Forschungsbericht 2006-07

Markov State Transition Models for the Prediction of Changes in Sleep Structure Induced by Aircraft Noise

Mathias Basner

Deutsches Zentrum für Luft- und Raumfahrt
Institute of Aerospace Medicine
Cologne

159 pages
57 figures
10 tables
53 references

Deutsches Zentrum
für Luft- und Raumfahrt e.V.
in der Helmholtz-Gemeinschaft

Markov State Transition Models for the Prediction of Changes in Sleep Structure Induced by Aircraft Noise

Deutsches Zentrum für Luft- und Raumfahrt (DLR)
Institute of Aerospace Medicine
Flightphysiology Department
51170 Cologne
Germany

Cologne, July 2006

Director of the institute:
Prof. Dr. med. R. Gerzer

Author:
Dr. med. Mathias Basner

Head of the department:
Dr. A. Samel

This report will be published simultaneously by the Faculty of Health Sciences of the University of Bielefeld, Germany as a Master Thesis for the graduation of the author as Master of Science in Epidemiology (MSc).

aircraft noise, sleep, EEG, Markov model, autoregressive multinomial logistic regression

Mathias BASNER
Institute of Aerospace Medicine at the German Aerospace Center (DLR), Cologne

Markov State Transition Models for the Prediction of Changes in Sleep Structure Induced by Aircraft Noise

DLR-Forschungsbericht 2006-07, 2006, 159 pages, 57 figures, 10 tables, 53 references

OBJECTIVE: To quantitatively assess the effects of the introduction of a noise-free period at Frankfurt Airport between 11 pm and 5 am on sleep structure. METHODS: A six state (Wake, S1, S2, S3, S4 and REM) Markov state transition sleep model was built. Transition probabilities between states were calculated with autoregressive multinomial logistic regression based on polysomnographic laboratory studies, where 128 subjects were investigated for 13 consecutive nights. First-order Monte Carlo simulation trials were performed for modelling a noise-free night and three noise scenarios: (1) traffic at Frankfurt Airport on 16 August 2005, (2) as (1), but flights between 11 pm and 5 am cancelled and (3) as (2), with flights between 11 pm and 5 am from (1) rescheduled to periods before 11 pm and after 5 am. RESULTS: The results of the models indicate that there will be a small benefit for airport residents in terms of sleep structure even if all traffic is rescheduled to periods before 11 pm and after 5 am. This benefit is likely to be outweighed by the increase of air traffic during shoulder hours, especially for those who choose to or have to go to bed before 10:30 pm or after 1 am. CONCLUSIONS: Alternative strategies might be necessary to both guarantee undisturbed sleep of airport residents and to minimize economic and legal disadvantages accompanied by a ban of air traffic between 11 pm and 5 am. The models developed in this thesis may serve as a valuable tool for optimizing air traffic patterns at Frankfurt Airport, and therefore guide political decision making.

Fluglärm, Schlaf, EEG, Markov Modelle, autoregressive multinomiale logistische Regression
(in englischer Sprache veröffentlicht)

Mathias BASNER
Institut für Luft- und Raumfahrtmedizin des DLR, Köln

Markov Prozesse zur Vorhersage fluglärmbedingter Schlafstrukturänderungen

DLR-Forschungsbericht 2006-07, 2006, 159 Seiten, 57 Bilder, 10 Tabellen, 53 Literaturstellen

ZIEL: Die quantitative Erfassung von Schlafstrukturänderungen hervorgerufen durch die Einführung eines Nachtflugverbotes von 23 Uhr bis 5 Uhr am Frankfurter Flughafen. METHODIK: Ein aus sechs Zuständen (Wach, S1, S2, S3, S4 und REM) bestehendes Markov Schlafmodell wurde entwickelt. Die Übergangswahrscheinlichkeiten zwischen den Zuständen wurden mit autoregressiven multinomial logistischen Regressionsmodellen basierend auf den Daten einer Laborstudie, in der 128 Versuchspersonen für 13 aufeinanderfolgende Nächte polysomnographisch untersucht wurden, berechnet. Monte Carlo Simulationen wurden zur Modellierung einer lärmfreien Basisnacht und von drei Fluglärmszenarien durchgeführt: (1) Flugverkehr am Frankfurter Flughafen am 16. August 2005, (2) wie (1), Flüge zwischen 23 Uhr und 5 Uhr wurden jedoch gestrichen und (3) wie (2), wobei Flüge, die in (1) zwischen 23 Uhr und 5 Uhr stattfanden, auf Zeiträume vor 23 Uhr und nach 5 Uhr verlegt wurden. ERGEBNISSE: Basierend auf den Ergebnissen der Modellrechnung wird die Schlafstruktur der Flughafenanrainer von der Einführung eines Nachtflugverbots profitieren, selbst wenn alle Flüge, die vor Einführung eines Nachtflugverbotes zwischen 23 Uhr und 5 Uhr stattfanden, auf Zeiträume vor 23 Uhr und nach 5 Uhr verlegt werden. Der Profit wird jedoch wahrscheinlich gering sein, und die Zunahme des Flugverkehrs in den Tagesrandzeiten wird diesen Profit schnell aufwiegen, insbesondere für diejenigen, die gewöhnlich vor 22:30 Uhr oder nach 1 Uhr ins Bett gehen (müssen). SCHLUSSFOLGERUNGEN: Es sollte über alternative Strategien nachgedacht werden, um sowohl einen ungestörten Schlaf der Flughafenanrainer zu garantieren als auch die rechtlichen und ökonomischen Nachteile, die mit der Einführung eines Nachtflugverbotes am Frankfurter Flughafen verbunden sind, zu minimieren. Die in dieser Arbeit entwickelten Modelle können als Werkzeug eingesetzt werden, um Flugverkehrsmuster am Frankfurter Flughafen zu optimieren und damit um den Prozess der politischen Entscheidungsfindung zu begleiten.

INDEX

1 Summary

INTRODUCTION Aircraft noise may cause changes in sleep structure and impair recuperation. People living in the vicinity of airports are very concerned about possible short- and long-term effects of a chronically disturbed sleep on health. Frankfurt Airport plans to build a new runway in order to meet increasing traffic demands. The airport applied for a ban of air traffic between 11 pm and 5 am in order to compensate people living in the highly populated vicinity of Frankfurt Airport for increased traffic volumes during the day. Some of the flights starting or landing between 23:00 and 05:00 today are likely to be rescheduled to periods before 23:00 or after 05:00. Therefore, it is unclear whether and to what extend sleep of airport residents will benefit from an air traffic ban in the night.

OBJECTIVES To quantitatively assess the impact of the introduction of a ban of air traffic at Frankfurt Airport between 23:00 and 05:00 on sleep structure using epidemiological and decision-analytic methods. The main question was, whether a ban of air traffic will still be beneficial for residents in terms of sleep structure if all nighttime air traffic is rescheduled to periods before 23:00 and after 05:00.

METHODS Analyses were based on a polysomnographic laboratory study on the effects of aircraft noise on sleep performed at the German Aerospace Center (DLR) in Cologne between 1999 and 2003. A six state (sleep stages Wake, 1, 2, 3, 4 and REM) Markov state transition model was used to simulate nights with and without aircraft noise. Transition probabilities for the Markov models were calculated with autoregressive multinomial logistic regression. Both regression and simulation results were validated with empirical data. Different times of falling asleep were considered in the analyses. The outcome variables "time spent in the different sleep stages", "number of sleep stage changes" and a newly introduced "sleep quality index" were used as typical indicators for sleep structure. Three noise scenarios were compared with a noise-free night and

with each other: (1) the current situation in Frankfurt with nocturnal air traffic, (2) a ban of air traffic between 23:00 and 05:00 and (3) as (2), but with flights that took place between 23:00 and 05:00 in (1) rescheduled to periods before 23:00 and after 05:00.

RESULTS A first-order autoregressive multinomial logistic regression model with elapsed sleep time as the only additional explanatory variable was used for the calculation of transition probabilities. Monte Carlo simulation trials showed that a ban of air traffic without rescheduling of flights lead to sleep structural improvements. These were diminished if traffic that currently takes place between 23:00 and 05:00 was rescheduled to periods before 23:00 and after 05:00, but still, a small benefit remained: Compared to the current situation without a ban of air traffic, average time spent awake decreased from 43.1 to 41.7 min (-3.2%), S1 decreased from 9.2 to 8.7 min (-4.6%), S2 decreased from 212.8 to 210.8 min (-0.9%), S3 increased from 37.2 to 38.3 min (+3.0%), S4 increased from 23.5 to 25.7 min (+9.2%), REM increased from 84.7 to 85.3 min (+0.6%), the number of sleep stage changes decreased from 121.3 to 118.3 (-2.5%) and sleep quality improved by 0.8%. These results were weighted according to the number of people falling asleep at specific times. In contrast to that, unweighted results showed that the impact of the time of falling asleep on sleep structure was much stronger than the traffic scenario itself. For example, the largest difference in time spent awake was observed within scenario 3, where it increased from 38.2 min when falling asleep at 23:15 to 57.9 min (+51.5%) when falling asleep at 20:45.

DISCUSSION If a ban of air traffic between 23:00 and 05:00 is introduced at Frankfurt Airport, it will be beneficial for sleep structure of affected people even if all traffic is rescheduled to periods before 23:00 and 05:00 (a worst case scenario). However, the expected benefits are rather small. At the same time, the results of the analyses stress the importance of air traffic during shoulder hours, which will increase in case of an expansion of Frankfurt Airport, both because of a general increase of traffic and because of flights rescheduled from the period between 23:00 and 05:00. Several limitations have to be borne in mind for the interpretation of the results. They are discussed in detail in the thesis.

CONCLUSIONS The results of this thesis indicate that the small sleep structural benefits of the introduction of a noise-free period between 23:00 and 05:00 are likely to be outweighed by far by the impact of air traffic during shoulder hours. Simultaneously, a ban of air traffic between 23:00 and 05:00 will be accompanied by severe economic and legal disadvantages. Therefore, alternative strategies might be necessary to both guarantee undisturbed sleep of airport residents and to minimize economic and legal disadvantages. The models developed in this thesis may serve as a valuable tool for optimizing air traffic patterns at Frankfurt Airport, and therefore guide political decision making.

2 Introduction

Nighttime air traffic generates noise, that may lead to sleep disturbances and therefore reduce the restorative power of sleep. This may result in short term (e.g. increased sleepiness, reduced performance) or even long term (e.g. higher risk of myocardial infarction [5, 49]) consequences for health.

In Germany, there are only a few airports with relevant amounts of nocturnal air traffic. One of them is Frankfurt Airport. Because of the ever increasing demand for air traffic, Frankfurt Airport authorities (Fraport AG) plan the construction of a new runway ("Landebahn Nordwest"). Residents living in the highly populated vicinity of Frankfurt Airport are very concerned about the health effects of increasing air traffic, that would be associated with the construction of a new runway. They are especially concerned about the effects of nighttime aircraft noise on sleep.

Therefore, a round table discussion (*Mediation*) consisting of representatives of the airport, carriers, air traffic control, residents, state departments, local cities, trade associations and labor unions was initiated in 1998 in order to establish the best way to proceed with a possible expansion of Frankfurt Airport. In January 2000, the Mediation group recommended in their final report the introduction of a noise-free period from 23:00 until 05:00 as one of five main concomitant measures in case of an expansion of the airport [2]. If the new runway is built, no airplane should be allowed to take off or to land during this period.

