Partnership for AiR Transportation
Noise and Emissions Reduction
An FAA/NASA/Transport Canada-
sponsored Center of Excellence

Design for a U.S. Field Study on the Effects of Aircraft Noise on Sleep

PARTNER Project 25B Year One Report

prepared by
Mathias Basner, MD, PhD, MSc

June 20, 2012

REPORT NO. PARTNER-COE-2012-003

Design for a US Field Study on the Effects of Aircraft Noise on Sleep

PARTNER Project 25B Year One Report

Mathias Basner, MD, PhD, MSc

PARTNER-COE-2012-003

June 2012

This work is funded by the US Federal Aviation Administration Office of Environment and Energy under FAA Award Nos. 07-C-NE-PU, Amendment Nos. 012, 021, 026, and 033, and 10-C-NE-UPENN, Amendment Nos. 001, 002,and 003. This project has been managed by Laurette Fisher and Natalia Sizov, FAA.

The Partnership for AiR Transportation Noise and Emissions Reduction — PARTNER — is a cooperative aviation research organization, and an FAA/NASA/Transport Canada-sponsored Center of Excellence. PARTNER fosters breakthrough technological, operational, policy, and workforce advances for the betterment of mobility, economy, national security, and the environment. The organization's operational headquarters is at the Massachusetts Institute of Technology.

The Partnership for AiR Transportation Noise and Emissions Reduction
Massachusetts Institute of Technology, 77 Massachusetts Avenue, 33-240
Cambridge, MA 02139 USA
http://www.partner.aero

Design for a US Field Study on the Effects of Aircraft Noise on Sleep

PARTNER Project 25B Final Report

Project Period: June 2010 - May 2011

Project Lead Investigator:

Mathias Basner, MD, PhD, MSc
Assistant Professor of Sleep and Chronobiology in Psychiatry
Unit for Experimental Psychiatry, Division of Sleep and Chronobiology
Department of Psychiatry
University of Pennsylvania Perelman School of Medicine
1013 Blockley Hall, 423 Guardian Drive
Philadelphia, PA, USA 19104-6021
Phone: (215) 573-5866
Fax: (215) 573-6410
basner@upenn.edu

Table of Contents

1 Abbreviations

AASM	American Academy of Sleep Medicine
ARAS	Ascending Reticular Activating System
ASDA	American Sleep Disorders Association
CNS	Central nervous system
DALY	Disability Adjusted Life Year
dBA	A-weighted Decibel
DLR	German Aerospace Center
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculogram
L_{Amax}	A-weighted (frequency) maximum sound pressure level
L_{ASmax}	A-weighted (frequency) slow-weighted (time) maximum sound pressure level
L_{max}	Maximum sound pressure level
min	minute
MSLT	Multiple Sleep Latency Test
MWT	Maintenance of Wakefulness Test
PST	Pupillographic Sleepiness Test
PVT	Psychomotor Vigilance Test
REM sleep	Rapid eye movement sleep
s	second
SD	Standard deviation
SEL	Single event level
SPL	Sound pressure level
SWS	Slow wave sleep

2 Summary

With the most recent US field study dating back to 1996 and compared to the efforts of other, especially European, countries, US research on the effects of aircraft noise on sleep has lagged over the past 15 years, while aircraft noise has continued to evolve. Within this period, air traffic has changed significantly, with substantial increases in traffic volume, on one hand, and significant improvements in noise levels of single aircraft, on the other.

The objective of the analyses presented here was to propose an "optimal" study design for a US field study on the effects of aircraft noise on sleep and recuperation that is based on the current scientific knowledge in both the noise effects research and the sleep research area. The report discusses:

- various methods for the assessment of sleep and noise-induced sleep disturbance,
- the assessment of short-term consequences of noise-induced sleep disturbance,
- the assessment of the acoustical environment,
- the synchronization of physiological and acoustical data,
- the assessment of non-acoustical extrinsic factors,
- inclusion criteria and sample size requirements,
- the selection of measurement sites, and
- study design related cost aspects.

Based on these analyses, it is recommended to use actigraphy (if possible with a higher than 1 per min sample rate) plus a single channel Electrocardiogram (ECG) plus the actigraph event marker to signal conscious awakenings plus brief questionnaires to be filled out by the subjects in the morning for a US field study on the effects of aircraft noise on sleep. This study should minimally include one airport with relevant amounts of nocturnal air traffic and one control site without aircraft noise exposure. This study design offers several benefits:

- The chosen design has a low methodological expense, as the subjects can apply the sensors and start and stop the measurements themselves. Also, the measurements could be running for several days without supervision of an experimenter. This will assure the investigation of large subject samples and many aircraft noise events per subject at low cost.
- This methodology will allow for comparisons with both former US and non-US field studies on the effects of aircraft noise on sleep.
- The ECG offers a unique opportunity to inexpensively and unobtrusively measure both subtle and more obvious changes in sleep physiology.

3 Background and Objectives

Undisturbed sleep of sufficient length is necessary to maintain daytime performance and health.[1] The human organism recognizes, evaluates and reacts to environmental sounds even while asleep.[2] These reactions are part of an integral activation process of the organism that expresses itself e.g. as changes in sleep structure or increases in blood pressure and heart rate. Environmental noise may decrease the restorative power of sleep by means of repeatedly occurring activations (so-called sleep fragmentation). Acute and chronic sleep restriction or fragmentation have been shown to affect, among others, waking psychomotor performance,[3] memory consolidation,[4] creativity,[5] risk taking behavior,[6] signal detection performance,[7] and accident risk.[8,9]

With the most recent US field study dating back to 1996 and compared to the sleep disturbance efforts of other, especially European, countries, US research on the effects of aircraft noise on sleep has lagged over the past 15 years, while aircraft noise has continued to evolve. Within this period, air traffic has changed significantly, with substantial increases in traffic volume, on one hand, and significant improvements in noise levels of single aircraft, on the other. Due to inter-cultural differences, results from studies performed outside the US may not be transferred 1:1 to US domestic airports. Therefore, it is important that US field studies be conducted to acquire sleep disturbance data for varying degrees of noise exposure.

The objectives of the analyses presented here is to propose an "optimal" study design for a field study on the effects of aircraft noise on sleep and recuperation. Criteria for the study design were:

- justification based on the current scientific knowledge in both the noise effects research and the sleep research area,
- to allow comparability to other similar past field studies performed both in- and outside of the United States,
- to allow for the investigation of subject samples that are representative for larger parts of the population (implying larger samples and different selection criteria compared to those that have typically been investigated and applied in field studies on the effects of aircraft noise on sleep in the past),
- to provide a sample size calculation that will identify the sample size required to derive precise dose-response relationships between single aircraft noise event metrics (like L_{Amax}) and sleep fragmentation indicators (like EEG awakenings).

4 Analysis of Study Design Components

The design stage of a field study on the effects of aircraft noise on sleep involves many decisions that will influence both the validity and generalizability of the findings and the methodological expense and therefore the cost of the study. In general, there is no perfect study design that maximizes all the desirable attributes of a study. Rather, every single decision made will influence some study aspects positively and others negatively. For example, choosing polysomnography (i.e., the simultaneous recording of the electroencephalogram, electrooculogram, and electromyogram) for measuring sleep will allow the detection of subtle physiological changes induced by aircraft noise. However, this measurement technique is somewhat disruptive and may therefore influence sleep itself. Also, its high methodological expense (and the associated costs) renders studies with large sample sizes impossible. It also very likely decreases response rates and therefore also decreases the generalizability of the findings.

In the following sections, several aspects that are involved in the decision making process for finding an (optimal) field study design will be discussed. Again, there will be no single optimal study design. What is considered "optimal" largely depends on the goal of the specific study, which in turn may be influenced by an analysis of the gaps in knowledge motivating the study. At the end of each section the findings will be summarized and general recommendations will be given.

4.1 Assessment of Sleep

Polysomnography, i.e. the simultaneous recording of the electroencephalogram (EEG), the electrooculogram (EOG), and the electromyogram (EMG) remains the gold standard to measure sleep. According to specific conventions,[10,11] the recorded night is usually divided into 30-second epochs. Depending on EEG frequency and amplitude, specific patterns in the EEG, muscle tone in the EMG, and the occurrence of slow or rapid eye movements in the EOG, different stages of sleep are assigned to each epoch. Wake is differentiated from sleep. Sleep is divided into rapid eye movement sleep (REM sleep) and non-REM sleep, which is again classified into light (stages S1 and S2) or deep sleep (stages S3 and S4, also called slow wave sleep - SWS). SWS and REM sleep seem to be very important for restoration and memory consolidation during sleep (see below).[4] Wake and S1 are typical indicators of disturbed or fragmented sleep, and they do not (or only very little) contribute to the recuperative value of sleep.[12] Even shorter activations (≥3 seconds) in the EEG and EMG, so-called arousals that would not qualify to be scored as an awakening, can be detected in the polysomnogram.[13] These arousals are usually accompanied by vegetative activations (see below).[14,15]

Sleep is a complex human behavior, integrating manifold vital physiological processes (e.g. protein biosynthesis, excretion of specific hormones, memory consolidation) that, in a broad sense, serve recuperation and prepare the organism for the next wake period. The human organism recognizes, evaluates and reacts to environmental sounds even while asleep.[2] As early as in 1939 Davis et al.[16] state, “The effectiveness of auditory stimuli during sleep may be no accident if we consider the general biological function of hearing in the role of watchman constantly on guard to signal danger”.

The so-called Ascending Reticular Activating System (ARAS) is part of the body's arousal system, and is most active during wakefulness. It receives input from several sensory systems (among them the auditory) and relays this information, e.g., to cardio-respiratory brainstem networks and through the Thalamus to the Cortex. The Thalamus has a gating function, i.e., based on the sensory information and the current state of the central nervous system (CNS), information may be relayed to or withheld from the Cortex.[17] If the information is passed on to the Cortex, it may lead to a Cortical arousal, that, if the subject is sleeping, may disturb or fragment sleep.

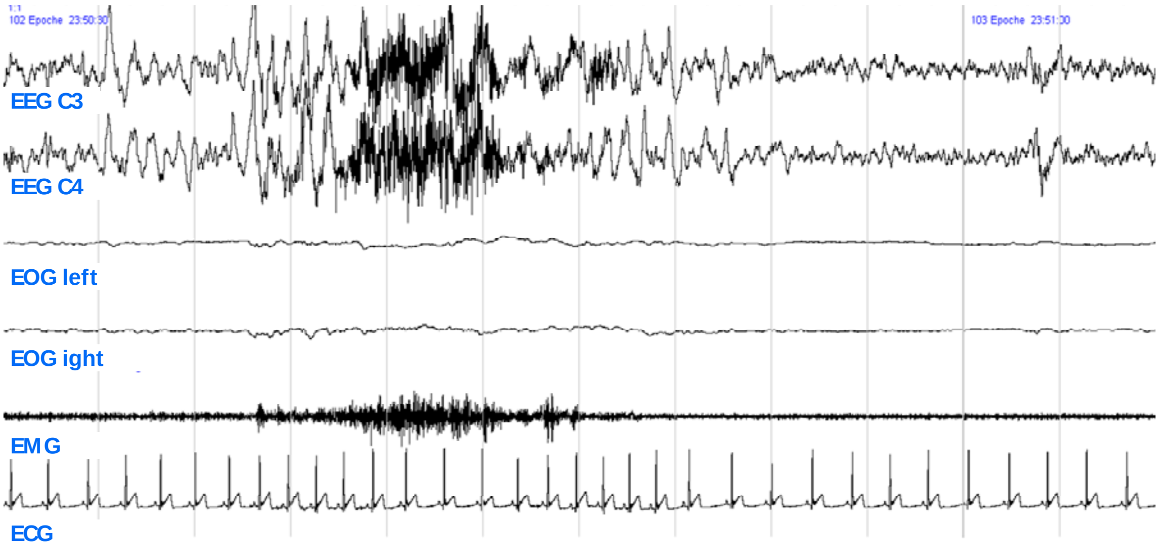


Figure 4-1: Example of an EEG arousal (according to ASDA 1992[18] or AASM 2007[10]). An increase in EMG amplitude (with corresponding artifacts in the EEG) and an increase in EEG frequency are the defining elements of the EEG arousal. The EEG arousal is accompanied by a vegetative arousal (i.e., an increase in heart rate). 36 s are printed in this record, and therefore the arousal lasts for approximately 12 s, which does not qualify the epoch to be scored as wake.[10,11]

Several important implications follow for the effects of aircraft noise on sleep:

(1) The organism's reaction to noise is not based on an all-or-nothing principle (i.e., not every noise event will lead to a conscious awakening). Rather, the reaction is fine-graded ranging from

(depending on the acoustic stimulus and the momentary state of the CNS) (a) no or minimal physiological reaction (not detectable with standard equipment), (b) to an isolated vegetative reaction (e.g., increase in heart rate and blood pressure), (c) to a cortical arousal of different degrees (subtle shift in EEG frequency, ASDA EEG arousal,[18] sleep stage shift to a lighter sleep stage, sleep stage shift to stage wake), and (d) to a full cortical arousal with regaining of waking consciousness (see Figure 4-2).

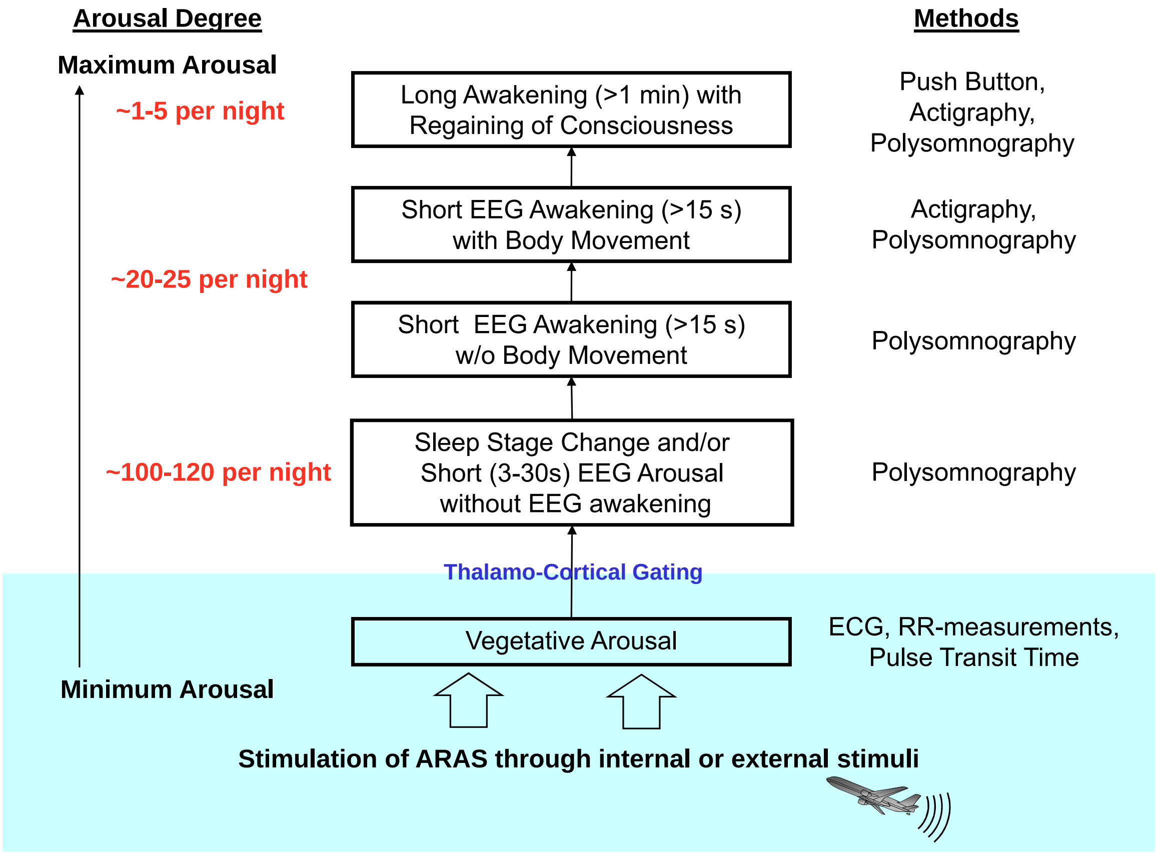


Figure 4-2: Simplified scheme of the body's reaction to external stimuli (as aircraft noise).

