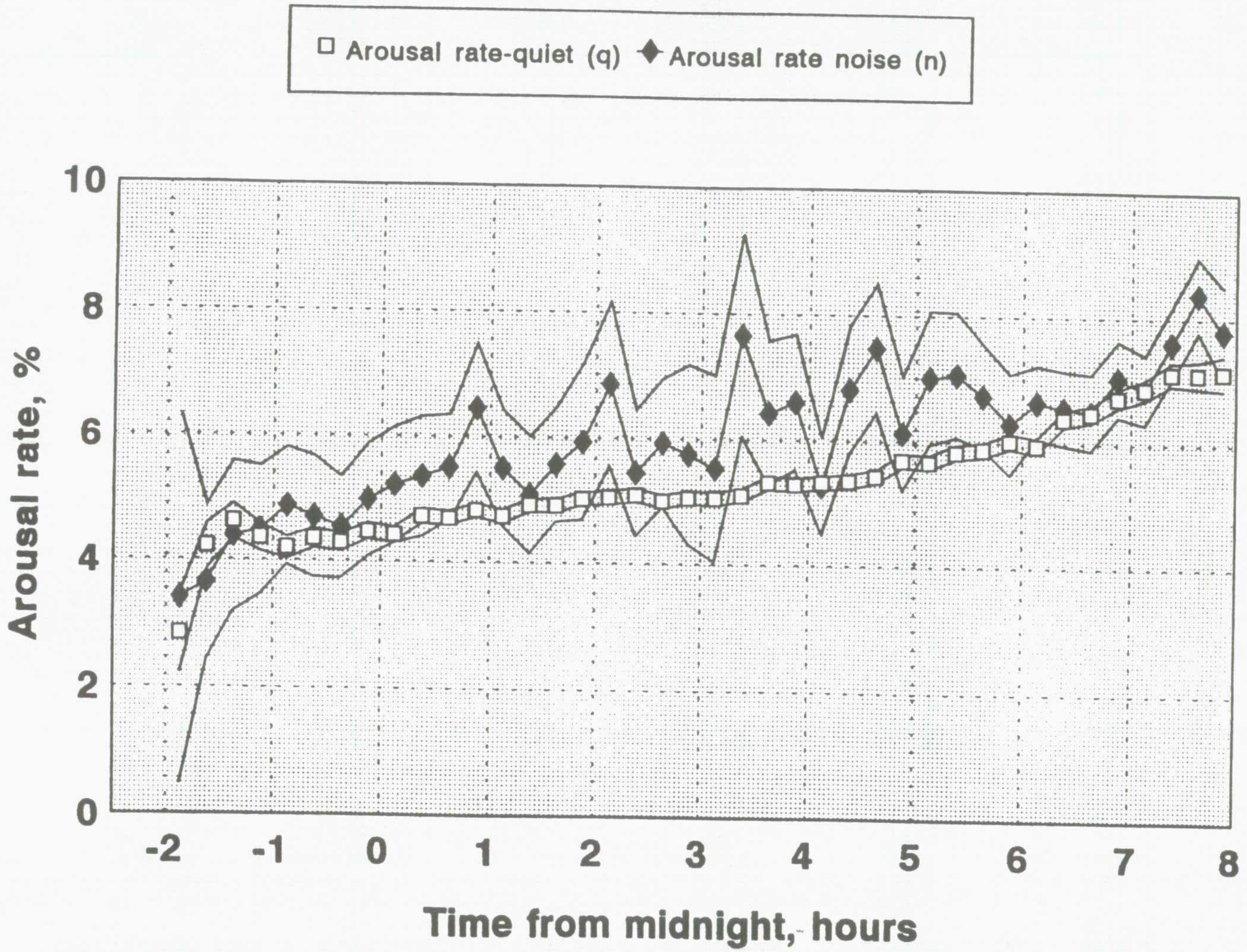
jopfig50

Figure 54 - Arousal rates in ANE (noise) epochs and 'quiet' epochs: variation with time of sleep (showing very approximate 95% confidence intervals)



jopfig51

Figure 55 - Arousal rates in ANE (noise) epochs: variation with time of night (showing very approximate 95% confidence limits)