In the final report, the following can be read on page 179 [2] (translation from German by the author): "The protection of airport residents against excessive noise is a priority. Therefore, the mediation committee thinks that a prohibition of nocturnal air traffic is indispensable. It recommends a period without air traffic from 23:00 until 05:00. Additionally, the mediation committee supports taking measures of noise reduction for the protection of other very sensitive periods. The prohibition of nocturnal air traffic demands a shift of mail-, freight and charter-flights. This may be accomplished by timetable changes or by relocations to other airports, e.g. to Hahn airport."

The authors of the final report must have been absolutely confident that a noise-free period from 23:00 to 05:00 constitutes a benefit for the residents affected by aircraft noise, as their suggestion is substantiated nowhere in the final report, although several expert hearings were conducted and expert opinions were gathered during the mediation process. These experts suggested several counter measures against the adverse effects of nocturnal aircraft noise, but, astonishingly, the introduction of a noise-free period from 23:00 to 05:00 was not among them [1].

The mediation process is continued by an institution called "Regionales Dialogforum". Several working groups have been established. One of them is called "Nachtflugverbot" (prohibition of nocturnal air traffic). Its main task is the elaboration of possible consequences of the introduction of a noise-free period during the night. Interestingly, the results that have been presented by the working group so far show that it seems to focus its work on economic and legal consequences rather than on the potential benefits or harms of a noise-free period for the affected population.

The expert opinions that have been gathered on the economic and legal consequences revealed several severe disadvantages of a period without air traffic between 23:00 and 05:00:

- Carriers that currently operate predominantly during the night will face major competitive disadvantages, especially those that use Frankfurt Airport as their home base (e.g. Lufthansa Cargo and Condor) [32].

- Frankfurt Airport will no longer be able to maintain a hub for mail deliveries [32]. Direct and indirect jobs will be lost and the existing infrastructure will no longer be needed. It will not be possible to participate in the steadily growing market of mail and freight traffic during the night. Collection and dispatch times of mail are likely to be impaired [23].

- Flights that formerly took place during the night have to be rescheduled, relocated or cancelled, with total costs of several 100 millions of Euros over the next years [23, 32].

- It is unclear whether it is legal to prohibit air traffic between 23:00 and 05:00 [27]. Although unlikely, Frankfurt may even lose its attribute "International Airport" because of a prohibition of nocturnal air traffic [3].

Altogether, there will be serious economic disadvantages if air traffic is banned between 23:00 and 05:00. Therefore, before such a decision is made, there should be absolutely no doubt that affected residents indeed benefit from a noise-free period and that no other superior alternative strategy exists which is as effective or more effective without major restrictions of nocturnal air traffic. Although there is considerable lack of knowledge about the consequences of the introduction of a ban of air traffic from 23:00 until 05:00, Fraport AG applied on 2 November 2004 for the expansion of Frankfurt Airport and for the introduction of a prohibition of air traffic between 23:00 and 05:00 [3], as suggested by the mediation committee [2].

On a hearing of the "Regionales Dialogforum" in February 2003, international scientists active in the field of noise effects research were asked about the consequences of the introduction of a noise-free period between 23:00 and 05:00 in Frankfurt. It was obvious that none was able to deliver more than speculations about what will happen because of the following reason: If a period without traffic is implemented during the night, air traffic during shoulder hours, i.e. before 23:00 and after 05:00, is likely to increase [23]. Up to now, it is not possible to assess which of both patterns, "noise-free period with increased traffic during shoulder hours" vs. "present situation with nocturnal air traffic", is more beneficial for the sleep of airport residents. Of course, the establishment of a noise-free period without increasing air traffic during shoulder hours would be undoubtedly beneficial, but this is unlikely to happen.

Therefore, a major goal of this thesis was to quantitatively assess the effects of the introduction of a noise-free period between 23:00 and 05:00 on sleep, and therefore add relevant information to the existing qualitative expert opinions. To accomplish this goal, the influence of noise on sleep was mathematically modeled. Different models were used to compare three noise scenarios with each other and with an undisturbed night: Noise scenario 1 consisted of the traffic that took place on a reference date in 2005 at Frankfurt Airport. Scenario 2 was similar to scenario 1, but flights between 23:00 and 05:00 were cancelled. Scenario 3 was similar to scenario 2, but flights that formerly took place between 23:00 and 05:00 in scenario 1 were rescheduled to periods before 23:00 and after 05:00.

3 Background and current knowledge

In the following paragraphs, important background knowledge for the understanding of the work that has bee accomplished in this thesis will be presented. Chapter 3.1 deals with the phenomenon "sleep" in general, explains its functions and how it is measured. Chapter 3.2 shows how traffic noise may interfere with sleep, leading to changes in sleep structure with possible implications for the recuperative value of sleep. Reasons for the growth of air traffic especially during the night are presented and past efforts of noise effects research are demonstrated. Chapter 3.3 gives a short introduction to Markov processes. Finally, Chapter 3.4 shows why Markov state transition models may be easily applied in the context of sleep structure modeling. Existing models of human sleep - including their theory - will be presented.

3.1 Sleep – its functions and its measurement

Humans spend about a third of their lives sleeping. Sleep itself is a condition where consciousness, i.e. the mind's perception of itself and of its surroundings, is reduced.

In order to distinguish whether someone is sleeping or just awake with eyes closed, a brain current diagram is needed. The German psychiatrist Hans Berger was the first person to measure electrical activity of the cerebral cortex by the use of electrodes attached to the scalp [11]. He called this method e*lectroencephalography* (EEG), establishing the foundation of modern sleep medicine: Soon, the difference between the EEG of waking and sleeping subjects was noticed. It was observed that while sleeping, the amplitude of the brain waves was higher than in those the EEG registered in persons being awake.

Today, in a modern sleep facility, parameters of sleeping persons are measured continuously throughout the night. It is necessary for the classification of sleep into different stages [42] to record the so-called *electrooculogram* (EOG) and *electromyogram* (EMG) additionally to the EEG (see Figure 3.1).

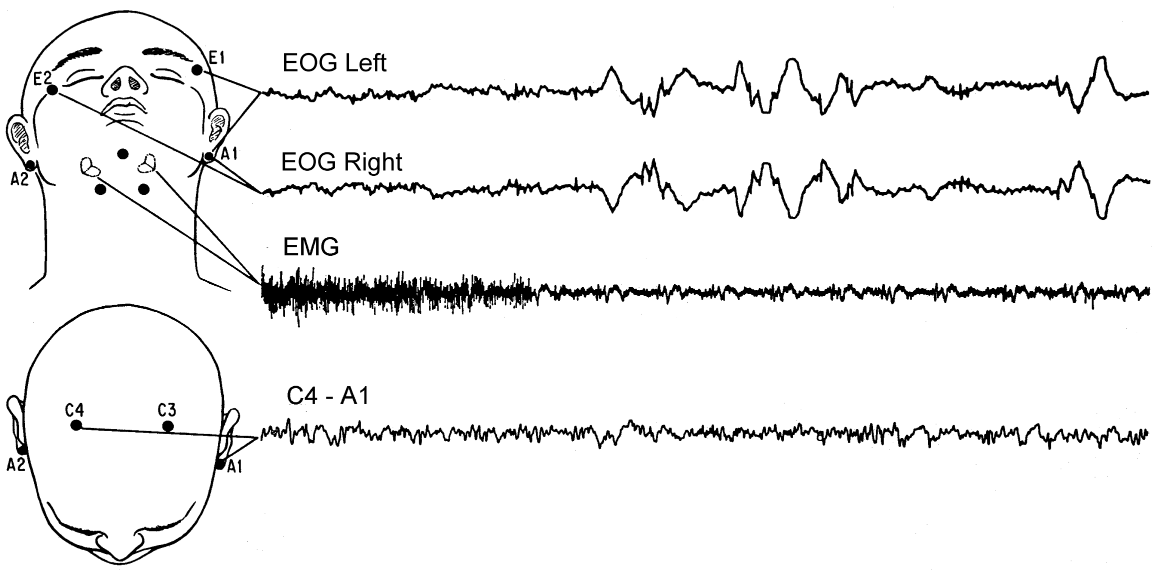


Figure 3.1: Signals and electrode positions used for the classification of sleep stages according to the criteria of Rechtschaffen und Kales [42].

The EOG registers the movement of the eyeballs. This method makes use of the fact that - unlike the cornea - the retina is charged negatively. As the electrodes are attached in the proximity of the eyeballs, changes in electrical potentials can be registered when the eyes move.

During the process of falling asleep, a slow and rolling movement of the eyes can be observed. The so-called *REM sleep* (rapid eye movement sleep) was named after fast conjugated movements of the eyeballs, which otherwise can only be observed while being awake. As the EEG of REM sleep and stage 1 sleep (see below) are very similar, the EOG is needed to differentiate between them.

The EMG is an indicator for the muscular tension. The latter decreases with increasing depth of sleep. The lowest muscle tension can be found during

REM sleep, whereas during wakefulness the muscular tension is highest. For the EMG recording, two electrodes are attached to the skin above the muscles of the chin.

Using the three parameters EEG, EOG and EMG, it is possible to distinguish three distinct conditions of consciousness. Consciousness can be classified into wakefulness and sleep. Sleep is subdivided into REM sleep and *NREM sleep* (non-REM sleep). NREM sleep is again subdivided into stage 1, stage 2, stage 3 and stage 4 sleep (in the following abbreviated as S1, S2, S3 and S4). The factors relevant for distinguishing between the different sleep stages are frequency and amplitude of the EEG, on the one hand, and the occurrence of special EEG signatures (e.g. spindles, vertex sharp waves etc.) typical for the appropriate stages of sleep, on the other hand (see Figure 3.2).

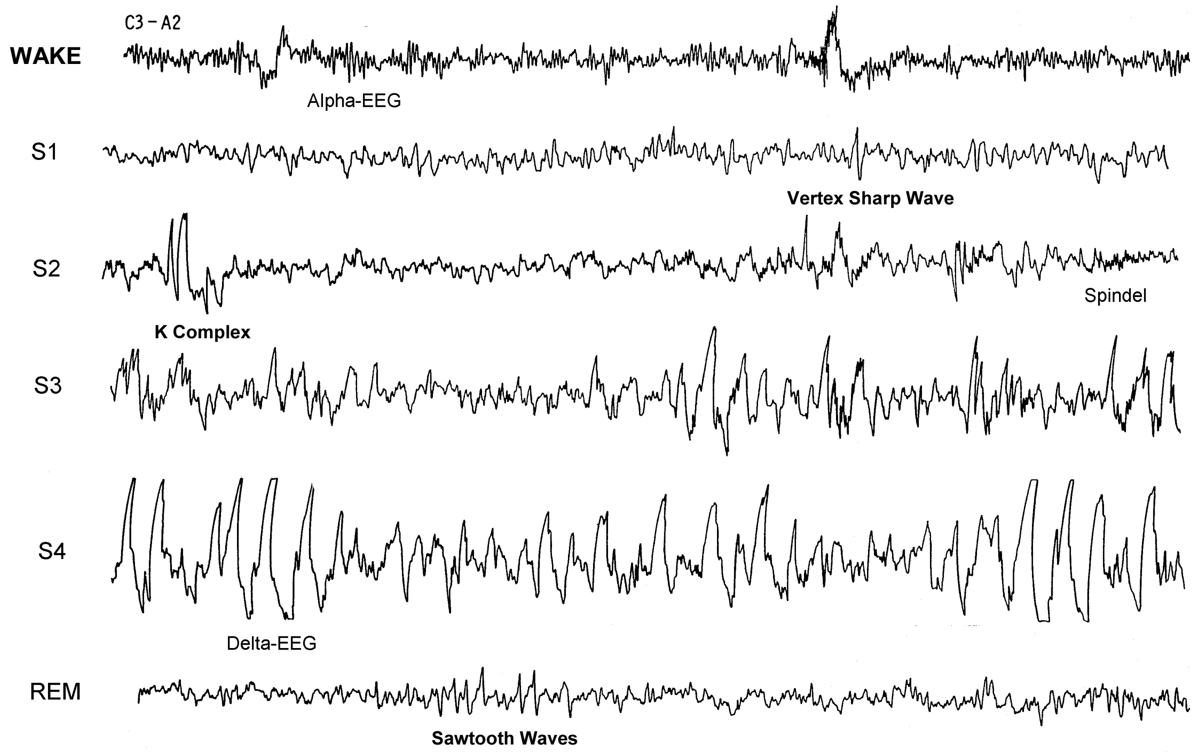


Figure 3.2: Typical EEG patterns of the different sleep stages [42].