It is generally accepted that stronger degree arousals (e.g., conscious awakenings) will have greater consequences for recuperation than those of lesser degree (e.g., vegetative arousals), especially since the former regularly include the latter, but not vice versa. For the same reason, in quiet nights conscious awakenings occur much less frequent than, e.g., brief EEG arousals. However, this does not mean that arousals of lesser degree are of no consequence for recuperation. In fact, it is assumed that many short CNS arousals will fragment sleep and impair recuperation even without relevant changes in sleep macrostructure (i.e., total sleep time, distribution of sleep stages),[19-21] although the independence of both processes is still a matter of debate.[12,22] In the end, the belief that shorter cortical activations are important for recuperation lead to the definition of EEG arousals (i.e., shorter activations of the EEG and EMG) by the American Sleep Disorders Association in 1992. EEG arousals are today routinely scored in sleep laboratories around the world.[18]

(2) CNS arousals are a physiological part of the sleep process and of no pathological consequence unless a certain physiological amount is exceeded (see red numbers given in Figure 4-2 for spontaneous arousal frequencies in quiet nights).[23] As a multitude of external and internal stimuli regularly induce CNS arousals during sleep (see Figure 4-3), the latter are unspecific (i.e., not specific for aircraft noise).

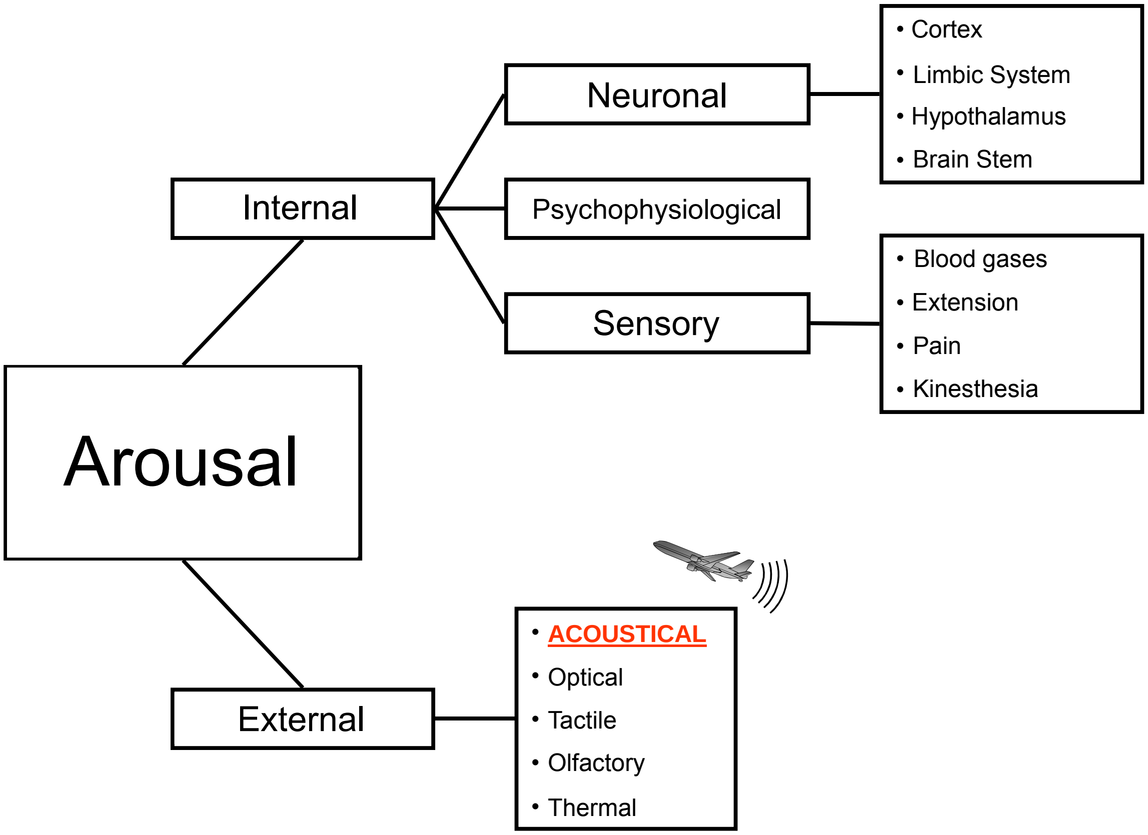


Figure 4-3: Internal and external pathways to CNS arousals (modified based on Raschke and Fischer[24])

This has several consequences. If a CNS arousal (of whatever degree) is observed in association with an aircraft noise event, we cannot be sure that this arousal was actually caused by the noise event, as it is possible that - by chance - it was induced by another external or internal stimulus at the same time. Therefore, only a certain fraction of CNS arousals will be attributable to the noise event, and there are different ways to calculate the magnitude of this fraction (see Brink et al.[25] for a detailed discussion). This will also affect the evaluation of the severity of one additional physiological reaction induced by aircraft noise. Clearly, one additional EEG arousal (with more than 100 spontaneous EEG arousals in undisturbed quiet nights) is very likely less harmful than one additional EEG awakening (20-25 in quiet nights) or one additional conscious awakening (1-5 in quiet nights).

All of the above have consequences for the choice of sleep measurement technique. Comparable to diagnostic tests, the different measurement techniques differ in their sensitivity and specificity. A very sensitive measurement technique (like polysomnography) will pick up even subtle physiological changes like EEG arousals. However, as noted above, EEG arousals are not very specific indicators for aircraft noise-induced sleep disturbance as they occur spontaneously more than 100 times in each (quiet) night. A very specific measurement technique (like the push button technique) may pick up events that otherwise occur only seldom in quiet nights. However, other forms of CNS activation, that may have consequences for sleep recuperation, will be missed by this insensitive method.

Hence, a measurement technique may be characterized as "optimal" if it has a favorable balance between sensitivity and specificity, i.e., if it detects all relevant noise-induced activations of the CNS. Unfortunately, there is no consensus among sleep researchers or noise effects researchers what exactly constitutes a relevant CNS activation. Guilleminault et al.[21] were able to demonstrate in a carefully designed study with, however, only six subjects that only cortical arousals were associated with increased sleepiness or reduced performance on the next day. Vegetative arousals alone did not lead to next day consequences. These results conflict with those of Martin et al.,[26] who claimed that vegetative arousals alone would significantly impair recuperation. However, taking a closer look at their experimental procedure, it is probable that the procedure itself induced some cortical arousals and even changes in sleep macrostructure, so that cortical arousals may indeed be a prerequisite for next day consequences, whereas vegetative arousals alone may suffice to increase the long-term risk of cardiovascular disease.

In the context of designing a field study on the effects of aircraft noise on sleep, this stresses the importance of defining *a priori* what is considered a relevant consequence of aircraft noise. Basner et al.[27] argue for EEG awakenings as adequate indicators for noise-induced sleep disturbance because (a) EEG awakenings demonstrate a good balance between sensitivity and specificity (see above and Figure 4-2), (b) they are, in contrast to briefer EEG arousals, accompanied by prolonged increases in heart rate[15,28] that may play a role in the development of high blood pressure and cardiovascular disease,[29,30] and (c) waking consciousness may be regained due to longer EEG awakenings.[31] These awakenings may be recalled the next morning and affect subjective assessments of sleep quality and quantity. At the same time, noise-events perceived during wake periods can result in annoyance and may prevent the subject from falling asleep again, especially in the early morning hours.[32] However, as mentioned above, this does not mean that shorter EEG activations (or even more subtle shifts in EEG frequency) are without consequences. Also, it was shown that EEG arousals habituate to a lesser degree than EEG awakenings, and, in contrast to EEG awakenings, that they do not replace spontaneous EEG arousals between noise events.[33] Therefore, focusing on EEG arousals may add relevant information, especially in chronic exposure situations (like in field studies) or in study regions with low noise exposure levels.[34]

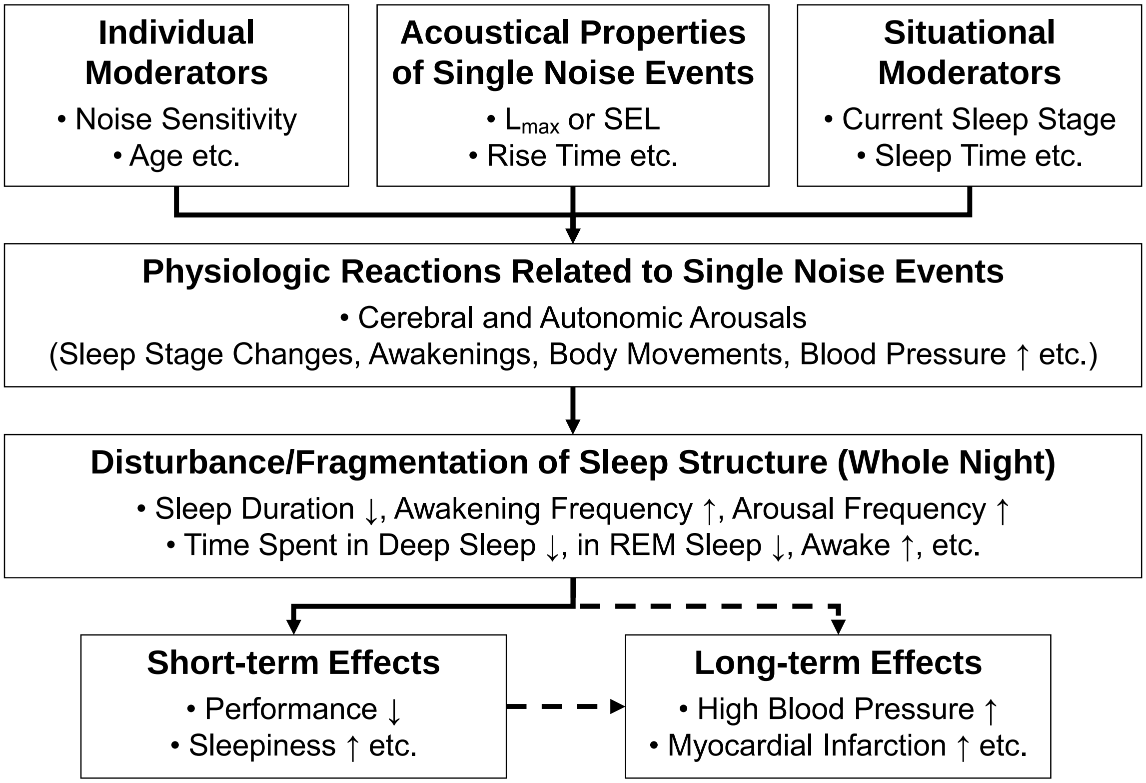


Figure 4-4: Flow chart on the effects of traffic noise on sleep. L_{max} = maximum sound pressure level, SEL = single event level (Reprinted from Appl Acoust 71(6) 2010, 518-22, Basner M, Müller U, Griefahn B, Practical guidance for risk assessment of traffic noise effects on sleep, Copyright 2010, with permission from Elsevier.)

Depending on their frequency, immediate noise effects on sleep (arousals, awakenings) cause a general elevation of the organism's activation level that consequently leads to a redistribution of time spent in the different sleep stages (i.e., changes in sleep macrostructure, see Figure 4-4) with an increase of the amounts of wake and stage S1, and a decrease of SWS and REM-sleep.[33-36] The changes in sleep macrostructure are, however, small and usually in the range of a few minutes. Not surprisingly, sleep fragmentation measures and sleep architecture variables are correlated. For example, in a laboratory study on the effects of aircraft noise on sleep the number of EEG awakenings per hour of total sleep time and the time spent in SWS were correlated with Pearson's correlation coefficient $\rho = -0.667$ ($p<0.0001$).[36]

Sleep can be assessed in several ways, ranging from questionnaire-based self-reports in the morning after nights with aircraft noise exposure to polysomnography, i.e. the simultaneous measurement of EEG, EOG, and EMG (even more elaborate techniques like functional neuroimaging can currently not be performed in the field). The different measurement techniques differ in their sensitivity and specificity for detecting noise induced sleep disturbances, in their invasiveness, in their methodological expense, and in monetary costs. In the following sections, the different measurement techniques are described and their advantages and disadvantages are discussed.

4.1.1 Polysomnography

Description: Polysomnography is the simultaneous recording of (at least[1]) the EEG, the EMG, and the EOG. According to specific conventions (International 10-20-system), electrodes are attached to the scalp and the skin of the face of the subject. The electrical potentials generated by the brain, chin muscles and eye movements are amplified, converted into digital signals and stored on digital media. The signals are later analyzed by trained personnel according to specific conventions (see above).[10,11]

Advantages: Polysomnography is the gold standard for measuring sleep, the evaluation of sleep structure and the degree of sleep fragmentation. As can be seen in Figure 4-2, it is the only method that covers all aspects of sleep (with the exception of conscious awakenings, as we cannot tell with certainty from the polysomnography signals whether a subject regained waking consciousness or not). It is thus a very sensitive method that will detect even subtle changes in sleep physiology. Also, the method itself is very well standardized.

Disadvantages: EEG, EOG, and EMG electrodes and wires are somewhat disruptive, may influence sleep, and thus at least one night is usually required for adaptation.[37] The measurement instruments are expensive and fragile. The instrumentation and de-instrumentation of subjects is cumbersome and has to be done by trained personnel. EEG and EMG electrodes are sometimes affected by movements or excessive sweating of the subjects, which may render the analysis of

---

[1] At least the EEG, EOG, and EMG are needed for sleep stage classification and arousal scoring. However, oftentimes (and certainly in clinical settings) additional sensors are applied to measure heart rate, movements of the rib cage and abdomen, limb movements, airflow, or esophageal pressure.

(part of) the data gathered during the night impossible. Finally, sleep stage classification requires trained personnel and is known to have high inter- and intra-observer variabilities.[38-40]

4.1.2 Actigraphy

Description: Actigraphs measure acceleration of body movements (in one or more dimensions), have the size of a watch, and are worn like wrist-watches (usually on the wrist of the non-dominant arm). Some products have additional features, e.g., light sensors measuring environmental light intensity (sometimes in different spectra), body position sensors, an event marker button (e.g., to signal lights out), or a display (e.g., for displaying clock time). Some devices even allow for sampling other physiological signals like the ECG, but this discussion shall focus on the basic feature of actigraphs to measure (wrist) movements during sleep. The devices usually sample at high rates internally (e.g., 256 Hz), but (user-defined) data storage rates are typically much lower (e.g., 1-2/min). Therefore, a 1 min or 30 s bin will store the degree of wrist movement in the respective period. Actigraphy was used in two large studies on the effects of aircraft noise on sleep in the vicinities of Heathrow[41] and Amsterdam Airport[42]. There are other methods related to actigraphy (e.g., seismosomnography[43]) that measure whole-body movements during sleep.

Advantages: Actigraphs are inexpensive and comparatively robust. After an initial orientation, subjects can wear the device for several days and nights unsupervised (i.e., the methodological expense is low). The movement activity data gathered with actigraphy are the measure of interest, so there is no need to visually score data. Actigraphs are less disruptive than the sensors applied for polysomnography, and it is unlikely that actigraphs substantially influence normal sleep.