jopfig52

Figure 56 - Arousal rates with different window conditions

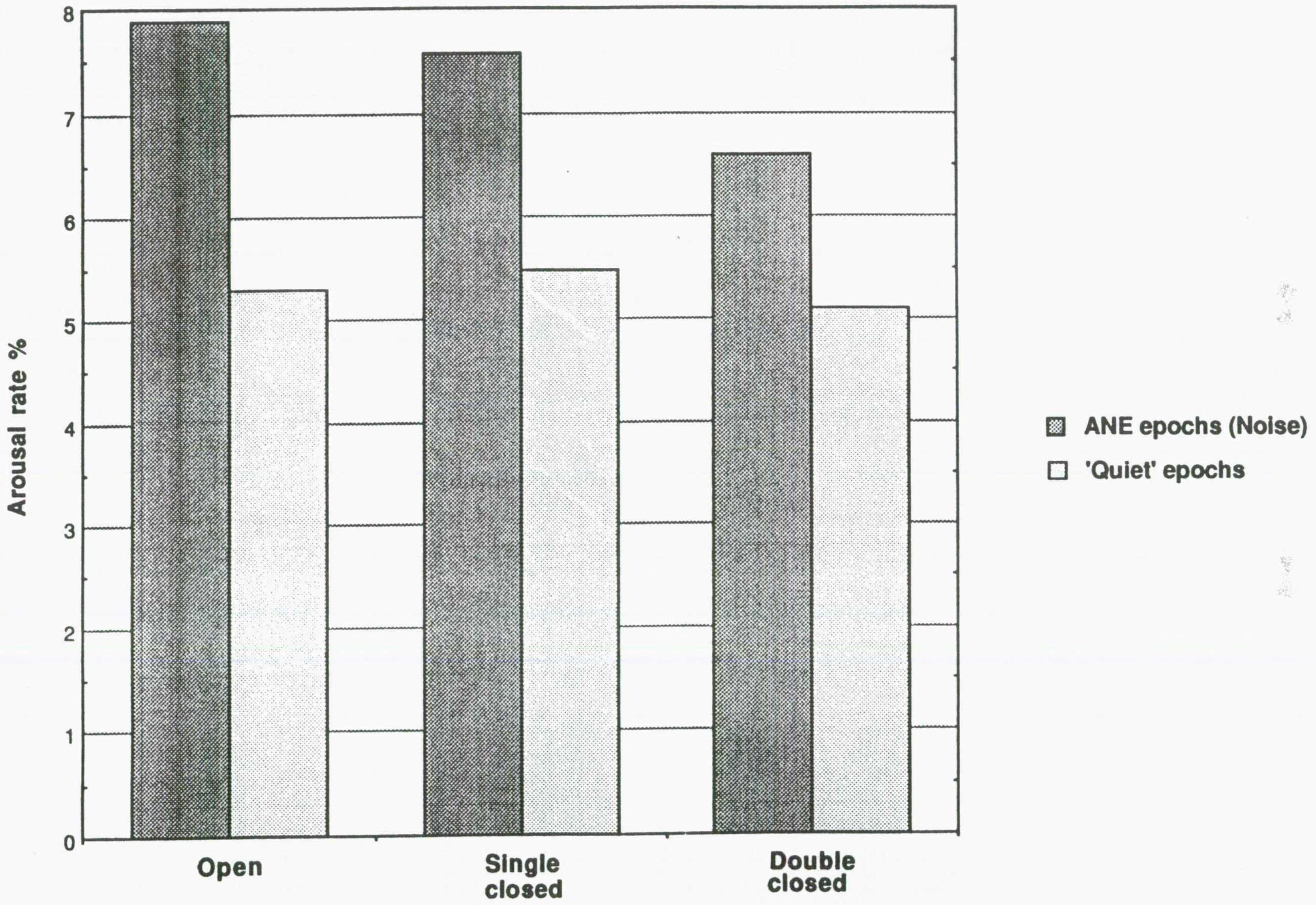