The sleep stage classification has been defined by a commission of experts led by A. Rechtschaffen and A. Kales in 1968 [42], and hence the results of different sleep facilities have been comparable since that time.

The night is subdivided into periods of 30 seconds (a so-called *epoch*). A trained evaluator then assigns one of the stages of sleep or wakefulness, i.e. stage Wake, to each of the 30 second periods. The organization into periods of 30 seconds is arbitrary and based on the fact that during the early days of sleep medicine, the EEGs were recorded on paper: With a paper speed of 10 mm/s, exactly 30 seconds fit on one page.

Based on this 30 second classification, a night that lasts 8 hours contains 960 epochs, which have to be assessed individually. The architecture of sleep can be illustrated by plotting the different sleep stages against time. This plot is called a *hypnogram*. A hypnogram of a healthy young volunteer is shown in Figure 3.3.

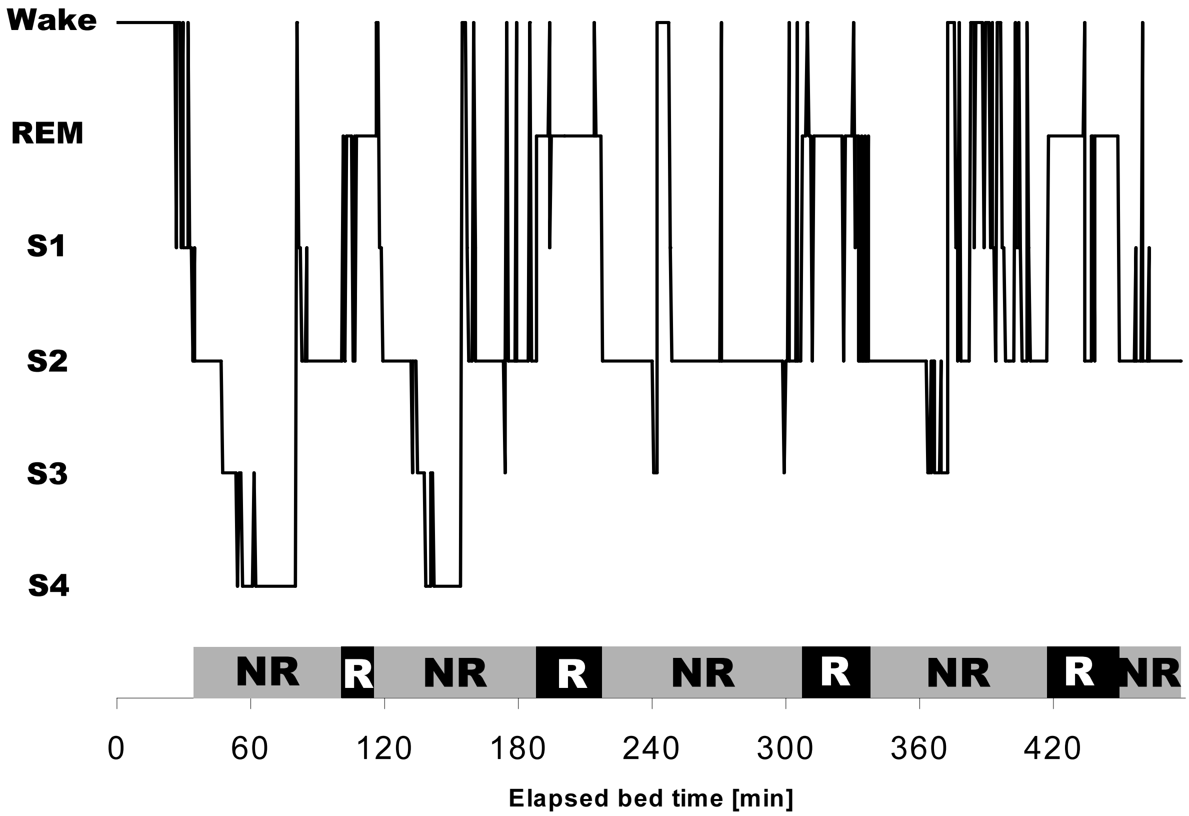


Figure 3.3: Typical hypnogram of a healthy young volunteer in the course of a noise-free baseline night (23:00 to 07:00). NREM-periods (NR) and REM-periods (R) are indicated by gray and black bars at the bottom of the hypnogram.

After falling asleep (sleep onset, defined as the first appearance of S2), the healthy young sleeper will usually quickly slide into sleep stages 3 and 4, also called *deep sleep* or *slow wave sleep (SWS)*. In S3 and S4, the EEG is synchronized with a high amplitude and a low frequency (see Figure 3.2). The designation *deep sleep* is based on the fact that the probability of being woken up by external stimuli in S3 and S4 is smaller than during so-called *light sleep* (stages 1 and 2) [48]. After a period of deep sleep (NREM-episode), human sleep will usually abruptly become lighter and be followed by a period of REM sleep (REM-sleep episode).

With the end of the first REM episode, the first sleep cycle is finished. One sleep cycle lasts approximately 90 to 120 minutes. 3 to 6 sleep cycles occur during the night. Deep sleep is more prominent in the first half of the night, whereas REM sleep dominates the second half. Even healthy subjects wake up frequently during the night (about 20 to 30 times) [9]. Usually, these periods are too short to be remembered in the morning. Very short (3 to 15 seconds) waking reactions are called *arousals* [12]. These can be observed several times per hour in the healthy sleeper, as well.

The restorative power of sleep is influenced by sleep duration as well as sleep structure. Present scientific knowledge assumes that sleep stages differ in their recuperative value, although the exact functions and mechanisms of the different sleep stages are still unknown. SWS is considered to be particularly important for the restorative power of sleep because of the following reasons [13]:

- its proximity to sleep onset
- its immediate rebound after sleep deprivation
- its association with high sensory thresholds and the excretion of growth hormones

Scientific results of the recent past indicate that SWS is involved in the consolidation of explicit memory contents, whereas REM sleep seems to be important for the consolidation of implicit memory contents [41, 44]. Wake and S1 do not contribute to the recuperative value of sleep or only very little, respectively [47], whereas sleep stage S2 takes an intermediate position.

Selective withdrawal of SWS or REM sleep can lead to severe performance deficits during the next day. This can be observed in patients with the so-called *obstructive sleep apnea syndrome* (OSAS). Because of a collapse of the upper airways, sleep of OSAS patients is frequently disrupted and amounts of SWS and REM sleep can be tremendously diminished up to the point where the night consists of light sleep only. Excessive daytime sleepiness, impairments of memory and concentration and mood swings are some of the complaints of these patients. Additionally, it is very likely that OSAS relevantly contributes to the prevalence of high blood pressure and the associated diseases of the cardiovascular system (e.g. stroke, myocardial infarction) [18, 28, 51].

In the newest past of noise effects research, there is the trend to minimize operating expenses for the single examined subject in order to raise sample sizes, e.g. by the substitution of the sumptuary polysomnography with actigraphy. The actigraph is a small device worn at the wrist of one arm. It records accelerations caused by body movements during the night. The amount and amplitude of these accelerations are said to be an indicator of sleep quality, an increasing number of movements indicating a lower quality of sleep. Major disadvantages of the actigraphic assessment of sleep are:

- the percentages of deep and REM sleep, which are very important for the restorative function of sleep, cannot be assessed properly,
- changes in the microstructure of sleep as brief arousals or sleep stage shifts, that are not accompanied by movements, cannot be detected,
- normal sleep may be accompanied by movements whereas longer periods spent awake without body movements may appear as undisturbed sleep.

3.2 Aircraft noise induced sleep disturbances

In Germany, air traffic has increased about sevenfold from the early sixties until now and even higher increments are predicted for the future. For safety reasons, a minimal interval between two starting or landing planes is crucial. Hence, airport capacities during the day become more and more depleted, leading to the tendency of air traffic to evade to late evening, early morning or even to the night. Integrators like UPS or TNT are often dependent on nocturnal air traffic. Therefore, the aircraft noise strain of residents living near airports has especially risen in these sensitive hours, and the exchange of very old and noisy with modern planes managed to counterbalance this effect only partially.

Sleep is vital for the recovery of physical and mental capacities. Environmental noise potentially disrupts the sleep process. Polysomnography (EEG, EOG, EMG) remains the gold standard for the measurement and classification of sleep. With this method, noise induced changes in sleep structure may be detected, for example, in form of reductions of total sleep time (TST), increments in the amounts of light sleep or in the number of sleep stage changes. As shown in the previous chapter, these primary sleep disorders may lead to deficits on the next day. Even long term health effects are controversially being discussed [4, 5, 38].

Sleep stage shifts are defined as the transition from one sleep stage to another. As can be seen in Figure 3.3, sleep stage shifts are a physiological characteristic of sleep and also a quite common phenomenon. E.g., transitions from S2 to S3 and S4 are essential after the initiation of sleep and part of the rhythmic changes between NREM- and REM-episodes within the night (so-called "ultradian rhythms"). Following stage shifts, sleep can either become lighter (e.g. S2→S1) or deeper (e.g. S2→S3).

Noise induced awakenings have always been the focus of research on the effects of noise on sleep. An awakening is a special case and the severest form of a sleep stage shift (REM→Wake, S4→Wake, S3→Wake, S2→Wake, S1→Wake). Stage Wake itself is not counted as sleep as it does not add to total sleep time. It is said not to contribute to sleep restoration. Frequent Wake episodes lead to a fragmentation of sleep and diminish its restorative power. Nonetheless, between 20 and 30 intermittent Wake episodes can be observed in a normal healthy eight hour sleep, and in that way they are also a physiological part of human sleep [6].

Although awakenings seem to be a reasonable choice for an indicator for noise induced sleep disturbances, other relevant changes in sleep structure, e.g. the time spent in the different sleep stages or the number of sleep stage changes, may be overlooked if the analysis is restricted to noise induced awakenings: Many ANEs may not lead to awakenings but nevertheless induce stage shifts (e.g. S4→S2), lead to a lightening of sleep and therefore reduce its restorative power.

3.3 Markov Processes

In this thesis, Markov processes are used to model human sleep in general and the influence of aircraft noise on sleep structure in particular. Markov models are part of the mathematical procedures used for formal decision analysis. A Markov process is a recursively-defined system that consists of a finite number of mutually exclusive and collectively exhaustive states, meaning that at any given time each person in the population being modeled must be in one of those states, and cannot be in more than one state at the same time.

Markov processes occur within a discrete time frame. Time progresses in intervals (Markov cycles). Cycle length is specified by the model builder. As time progresses, *transitions* take place from one cycle to the next. Figure 3.4 shows a state-transition diagram (so-called bubble diagram) of a simple Markov model.