Disadvantages: Although actigraphs are an accepted measure to determine rest-activity cycles,[44] more subtle physiological changes cannot be detected by actigraphy. Unfortunately, the degree of standardization overall is relatively low. Different models (i.e., hardware) will give slightly different results, there are several methods to determine activity counts (time above threshold, zero crossing, digital integration),[44] and each company has its own algorithm to differentiate wake from sleep periods. Therefore, it is not surprising that the results of comparisons between polysomnography and actigraphy vary widely.[44-50] Although CNS activations and body movements often occur simultaneously, both may occur independently from each other, and thus one cannot expect a 1:1 agreement. Rather, some misclassifications are obvious: e.g., someone lying awake and not moving but trying to fall asleep would be misclassified as being asleep by actigraphy. A data storage rate of 1-2/min is too low to be useful for an event-related analysis.[14,27] Newer devices allow for higher data storage rates, however, that capability may, in turn, limit the duration of continuous recording.

4.1.3 Electrocardiography

Description: Noise induces activations of the autonomic nervous system, like increases in blood pressure and heart rate, which can be easily measured with electrocardiography (ECG) or plethysmography.[15,51] Methods allowing for the measurement of vegetative arousals (like the ECG) are somewhat unique as they measure early stage CNS arousals (i.e., before thalamo-cortical gating, see Figure 4-2) that may or may not evolve into cortical arousals of different degrees. As cortical arousals are regularly associated with vegetative arousals, and stronger cortical activations are associated with longer and more severe vegetative activations,[15] this offers a unique opportunity to measure both subtle and more obvious changes in sleep physiology with less invasive and less expensive methods than polysomnography.

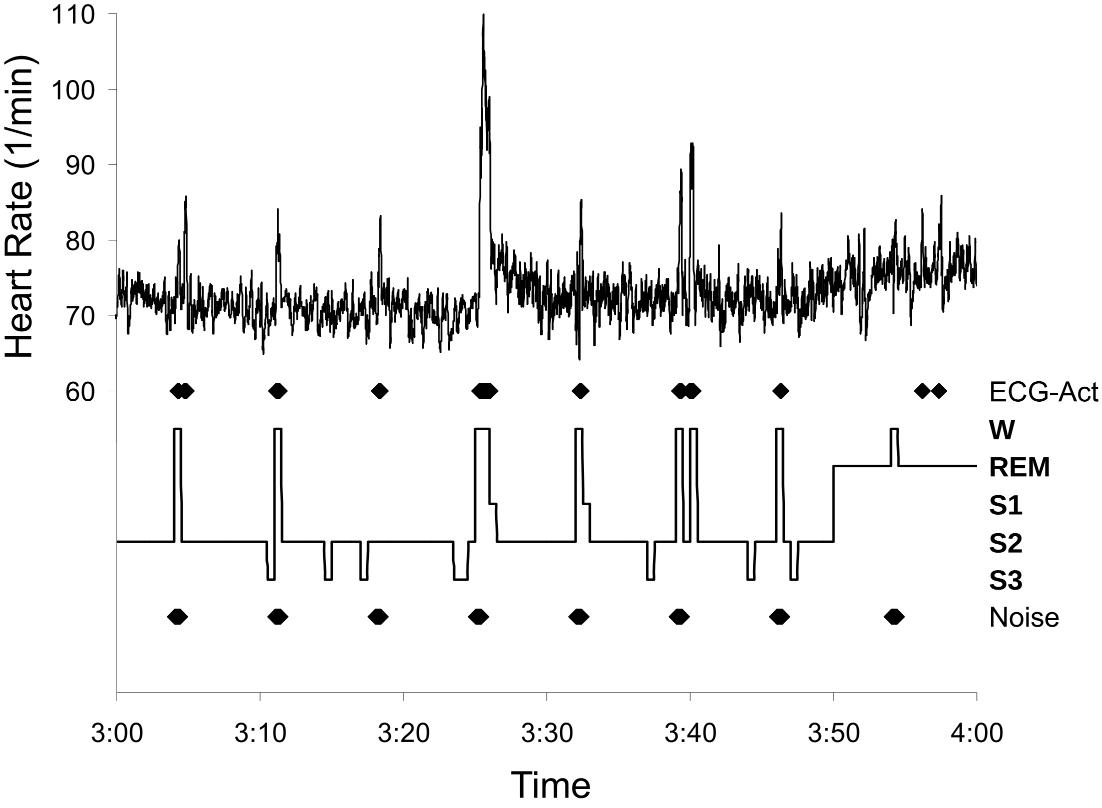


Figure 4-5: In this one hour period, 8 aircraft noise events with maximum sound pressure levels of 65 dBA were played back with seven or eight minute intervals between playbacks. The heart rate trace is shown in the upper part of the figure, the hypnogram in the lower part. Rechtschaffen and Kales[11] sleep stages are given on the right side of the figure (W=wake, REM=rapid eye movement sleep, S1/S2/S3=sleep stages 1/2/3). Times of aircraft noise event playbacks (Noise) are given below the hypnogram as black diamonds. Times of automatic ECG activation recognition by the ECG algorithm (ECG-Act) are given above the hypnogram also as black diamonds. Adopted from Basner et al.[14].

Based on this idea, Basner et al. developed an ECG-based algorithm for the automatic identification of autonomic activations associated with cortical arousals.[15] Heart rate information of five consecutive heart beats relative to a moving median heart rate is used to estimate the probability that the current heart beat is associated with a cortical arousal. A cardiac arousal was defined as 4 consecutive beats with the above mentioned probability exceeding 35% (see Figure 4-5 for an example).

This algorithm was later validated on extensive laboratory data from a study on the effects of aircraft noise on sleep.[14] EEG awakenings and cardiac arousals showed good agreement, especially if spontaneous reaction probabilities were taken into account (see Figure 4-6 A).

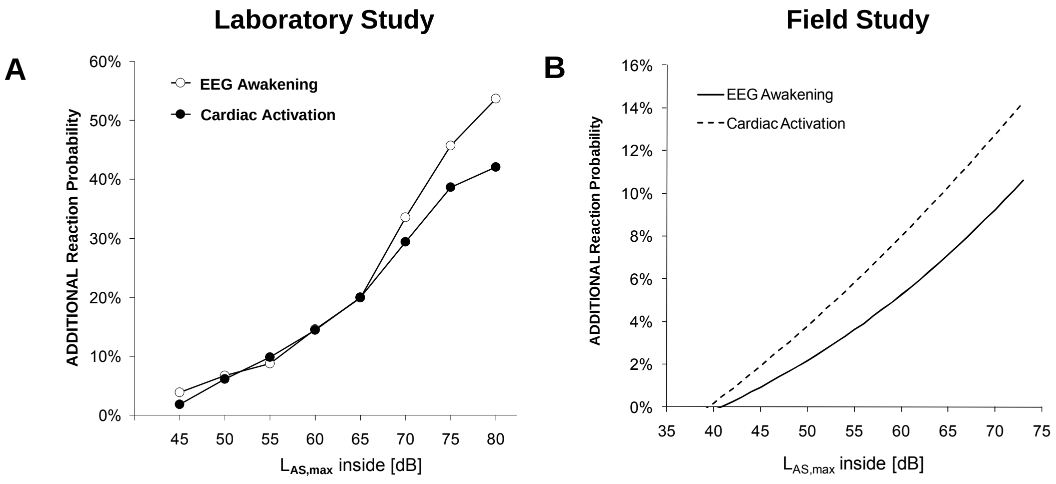


Figure 4-6: Comparison between the probability of EEG awakenings (determined by polysomnography) and cardiac activations (determined automatically based on the ECG according to Basner et al.[15]) depending on maximum sound pressure level $L_{AS,max}$ inside for a laboratory setting (A, adopted from Basner et al.[14]) and in a field setting (B, see Basner et al.[52]). The plotted probabilities (additional reaction probability) take spontaneous reaction probabilities into account (see Basner et al.[14] and Brink et al.[25] for further explanations).

For this report, the comparison between EEG awakenings and cardiac activations was repeated based on data of the DLR field study.[52] The solid line in Figure 4-6 B shows the exposure-response relationship between maximum SPL and EEG awakenings. This exposure response relationship is based on the regression model published in Basner et al.[52], with the distinction that only awakenings (not awakenings and sleep stage changes to S1) were used as the dependent variable. The dashed line shows the exposure-response relationship between maximum A-weighted indoor SPL and cardiac activations. In general, there is a good agreement between both exposure-response relationships with a similar threshold around 40 dBA. However, cardiac arousal probability somewhat exceeded EEG awakening probability. This is not surprising, as (a) the ECG algorithm was designed to detect cardiac activations associated with cortical arousals, and EEG awakenings are only a subset of the > 3 s cortical arousals, and (b) habituation may primarily relate to EEG awakenings, and only to a lesser extent to EEG arousals.[33] However, it is likely that those events detected by the algorithm that were not EEG awakenings still represent relevant cortical activations.

Advantages: Similar to actigraphy, devices measuring the ECG are relatively inexpensive and robust. After an initial orientation, subjects can attach and detach the ECG electrodes themselves and (depending on storage capacity) can wear the device for several days and nights unsupervised (i.e., the methodological expense is low). The data are scored automatically by the algorithm described above, so there is no need to visually score data. The ECG is less invasive than the sensors applied for polysomnography, and it is unlikely that the ECG alone substantially influences normal sleep. Repeated noise induced autonomic activations may play a key role in the genesis of hypertension and associated cardiovascular diseases, and therefore measuring autonomic activations may be an advantage from a conceptual standpoint. In the recent past, the utility of specific aspects of the ECG signal (like heart rate variability[53] or cardiopulmonary coupling[54]) for sleep research has been acknowledged in the field. For this reason alone it will be worthwhile to sample the ECG in a field study on the effects of aircraft noise on sleep.

Disadvantages: The method is relatively new and there are no published studies that could be used for comparison (except the DLR laboratory and field study results shown in Figure 4-6). Therefore, further validation of the ECG algorithm would be desirable. This validation should investigate whether the algorithm performs equally well in every subject and in all sleep stages (the greater heart rate variability in REM sleep may pose a problem for the algorithm). Also, the algorithm was primarily developed for the detection of EEG arousals from sleep. However, a certain period throughout the night is spent awake, and it is unclear how to interpret heart rate increases during wakefulness (the same is true for actigraphy, see above). Basner et al.[14] discuss this the following way: "Situations where the subject was already awake before playback of the ANE started (10.3% of all events) were excluded from the analysis in this study […]. Comparable to actigraphy, the ECG algorithm is not able to differentiate between wake and sleep unless polysomnography is performed simultaneously. If the ECG is sampled alone, cardiac activations during wakefulness may be misinterpreted as awakenings, potentially overestimating the number of traffic noise induced awakenings. However, in situations where the subject is already awake traffic noise may nevertheless adversely affect sleep by preventing the subject from falling asleep again, and therefore prolonging spontaneous or noise induced awakenings.[31] In these situations, noise induced cardiac activations may indicate an increased state of arousal and, therefore, a decreased likelihood of falling asleep again. Hence, although cardiac activations during wake periods may overestimate the number of EEG awakenings, they may nevertheless be a useful indicator of noise induced sleep disturbance. Further analyses on the association of cardiac activations during wakefulness and the time needed to fall asleep again should be performed in the future".

4.1.4 Signaled Awakening

Description: Several studies investigated the influence of traffic noise on signaled awakenings.[55,56] Here, the subject has to give an agreed upon signal (e.g. pressing a button) to indicate the awakening.

Advantages: This method is very easy to use and inexpensive. Signaled awakenings are very specific (i.e., there are only few spontaneous conscious awakenings in undisturbed nights, see above).

Disadvantages: The method has a low sensitivity, i.e., relevant aircraft noise induced physiological activations will be missed. It is also not well standardized in the sense that there are no standardized instructions. These instructions may greatly influence whether or not the subject will press the response button or not. On the same note, by demanding an active cooperation of the subject the importance of the signal, reaction probability, and sleep itself may be altered.[57,58] Also, subjects may forget or be too tired or languid to give the signal.

4.1.5 Questionnaires

Description: Typically, subjects answer questions concerning sleep quality and quantity, number and duration of nocturnal wake episodes, noise annoyance, and momentary state (e.g., tired or refreshed) after waking up in the morning.

Advantages: Questionnaires represent the easiest and probably cheapest way to gather information on sleep. This method is non-invasive and unlikely to substantially alter sleep (aside from the Hawthorne effect, i.e. subjects modifying an aspect of their behavior simply in response to the fact that they are being studied, inherent to all the measures described here).

Disadvantages: The validity of assessing the effects of aircraft noise on sleep is at least questionable, as during most of the night the sleeper is unconscious and not aware of herself/himself or her/his surroundings. This provokes false negative assessments (i.e., subjects may not be aware of relevant physiological activations that do not lead to conscious awakenings, very much like in obstructive sleep apnea). The process of falling asleep and longer wake periods during the night contribute exceptionally to subjective estimates of sleep quality and quantity, which may therefore differ substantially from objective measures.[59] Also, these subjective assessments are prone to manipulation (i.e., subjects may answer in a certain way to, e.g., make a political statement or use the questionnaire as a means to express their frustration with the current noise policy). However, in a laboratory study on the effects of air, road, and rail traffic noise on sleep Basner et al.[33] described the following: "Although most of the night is spent in an unconscious state, subjects were not only able to differentiate between nights with and without noise, but also between nights with low and high degrees of traffic noise exposure. Hence, if these findings extend to the field, morning questionnaires, although prone to manipulation, may be a very cost-effective way for the investigation of traffic noise effects on sleep."

4.1.6 Recommendations

Based on the above descriptions, it is recommended to assess the effects of aircraft noise on sleep jointly with actigraphy and a single channel ECG. These robust, inexpensive, and non-invasive methods should be complemented with questionnaires retrospectively assessing sleep quality and quantity and aspects of aircraft noise exposure of the last night. As outlined above (see Background and Objectives), a potential US field study should investigate large representative subject samples (probably including risk groups). Due to its high methodological expense, this will not be possible with standard polysomnography, which is why it is not recommended to use polysomnography. Rather, already existing data from polysomnographic field studies on the effects of aircraft noise on sleep should be used to increase validation of the methods recommended here, which are detailed and substantiated as follows:

Actigraphy is a well-established method in research on the effects of aircraft noise on sleep. It was used in studies around Heathrow,[60] Amsterdam,[42] and Cologne-Bonn Airport.[52] Therefore, using actigraphy ensures comparability of the results of a US field study with those of the above mentioned European studies. The actigraph's event marker could be used to gather information on conscious awakenings. This, on the one hand, would ensure comparability with earlier US field studies on the effects of aircraft noise on sleep.[55,56] On the other hand, the instruction to press a button when awake would likely influence sleep itself (see 4.1.4). Finally, it is recommended that higher data storage rates are used than those commonly applied (1-2/min) in order to allow for an event-related analysis. If supported by the hardware, data storage rates of 0.5-1 Hz would be desirable.

As described above, the ECG offers a unique opportunity to measure both subtle and more obvious changes in sleep physiology with less disruptive and less expensive methods than polysomnography. Self-instrumentation and automatic data analysis make this an inexpensive and objective method. Nocturnal vegetative activations may play an important role in the genesis of cardiovascular disease, and therefore the analysis of heart rate information alone delivers important insights, but conclusions on the frequency of EEG awakenings also seem possible.[14] Obviously, one cannot expect a 1:1 agreement between EEG arousals/awakenings and cardiac activations, as there are instances of cortical arousal without cardiac arousal (e.g., arousals originating primarily from the cortex, see Figure 4-3) as well as instances of cardiac arousal without cortical arousal (e.g., due to thalamo-cortical gating, see Figure 4-2). Optimally, the ECG signal would be both sampled and stored at a high rate (if possible with at least 256 Hz[61,62]). This would allow for offline analysis of the ECG signal, including its spectral analysis and the analysis of heart rate variability. However, this both depletes data storage and battery capacity and may restrict the maximum sampling duration. For the ECG algorithm, it would suffice if heart rate (i.e., R-R-intervals) was determined internally by the measuring device and stored at a rate of 4 Hz.