Figure 57 - Causes of awakenings reported in sleep logs

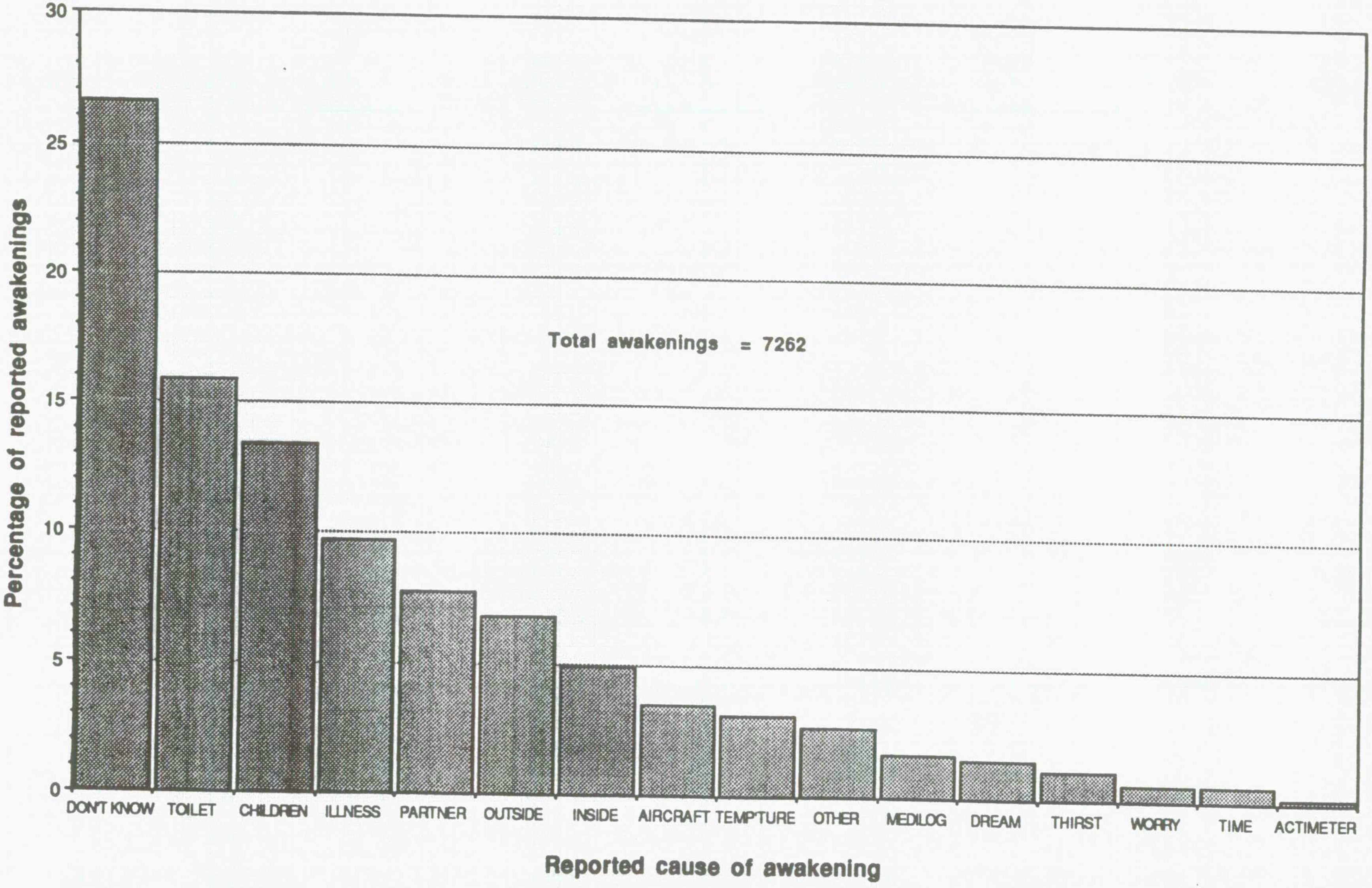
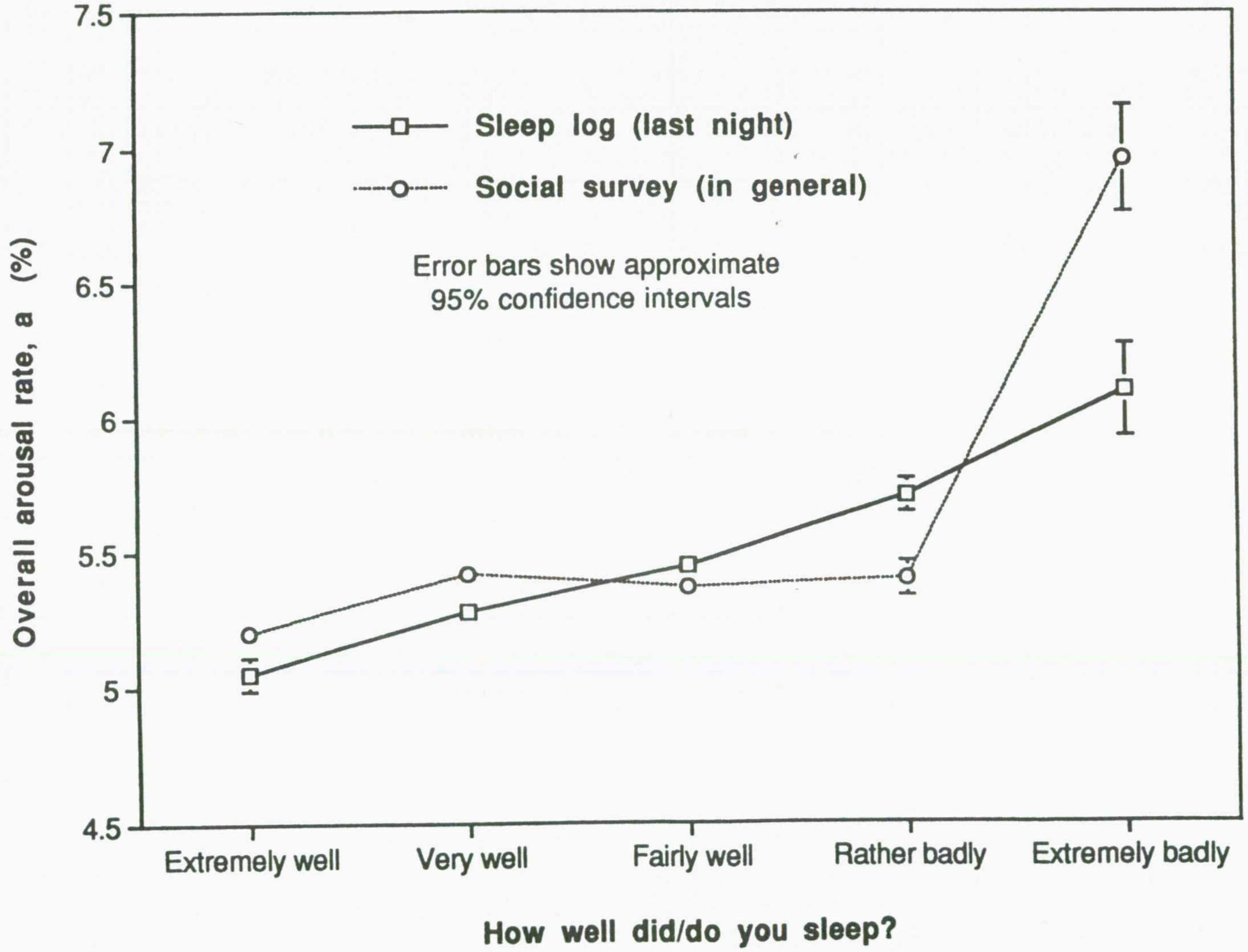