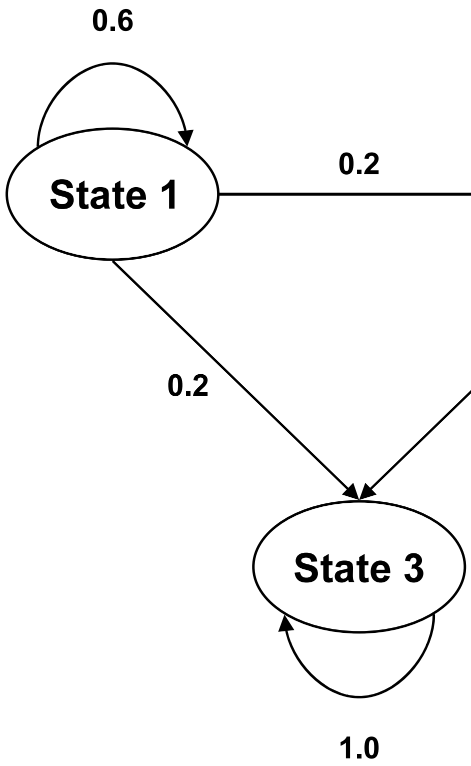


Figure 3.4: Bubble diagram of a simple Markov model.

The arrows represent transitions. A person who is in State 2 at cycle T may be either in State 2 or in State 3 at cycle T+1. Since there is no arrow from State 2 to State 1, a transition from State 2 to State 1 is not possible in this model. Arrows which are drawn to and from the same state (such as State 1 → State 1) represent the possibility of remaining in that state for the succeeding cycle. Transitions are considered instantaneous for modeling purposes, even though this may not be true in reality. From State 3, the only allowable transition is back to State 3. State 3 is therefore called an "absorbing state", because once it is entered, there is no exit.

The transition probabilities are also specified in Figure 3.4. The transition probability p(j|i) is the probability of going from state i to state j. The transition probabilities for exiting a particular state (including the probability of remaining in that state) at a particular cycle must sum to 1. Transition probabilities may be constant over time (Markov chains), or they may be time-dependent (Markov processes).

An important assumption, also called Markovian assumption, states that transition probabilities depend only on the current state and not on past states. In this way, the Markov model has no memory for states preceding the current state.

3.4 Modeling human sleep

After this brief introduction to Markov state transition models, one easily notices that the sleep process itself fulfills several properties that facilitate the application of Markov models for the modeling of human sleep: Using standard criteria [42], the sleep period is usually divided into 30 second segments, representing the Markov cycle length, and each segment is classified in one of six mutually exclusive and collectively exhaustive states: Wake, S1, S2, S3, S4 and REM[1].

Because of the homeostatic sleep process and circadian and ultradian rhythms, that is, periodic fluctuations of external *Zeitgebers* and internal circadian timing systems [16, 43], the different sleep stages are not equally distributed over the night. Therefore, transition probabilities do not remain constant during the night. Hence, Markov processes, allowing for changing transition probabilities with time, are suitable for modeling sleep, as opposed to Markov chains with fixed transition probabilities. However, one of the very first applications of Markov models on sleep used Markov chains [53].

---

[1] Body movements accounting for more than half of the epoch, that would have otherwise been scored as *Movement Time* according to Rechtschaffen and Kales [42], were classified as *Wake*, because it was assumed that these kinds of movements do not occur without respective cortical activation.

Historically, there have been two categories of human sleep models [34]:

Deterministic models, which used time-varying functions to describe systematically occurring characteristics like the NREM-REM periodicity, the decreasing amount of SWS in the course of the night or the relationship between sleep and the circadian rhythm [10, 16, 19, 29, 35, 36, 50].

Probabilistic models, which used stochastic processes to describe statistical properties of hypnograms like the variability of the NREM-REM cycle, the short interruptions within REM episodes or the random (in time and in direction) transitions between stages [17, 46, 52, 53].

As shown above, human sleep may easily be implemented in and described by Markov state transition models. Nevertheless, these models have only relatively sparsely been used for this purpose in the past:

In 1965, Zung et al. reported the use of a first-order Markov chain for the computer simulation of sleep EEG patterns [53], which was corrected and extended by Yang and Hursch in 1973 [52]. It was the primary goal of Zung et al. to achieve a high degree of data reduction as well as a meaningful description of sleep patterns in terms of transition probabilities. Hence, they wanted to evaluate the significance of stage shifts and stress the advantages of this procedure opposed to visual comparisons of hypnograms or the evaluation of summary statistics of sleep stages. One of their first analyses showed that transition probabilities do not remain constant during the night. Therefore, they divided the night into 30 min segments and assumed constant transition probabilities during each segment.

On page 355 they conclude [53]: "*Perhaps the most important aspect of our Markov chain model of sleep EEG patterns stems from the fact that it can be used to conduct simulated experiments on a digital computer. That is, if we postulate the effects of psychiatric disorders, or of psychotropic drugs on sleep, and transpose these mathematically onto the transition* *matrices of the subject population under consideration, we can use a computer to generate a sample of sleep patterns under a variety of different experimental conditions. Using these computer generated patterns as a guide, we may be able to predict certain changes, and use our computer model as an aid in evaluating clinical observations.*"

In that way, their objectives were very similar to those pursued in this thesis, aiming at the prediction of the effects of aircraft noise on sleep structure.

Kemp and Kamphuisen proposed a continuous-time Markov chain process for the description of human sleep [34]. In contrast to Markov processes with fixed cycle lengths, they modeled the probability of a sleep stage change after an arbitrary time interval Δt. They assumed that transition rates were constant within small intervals Δt, compared to the length of ultradian periodicity, and divided the eight hour night into 32 fifteen minute segments. They derived the average time spent in the same sleep stage from the transition rates, and, using a random number generator, were the first to simulate single human hypnograms. Compared to constant transition probabilities, they found significant inhomogeneities across the night in five transition rates: $p(1|0)$ and $p(2|1)$ tended to be high in the beginning of the night and remain constant or decrease during the night, which could be attributed to the process of falling asleep[2]. $p(3|2)$, $p(5|2)$ and $p(5|1)$ showed periodic changes during the night, accounting for the NREM-REM periodicity. They noted that these periodicities may be obscured by averaging over different subjects with varying periodicities. On page 413, they conclude: "*It is possible to combine the Markov model and some deterministic models by interpreting the deterministic models as mechanisms that generate the rates for the Markov model. The result might be a model with few parameters that still simulates most probabilistic and deterministic aspects of sleep.*"

[2] Sleep stages are abbreviated as follows: Wake = 0, S1 = 1, S2 = 2, S3 = 3, S4 = 4 and REM = 5.

Karlsson et al. developed a first-order Markov model that described the probability of sleep stage changes as a function of time after intake of 20mg temazepam or placebo, based on data of polysomnograms of 21 patients suffering from insomnia [33]. They used a mixed-effects model with *nighttime* (elapsed time since lights out), *stage time* (time since last sleep stage change), which was divided into segments, and *drug exposure* (dose of temazepam) as independent variables. They found that temazepam facilitated transitions to deeper sleep, whereas transitions to lighter sleep were inhibited. Also, Temazepam inhibited transitions to the wake state, regardless of sleep stage, and return to sleep was facilitated. According to the authors, the model was able "to simulate several well-known patterns that are also seen in the present data", but simulation results of single hypnograms are not shown. As Karlsson et al. used a Markov state transition model to investigate the influence of an extrinsic factor, i.e. a hypnagogic drug, on sleep, their objectives were very similar to the objectives of this thesis.

Karlsson et al. [33] used a statistically sophisticated approach to estimated the influence of a drug on transition probabilities between sleep stages, and therefore on the mechanisms leading to changes in sleep structure. A potential drawback is that several models had to be built to describe the sleep process and that the methods applied may be understood by few people with very advanced statistical knowledge only.

It would be desirable to describe transition probabilities, and therefore the whole sleep process, by a relatively simple and single model, that nevertheless provides unbiased predictions of transition probabilities and simulation outcome variables. Although many authors mention that their models could be used to predict the effect of unprecedented situations, to the knowledge of the author of this thesis such predictions have not been reported in the literature yet.

4 Objectives

The main objective of this thesis was to quantitatively assess possible effects on several aspects of sleep structure of an air traffic ban between 23:00 and 05:00 at Frankfurt Airport, providing important information additional to the effects of noise on the probability and number of awakenings. In order to achieve this goal, the following objectives were followed:

(1) Development of a mathematical model capable of representing sleep structure and its typical features, based on noise-free baseline nights of 125 subjects. The model should be able to reproduce typical features of undisturbed human sleep without bias and as closely as possible. Additionally, the characteristic sequence of sleep stages seen in hypnograms simulated by the model should resemble the natural sequence observed in human hypnograms.

(2) Extension of the model in order to allow for the prediction of changes in sleep structure in the presence of aircraft noise, based on the data collected in a polysomnographic laboratory study on the effects of aircraft noise on sleep.

(3) Comparison of three different noise scenarios, derived from the current and from possible future constellations at Frankfurt Airport, with each other and with noise-free nights. The results of this comparison should support the decision making process of policy makers and legislative bodies.

5 Methods

Before methodological details are introduced, the reader should be able to understand the basic ideas and methodological approaches followed in this thesis.

Based on empirical data, models for the simulation of human sleep were developed. These models should be able to

(1) simulate physiological human sleep without bias and

(2) estimate the effects of nocturnal aircraft noise on sleep structure.

For this purpose, a Markov state transition model with sleep stages Wake, S1, S2, S3, S4 and REM represented in six Markov states was used (see Chapter 5.1). The parameters of the model were estimated based on data sampled at the DLR-Institute of Aerospace Medicine in Cologne between 1999 and 2003 in polysomnographic laboratory studies on the effects of aircraft noise on sleep. Here, subjects were investigated in noise-free baseline nights as well as in aircraft noise exposure nights for several consecutive nights (see Chapter 5.2). As several statistical analyses have been performed on the DLR data set in order to derive model parameters for the sleep Markov model, some of these statistical calculations will already be presented in the methods section rather than in the results section to improve readability.

The model was used to investigate the impact of different noise scenarios on sleep. Outcome variables were chosen in order to reflect the most important aspects of sleep structure (see Chapter 5.5). Three noise scenarios were derived from the flight schedule of Frankfurt Airport from August 2005 and from prognosticated traffic scenarios that are likely to occur in the event of an expansion of Frankfurt Airport (see Chapter 5.4.2.1). The ultimate goal was to compare three noise scenarios with each other and with a noise-free baseline night in order to assess the magnitude of the impact of the different noise scenarios on sleep structure.