Preferentially, both actigraphy and the ECG will be recorded with the same device. This avoids data synchronization problems (see 4.4).

Regardless of the disadvantages described above, the subjective assessment of sleep will always be of interest not only from a scientific but also from a political and legislative point of view. WHO started adopting the concept of Disability Adjusted Life Years (DALYs) based on the number of subjects "highly sleep disturbed" (i.e., checking off the highest 28% of the answering scale).[63] The percentage of subjects being highly sleep disturbed depending on noise exposure is derived from questionnaire data gathered in surveys. Also, despite the disadvantages described above, average subject responses accurately mirrored the degree of noise exposure in a laboratory study on the effects of traffic noise on sleep.[33] If this can be replicated for the field, questionnaires may be a very cost-effective way to investigate the effects of aircraft noise on sleep. For these reasons, questionnaires should be a part of the methodological array assessing aircraft noise-induced sleep disturbance.

Finally, one alternative approach shall be briefly discussed here: It may be possible to gather information on cortical arousals without a full polysomnography, i.e. the application of electrodes on the scalp (positions F_4, C_4, O_2), next to the eyes (EOG-L, EOG-R), above chin muscles (EMG-1, EMG-2), and above the mastoids (M-1, M-2). One scalp electrode (e.g., F_z) together with a reference electrode (M-1) and a mass electrode may suffice to reliably detect cortical arousals. These electrodes could probably be self-administered by the subjects (although adequate skin preparation is more important for the EEG signal than for the ECG signal due to lower electrical potentials generated by the Cortex). However, even this minimal EEG montage may still disturb sleep to a certain degree. Also, sleep stage classification will not be possible based on this minimal montage, as the relevant standards[10,11] require additional EEG electrodes in central and occipital positions.

There is one commercially available wireless system that uses proprietary dry silver-coated fabric sensors in a headband and requires no additional electrodes. A first validation study under controlled laboratory conditions in N=26 subjects undergoing two nights of full polysomnography was recently published.[64] The automatic scoring algorithm of the device showed substantial chance-corrected agreement with two separate visual scorings and a visual consensus scoring, but fell behind the level of agreement of the two visual scorings (the device only differentiates light sleep, deep sleep, and REM sleep). The authors conclude: "We conclude that the wireless system shows promise as a relatively accurate system for scoring sleep." It should also be noted that because of the temporal smoothing of the algorithm, the wireless system is not suitable for scoring single epoch intervals of wakefulness or arousals in the current configuration (i.e., a noise event-related analysis would not be possible). Future development and field validation of this method will have to demonstrate its potential usefulness for investigating the effects of aircraft noise on sleep.

4.2 Assessment of the Consequences of Aircraft Noise-Induced Sleep Disturbance

Not only the sleep disturbing impact of aircraft noise per se is of interest, but also short-term and long-term after effects of (chronically) noise-disturbed sleep. Therefore, additional psychological and physiological measurements may be performed in the evening before or in the morning after a night with aircraft noise exposure. The different methods that have been applied in the past will be briefly discussed below. The use of questionnaires has already been discussed in 4.1.5.

4.2.1 Performance Tests

A few past studies have used objective cognitive tests to assess the influence of aircraft noise on sleep recuperation.[33,65,66] The size of the observed effects was in general small. Furthermore, the tests are usually performed shortly after subjects wake up in the morning, and thus varying degrees of sleep inertia may affect test results more strongly than the degree of aircraft noise-induced sleep disturbance.[67] Finally, most of the tests show practice effects with learning curves (i.e., the tests would need to be performed several times prior to the start of the study to achieve stable performance levels). The Psychomotor Vigilance Test (PVT) is an exception in the sense that it has negligible aptitude and learning effects.[68] It may be possible to administer a modified brief (3-min.) version of the PVT in the morning after exposed nights, that has been shown to be sensitive to the effects of acute total and chronic partial sleep loss,[69] but otherwise the increase in the burden of subjects may be too high potentially decreasing participation rates.

4.2.2 Memory Tests

Memory consolidation seems to be a very important function of sleep[4]. Sleep restriction and fragmentation can impair the consolidation of content learned during the day. However, in a laboratory study on the effects of traffic noise on sleep, Basner et al. found no significant influence of traffic noise exposure on the number of word pairs recollected on the next morning.[33] As the effects of traffic noise are usually stronger in the laboratory compared to the field, it is unlikely that relevant effects on memory consolidation would be found in a field study on the effects of aircraft noise on sleep (at least if the same test paradigm is used as in Basner et al.[33]).

4.2.3 Objective Measurement of Sleepiness

Sleepiness is usually defined as the propensity to fall asleep. In contrast to subjective assessments of sleepiness via questionnaires (e.g., via the Karolinska Sleepiness Scale or visual analogue scales), objective measurements of sleepiness involve some sort of physiological measurement. The gold standard for the measurement of sleep propensity is the Multiple Sleep Latency Test (MSLT), where, in a quiet environment, subjects are measured polysomnographically and asked to fall asleep. Short sleep latencies indicate a high degree of sleepiness[70]. The MSLT is performed repeatedly during the morning and afternoon. The maintenance of wakefulness (MWT) is a test related to the MSLT, but here the subjects are asked to stay awake as long as possible. Thus, while the MSLT measures the ability to fall asleep, the MWT measures the ability to remain awake.

The Pupillographic Sleepiness Test (PST) measures spontaneous fluctuations of pupil size in darkness over a period of 11 min,[71] that have been shown to increase in sleepy subjects. The measurement environment has to be dark and quiet in order to assure valid results. Basner was able to show that nocturnal aircraft noise exposure in the laboratory increased sleepiness assessed with the PST[72]. However, all three methods (MSLT, MWT, PST) are time consuming and require controlled environmental conditions. Their implementation in field studies may thus be infeasible.

4.2.4 Blood Pressure Measurements

A decrease in blood pressure during the sleep period (so-called dipping) seems to be important for the cardiovascular system and for the prevention of cardiovascular disease[73]. Extrinsic sleep fragmentation was shown to increase blood pressure levels both during the night[74] and during the day.[75] It is plausible that these increases in blood pressure, that may initially be transient, contribute to long-term increases in blood pressure and thus to the genesis of cardiovascular disease.[76] For this reason, blood pressure measurements in the morning may be valuable to assess the effects of aircraft noise induced sleep disturbance. Automatic blood pressure measurement devices are available, but the correct placement of the sensor and standardized measurement conditions (e.g., body position and a certain rest period before the measurements) are a prerequisite for meaningful measurement results. Thus, in the case of unsupervised measurements by study participants both thorough instructions and highly compliant subjects are needed to achieve valid results.[77]

4.2.5 Recommendations

In order to minimize methodological expense (and thus maximize response rates and generalizability of results), it is recommended that the assessment of the consequences of aircraft noise induced sleep disturbance be restricted to brief morning questionnaires and blood pressure measurements. The feasibility of unsupervised blood pressure measurements in the morning should be assessed in a pilot study prior to implementation in the final study protocol. The objective measurement of daytime sleepiness (via MSLT, MWT, or PST) will not be feasible in a field study on the effects of aircraft noise on sleep. Although the administration of the 3-min. Psychomotor Vigilance Test would be feasible in the field setting, significant changes in PVT performance are not expected after nights with noise exposure in the field, and thus the increase in methodological expense is not justified. The same is true for other cognitive and memory tests.

4.3 Assessment of the Acoustical Environment

In contrast to noise mapping, where average noise levels outside the dwelling are usually calculated based on historical, current, or projected traffic volumes, aircraft types, and flight paths, the sound immissions of single passing aircraft inside the bedroom are usually of primary interest in field studies on the effects of aircraft noise on sleep.

Depending on the sound insulating properties of the dwelling (including the nature of the windows), the location of the dwelling relative to the flight path, and current atmospheric conditions aircraft noise will be attenuated from outside to inside to different degrees. Attenuation differs both across different types of dwelling (and therefore across countries) as well as within the same dwelling across nights (e.g., depending on window position). Additionally, the use of ventilators or air condition units will influence background noise levels and the difference between aircraft noise and background noise levels (so-called emergence).

From a scientific point of view, it is desirable to measure the acoustic environment both inside the bedroom and outside the house. This would deliver information on the sound attenuating properties of the building at the measuring site depending on window position, which will be important to inform limit values based on outside noise level predictions. The outside recording could be used to identify those noise events that are generated outside the house (e.g., by a passing aircraft or truck). It can be difficult to differentiate between internal and external noise generation if only a recording inside the bedroom is available, particularly if it is a recording of time-averaged sound pressure level as opposed to an audio-quality recording that can be listened to afterwards. However, if the time of occurrence of aircraft noise events is known (e.g., due to cooperation with the airport providing detailed flight records), it may be possible to record only inside the bedroom, which would lower the methodological expense. Also, if multiple measurement sites are in close proximity, it may suffice to record outside at one central site only for all measurement sites for the purpose of identifying outside noise events[42]. From a political point of view, one could argue that it suffices to record only outside the bedroom, as legislation is based only on calculated outside sound pressure levels, anyway. However, the level of outside to inside sound attenuation may vary considerably both between and within sites (e.g., depending on weather conditions), and, as the reaction of the sleeper is determined by the noise level inside the bedroom, this introduces unnecessary variance on the side of the independent variable. Therefore, it may be advantageous to base exposure-response relationships on inside noise levels, and rather assume a constant (and perhaps conservative) attenuation between outside and inside noise levels.

Calibrated class-1 sound level meters should be used for the sound pressure level recordings, and at least the A,S-weighted sound pressure level should be stored at a high temporal resolution (e.g. 8 Hz). However, not only the sound levels, but the actual sounds should be continuously recorded, too, for the following reasons. First, listening to sound events will be necessary to identify different noise events generated inside and outside the bedroom and to determine whether other noise sources interfered with an aircraft noise event (e.g., simultaneous pass-by of a truck). Second, recent research has demonstrated the importance of the spectral composition of noise events for their effects on sleep.[33] Thus, a continuous recording of the noise events will guarantee the capability of offline acoustical analysis of noise events. Both noise levels and the actual noise events should be preferentially recorded with the same class-1 sound level meter.

Recommendations: It is recommended that actual sounds are continuously measured along with noise levels with class-1 noise level meters inside the bedroom. Also, it is recommended that the study is done in cooperation with the airport that provides detailed information on flight data with a high temporal resolution. The combination of interior acoustical measurements and operations data should be sufficient for the identification of aircraft noise events. If this is not possible, recording of outside sounds may be necessary to correctly identify outside noise events. If there are simultaneous measurements at several measurement sites, it may be sufficient to record outside sounds at one central site.

4.4 Data Synchronization

It may be possible to record actigraphy and the ECG with the same device. However, physiological and acoustical data will be recorded with different devices, so these devices have to be synchronized in order to allow for an event-related analysis.[52] There are several ways to assure synchronization:

(1) A trigger signal will be recorded on all measuring devices at regular or irregular intervals. In the DLR field study, the trigger signal was generated by the sound level meter once a certain background level was exceeded and recorded on both the sound level meter and the physiological measurement equipment. However, this kind of synchronization requires a connection between different measurement devices, that, unless established wirelessly, may complicate the measurement and eventually affect sleep. The minimum requirement would be to record a trigger at the beginning and the end of the measurement period. With these triggers, the time drift of the different measuring devices can be determined and corrected for.

(2) In a Swiss study on the effects of church bell noise on sleep (not yet published), the DCF77 longwave time signal (generated from atomic clocks) transmitted from the city of Mainflingen, Germany, was recorded with both the acoustical and physiological measurement equipment for the purpose of synchronization. This does not require the different devices to be physically connected. A similar signal is generated at Fort Collins, Colorado, and can be received through most of mainland USA.

(3) Wireless technology (like Bluetooth) may be used to transmit data from several devices (like an actigraph and a sound level meter) to a laptop computer that simultaneously records all signals. If there is no delay in data transmission (or the delay is known), this would assure synchronization of the different signals. However, wireless technology often drains battery capacity quickly, and will thus limit the duration of continuous, unsupervised recording without recharging the battery.

(4) If there is no synchronizing signal, it is possible to synchronize the internal clocks of all measuring devices immediately before the start of the measuring period and correct the data for the time drift of each individual device. The drift of each device can be determined before the study and preferentially at different time points throughout a typical measurement period, as the drift may not be linear. As the drift may depend on whether or not data are recorded and on the sampling and data storage rate, the conditions should resemble those of the actual study (e.g., actigraphs should be worn and have the same settings for sampling and data storage rate).

<u>Recommendations</u>: It is recommended that actigraphy and the ECG are recorded with the same device. If it is not feasible to use wireless technology, the internal clocks of all measuring devices should be synchronized immediately before the start of the measuring period and the data corrected for the time drift of each individual device (that was established before the start of the measurement period), in order to assure synchronization between acoustic and physiological variables.

4.5 Assessment of Non-Acoustical Extrinsic Factors Influencing Sleep

Environmental influences other than aircraft noise may impact sleep and mediate the influence of aircraft noise on sleep. The most important factors are room temperature, humidity, and light intensity. Systematic research on the effects of room temperature on sleep is surprisingly scarce. However, recent research with temperature and humidity measurements in the duvet and the mattress stresses the importance of the bed as part of the thermal sleep environment, and that room temperature and humidity may be bad proxies for the actual exposure.[78] However, more elaborate measurements of duvet and mattress temperature may not be feasible within a large-scale field study. In the DLR field study,[52] room temperature was measured continuously inside the bedroom, but no effect on awakening probability was found (unpublished data). Light exposure influences the excretion of the hormone melatonin, which itself influences sleep initiation, quality, and duration. Some actigraphs measure light intensity, but the sensor may be covered by the pajamas. Again, continuous measurements of light intensity may be advantageous to be able to control for light exposure in the statistical analysis.

Recommendations: It would be advantageous to measure temperature, humidity and light intensity in the bedroom in a field study on the effects of aircraft noise on sleep, in order to be able to control for the effects of these variables on sleep in the statistical analysis. However, the measurement of these variables does not seem to be a prerequisite for the success of the study, and it would increase the methodological expense. It is, however, recommended to measure both exposure and control groups during the same season of the year (e.g., only during the summer), which would assure comparable environmental conditions for both groups.

4.6 Selection Criteria and Sample Size

In this chapter selection criteria for study participants and present sample size calculations will be described.

4.6.1 Selection Criteria

Selection criteria are usually applied to increase the internal validity of a study. For example, subjects have to be able to understand study instructions and perform study specific procedures to be eligible to participate. In a study on aircraft noise effect on sleep subjects with intrinsic sleep disorders (e.g., Obstructive Sleep Apnea, Periodic Limb Movements in Sleep) may be excluded because it will be hard to disentangle extrinsic (noise) and intrinsic (apnea) reasons for the sleep disturbance. Maximizing internal validity will assure that the results are valid for the investigated group of subjects, but it also decreases external validity and generalizability of results (i.e., results are transferable to a smaller group of subjects).

In past studies on the effects of aircraft noise on sleep, a wide range of selection criteria have been applied. Participants were often restricted to young healthy adults with no intrinsic sleep disorders, no other relevant diseases, no shift-work history, normal hearing thresholds according to age, and no medication use that could affect the sleep process, etc. At the same time, sample sizes of the investigated populations were often low, bringing into question how representative the obtained results are of the response of the general population, which makes it hard to use the findings for legislative purposes[27].