Figure 58 - Relationships between measured sleep disturbance and reported sleep quality



APPENDIX A

SIMPLE STATISTICAL TEST OF THE DIFFERENCES BETWEEN MEASURED DISTURBANCE RATES

Method

A simple index of sleep disturbance based on the actimetric measurements is the disturbance proportion:

$$g = \frac{\text{actual number of disturbances}}{\text{maximum possible number of disturbances}}$$

$$= \frac{\text{number of disturbed epochs}}{\text{total number of epochs}}$$

For large samples (size N) drawn from a normally distributed population, the associated 95% confidence limits would be

$$\pm 1.96\sqrt{\frac{g(1-g)}{N}}$$

Two proportion statistics quoted frequently in the text are:

$$n = \frac{\text{number of disturbed ANE epochs}}{\text{total number of ANE epochs}}$$

$$q = \frac{\text{number of disturbed non-ANE epochs}}{\text{total number of non-ANE epochs}}$$

An ANE epoch is one which includes an aircraft noise event - timed at the instant when the sound level reaches its maximum value, Lmax. The summations may be carried out upon particular sub-sets of the actimetry data, eg divided by time of night, subject age group etc.

In many analyses, it is desirable to determine whether the difference between two disturbance ratios, eg n and q, is statistically significant, and not merely the result of chance. Again assuming that normal probability theory is valid, a direct approach is to calculate the z-statistic for two proportions g_1 and g_2:

$$z = \frac{g_1 - g_2}{\sigma\sqrt{\frac{N_1 + N_2}{N_1 N_2}}}$$

where $\sigma = \sqrt{g(1 - g)}$

$$g = \frac{N_1 g_1 + N_2 g_2}{N_1 + N_2}$$

and N_1, N_2 = number of observations contributing to g_1, g_2.

This statistic can be tested in the standard way using normal probability tables. However, when applied to the actimetry results, this test is flawed because it violates of one of the essential assumptions of normal probability theory, ie that all observations are independent of each other. Although individual subjects are independent, the arousal measurements are not independent because (a) in comparisons the number of ANE events effectively determines the number of non-ANE events and (b) different observations from a single subject are 'repeated measures' subject to serial correlation effects (see text, Section 6.3). Similar limitations apply to the simple expression above for the 95% confidence interval associated with g.

The flaw should not be too serious when the sample sizes are large enough to detect differences with levels of significance $p << 0.05$. However, caution is required when $p > 0.01$, say. This means that the test can be useful for assessing broad trends in large data sets, but unreliable for testing differences between means of small sub-sets of data.

A more elaborate procedure which overcomes the flaws of this test is outlined in Appendix B.

APPENDIX B

LOGISTIC REGRESSION ANALYSIS

Method

Ordinary linear regression is widely used to determine mathematical relationships between a dependent variable y and a number of independent variables x_k (ie x_1, x_2, x_3, ...). For example, in social survey studies of aircraft noise effects, the dependent variable y has been defined as annoyance (Guttman Annoyance Score) while the independent variables x_k included noise exposure level, L_{eq} and numerous individual variables such as age, sex, years of residence, fear of aircraft crashing etc.

Linear regression gives a relationship of the form

$$y' = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots$$

$$= b_0 + \sum_k b_kx_k \qquad \ldots (1)$$

where the *regression coefficients* b_k define how strongly the different factors x_k influence the dependent variable y. The coefficients are computed by fitting the equation to a suitably large sample of experimental data (ie a number of cases, each consisting of one set of values: y, x_k). The method used is that of *ordinary least squares* (OLS); ie the best fitting equation is that which minimises the variance of the residual error (see below). Practical regression methods give estimates of the coefficients together with their confidence intervals (and whether or not the associated variable is a significant contributor).

Once the coefficients have been determined, the *regression equation* (1) can be used to calculate the expected value y' of the dependent variable corresponding to any particular set of independent variables x_k. The difference between the actual measured value y and the expected value y' is the residual error e. Thus

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + e \qquad \ldots (2)$$

This error may arise as a result of inadequate measurement and/or because important explanatory variables have been overlooked or simply as a result of random fluctuations. Often, in practical regression analysis, detailed examination of the residual errors point to other independent variables which should be considered (success is very much a question of trial and error).