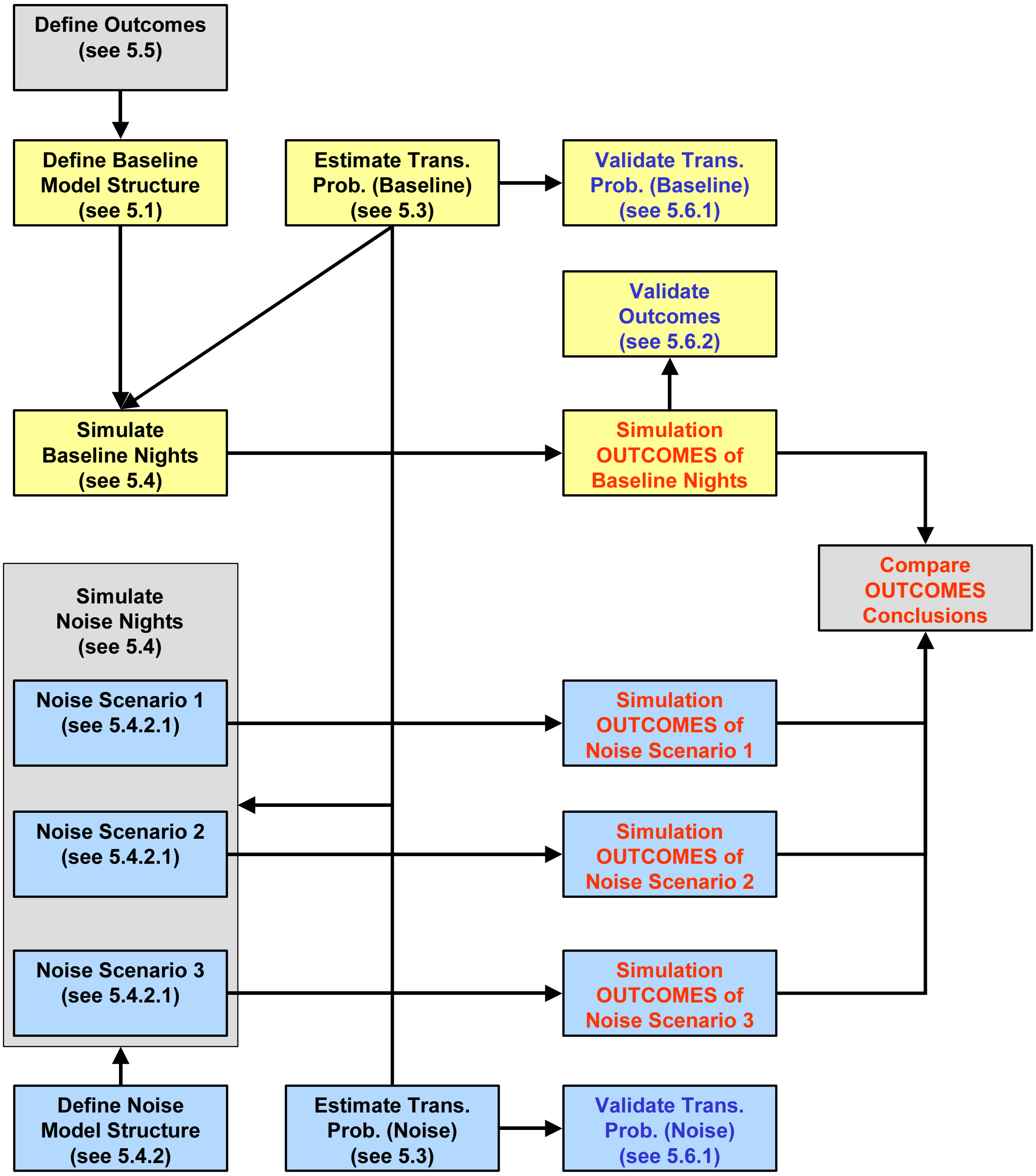


Figure 5.1: Schematic overview of strategy and methods applied for the simulation of noise-free baseline nights (yellow) and of nights with aircraft noise (blue). Chapter numbers containing a detailed description of the applied methods are given in parenthesis. Trans. Prob. = Transition Probabilities

In order to achieve the goals mentioned above (see Figure 5.1),

(1) relevant outcomes for the description of sleep structure had to be defined (see Chapter 5.5),

(2) the structure of the Markov model had to be developed for the simulation of noise-free baseline nights (see Chapter 5.1) and for nights containing aircraft noise (see Chapter 5.3.2.2),

(3) for <u>noise-free baseline nights</u>, transition probabilities between sleep stages had to be calculated (see Chapter 5.3.2.1) based on and validated with (see Chapter 5.6.1) empirical data,

(4) based on the calculated transition probabilities, Markov model simulations of undisturbed human sleep had to be performed (see Chapter 5.4.1), and the outcomes of the simulations had to be validated with empirical data (see Chapter 5.6.2),

(5) for <u>nights containing aircraft noise</u>, transition probabilities between sleep stages had to be calculated (see Chapter 5.3.2.2) based on and validated with (see Chapter 5.6.1) empirical data,

(6) based on the calculated transition probabilities, Markov model simulations of sleep disturbed by aircraft noise had to be performed (see Chapter 5.4.2) for three different noise scenarios (see Chapter 5.4.2.1) and

(7) the outcomes of the simulation of the three noise-scenarios had to be compared with the outcomes of the simulation of noise-free baseline nights and with each other.

5.1 Overview of model structure and modeling process

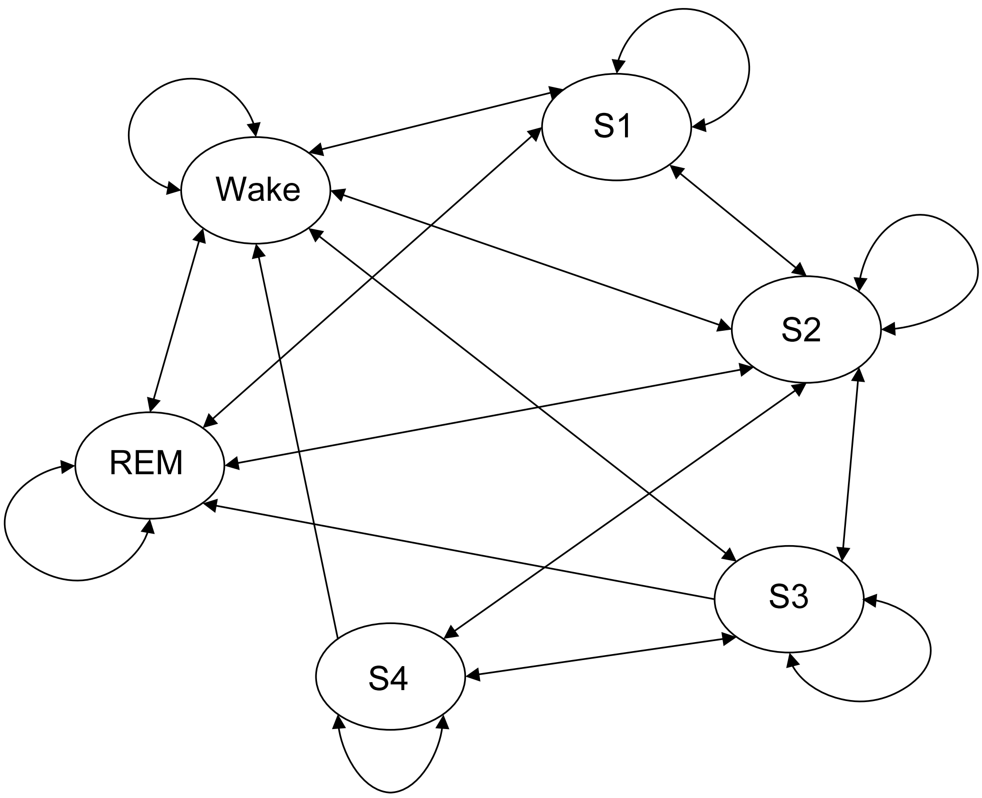


Figure 5.2: Bubble diagram of the sleep model used in this thesis. Some transitions occur never (e.g. S1→S4 or S4→REM). Average transition probabilities are given in Table 6.1.

Figure 5.2 shows the Markov state transition diagram, and therefore the model structure, of the Markov model of human sleep used in this thesis. Theoretically, transitions from one sleep stage to every other sleep stage are possible, but some transitions are almost never (Wake→S4, S1→S3, S3→S1, S4→S1/REM, REM→S3) or never (S1/REM→S4) seen (for details also see Table 6.1).

An overview of the different methodological aspects involved in modeling human sleep with Markov state transition models will be given below. The modeling process and the methodological aspects involved are summarized in Figure 5.3.

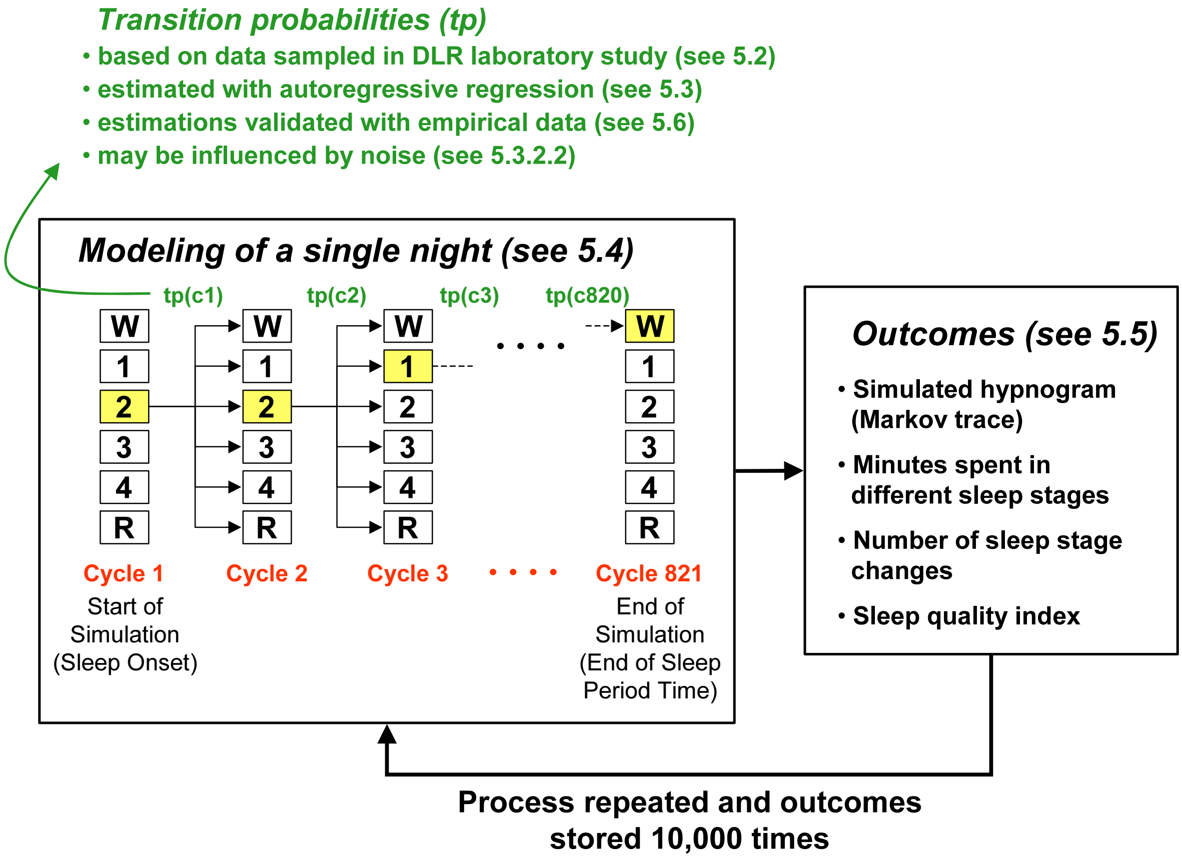


Figure 5.3: Overview of modeling process and methodological aspects involved (W = Wake, R = REM, tp = transition probability).

All outcomes are based on Markov state transition models and the simulation of single nights. A simulation always starts with sleep onset (cycle 1) and ends after 820 transitions in cycle 821, where each cycle has a fixed length of 30 seconds. The rationale for modeling a sleep period time of exactly 821 cycles or 410.5 min (= 6.8 hours) will be given in Chapter 5.2. The simulation process is described in detail in Chapter 5.4.