There are several ways to increase response rates to advertisements asking for study participation (beside increasing reimbursement). Studies with low methodological expense that are less demanding for the subjects will achieve higher response rates. For example, a polysomnographic study that requires measuring for 9 consecutive nights (like the DLR field study) will likely achieve lower response rates than an ECG/actigraphy study for 2 consecutive nights. Also, fewer eligibility criteria will increase response rates.

For a US field study on the effects of aircraft noise on sleep it is suggested to keep methodological expensed as low as possible and to apply as few eligibility criteria as possible to increase response rates and generalizability of the study results[42]. It is possible to address some of the typical eligibility criteria during the analysis phase of the study. For example, it could be investigated whether subjects that consume certain medication show higher or lower reaction probabilities compared to those free of medication, and if so, the analyses could be adjusted for this effect. On the other hand, assuming that the investigated sample represents the exposed population well, it may be advantageous not to adjust for any of these effects, as the exposure-response relationship will then better reflect the response of the entire exposed population. In this case, a control group of subjects not exposed to aircraft noise would be very helpful to assess the effects observed in the exposed subject sample.

4.6.2 Sample Size Estimation

Sample size determination is a crucial step in the design of every study. The sample size should just be large enough to detect an effect that is regarded as being relevant. Both too small and too large sample sizes are regarded unethical. Underpowered studies (sample size too small) will not be able to achieve a significant effect, while overpowered studies (sample size too large) are a waste of funds that could be allocated more appropriately.

The primary outcome of a field study on the effects of aircraft noise on sleep will be an exposure-response relationship between acoustical characteristics of single aircraft noise events (e.g.,

$L_{AS,max}$, SEL) and the probability to react to the noise event (e.g., EEG awakening, body movement). In this case, there are two criteria that can guide sample size determination:

(a) The precision of the exposure-response relationship. This can be quantified by using the width of the 95% confidence interval. The confidence interval surrounds the point estimate, and if the study was repeated multiple times, the true mean population value would fall in 95% of all the 95% confidence intervals constructed around the point estimates. We could say that we are 95% confident that the true mean population value lies within this interval. Although there are no general rules on the required degree of precision of an exposure-response relationship, narrow confidence intervals (i.e., a high precision of the estimate) are highly desirable. The precision of the exposure-response relationship should be high enough to be useful for noise-effects predictions in policy contexts.

(b) The statistical power to detect a significant influence of the acoustical characteristic on physiological reaction probability. According to statistical conventions, the power of a test (i.e., the probability to detect a statistically significant effect if in reality there is an effect) should be at least be 80% at a Type-I error rate (i.e., the probability to find an effect if in reality there is no effect) of 0.05.

4.6.2.1 Methods for Sample Size Estimation (Power Calculation)

The power calculation for a field study on the effects of aircraft noise on sleep is not straightforward, as the reaction of individual subjects to single noise events are the primary outcome. Each subject is exposed to multiple noise events, and reactions to noise events are correlated within each subject (i.e., each subject may react with a specific individual probability determined, e.g., by individual traits or environmental/situational conditions specific for that subject). The power of the study and the precision of the exposure-response relationship will therefore depend on:

(a) the number of investigated subjects

(b) the number of noise events per subject

(c) the intra-class correlation (i.e., how strongly reactions vary within subjects relative to between-subject variance).

As mixed effect logistic regression models within a repeated measures design are used to determine the exposure-response relationship, there is no closed form solution for calculating the power (i.e., power calculations have to be based on simulations). Thus, a bootstrap approach based on Monte Carlo simulations was used to perform the power calculations.

The power calculation simulations presented here are based on 63 subjects participating in the DLR field study on the effects of aircraft noise on sleep near Cologne/Bonn airport between 2001 and 2002.[52] This study was approved by the local ethics committee. Subjects gave written informed consent prior to study participation. The re-analysis of the data within this project was approved by the Institutional Review Board of the University of Pennsylvania.

The primary outcome was an EEG awakening defined according to the criteria of Rechtschaffen et al.[11] in the 90-s period (3 sleep stage epochs) following the start of an aircraft noise event.[52] Only data sampled during study nights 2 to 9 were used (the first night served as adaptation to the measurement equipment). Two subjects were excluded from the simulations as they were exposed to fewer than 60 aircraft noise events during the whole study period. The remaining 61 subjects (average age 37.0±12.9 years, range 19-61 years, 27 male, 34 female) were exposed to a total of 16,253 noise events with an average maximum sound pressure level inside the bedroom of 43.7 dBA (SD 8.7 dBA, range 13.8-73.2 dBA) during the 8 study nights. On average each subject was exposed to 266 noise events (range 60-632 events) during the 8 study nights, corresponding to 33 noise events per night.

Simulations were performed for sample sizes of 10, 20, 30, 40, 50, and 60 subjects (6 classes) and for 20, 40, 60, 80, 100, 150, and 200 noise events per subject (7 classes). For each of the 6 x 7 = 42 combinations, 200 Monte Carlo simulations were performed by following this procedure:

(1) Depending on the investigated "number of subjects" class, between 10 and 60 subjects were randomly drawn without replacement (i.e., each subject could only be drawn once) from the set of the 61 subjects.

(2) For each of the randomly drawn subjects and depending on the "number of noise events" class, between 20 and 200 noise events were randomly drawn without replacement from the subject selected under (1). If the subject was exposed to fewer aircraft noise events in the original Cologne/Bonn study relative to the current "number of noise events" class, the remaining noise events were drawn from the next subject in addition to the events dictated by the current "number of noise events" class. This way, some variance in the number of noise events per subject was introduced into the randomly drawn datasets, mimicking the variance observed in the Cologne/Bonn study. If the last subject was exposed to fewer aircraft noise events in the Cologne/Bonn study than dictated by the current "number of noise events" class plus any remaining noise events from prior subjects, the current dataset was discarded and the process described in (1) and (2) was repeated until no noise event remained.

(3) The process described in (1) and (2) was repeated 200 times to produce 200 data sets for each "number of subjects" class by "number of noise events" class combination.

This procedure produced data sets with sizes of N=200 up to N=12,000 observations that were imported into the statistics software SAS (SAS Institute, Version 9.3). Random subject effect logistic regression models with aircraft noise event maximum SPL as the only explanatory variable (beside the intercept) were performed on each data set (Proc NLMIXED). If the effect of maximum SPL was found to be statistically significant ($P<0.05$) in the regression models it was noted. The proportion of models with $P<0.05$ is equivalent to the estimated power of the investigated "number of subjects" class by "number of noise events" class combination.

Finally, with an estimate statement within Proc NLMIXED point estimates and 95% confidence intervals were calculated for awakening probability for 11 maximum SPLs surrounding the average $L_{AS,max}$ of 43.7 dBA (i.e., at 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, and 60 dBA). The average width of the 95% confidence intervals at the 11 maximum SPLs was calculated as a measure of precision for each data set. The SPL range from 27 to 60 dBA covered 94.8% of the 16,253 noise events in the original data set. It was decided not to extend this range, as wide confidence intervals at extreme (both low and high) sound pressure levels would have dominated the average width of the confidence intervals.

4.6.2.2 Results of the Power Calculations

The results of the power calculations are shown in Figure 4-7.

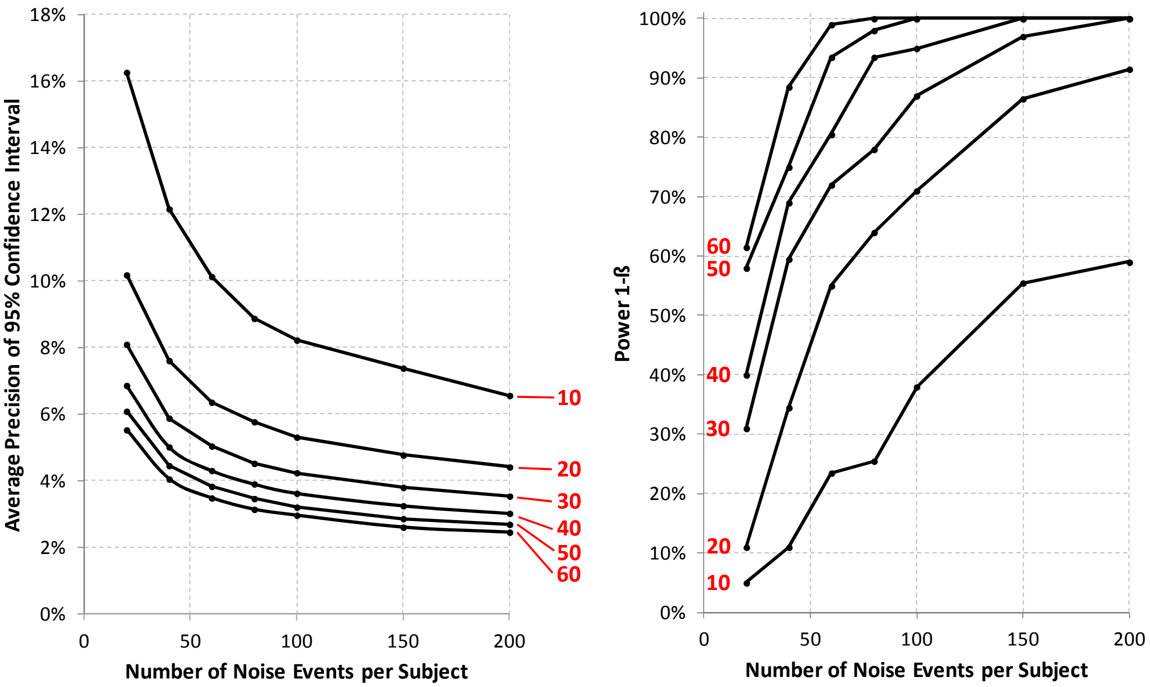


Figure 4-7 Results of the power calculations (the number of subjects used for the simulations are shown in red).

As expected, the precision of the dose response relationship increased both with the number of investigated subjects and the number of noise events per subject (Figure 4-7, left). However, the gain in precision per subject dropped markedly if the number of subjects was increased above 50. Likewise, the gain in precision was less pronounced for >100 noise events per subjects relative to <100 noise events per subjects.

The statistical power was >80% for 20 subjects exposed to >150 noise events each, for 30 subjects exposed to >100 noise events each, for 40 and 50 subjects exposed to >60 noise events each, and for 60 subjects exposed to >40 noise events each. 10 subjects failed to reach a power of 80% even if exposed to 200 noise events each.

Figure 4-8 shows two examples of dose-response relationships with average precisions of 3.0% and 6.3%. The precision is highest near the average maximum SPL of the original data set (43.7 dBA). It decreases with both increasing and decreasing maximum SPL, as the number of noise events with either low or high SPLs decreases simultaneously. Clear advantages in precision can be seen for the example of 60 subjects with 100 noise events per subject relative to the example of 40 subjects with 20 noise events per subject.

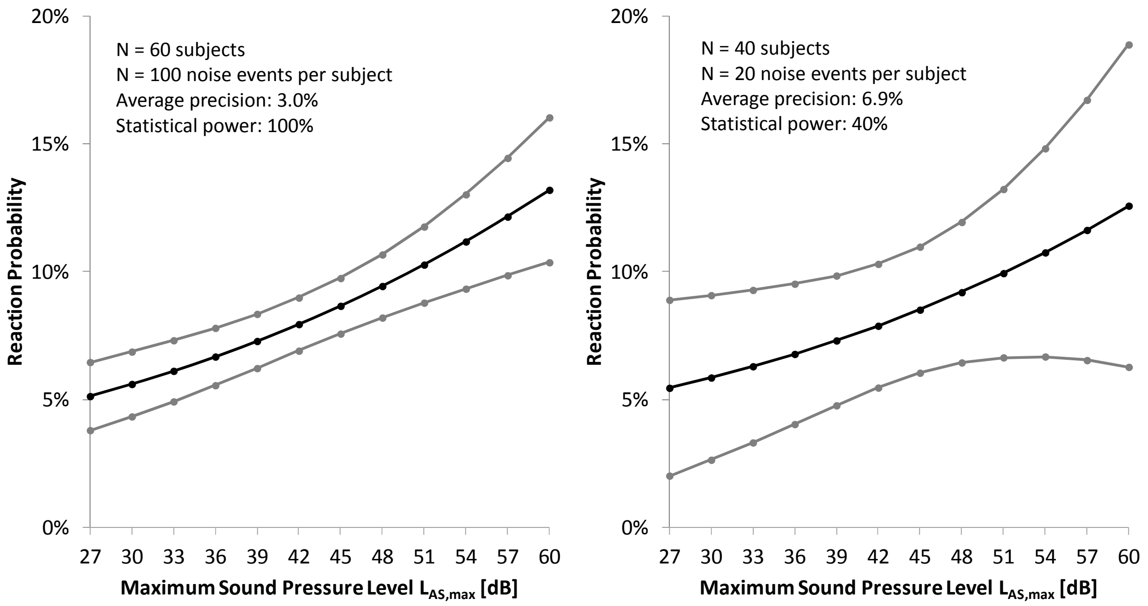


Figure 4-8: Two examples of exposure-response relationships and 95% confidence intervals depending on the number of subjects (60 vs. 40) and the number of noise events per subject (100 vs. 20). Each data point for both point estimates (black) and confidence limits (gray) represents the average of the 200 simulation results.

4.6.3 Recommendations

As far as selection criteria are concerned, it is recommended to use as few criteria as possible to increase response rates and the generalizability of results (with representation of both sexes and a wide age range, probably including children). It will be possible to adjust for some of the standard selection criteria in the analysis phase of the study. Another strategy may be not to adjust for them but rather use a control group not exposed to aircraft noise but with similar properties as the exposed group to assess the effects seen in the exposed group of subjects.

As the power calculations have demonstrated, the power of the study and the precision of the exposure-response relationship depend on both the number of investigated subjects and the expected number of noise events per subject (assuming the intra-class correlation is similar to the subjects investigated in the DLR field study). The latter will depend on the traffic volume at the study site. Therefore, at busy airports it may be sufficient to investigate subjects for a single or a few nights, whereas at airports with low traffic volumes or traffic curfews it may be necessary to measure for several nights. Figure 4-7 shows that different combinations of "number of subjects" and "number of noise events per subject" can lead to the same power/precision. In this case, the preference should be given to increase the number of subjects, as we are more interested in getting precise information on a representative group of subjects than very precise information on a smaller group of subjects. The investigated number of subjects and number of nights per subject should be chosen in a way that at least 80% power is achieved even with some attrition or a lower than expected number of noise events per night. Based on the examples shown in Figure 4-8 and depending on the number of noise events per subject, an average precision of the exposure-response function of 3% or higher can be achieved with subject samples of N=40 or higher, given a high enough number of noise events per subject.

4.7 Measurement Sites

The measurement site can be characterized in various ways:

(1) Traffic density: Traffic density affects the intermitted character of aircraft noise events. During times of high traffic density, the noise-free interval between two planes will be shorter, and it is more likely that reactions to the prior noise event will affect reactions to the current noise event (i.e., dependency between noise events increases). For example, a subject that wakes up due to a noise event may be prevented from falling asleep again by the next noise event, or a subject that transitioned from a deep to a light sleep stage due to a noise event is more likely to wake up due to the next noise event. Traffic density will typically fluctuate both seasonally and during the course of the day, with high traffic density periods during the morning and the evening.