Equation (1) can be used to express the relationship between y and any of the x variables, holding the others constant. But because the model is linear, this relationship can only be a straight line. This model is therefore really only appropriate if a straight line outcome is expected within the range of interest. It is unlikely to be suitable where the dependent variable is a *proportion* (eg the percentage of the sample exhibiting a certain characteristic) and where that proportion is small, eg less than 10%. In this case, a sigmoidal curve is to be preferred, ie an elongated S-curve.

This is the principle of *logistic regression* which is based on a relationship of the form:

$$\mathrm{Ln}\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + e$$

$$= b_0 + \sum_k b_kx_k + \ldots + e \qquad \ldots (3)$$

The relationship between the proportion p and any one of the x_k (the others held constant) is an S-curve which is asymptotic to 0% at low x-values and 100% at high x-values. Many studies of public reaction to noise have yielded 'dose-response' curves having such a shape. Given that the variable p is restricted to the range 0 to 100%, it is much more logical to fit an empirical logistic curve to experimental data than a linear one (which extends beyond the 0% to 100% range). Also, the necessary analytical tools and statistical tests are well developed for this particular function. It should be noted that the ratio $\frac{p}{1-p}$ is the 'odds', ie the probability of a positive response divided by the probability of a negative response. The function $Ln(\frac{p}{1-p})$, sometimes written Logit(p), is referred to as the 'log odds'.

Because, in any single observation, p can only take one of two values, (ie 1 = characteristic present or positive response, 0 = characteristic absent or negative response) the method of OLS is not suitable for fitting Equation (3) to measured data because the residual error e is usually large. Instead, a method is used which maximises L, the likelihood of observing the measured data sample under the assumption that the logistic model is correct. Theoretically, all possible combinations of the b_k are considered - the one is picked which maximises L. Thus the method is called *maximum likelihood.*

In linear regression, the coefficient b_k defines the change in y that accompanies unit change in x_k. In logistic regression, the coefficient b_k defines the change in the *log odds* that accompanies a unit change in x_k. But the physical significance of log odds is difficult to grasp. Fortunately, it is readily shown that the exponential of the coefficient b_k is an *odds ratio* - which may be interpreted as the probability that a unit increase in x_k will cause a change from a negative to a positive response.

The convenience of this technique for analysing the sleep disturbance data is immediately apparent. For any particular epoch, the x_k represent all the factors that may be expected to influence whether or not sleep will be aroused. They would include, for example, aircraft noise level, age, sex, time of night, window state, etc. The proportion p is then the probability that an individual will be aroused in that epoch. If the regression coefficients b_k can be determined, then the influence of any variable x_k can be quantified in terms of its odds ratio - the probability that it will cause sleep disturbance.

However, there is a complication. Generally, regression methods, like most statistical tests, are only valid if each and every observation is independent of all others (in the language of statistics, the sample is randomly drawn from a much larger population). Whilst it may be reasonable to assume that all subjects behave independently, the same cannot be assumed of different observations ('repeated measures') from the same subject. Because a subject's sleep state at any particular time must depend to some extent on his/her previous sleep state, the observations are said to be serially correlated.

This difficulty has been overcome by using a special *random effects* version of logistic regression. In this model the residual error in Equation 3 is split into two parts:-

$$e = \varepsilon + u_i$$

where ε is the 'true' random error and u_i is a systematic mean error for the i^{th} subject which does not vary with time but does differ randomly between subjects. This extra random error term u_j accounts for the serial correlation between the 'within subjects' observations.

Theoretically, the u_i could be determined broadly by introducing a dummy variable for each subject i into Equation 3 and calculating its coefficent in the same way as the others. But since the number of dummy variables would have to equal the number of subjects minus one, the analysis would, in practice, be intractable. Instead, it is assumed that the u_i, although essentially unknown, follow a normal distribution with standard deviation σ, ie

$$u_i = \mathbf{N}(0,\sigma), \text{ or } u_i = \sigma\delta_i \text{ where } \delta_i = \mathbf{N}(0,1)$$

N is a normal distribution with zero mean, standard deviation σ. Adding this extra random error term in Equation 3, we get:-

$$\mathrm{Ln}\left(\frac{p}{1-p}\right) = b_0 + \sum_k b_k x_k + \ldots + \varepsilon + \sigma\delta_i \qquad \ldots (4)$$

Under the assumption of a normally distributed variable δ_i, the maximum likelihood analysis yields an estimate of its standard deviation σ. However, it does not give the values of δ_i for individual subjects; once determined, Equation 4 is used (with $\varepsilon = 0$) simply to estimate 'average' odds ratios for all subjects (with respect to the distribution for δ_i). The estimate for σ, which expresses the extent of 'within subject' serial correlation, defines the expected range of proportions p across the different subjects. For example, the range $\pm 2\sigma$, which encompasses around 95% of all subjects, defines upper and lower limits for p, corresponding to the most and least sensitive people respectively.