The transition probabilities were estimated with autoregressive multinomial logistic regression (see Chapter 5.3) based on data sampled during a large polysomnographic laboratory study (see Chapter 5.2). Transition probability estimates of autoregressive multinomial logistic regression were validated with transition probabilities derived from the original data (see Chapter 5.6).

Modeling a single night produces several direct and indirect outcome variables (see Chapter 5.5). Plotting the realizations of the different Markov states, i.e. sleep stages, against time yields a so-called Markov trace, which is the model equivalent of the human hypnogram (see Figure 3.3). These Markov traces were used to check the face validity of the model (see Chapter 5.6.2). The time spent in the different sleep stages during the simulation as well as the number of changes to other sleep stages than the current sleep stage were two other important outcomes (see Chapters 5.5.2 and 5.5.3). Additionally, a sleep quality index (SQI) was introduced. Here, the time spent in the different sleep stages during the simulation was weighted with the proposed recuperative value of the respective sleep stages resulting in a single value index (see Chapter 5.5.4).

The process of modeling a single night was repeated 10,000 times (see Chapter 5.4). Outcome variables were stored for each of the simulation realizations. Average values of the outcome variables and their distribution were validated with empirical data (see Chapter 5.6).

Noise may alter transition probabilities between sleep stages, and therefore lead to changes in the distribution of outcome variables. Transition probabilities were, again, estimated with autoregressive multinomial logistic regression based on periods where aircraft noise was played back during the laboratory studies (see Chapter 5.3.2.2). Outcome variables of simulations of the three different noise scenarios (see Chapter 5.4.2.1) were compared to outcome variables of simulations of a noise-free baseline night and with each other.

5.2 Data basis

From 1999 until 2003, the DLR Institute of Aerospace Medicine (Cologne, Germany) investigated the influence of nocturnal aircraft noise on sleep. The study protocol was approved by the ethics committee of the Medical

Board of the district North Rhine. Subjects were instructed according to the Helsinki declaration, participated voluntarily and were free to discontinue their participation at any time without explanation.

128 subjects (mean age 38 ± 13 years, range 19-65, 75 female, 53 male) were selected in a stepwise manner from more than 1,000 applicants responding to appeals in newspaper articles or flyers. Selection was performed by questionnaires and a medical screening. Subjects had to have normal hearing thresholds according to age and be free of cardiac arrhythmias and intrinsic sleep disorders. They were not allowed to snore loudly. As there were far more applicants than subjects needed for the study, final participants were randomly drawn from eligible candidates. Two subjects had to be replaced with backups after the first night because of a sleep-related breathing disorder. For a detailed description of the selection process see [6] or [9].

Subjects were investigated in a 300 m^2 sound-proof and air-conditioned sleep facility situated in the basement of the Institute of Aerospace Medicine. The sleep area of the facility consists of eight separate sleep cabins that allow for the simultaneous investigation of eight subjects. Therefore, 16 study periods were needed for the investigation of 128 subjects.

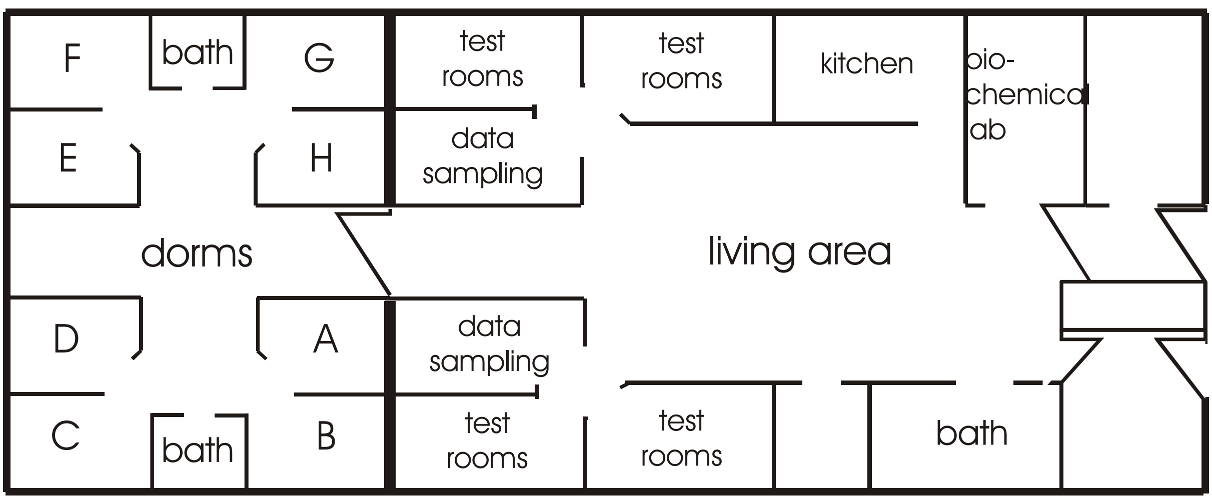


Figure 5.4: Isolation facility of the institute (about 300 m^2).

Each study period started on a Monday evening and lasted 13 consecutive nights (including the weekend). Subjects arrived at the lab at 19:00. They filled in questionnaires, performed computer assisted performance tests, had supper and were equipped for the night. Lights were turned off at 23:00 and on again at 07:00. In the morning, subjects filled in questionnaires and carried out the performance tests again. Afterwards, they were permitted to have breakfast and leave the facility until the evening of the same day. Subjects were not allowed to nap during the day. The general activity of the subjects during the day was monitored with actigraphs, that had to be worn all day and night.

16 (mean age 40 ± 13 years, range 24-59, 10 female, 6 male) subjects served as a control group in order to investigate the influence of the laboratory situation on otherwise undisturbed sleep and therefore did not receive any noise at all. The other 112 subjects (mean age 38 ± 13 years, range 19-65, 65 female, 47 male) served as the experimental group. Here, the first night served as adaptation, the second as baseline and nights 12 and 13 as recovery. These nights were noise-free, with the exception that the recovery nights of four out of 16 study periods were used to investigate the influence of special noise patterns. During nights three to eleven 4, 8, 16, 32, 64 or 128 ANEs of starting and landing planes with maximum SPLs of 45, 50, 55, 60, 65, 70, 75 or 80 dB were equidistantly played back during the night (see Table 5.1). In noise effects research, linear SPLs are usually filtered with the so-called A-filter, in order to account for the frequency dependency of the sensitivity of the ear. All SPLs presented in this thesis are A-weighted SPLs. The appendix "(A)" will be left out for reasons of simplicity from now on.

Table 5.1: Number of exposure nights depending on maximum SPL and number of aircraft noise events per night (e.g. 56 subject nights with 64 x 65 dB).

		Number of ANEs					
		4	8	16	32	64	128
Maximum SPL in dB	45					32	
	50			32	32	32	32
	55	40	40	32	32	32	32
	60	40	40	32	32	32	
	65	32	32	32	32	56	
	70	32	32	32	32		
	75	32	32	32			
	80	32	24				

The exposure time depended on the number of ANEs per night. In exposure nights with four ANEs per night, playback started at midnight and stopped at 06:00. In exposure nights with eight ANEs per night, playback commenced at 23:30 and ended at 06:30. In exposure nights with 16 or more ANEs per night, playback started at 23:15 and stopped at 06:45. The noise-free time interval between two ANEs ranged from two hours (four events per night) to three minutes (128 events per night). As the participants did not know of the equal distances between ANEs, an anticipation of the time of occurrence of the next ANE was not possible. Watches and alarm clocks were not allowed in the sleep cabins.

The combinations 4 x 50 dB and 8 x 50 dB were not used because relevant reactions were not expected at this low level of exposure. On the other hand, the combinations 128 x 60 dB up to 128 x 80 dB, 64 x 70 dB up to 64 x 80 dB, 32 x 75 dB, 32 x 80 dB and 16 x 80 dB were not played back because these exposures were considered to be unrealistically high and may not have been tolerated by the subjects.

Each category was planned to consist of at least 32 subject nights, which was accomplished for all categories but 8 x 80 dB. The increased number of nights with combinations 64 x 65 dB resulted from a special sub-experiment, in which it should be compared to the combination 64 x 45 dB. No other exposure patterns were planned with maximum SPLs of 45 dB. The number of nights with the combinations 4 x 55 dB, 8 x 55 dB, 4 x 60 dB and 8 x 60 dB were increased because they surrounded the so-called *Jansen criterion [25]*, which states that maximum SPLs of 60 dB or more should not be exceeded more often than six times per night. Therefore, these combinations should be covered more intensively.

The different exposure patterns were randomly assigned to nights three to eleven of each study period. There was no wash-out period in form of noise-free nights interposed between two exposure nights. As there were 30 exposure patterns but only nine exposure nights per subject, the study had a randomized incomplete block design.

ANEs played back during the exposure nights were recorded with class-1 sound level meters (NC-10, Cortex Industries) in the vicinity of Düsseldorf Airport with closed or tilted windows. The microphone was installed near pillow position, i.e. "at the sleeper's ear". Playback of ANEs was realized with an Acoustic Workstation CF85 (Cortex Industries). Before each study period, every sleep cabin was acoustically calibrated with class-1 sound level meters in order to guarantee realistic playback of ANEs. There was a constant background noise level of about 30 dB caused by the air condition system of the sleep facility.

During a single study night, always the same ANE was played back (e.g. 50 dB only), i.e. there was no mixing of different ANEs in one single night. All eight subjects of one study period received the same noise pattern, i.e. the same ANE was played back in all sleep cabins at the same time. As sound insulation of the sleep cabins was not absolute, a temporal offset of playback of ANEs was avoided, as it might have lead to the perception of ANEs from neighboring sleep cabins.

The SPL in each sleep cabin was recorded continuously during each study night and allowed for the control of the correct playback of each ANE. Additionally, it was possible to identify loudly snoring subjects.

The subjects were only informed about the first two study nights being noise-free. They were otherwise blinded with respect to noise exposure, i.e. they did not know when, how many and what kind of ANEs were played back after the second night. In order to avoid subconscious manipulations, the investigators were also blinded for the noise pattern of the specific night. Only after the beginning of data sampling, i.e. after 23:00, they were informed about the noise pattern of the specific night, and thus were able to monitor the correct playback of ANEs. Altogether, 30,588 ANEs were played back in the laboratory studies during nights three to eleven.

Sleep was polysomnographically recorded using the standard setup (EEG C3-A2, C4-A1, left and right EOG and submental EMG). Sleep stages were classified by two experienced scorers following standard criteria [42] and using 30-second epochs. The inter-rater reliability was 88.1% on average. Body movements accounting for more than half of the epoch, that would have otherwise been scored as *Movement Time*, were classified as *Wake*, because it was assumed that these kinds of movements do not occur without respective cortical activation.

Prior to sleep stage classification, sleep files were renamed and their order was randomly changed. In this way, the person analyzing the file was blinded as he/she did neither know which file he/she was scoring and therefore nor whether and how many ANEs were played back during that night. This process aimed at avoiding a systematic bias. All nights of one subject were scored by the same person in order to avoid the problem of inter-rater variability in within-subject comparisons. By means of a randomized sequence of the nights to be analyzed, a systematic habituation of the scorer to the specific EEG-patterns of one subject always in the same order (i.e. nights 1 to 13) was avoided.