(2) Traffic curfews: Some airports have nocturnal traffic curfews in place (i.e., planes are not allowed to take off or land during a certain time period, except for emergencies). These traffic and noise-free periods are obviously beneficial for sleep. However, at airports with traffic curfews, traffic densities before and after the curfew may be higher compared to airports without curfews due to re-scheduling of flights that would otherwise have taken place in the curfew period. Basner and Siebert showed that the beneficial effects on sleep of a 5-hour curfew at Frankfurt airport are likely to be comparatively minor, and that sleep quality of subjects who (have to) go to bed very early or very late likely decreases if the curfew is in place.[79]

(3) Exposure to other noise sources: Parts of the population are exposed to more than one traffic noise source (e.g., subjects living close to an airport and close to a busy road). Obviously, traffic noise other than aircraft noise potentially also disturbs sleep. At measurement sites with high rail or road traffic volumes, it will be difficult to disentangle the contribution of each of the different traffic noise sources (including air traffic) to sleep disturbance. Also, many of the nocturnal aircraft noise events will coincide with road or rail noise events, and in these instances it will be unclear what noise source caused a possible physiological reaction. One possibility is to identify all non-aircraft traffic noise events and discard those aircraft noise events contaminated by other traffic noise from the analysis. However, it is sometimes hard to identify other traffic noise events, especially if acoustic measurements are only performed inside the bedroom. At the same time,

contamination of aircraft noise events with noise from other sources is quite common, and discarding contaminated noise events would decrease the generalizability of results. For these reasons, it may be advantageous to disregard other noise sources in the analysis, if it can be guaranteed that exposure and control sites are comparable in terms of road and rail traffic noise exposure.

(4) Degree of steady-state: It has been shown that, at the same Leq, annoyance at newly opened or extended airports is higher than would be expected based on exposure-response relationships that have been generated at steady-state airports (i.e., those with no recent changes in infrastructure and/or air traffic)[80]. It would be interesting to investigate whether and if, how long, the degree of sleep disturbance exceeds that observed at steady-state airports.

Recommendations: It is recommended that a US field study on the effects of aircraft noise on sleep should be performed at least at five measurement sites:

(1) an airport with high traffic densities during the night (e.g., a freight hub) and no nocturnal traffic curfew

(2) an airport with low traffic densities during the night and no nocturnal traffic curfew

(3) an airport where a nocturnal traffic curfew is in effect,

(4) an airport that recently has been extended (i.e., experienced a significant change in air traffic), and

(5) at least one control site without aircraft noise exposure.

The exposure to road and rail traffic noise at the control sites should be comparable to that at the airport measurement sites. Runway use depends on wind direction, as planes have to take-off and land with headwind (tailwind can be tolerated up to a certain degree). This necessarily means that there will be pauses in noise exposure if the wind changes direction and the direction of runway use changes. This will allow for interesting within-subject comparisons (i.e., comparing nights with and without noise exposure for the same subject), but it will not replace the necessity of control sites.

If it is infeasible to study at five locations, measurements sites should minimally include one airport site (preferentially one with high traffic densities during the night) and one control site. The choice of study regions around the airport should also reflect varying degrees of aircraft noise exposure (i.e., high exposure regions in close proximity to the runways and low exposure regions farther away from the runways). These study regions may also be classified according to additional criteria (e.g., socio-economic structure).

In the Appendix, we report air traffic volumes depending on time of day for the 35 Operational Evolution Partnership (OEP) airports (i.e., commercial U.S. airports with significant traffic activity). These 35 airports serve major metropolitan areas and also serve as hubs for airline operations. More than 70 percent of passengers move through these airports. These data can be used for the selection of study sites (although it may be necessary to acquire additional more comprehensive data).

4.8 Study Costs

The costs associated with a US field study on the effects of aircraft noise on sleep will be briefly discussed below. This is more a list that provides information on the most important study cost related aspects, not a detailed budget (the latter will depend on the details of a final study design and is out of the scope of this report).

Hardware and software: The number of required devices/licenses depends on the number of concurrent measurement sites and the number of investigators concurrently analyzing data. There needs to be some redundancy as measurement devices may fail and need to be replaced. Hardware and software requirements include:

- Sound level meters
- Actiwatches and ECG measuring devices (if possible combined in a single device)
- Automatic blood pressure measuring devices
- Laptops for the administration of questionnaires and data backup
- Electrodes
- Data analysis software (for actigraphy, sound pressure levels, questionnaires, statistical analyses, etc.)

Personnel: Qualified and trained personnel is required for all aspects of study design, data acquisition, data analysis, and publication. This includes

- Preparing documents to receive IRB approval,
- Subject recruitment,
- Informed consent briefing and instruction of subjects,
- Set up and de-installation of measurement equipment,
- On-call duty overnight,
- Data inspection and backup,
- Data preparation (including acoustical analysis) for statistical analysis,
- Statistical data analysis,
- Generation of intermediate reports and a final report, and

- Publication in the peer-reviewed literature.

Some of the tasks (e.g., measurement equipment set-up or data backup) can be performed by well-trained students/staff that are less expensive than scientists. One advantage of the proposed study design is that ECG electrodes can be easily attached and detached by the investigated subjects themselves, so personnel will only infrequently be required at the measurement site. As the ECG algorithm is automatic and objective, no trained personnel is required for this data analysis step.

Travel and accommodation of personnel: Personnel needs to travel to measurement sites to instruct subjects and install and de-install equipment. This will generate costs associated with rental cars and fuel. Depending on the measurement site, accommodation for personnel has to be provided (e.g., hotel room).

Subject reimbursement: Subjects are typically reimbursed for their study participation.

5 Acknowledgements and Disclosures

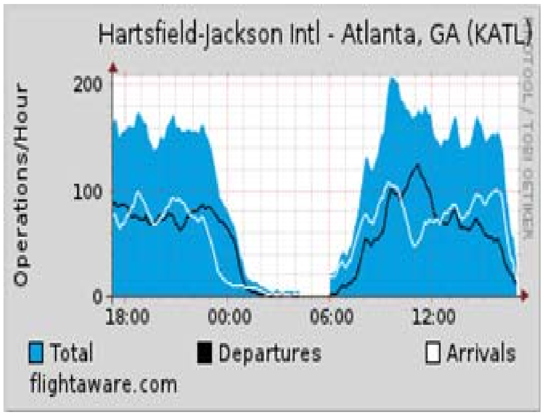
I would like to thank Mark Brink, Patricia Davies, Eva-Maria Elmenhorst, Barbara Griefahn, Ken Hume, Sarah McGuires, and Uwe Müller for reviewing a draft version of this report and providing me with helpful comments.

Dr. Basner is Associate Editor of the Journal SLEEP. Dr. Basner has received compensation for consulting from Purdue University for the FAA PARTNER Center of Excellence Project 25B in 2009. Dr. Basner has made paid presentations to the World Health Organization (WHO, German office).

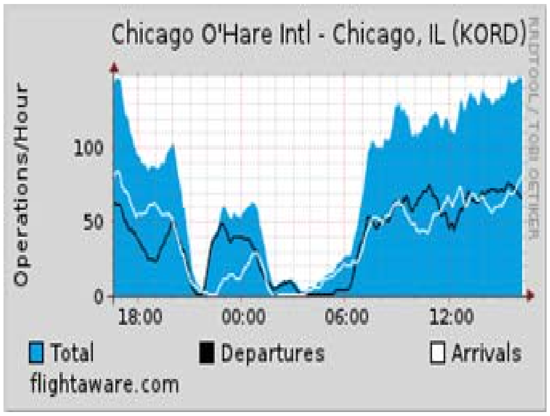
6 Appendix

In this appendix air traffic volumes are reported as a function of the time of day for the 35 Operational Evolution Partnership (OEP) airports (i.e., commercial U.S. airports with significant traffic activity). These 35 airports serve major metropolitan areas and also serve as hubs for airline operations. More than 70 percent of air passengers move through these airports.

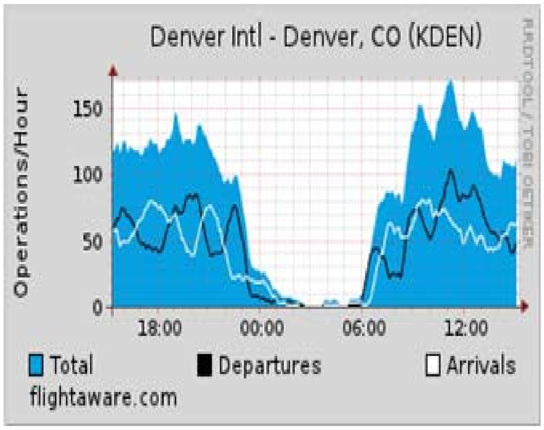
The number of flight operations on 06/22/2011 were extracted for each of these airports from FAA's Air Traffic Activity System (ATADS - http://aspm.faa.gov/opsnet/sys/Airport.asp). They are reported below in each of the graphs on the following pages. These graphs show the number of flight operations per hour for each of the 35 OEP airports within the past 24 hours relative to the time of data extraction. The graphs were extracted from flightaware.com on 06/22/2011 and are reproduced here with permission.