Analysis

Two analyses have so far been carried out using this approach.

The first, a preliminary sift, involved 120 cases each comprising all epochs from 50 subjects. Five variables were included: aircraft noise, time of night, age, sex and individual arousability (the random error). Sleep disturbance was significantly related to time of night on all but 4 nights, to subject's age on 8 nights, to subject's sex on 24 nights, and to aircraft noise on 14 nights (11 of those being Manchester/HGN cases). Because the individual data sets were small, the statistical power of this analysis was limited but it gave clear indications (a) that time of night is a most important factor and (b) that aircraft noise effects were likely to be small except at the Manchester sites.

The second analysis covered all subjects and all nights but was limited, again to restrict the amount of computation, (a) to the two Manchester sites HGN and EDG and (b) to the three time periods 0100-0130, 0300-0330 and 0500-0530. The same variables were included although ANE levels Lmax were divided into 3 categories and several cases were run using different category boundaries. The results are given in Table B1 for the case with noise categories Lmax < 75 dBA (including 'quiet'), 75-79dBA and ≥80 dBA. The entries in Table B1 show the effects of the various factors on the expected probabilities of disturbance *when the other factors are controlled* (ie 'averaged out').

Noise epochs for ANEs in the Lmax range 60-74 dBA were combined with quiet epochs after no significant differences were found between the separate disturbance rates in these categories. Nor are the differences between the <75 dBA and 75-79 dBA categories significant at the 5% level. However the differences between those above 80dBA and those below 75dBA Lmax were significant. For the subject of average sensitivity ($\delta_i = 0$), this difference is 8.36-5.04 = 3.32%. This may be interpreted as the probability that an ANE with Lmax ≥ 80dBA will disturb the average subject. Put another way, since 8.36/5.04 = 1.66, there is a 66% greater probability of being disturbed during a noise epoch than during a quiet one.

'Noise sensitivity' defines the $\pm 2\sigma$ range of δ_i described above which encompasses approximately 95% of the expected distribution of individual arousability; ie 'high' applies to the 2.5% most sensitive, 'low' to the 2.5% least sensitive. Table B1 indicates that the former are more than 2.5 times more susceptible to disturbance than the latter in all three noise categories. Equally, they are more likely to be disturbed by aircraft noise; the difference between their disturbance probabilities for high noise epochs (Lmax ≥ 80 dBA) and quiet epochs being 13.20-8.12=5.08%. However, the same people are more likely to be disturbed during quiet epochs anyway and, since 13.20/5.08 = 1.63, they are 63% more likely to be disturbed in noise than in quiet. This is very similar to the result of 66% for people of average sensitivity.

Table B1 indicates that sleep disturbance is only weakly related to the subjects' age and sex. However, time of night has a strong influence; expected disturbance rates between 0500 and 0530 being 26% greater than between 0100 and 0130.

TABLE B1 RESULTS OF INITIAL LRA ANALYSIS: EXPECTED PROBABILITY OF DISTURBANCE (%) CONTROLLED FOR AVERAGE PERSON/CONDITIONS

Time	**Age, years**		
	20-34	**35-49**	**≥ 50**
0100-0130	7.79	7.66	7.58
0300-0330	8.60	8.76	7.39
0500-0530	10.10	9.68	8.13

Sex	**Lmax, dBA**		
	< 75	**75-79**	**≥ 80**
Male	6.71	7.06	10.10
Female	6.12	6.44	10.10

Lmax	**Sensitivity**		
	High	**Average**	**Low**
< 75	8.12	5.04	3.10
75-79	8.52	5.31	3.27
≥ 80	13.20	8.36	5.21

Lmax	**Age, years**		
	20-34	**35-49**	**≥ 50**
< 75	5.30	5.21	4.99
75-79	5.57	5.48	5.24
≥ 80	8.76	8.62	8.25

APPENDIX C

THE WILKINSON-DIAMOND ANALYSIS

Method

The random effects model described in Appendix B permits efficient estimation of the probability of being disturbed in any epoch by controlling for the problem of repeated measurements - the fact that each individual contributes many observations to the data. This model is used to distinguish between 'quiet' and 'noisy' epochs through the use of a dummy (or indicator variable). In this way it is possible to establish whether there is a significant difference between the probabilities of being disturbed in noisy and quiet epochs.

It is very likely that the probability of being disturbed by an aircraft noise event will depend on the extent to which the individual has been disturbed in the quiet period before that event. A strategy to control for this was proposed independently by Dr R Wilkinson and Professor Diamond and is described below. In the text, this has been termed the Wilkinson-Diamond or W-D analysis.

In this analysis, the data set is restricted to noise events alone and the extent to which the individual is disturbed in the quiet epochs may be controlled in two ways:

(i) According to whether the individual is disturbed in a random epoch. Here, an epoch in the quiet period before the ANE is chosen at random and a dummy variable is formed to indicate whether or not the individual is disturbed.

(ii) According to the rate of disturbance in the quiet period. This is a continuous variable formed from the ratio of the number of disturbances in the quiet period to the length of the quiet period (in epochs).