It was not possible to completely analyze the data of three noise-free baseline nights. For the remaining 125 nights, subjects were synchronized to sleep onset, i.e. the process of falling asleep was not investigated in this analysis. Sleep onset was defined as the first appearance of sleep stage S2. The analysis was restricted to the first 410.5 min (= 6.8 hours or 821 epochs) of sleep period time, in order to avoid censored data of subjects with very long sleep onset latencies (see Figure 5.5). Therefore, 102,500 transitions of 125 baseline nights were used to calculate regression coefficients altogether (see Chapter 5.3).

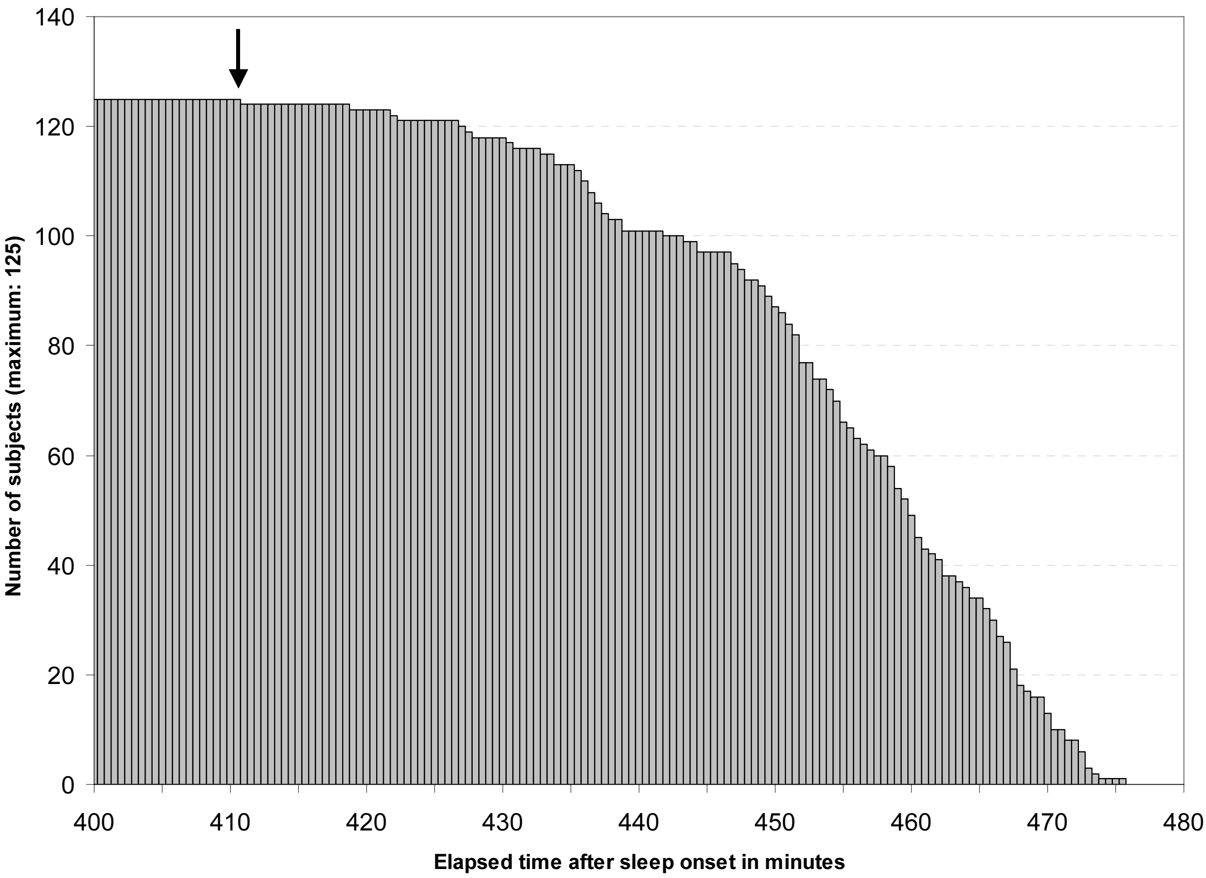


Figure 5.5: The analysis was restricted to the first 410.5 minutes of SPT in order to assure uncensored data for all 125 subjects.

5.3 Estimation of transition probabilities for the Markov model

Transition probabilities between sleep stages for the sleep Markov model were calculated using autoregressive multinomial logistic regression. Autoregressive logistic regression is a technique to model dependent binary responses (i.e., dependent variable) with repeated observations of an outcome variable within the same individual as a special case of dependence [15]. With repeated measurements, previous responses are used as covariates predicting future responses. By decomposing the joint probability into a product of successive conditional logistically modeled probabilities, the initially multivariate problem can be solved with ordinary univariate logistic regression, using standard statistical software.

Therefore, a prominent feature of autoregressive models is that conditions of the present (T0) and the past (T-1, T-2, T-3, etc.) are directly incorporated in the model in order to predict future conditions (T1). A first-order autoregressive transitional model facilitates only information of the present to predict future states. Higher order models use information of states further in the past to predict future outcomes (e.g. a third-order model uses the information of responses at T0, T-1 and T-2 to predict the response at T1). In contrast to other model types, where the correlation structure of dependent data is modeled indirectly, the conditional distribution of each response is expressed as an explicit function of past responses (the history) and covariates. Therefore, the correlation among the repeated responses can be thought of as arising due to past values of the responses explicitly influencing the present observation (i.e. the present depends on the past). In autoregressive transitional models, the covariates and lagged responses are treated on an equal footing.

Although autoregressive regression was originally defined for modeling binary responses only, it can be easily extended for the multinomial case [20]. For this analysis, the autoregressive multinomial logistic regression was computed with the SAS procedure CATMOD (SAS Institute Inc., Version 8.2).

A potential drawback, especially of autoregressive models of higher order, is that there is no information on past events for the prediction of the very first transitions. Nevertheless, this posed no problem in this thesis, as the simulation started with sleep onset, and therefore, information on sleep stage history prior to sleep onset, derived from the process of falling asleep, was available (see Table 5.2).

5.3.1 Data management

Sleep stage classification data were stored together with the starting time of each sleep epoch in separate ASCII-files for each night. Data were read from these files and arranged in Microsoft Excel spreadsheets according to Table 5.2.

Table 5.2: Data structure for input in PROC CATMOD (SAS). Cells containing sleep stage information prior to sleep onset are shaded gray. NoChange = Number of cycles since last sleep stage change, ID = unique identifier of each subject

Future State (to be predicted)	Present State	Past States					
T1	**T0**	**T-1**	**T-2**	**Transition**	**NoChange**	**Age**	**ID**
2	2	1	1	1	1	24	205
2	2	2	1	2	2	24	205
2	2	2	2	3	3	24	205
2	2	2	2	4	4	24	205
2	2	2	2	5	5	24	205
2	2	2	2	6	6	24	205
0	2	2	2	7	7	24	205
0	0	2	2	8	1	24	205
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
5	5	5	5	819	3	24	205
5	5	5	5	820	4	24	205
2	2	1	1	1	1	46	285
2	2	2	1	2	2	46	285
2	2	2	2	3	3	46	285
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

T1 refers to the sleep stage that is predicted by the CATMOD procedure with multinomial logistic regression, i.e. the dependent variable. T0 refers to the present sleep stage, T-1 and T-2 refer to past sleep stages, i.e. to the sleep stages one (T-1) or two (T-2) epochs before the present epoch. The sleep period by definition starts with the first occurrence of sleep stage 2 (T0 at transition #1 = sleep stage 2). Cells containing information from epochs prior to sleep onset are shaded gray. The number of cycles since the last sleep stage change, the age of the subject and a unique subject identifier are also part of the table. A sleep period of 410.5 min consists of 821 sleep epochs and contains 820 transitions between these epochs, which are listed in the spreadsheet for each subject. Altogether, the table consists of 125 x 820 transitions = 102,500 rows.

These rows were divided over two spreadsheets and imported into SAS. In a data step, indicator variables for sleep stages of the present state and past states were generated with reference coding (Wake being the reference state). By default, PROC CATMOD assumes the highest level of the outcome variable is the reference group, but unlike in PROC LOGISTIC there is no DESCENDING option. As sleep stage 2 should be the reference group, values of 2 at T1 were changed to a value of 6, data were sorted in descending order with a PROC SORT statement and the ORDER = DATA option was used in PROC CATMOD. As there are six possible outcomes, PROC CATMOD produces five parameter estimates for each independent variable (for a detailed description see Chapter 10).

Regression results were written to a Microsoft Excel file. During the simulations described in Chapter 5.4, transition probabilities were calculated based on these regression results.

5.3.2 Regression model building process

The regression model building process is described separately for the baseline model and for the noise models.

5.3.2.1 Baseline model

Several models were built and compared in order to finally choose the most parsimonious, yet biologically reasonable model. The models always included indicator variables for sleep stage history, resulting in a first (only T0) or higher (T0, T-1, etc.) order autoregressive model. The models also always included a continuous variable representing the number of epochs elapsed after sleep onset (termed "Transition" in Table 5.2), as sleep stages are not uniformly distributed over the night and transition probabilities clearly depend on elapsed sleep time.

The model building process started with a simple first-order autoregressive model, i.e. only the present sleep stage was used to predict the sleep stage of the following epoch, containing elapsed sleep time as the only additional explanatory variable. This model is shown in equation (1):

$$p(stage\_T1)=Intercept+\beta_1 \cdot S1\_T0+\beta_2 \cdot S2\_T0+\beta_3 \cdot S3\_T0+\beta_4 \cdot S4\_T0+\beta_5 \cdot REM\_T0+\beta_5 \cdot Transition \quad (1)$$

The ß's stand for the regression coefficients estimated with maximum likelihood techniques based on the empirical data. S1_T0, S2_T0, S3_T0, S4_T0 and REM_T0 are five indicator variables, indicating presence (1) or absence (0) of each of the five sleep stages at T0. If the present sleep stage is S2, S2_T0 will be 1, and the other four indicator variables will be 0. If the present sleep stage is Wake, all five indicator variables equal zero. "Transition" is a continuous variable representing the number of the current transition. As a constant SPT of 410.5 min equaling 821 epochs was modeled, the variable "Transition" varied between 1 and 820.

Then, other potentially important variables were added to the model, i.e. time spent in the same sleep stage in epochs (NoChange in Table 5.2) and age. Also, interactions of prior sleep stages with elapsed sleep time and higher order terms of elapsed sleep time (T^2 and T^3) were tested for their relevance. Finally, a second-order (T0 and T-1) and a third-order (T0, T-1 and T-2) autoregressive model, with no additional variables but elapsed sleep time, were built.

All models were compared to the simple first-order model according to

(1) goodness-of-fit of transition probabilities estimated by the model compared to transition probabilities observed in the original data (see Chapter 5.6.1),

(2) the relative bias of outcome variables between Monte-Carlo simulations and observed empirical data (see Chapter 5.6.2) and

(3) single Markov traces produced by first-order Monte-Carlo simulations (see Chapter 5.6.2).

5.3.2.2 Noise models

In terms of transitional models, aircraft noise may alter transition probabilities, e.g. the probability for changes to lighter sleep stages (e.g. p(1|2)) may be increased, while the probability for changes to deeper sleep stages (e.g. p(3|2)) may be decreased simultaneously. During nights 3 to 11 of the experimental laboratory study on the influence of nocturnal aircraft noise on sleep, ANEs with different maximum SPLs and frequencies of occurrence were played back (see Chapter 5.2). The aircraft noise events of all exposure nights, irrespective of maximum sound pressure level, were used for estimating transition probabilities under the influence of aircraft noise.