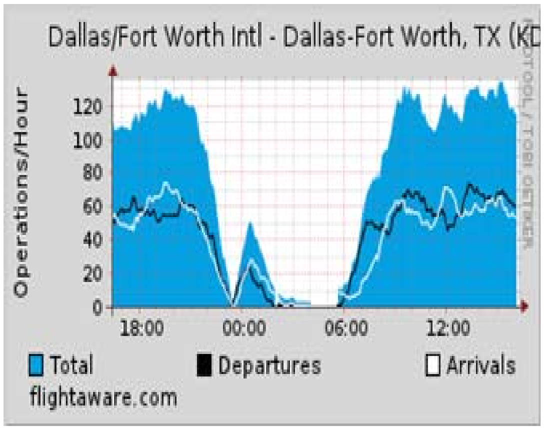
Operations on 06/22/2011: 2,600



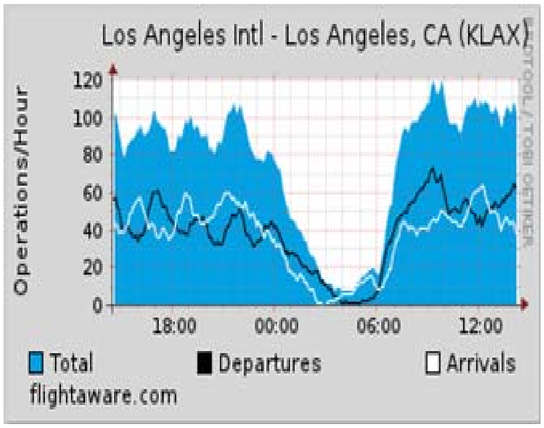
Operations on 06/22/2011: 2,278



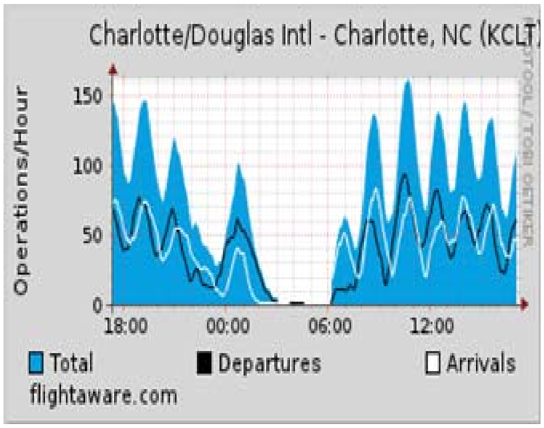
Operations on 06/22/2011: 1,945



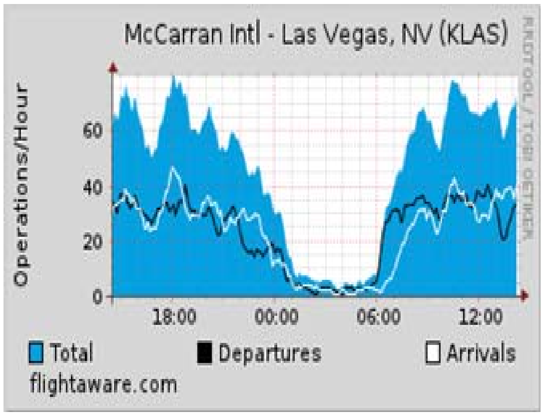
Operations on 06/22/2011: 1,932



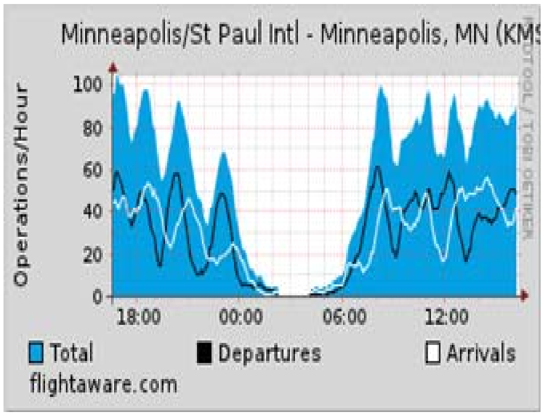
Operations on 06/22/2011: 1,832



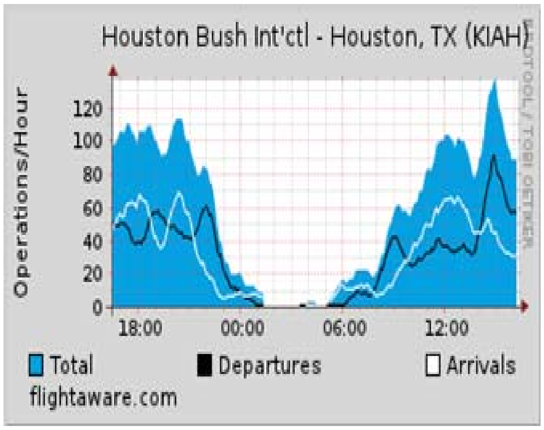
Operations on 06/22/2011: 1,594



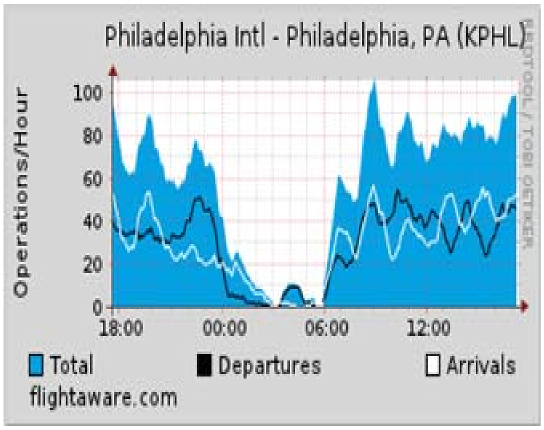
Operations on 06/22/2011: 1,567



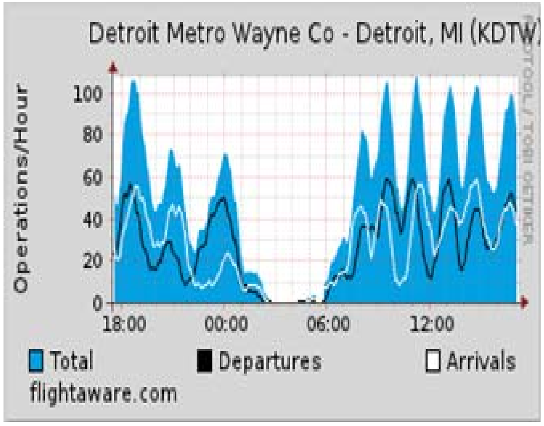
Operations on 06/22/2011: 1,401



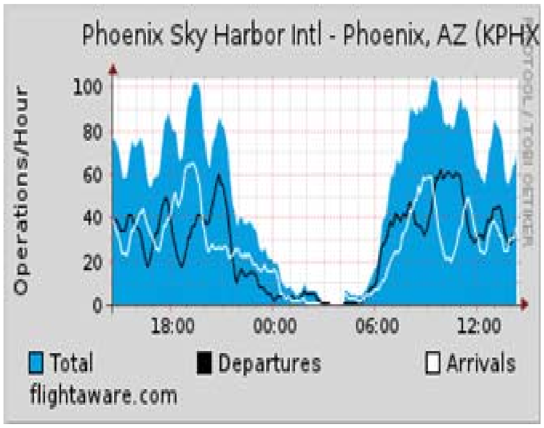
Operations on 06/22/2011: 1,398



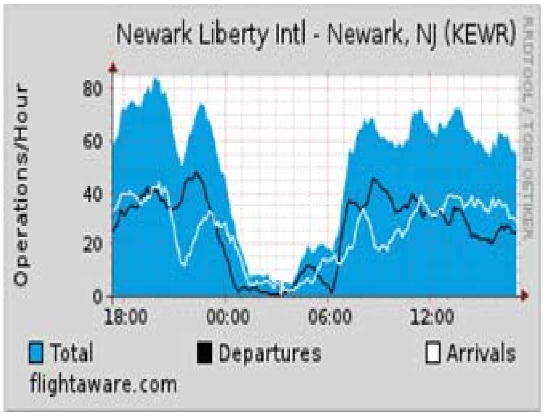
Operations on 06/22/2011: 1,345



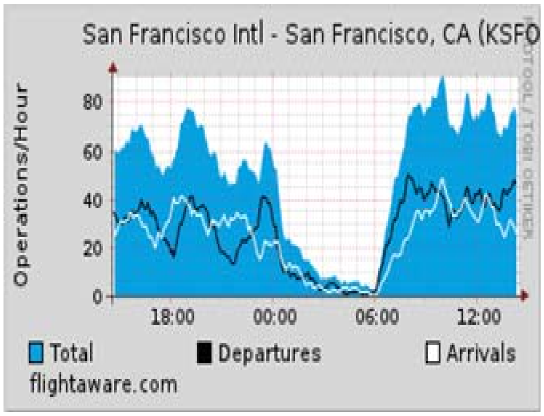
Operations on 06/22/2011: 1,331



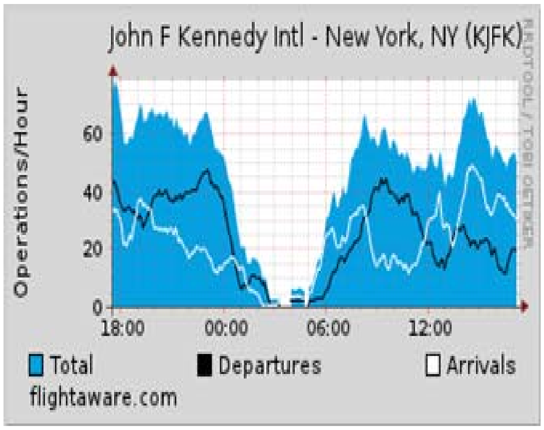
Operations on 06/22/2011: 1,300



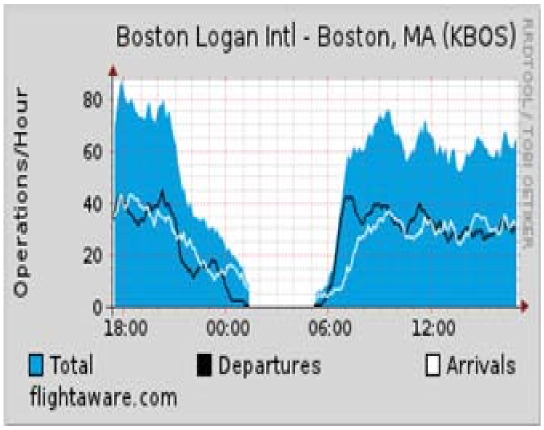
Operations on 06/22/2011: 1,226



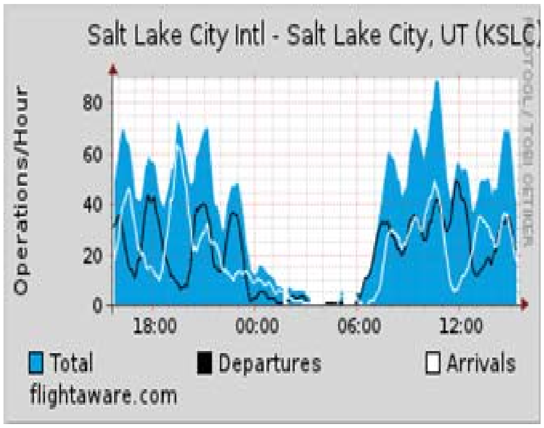
Operations on 06/22/2011: 1,222



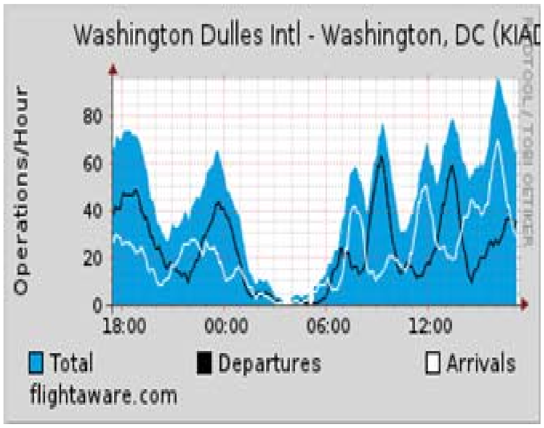
Operations on 06/22/2011: 1,190



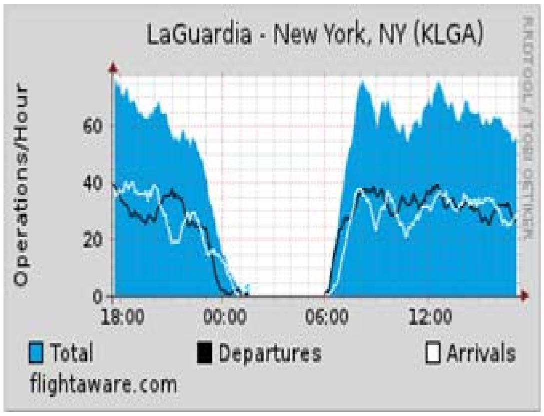
Operations on 06/22/2011: 1,174



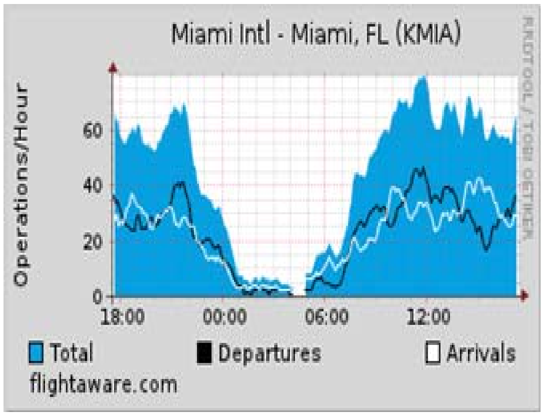
Operations on 06/22/2011: 1,174



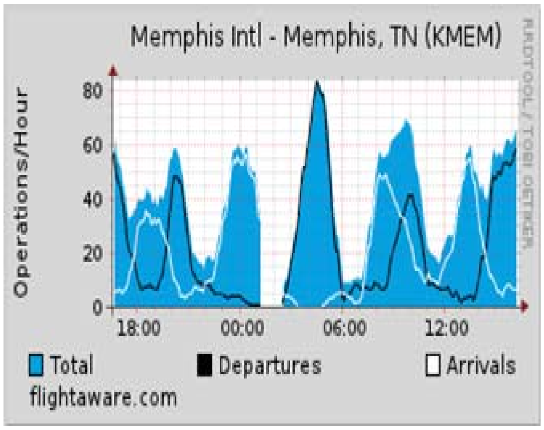
Operations on 06/22/2011: 1,127



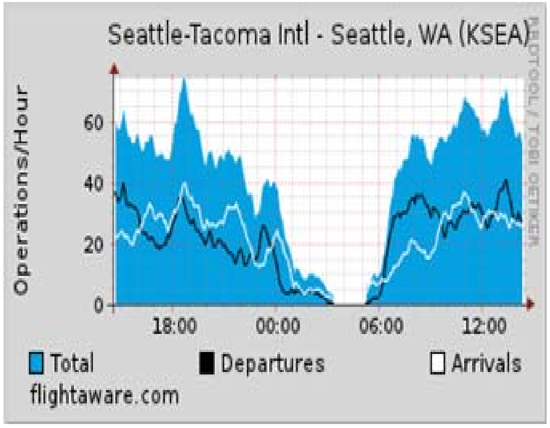
Operations on 06/22/2011: 1,102



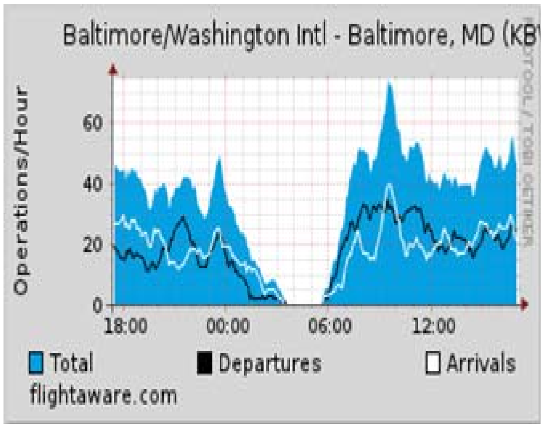
Operations on 06/22/2011: 1,101



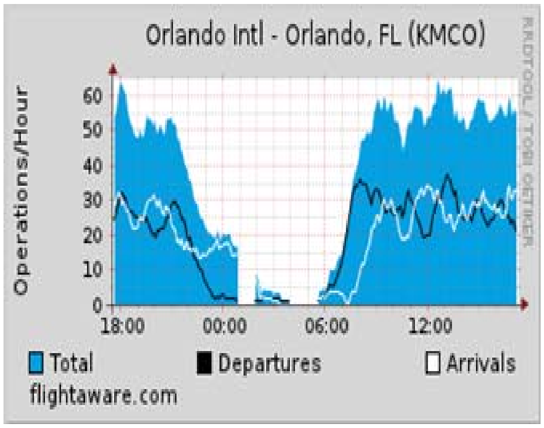
Operations on 06/22/2011: 1,017



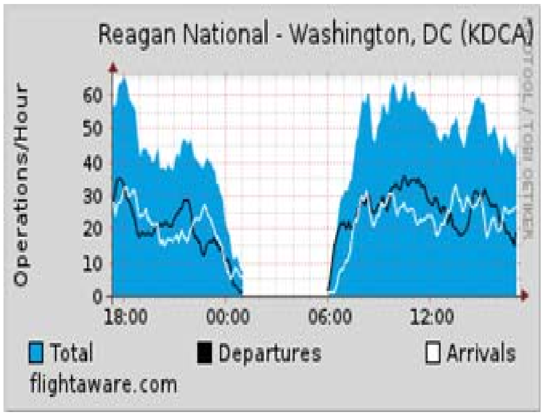
Operations on 06/22/2011: 992



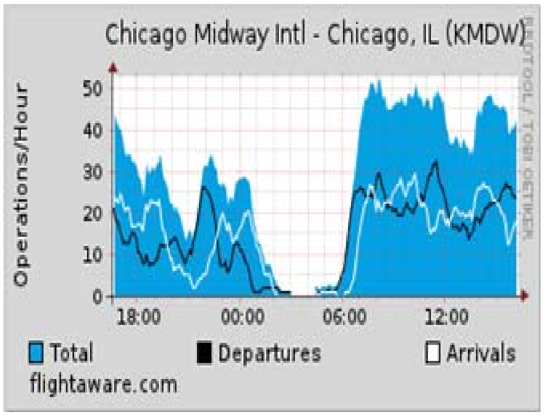
Operations on 06/22/2011: 890



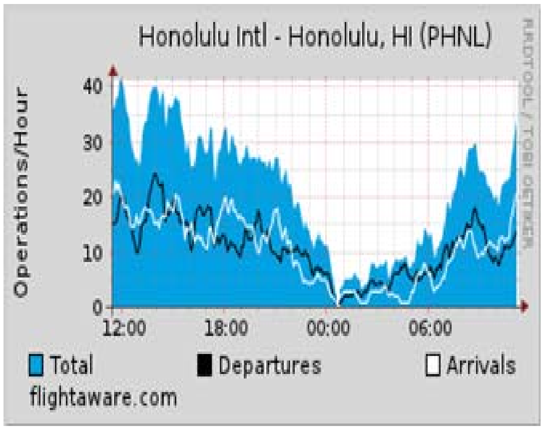
Operations on 06/22/2011: 886



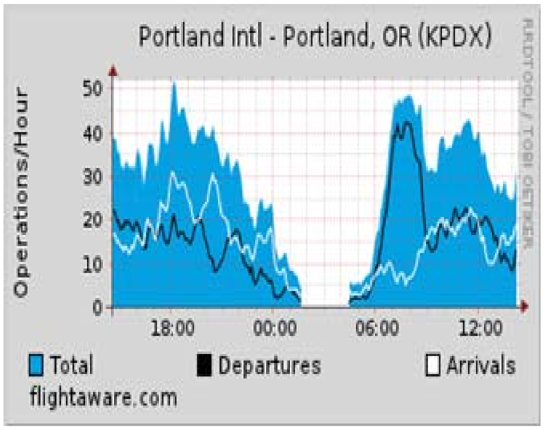
Operations on 06/22/2011: 868



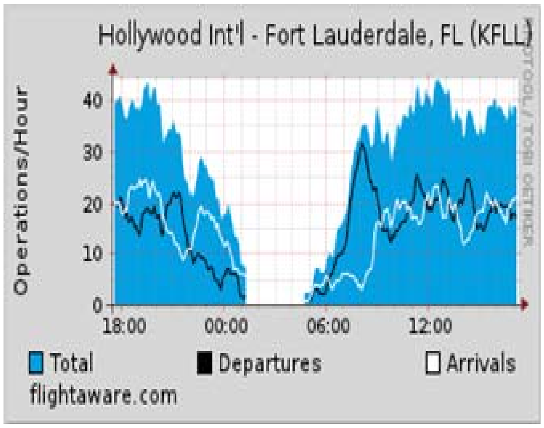
Operations on 06/22/2011: 833



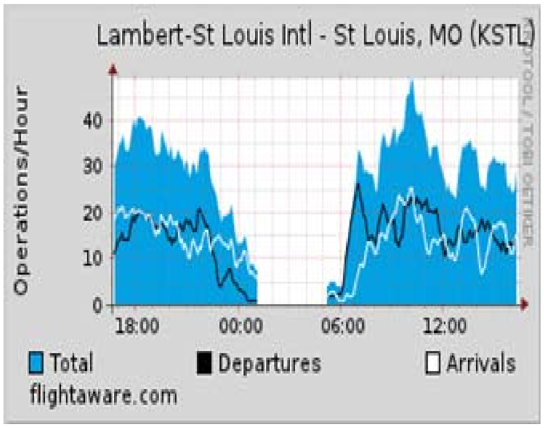
Operations on 06/22/2011: 737



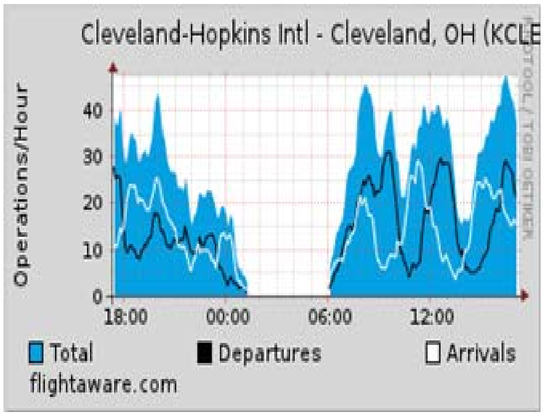
Operations on 06/22/2011: 693



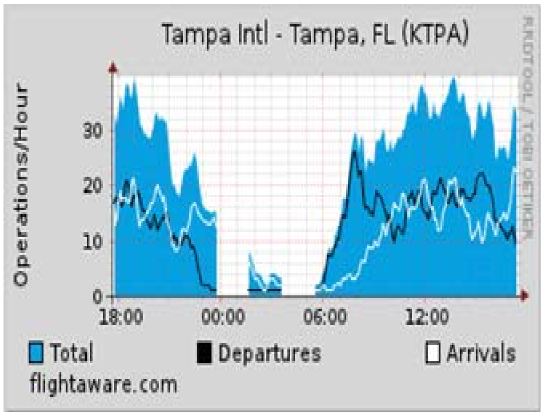
Operations on 06/22/2011: 673



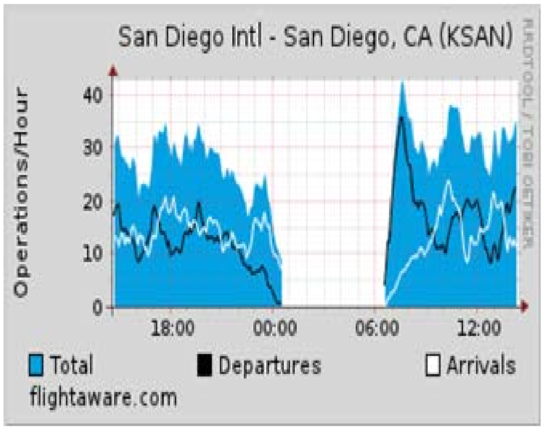
Operations on 06/22/2011: 616



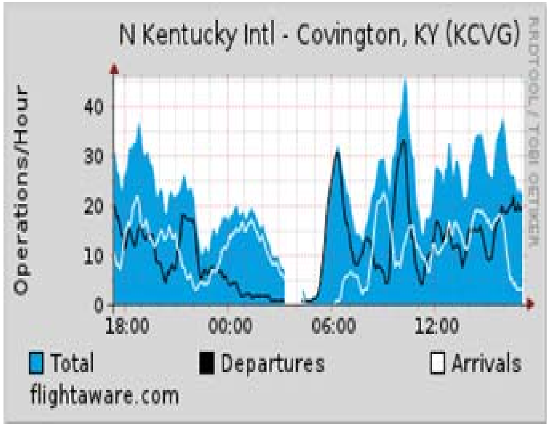
Operations on 06/22/2011: 586



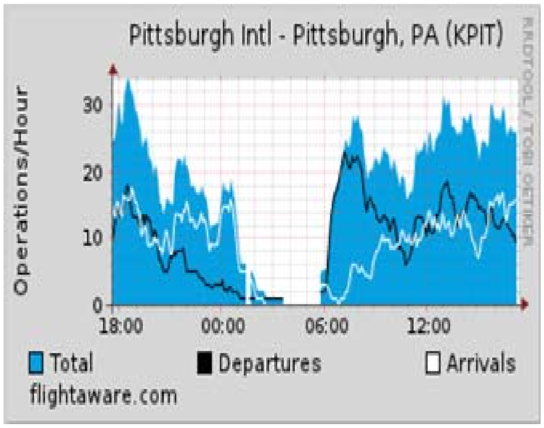
Operations on 06/22/2011: 548



Operations on 06/22/2011: 544



Operations on 06/22/2011: 512



Operations on 06/22/2011: 476

7 Literature

1. Banks S, Dinges DF. Behavioral and physiological consequences of sleep restriction. J.Clin.Sleep Med. 2007;3:519-28.
2. Oswald I, Taylor AM, Treisman M. Discriminative responses to stimulation during human sleep. Brain 1960;83:440-53.
3. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003;26:117-26.
4. Stickgold R. Sleep-dependent memory consolidation. Nature 2005;437:1272-8.
5. Wagner U, Gais S, Haider H, Verleger R, Born J. Sleep inspires insight. Nature 2004;427:352-5.
6. McKenna BS, Dickinson DL, Orff HJ, Drummond SPA. The effects of one night of sleep deprivation on known-risk and ambiguous-risk decisions. J.Sleep Res. 2007;16:245-52.
7. Basner M, Rubinstein J, Fomberstein KM, et al. Effects of night work, sleep loss and time on task on simulated threat detection performance. Sleep 2008;31:1251-9.
8. Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N.Engl.J Med. 2005;352:125-34.
9. Scott LD, Hwang WT, Rogers AE, Nysse T, Dean GE, Dinges DF. The relationship between nurse work schedules, sleep duration, and drowsy driving. Sleep 2007;30:1801-7.
10. Iber C, Ancoli-Israel S, Chesson A, Quan SF. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Westchester, Illinois: American Academy of Sleep Medicine, 2007.
11. Rechtschaffen A, Kales A, Berger RJ, et al. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Washington, D.C.: Public Health Service, U.S. Government, Printing Office, 1968.
12. Wesensten NJ, Balkin TJ, Belenky G. Does sleep fragmentation impact recuperation? A review and reanalysis. J.Sleep Res. 1999;8:237-45.
13. Bonnet MH, Doghramji K, Roehrs T, et al. The scoring of arousal in sleep: reliability, validity, and alternatives. J Clin.Sleep Med. 2007;3:133-45.
14. Basner M, Müller U, Elmenhorst EM, Kluge G, Griefahn B. Aircraft noise effects on sleep: a systematic comparison of EEG awakenings and automatically detected cardiac activations. Physiol.Meas. 2008;29:1089-103.
15. Basner M, Griefahn B, Müller U, Plath G, Samel A. An ECG-based algorithm for the automatic identification of autonomic activations associated with cortical arousal. Sleep 2007;30:1349-61.
16. Davis H, Davis PA, Loomis AL, Harvey EN. Electrical reactions of the human brain to auditory stimulation during sleep. J.Neurophysiol. 1939;2:500-14.
17. Dang-Vu TT, McKinney SM, Buxton OM, Solet JM, Ellenbogen JM. Spontaneous brain rhythms predict sleep stability in the face of noise. Curr.Biol. 2010;20:R626-R7.
18. Bonnet M, Carley DW, Carskadon MA, et al. EEG arousals: Scoring rules and examples. A preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep 1992;15:173-84.
19. Bonnet MH. Effect of sleep disruption on sleep, performance, and mood. Sleep 1985;8:11-9.

20. Bonnet MH. Performance and sleepiness as a function of frequency and placement of sleep disruption. Psychophysiology 1986;23:263-71.
21. Guilleminault C, Abad VC, Philip P, Stoohs R. The effect of CNS activation versus EEG arousal during sleep on heart rate response and daytime tests. Clin.Neurophysiol. 2006;117:731-9.
22. Bonnet MH. Differentiating sleep continuity effects from sleep stage effects (Letter to the editor). J.Sleep Res. 2000;9:403-4.
23. Bonnet M, Arand DL. EEG arousal norms by age. J.Clin.Sleep Med. 2007;3:271-4.
24. Raschke F, Fischer J. "Arousal" in der Schlafmedizin. Somnologie 1997;2:59-64.
25. Brink M, Basner M, Schierz C, et al. Determining physiological reaction probabilities to noise events during sleep. Somnologie 2009;13:236-43.
26. Martin SE, Wraith PK, Deary IJ, Douglas NJ. The effect of nonvisible sleep fragmentation on daytime function [see comments]. Am.J.Respir.Crit Care Med. 1997;155:1596-601.
27. Basner M, Van den Berg M, Griefahn B. Aircraft noise effects on sleep: Mechanisms, mitigation and research needs. Noise and Health 2010;12:95-109.
28. Griefahn B, Bröde P, Marks A, Basner M. Autonomic arousals related to traffic noise during sleep. Sleep 2008;31:569-77.
29. Huss A, Spoerri A, Egger M, Roosli M. Aircraft noise, air pollution, and mortality from myocardial infarction. Epidemiology 2010;21:829-36.
30. Sorensen M, Hvidberg M, Andersen ZJ, et al. Road traffic noise and stroke: a prospective cohort study. Eur Heart J 2011;32:737-44.
31. Basner M, Siebert U. Markov-Prozesse zur Vorhersage fluglärmbedingter Schlafstörungen. Somnologie 2006;10:176-91.
32. Quehl J, Basner M. Annoyance from nocturnal aircraft noise exposure: Laboratory and field-specific dose-response curves. J.Env.Psychol. 2006;26:127-40.