Variables representing each of these models were formed and were considered separately in the analysis (they could not be considered contemporaneously as they are essentially measuring the same thing). There is also a random effects model which controls for the fact that each individual experiences all the noise events at a site. The model is then

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \sum_k b_{ijk}x_{ijk} + c_{ij}WD_{ij} + u_i + \varepsilon_{ij}$$

where

$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right)$	is a function of the probability of being disturbed in a noise epoch
$\sum_k b_{ijk}x_{ijk}$	is the sum of the products of the covariates and their estimated regression coefficients. Note that x_{ijk} refers to the kth covariate for the jth epoch for the ith individual.
WD_{ij}	is the Wilkinson-Diamond disturbance variable for the quiet period before the jth epoch for the ith individual
c_{ij}	is the estimated regression coefficient for the Wilkinson-Diamond disturbance variable
u_i	is the random effect for the ith individual across all epochs
ε_{ij}	is an overall error term

These models cannot be fitted straightforwardly and require specialised software. For these analyses, adaptations were made to the EGRET statistical package. After much experimentation, it was concluded that approach (ii) described the data most effectively.

Analysis

Because this application of LRA is novel, progress has been cautious and the work so far completed has involved two main stages.

First, in order not to impose excessive demands upon available computational facilities, data from the two Manchester sites (HGN and EDG) were analysed as a separate data set. This case was chosen because, together, those two sites provided the greatest range of ANE levels, distributed evenly over the night (see text, Fig 6). The analysis was restricted to the period 2330 to 0530. More than 40 variables were included; these are listed in Table C1. They were selected as potential 'predictors' of sleep disturbance, ie as physical, psychological and personal factors which may influence the way people respond to aircraft noise at night.

In fact, only four of these variables were found to have statistically significant effects ($p<.05$) upon ANE epoch arousals. These were:

- noise level of the ANE,
- the arousal rate in the interval since the last ANE,
- the time of night, and
- individual arousability (susceptibility to sleep disturbance).

The last of these is the 'random effect', a systematic variation in arousability from person to person (over and above any random 'measurement' errors - see Appendix B) which cannot be explained by variations in the potential predictor variables - although it is of course possible that other significant factors may have been overlooked.

Table C2 summarises the effects of the four significant variables. These are expressed in terms of the relative chance of being aroused. Several levels, or categories, are assigned to each variable; eg 'pre-ANE arousal' is defined as 'none, low, medium or high'. These categories relate to the incidence of arousals during the 'quiet' interval since the previous ANE, ie the number of disturbed epochs expressed as a proportion, α, of the total number of epochs. The probability of arousal associated with any particular category of a variable is defined in relation to that of the reference category. Thus, a person who experiences a medium arousal rate has a probability of being aroused which is 130% of the 'reference' individual's, ie one with no arousals during the quiet period. In other words, he or she would be disturbed 30% more often.

In the second main analysis, data from all sites for the period 2300-0530, a total of more than 31,000 subject-ANEs, was analysed together. The following variables were included:

Variable	*Categories*
Time of night	4 (2330-0100, 0100-0230, 0230-0400, 0400-0530)
ANE level (SEL, dBA)*	6 (<74, 75-79, 80-84, 85-89, 90-94, ≥95)
Subject's age, years	3 (20-34, 35-49, ≥50)
Subject's sex	2
Time since last event, epochs	8 (10-14, 15-19, 20-24, 25-29, 30-44, 45-59, 60-89.>90)
Arousal rate since last event (%)	4 (0,1-4, 5-9, ≥10)
Site	8
Random effects (individual arousability)	Assumed normally distributed

Table C3 summarises the results. Because the data set here was very much larger than in the initial analysis, the statistical power was greater, thus allowing more noise level categories to

---

* These were analysed in both categorical (5dB bands) and continuous form

Table C3 summarises the results. Because the data set here was very much larger than in the initial analysis, the statistical power was greater, thus allowing more noise level categories to be used. Also, the 95% confidence interval is tabulated along with the estimated probabilites of arousal. The table shows that subjects' sex also emerged as a variable of significance. However, ANE disturbance showed no significant dependence upon site, subject's age, and the duration of the the preceding 'quiet' interval.