33. Basner M, Müller U, Elmenhorst E-M. Single and combined effects of air, road, and rail traffic noise on sleep and recuperation. Sleep 2011;34:11-23.
34. Basner M, Glatz C, Griefahn B, Penzel T, Samel A. Aircraft noise: effects on macro- and micro-structure of sleep. Sleep Med. 2007;9:382-7.
35. Griefahn B, Marks A, Robens S. Noise emitted from road, rail and air traffic and their effects on sleep. J.Sound Vib. 2006;295:129-40.
36. Basner M, Samel A. Effects of nocturnal aircraft noise on sleep structure. Somnologie 2005;9:84-95.
37. Agnew HW, Jr., Webb WB, Williams RL. The first night effect: an EEG study of sleep. Psychophysiology 1966;2:263-6.
38. Loredo JS, Clausen JL, Ancoli-Israel S, Dimsdale JE. Night-to-night arousal variability and interscorer reliability of arousal measurements. Sleep 1999;22:916-20.
39. Drinnan MJ, Murray A, Griffiths CJ, Gibson GJ. Interobserver variability in recognizing arousal in respiratory sleep disorders. Am.J.Respir.Crit Care Med. 1998;158:358-62.
40. Basner M, Griefahn B, Penzel T. Inter-rater agreement in sleep stage classification between centers with different backgrounds. Somnologie 2008;12:75-84.
41. Horne JA, Pankhurst FL, Reyner LA, Hume K, Diamond ID. A field study of sleep disturbance: effects of aircraft noise and other factors on 5,742 nights of actimetrically monitored sleep in a large subject sample. Sleep 1994;17:146-59.

42. Passchier-Vermeer W, Vos H, Steenbekkers JHM, Van der Ploeg FD, Groothuis-Oudshoorn K. Sleep disturbance and aircraft noise exposure - exposure effect relationships. Netherlands: TNO,. Report 2002.027, 2002.

43. Brink M, Muller CH, Schierz C. Contact-free measurement of heart rate, respiration rate, and body movements during sleep. Behav.Res.Methods 2006;38:511-21.

44. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26:342-92.

45. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. Sleep 2007;30:1362-9.

46. Hedner J, Pillar G, Pittman SD, Zou D, Grote L, White DP. A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. Sleep 2004;27:1560-6.

47. Stanley N. Actigraphy in human psychopharmacology: a review. Hum.Psychopharmacol. 2003;18:39-49.

48. Pollak CP, Tryon WW, Nagaraja H, Dzwonczyk R. How accurately does wrist actigraphy identify the states of sleep and wakefulness? Sleep 2001;24:957-65.

49. Blood ML, Sack RL, Percy DC, Pen JC. A comparison of sleep detection by wrist actigraphy, behavioral response, and polysomnography. Sleep 1997;20:388-95.

50. Sadeh A, Sharkey KM, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 1994;17:201-7.

51. Pitson D, Chhina N, Knijn S, van Herwaaden M, Stradling J. Changes in pulse transit time and pulse rate as markers of arousal from sleep in normal subjects. Clin.Sci.(Colch.) 1994;87:269-73.

52. Basner M, Isermann U, Samel A. Aircraft noise effects on sleep: Application of the results of a large polysomnographic field study. J.Acoust.Soc.Am. 2006;119:2772-84.

53. Sforza E, Pichot V, Cervena K, Barthelemy JC, Roche F. Cardiac variability and heart-rate increment as a marker of sleep fragmentation in patients with a sleep disorder: a preliminary study. Sleep 2007;30:43-51.

54. Thomas RJ, Mietus JE, Peng CK, Goldberger AL. An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. Sleep 2005;28:1151-61.

55. Fidell S, Pearsons K, Tabachnick BG, Howe R. Effects on sleep disturbance of changes in aircraft noise near three airports. J.Acoust.Soc.Am. 2000;107:2535-47.

56. Fidell S, Pearsons K, Tabachnick BG, Howe R, Silvati L, Barber D. Field study of noise induced sleep disturbance. J.Acoust.Soc.Am. 1995;98:1025-33.

57. Williams HL. Effects of noise on sleep: a review. In: Ward WD, editor: Proceedings of the International Congress on Noise as a Public Health Problem, 1973; Washington, D.C., p. 501-11, 1973.

58. Born J, Hansen K, Marshall L, Molle M, Fehm HL. Timing the end of nocturnal sleep. Nature 1999;397:29-30.

59. Silva GE, Goodwin JL, Sherrill DL, et al. Relationship between reported and measured sleep times: The Sleep Heart Health Study (SHHS). J.Clin.Sleep Med. 2007;3:622-30.

60. Ollerhead JB, Jones CJ, Cadoux RE, et al. Report of a Field Study of Aircraft Noise and Sleep Disturbance. London, United Kingdom: Department of Transport, 1992.

61. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Eur Heart J 1996;17:354-81.

62. Pizzuti GP, Cifaldi S, Nolfe G. Digital sampling rate and ECG analysis. Journal of Biomedical Engineering 1985;7:247-50.

63. Miedema HM, Vos H. Associations between self-reported sleep disturbance and environmental noise based on reanalyses of pooled data from 24 studies. Behav.Sleep Med. 2007;5:1-20.

64. Shambroom JR, Fabregas SE, Johnstone J. Validation of an automated wireless system to monitor sleep in healthy adults. Journal of Sleep Research 2011;21:221-30.

65. Elmenhorst EM, Elmenhorst D, Wenzel J, et al. Effects of nocturnal aircraft noise on cognitive performance in the following morning: dose-response relationships in laboratory and field. Int Arch Occup Environ Health 2010;83:743-51.

66. Marks A, Griefahn B, Basner M. Event-related awakenings caused by nocturnal transportation noise. Noise Contr.Eng.J. 2008;56:52-62.

67. Tassi P, Muzet A. Sleep inertia. Sleep Medicine Reviews 2000;4:341-53.

68. Basner M, Dinges DF. Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. Sleep 2011;34:581-91.

69. Basner M, Mollicone DJ, Dinges DF. Validity and sensitivity of a brief Psychomotor Vigilance Test (PVT-B) to total and partial sleep deprivation. Acta Astronautica 2011;69:949-59.

70. Wise MS. Objective Measures of Sleepiness and Wakefulness: Application to the Real World? J Clin.Neurophysiol. 2006;23:39-49.

71. Wilhelm B, Wilhelm H, Ludtke H, Streicher P, Adler M. Pupillographic assessment of sleepiness in sleep-deprived healthy subjects. Sleep 1998;21:258-65.

72. Basner M. Nocturnal aircraft noise increases objectively assessed daytime sleepiness. Somnologie 2008;12:110-7.

73. Sayk F, Becker C, Teckentrup C, et al. To dip or not to dip: on the physiology of blood pressure decrease during nocturnal sleep in healthy humans. Hypertension 2007;49:1070-6.

74. Carrington MJ, Trinder J. Blood pressure and heart rate during continuous experimental sleep fragmentation in healthy adults. Sleep 2008;31:1701-12.

75. Morrell MJ, Finn L, Kim H, Peppard PE, Badr MS, Young T. Sleep fragmentation, awake blood pressure, and sleep-disordered breathing in a population-based study. Am.J.Respir.Crit Care Med. 2000;162:2091-6.

76. Babisch W, Kamp I. Exposure-response relationship of the association between aircraft noise and the risk of hypertension. Noise.Health 2009;11:161-8.

77. Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Circulation 2005;111:697-716.

78. Verhaert V, De Bruyne G, De Wilde T, et al. Bed temperature and humidity during sleep in mild thermal conditions. Sleep 2011;34:A109.

79. Basner M, Siebert U. Markov processes for the prediction of aircraft noise effects on sleep. Med.Decis.Making 2010;30:275-89.

80. Guski R. Neuer Fluglärm gleich alter Fluglärm? Zeitschrift für Lärmbekämpfung 2003;50:14-25.