TABLE C1 VARIABLES INCLUDED IN INITIAL W-D ANALYSIS

Case variables:	
Site	
Night	
Subject	
ANE variables:	
Disturbed in noise epoch	
Disturbed in random epoch	
Number of disturbances since last ANE	
Time of night	
ANE level, SEL	
ANE level, Lmax	
Time since last ANE	
General subject variables from social survey questionnaire:	
Subject	
Age group	
Sex	
Occupational group	
ANGEN	Very much annoyed by aircraft noise
HEARNT	Very much annoyed by aircraft noise at night
ANWK	Awakened at night by aircraft noise
COMPLAIN	Has made a formal complaint about noise
DIFFGET	Has difficulty getting to sleep
WOKENREG	Regularly woken up once asleep
GETBACK	Has difficulty getting back to sleep once woken
HOWSLEEP	How well/ badly sleeps at night
SLPART	Sleeps with partner
HOWFEEL	Refreshed or sleepy in morning
ACMAINS	Aircraft noise main cause of difficulty getting to sleep
ACMAINW	Aircraft noise main cause of awakening during night
General subject variables from subject selection interview:	
Q8	Number of cups of coffee per day
Q22	Number of times awakens per night
Q24a	Wakes up at night, cause unknown
Q24h	Wakes up at night, noise is cause
Q25	Feels refreshed just after getting up
Q26	How feels 15 min after waking
Daily subject variables from sleep log:	
SLQ3	Windows open or shut
SLQ4	Secondary glazing open or shut
WOPEN	Open, single glazing shut, double glazing shut
SLQ7	Sleepy, refreshed 15 minutes after awakening
SLQ8	Slept well badly last night
General subject variables from debrief questionnaire:	
DB Q5	How much aircraft noise bothers after going to bed
DB Q7	Aircraft noise disturbs sleep
DB Q9	Aircraft noise causes wakens
DB Q13	Considers air transport dangerous
STA1	State-Trait anxiety score
BORTNER	Bortner personality score

TABLE C2

FACTORS AFFECTING THE CHANCE OF BEING AROUSED BY AN AIRCRAFT NOISE EVENT: MANCHESTER DATA

(2330-0530)

Variable	**Category**	**relative % chance of arousal***
Arousability	Low†	50
	Average	**100**
	High†	197
Noise level (Lmax, dBA)	**< 75**	**100**
	75-79	115
	≥ 80	146
Arousal rate in 'quiet' period since last ANE	**None**	**100**
	Low	138
	Medium	130
	High	149
Time of night	**2330-0100**	**100**
	0100-0230	127
	0230-0400	138
	0400-0530	136

* Chance of being aroused during the ANE epoch relative to that in reference category denoted in bold (with 100% chance)

† High/low arousability is 2 standard deviations above/below mean

TABLE C3

FACTORS AFFECTING THE CHANCE OF BEING AROUSED BY AN AIRCRAFT NOISE EVENT: ALL DATA
(2330-0530)

Variable	**Category**	**relative % chance of arousal*** (with 95% confidence interval)
Arousability	Low†	56 (52-60)
	Average	**100**
	High†	178 (166-192)
Sex	**Male**	**100**
	Female	87 (77-100)
Noise level (SEL, dBA)	**< 75**	**100**
	75-79	82 (67-100)
	80 -84	92 (79-108)
	85-89	104 (87-125)
	90-95	129 (106-156)
	≥ 95	141 (115-173)
Arousal rate in 'quiet' period since last ANE	**None**	**100**
	Low	122 (106-141)
	Medium	139 (121-160)
	High	164 (143-189)
Time of night	**2330-0100**	**100**
	0100-0230	120 (104-138)
	0230-0400	133 (115-154)
	0400-0530	137 (119-157)

* Chance of being aroused during the ANE epoch relative to that in reference category (= bold entry with 100% chance). Parameter estimate for reference category is constrained to a fixed value - confidence interval cannot be estimated.

† High/low arousability is 2 standard deviations below mean

APPENDIX D

PREDICTION INTERVALS FOR THE PROPORTION DISTURBED BY ANEs

In order to calculate reliable prediction (confidence) intervals for the in-noise arousal rate n (the proportion of people aroused by an ANE) it is necessary to control for individual arousability. This can be done using random effects logistic regression analysis (Appendix B) using the model:

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = a + b.\mathrm{SEL} + \sigma\upsilon_i$$

where SEL is the ANE level, categorised in 5dB bands
υ_i is the random effect (individual arousability)
a, b, σ are estimated regression coefficients

Prediction intervals are calculated here for individuals with 'average' arousability; ie with $\upsilon_i = 0$. Thus, the linear predictor on which the prediction interval is to be based is given by

$$\eta = \alpha + \beta.\mathrm{SEL}$$

To calculate prediction intervals for n, prediction intervals are first calculated for η. They are then transformed to prediction intervals for n using the transformation

$$n = \frac{1}{1 + e^{-\eta}}$$

This is possible because the transformation is a monotonically increasing function of η (Ref D1).

The prediction intervals for η are given by the expression

$$\eta \pm 1.96\sqrt{\mathrm{var}(\eta)}$$

where $\mathrm{var}(\eta) = \mathrm{var}(a) + \mathrm{var}(b) + 2.\mathrm{cov}(a,b)$

Therefore, if U, L and PE are the upper and lower limits and point estimates respectively for the prediction intervals of η, then the prediction interval for n is given by

Upper limit: $n_U = \dfrac{1}{1 + e^{-U}}$

Point estimate: $n_{PE} = \dfrac{1}{1 + e^{-P}}$

Lower limit: $n_L = \dfrac{1}{1 + e^{-L}}$

Reference D1: Cox, D R and Hinckley, D (1989), *Theoretical Statistics*

Printed in the United Kingdom for HMSO
Dd P8175575 1/93