Past analyses of noise induced awakening probability indicated that transition probabilities may not only be influenced in the epoch where playback of the ANE is started, but also in cycles following that epoch [6]. On the one hand, it is likely that the probability of transitions to lighter sleep stages or Wake increases under the influence of aircraft noise (e.g., p(0|2) or p(2|3)). On the other hand, transition probabilities back to deeper sleep stages should increase after the noise event is finished (e.g., p(1|0) or p(2|0)). Both mechanisms should be captured by a noise model.

Therefore, a first analytical step investigated whether there were differences in transition probabilities between aircraft noise and noise-free conditions, and if so, for how many epochs after the start of an ANE these differences persisted. For this purpose, data from each of ten epochs following the onset of an ANE as well as from the corresponding epochs in the noise-free baseline night of the same subject, i.e. after the same elapsed sleep time, were extracted. An epoch was defined as "the first epoch under the influence of aircraft noise" if playback of the ANE started before the middle of the epoch, i.e. latest after 15 seconds. If an ANE starts after the middle of an epoch, it is very unlikely that the classification of the epoch will be influenced by the noise event, as more than half of the epoch have to show typical features of the sleep stage in order to be scored as such [42].

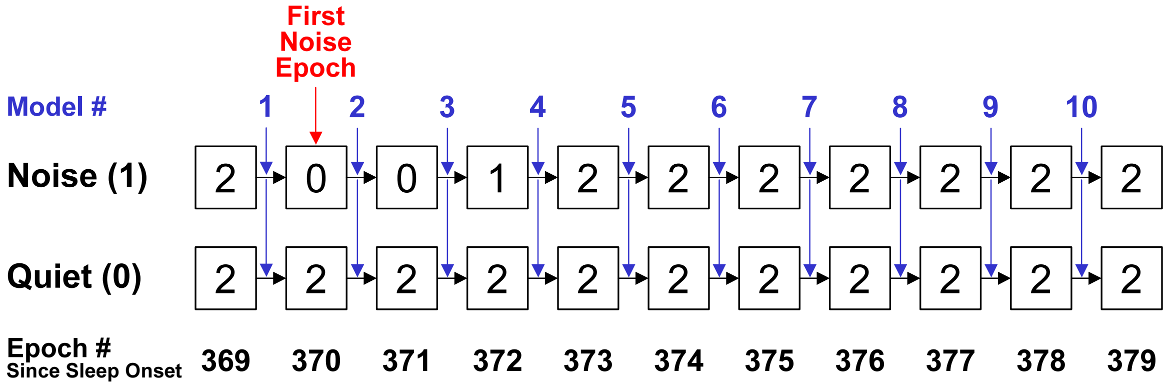


Figure 5.6: Differences in transition probabilities between quiet and noisy conditions. Blue arrows indicate the transitions investigated by the ten models. In this example, the noise event starts in epoch #370 after sleep onset. Sleep stages of epochs #369 to #379 are shown for the noise night (Noise) and the baseline night (Quiet) of the same subject.

For each of the ten epochs, a first-order autoregressive multinomial logistic regression model (see Chapter 5.3) was built with a noise indicator variable (1 or 0) as the only additional explanatory variable (elapsed sleep time was per definition identical between both conditions, see Figure 5.6). In this case, the regression output of the noise indicator variable indicates direction, magnitude and statistical significance of the difference in transition probabilities between noise-free epochs and epochs with ANEs.

Between 1 and 10 "noise models" shall be built additionally to the model of noise-free baseline nights: One for the first epoch under the influence of aircraft noise, one for the second epoch and so on, until results of the regression analyses show that there is no relevant difference in transition probabilities between noisy and noise-free conditions. The structure of the noise models should be otherwise equal to the final baseline model, which resulted from the model building process described in 5.3.2.1. In the simulation of noise nights, transition probabilities derived from the noise models were only used in those epochs containing aircraft noise, otherwise transition probabilities derived from the baseline night regression model were used.

In this thesis, different maximum SPLs of ANEs were not differentiated. Rather, all noise events were treated as if they consisted of the same maximum SPL, and this maximum SPL can be thought of as an "average" maximum SPL of all 26,135 ANEs played back during the first 410.5 min of sleep period time (see also Table 5.1).

5.4 Monte Carlo Simulations

There are two ways of analyzing Markov models: Deterministic cohort analyses or Monte Carlo simulations with individual trials. In cohort analyses, a cohort of subjects is simultaneously moved through the model. Depending on the transition probabilities and on the distribution of cohort members in the present cycle, the distribution of cohort members in the next cycle is deterministically calculated. One of the important assumptions in a Markov cohort analysis is that the model maintains no memory of previous events. Since cohort calculations are used, there are no “individuals” to carry memories with them through the process. Without memory of earlier stages of the Markov process, a given state’s transition probabilities and other values cannot depend on knowledge of prior events. Similarly, prior events cannot be directly inferred from current state membership. In a Markov cohort analysis, the only way to remember where a particular portion of the cohort has been is to create many additional states to keep separate those cohort members who experience different events. Even when some form of Markov memory can be accomplished in this manner, it usually requires an unwieldy model.

The most common (and efficient) method for introducing detailed memory into a Markov process is to analyze it using Monte Carlo simulation trials, rather than cohort analyses. In a single first-order trial of a Markov model, one member of the cohort is randomly stepped through the process, based on the probabilities in the model. In a trial, special tracker variables can be used to track each individual's particular steps through the process, creating a flexible form of memory that can be used in determining transitions. By repeating a large number of such trials, an expected value calculation of the Markov process can be simulated.

As first-, second- and third-order autoregressive multinomial logistic regression models had to be compared according to their Markov model outcomes, Monte Carlo simulation trials had to be performed, where one member of the cohort was randomly stepped through the process, based on the probabilities calculated by the regression model. The first-order simulation trial is described in detail in Chapter 5.4.1 for the baseline nights and in Chapter 5.4.2 for the noise nights.

5.4.1 Baseline model

The simulations were programmed in a Visual Basic for Applications environment (Microsoft Excel 2000). The first appearance of sleep stage 2 was defined as sleep onset. Hence, the simulation always started with sleep stage S2 in cycle #1 and ended after 820 transitions in cycle #821. In that way, it was possible to compare the results of the model with the raw data of all 125 subjects (see Chapter 5.6). The history of prior sleep stages was stored in so-called *tracker variables*, which can only be used in simulation trials. The regression analyses described in Chapter 5.3 provided the transition probabilities depending on past sleep stages and other covariates.

The sleep stage realizations were stored for each cycle. Sleep stage realizations plotted against cycle number are also called "Markov trace", which is the model equivalent to the hypnogram of human subjects (see Figure 3.3). For every regression model, 10,000 first-order Monte-Carlo trials were performed. Averaging over all simulated Markov traces leads to a probability distribution of sleep stages depending on cycle number, i.e. elapsed sleep time. The results of the single trials were averaged and validated with raw data of the 125 subjects (see Chapter 5.6).

Figure 5.7 exemplary shows one realization of first-order Monte Carlo trial of a noise-free baseline night. In this example, transition probabilities are based on a first-order autoregressive model with elapsed sleep time as the only additional explanatory variable.

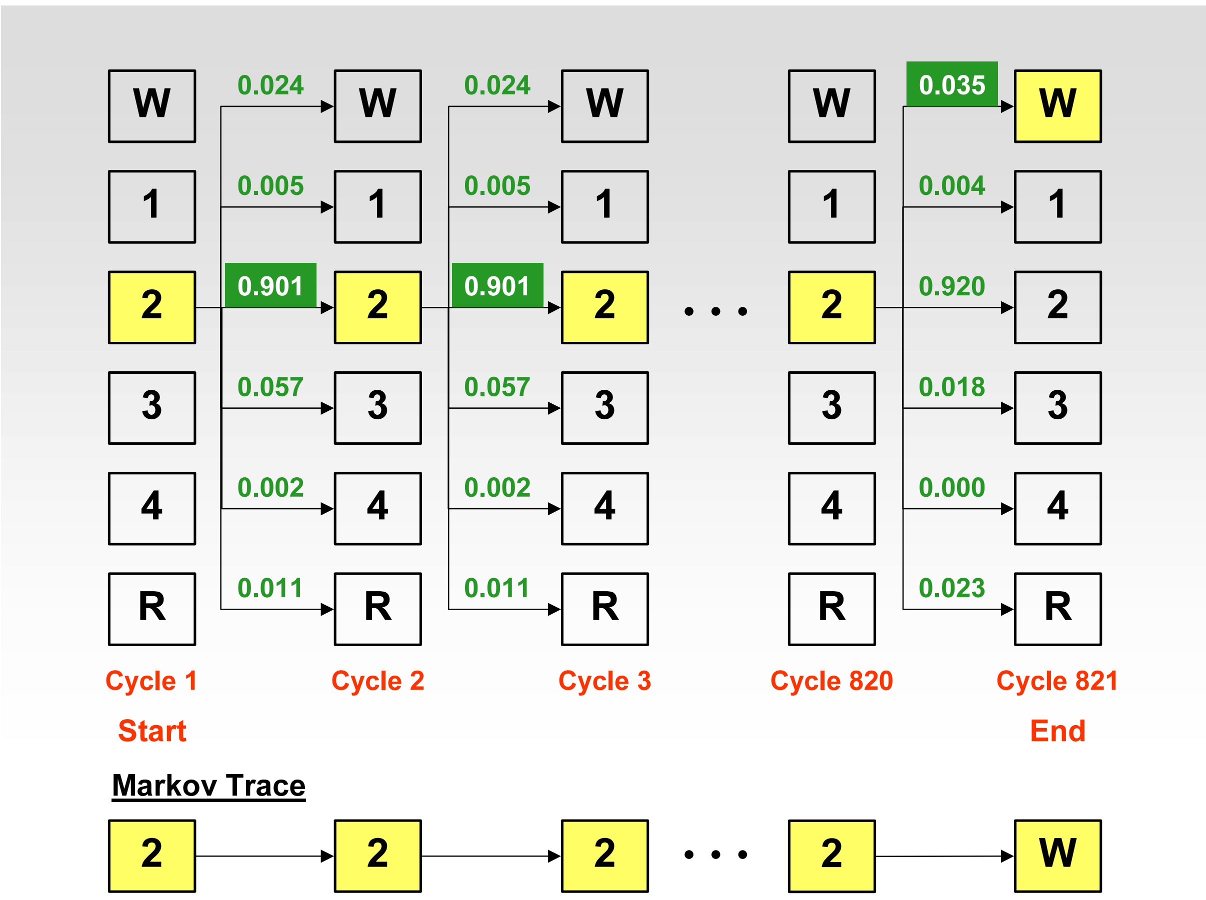


Figure 5.7: One possible realization of a first-order Monte Carlo trial of the baseline nights.

The simulated night starts with stage 2 (sleep onset) in the first cycle. The sleep stage in cycle #2 is chosen by Excel's random number generator based on the transition probabilities provided by the regression model. In this example, the random number generator chose stage 2 for cycle #2, which was likely to happen in 901 out of 1,000 trials. Again, stage 2 is randomly chosen for cycle #3.

Sleep stages are not uniformly distributed over the night. Hence, transition probabilities also change during the night. Transition probabilities from cycle #1 to cycle #2 and from cycle #2 to cycle #3 do not differ as long as three digits after the decimal point are taken into account, because elapsed sleep time is practically the same. If these transition probabilities are compared to the transition probabilities from cycle #820 to cycle #821, there are obvious differences: Changes to stages Wake and REM are more likely at the end of the night, while changes to slow wave sleep stages 3 and 4 are less probable. The Markov trace, i.e. the sleep stage realizations of each cycle are shown at the bottom of Figure 5.7.

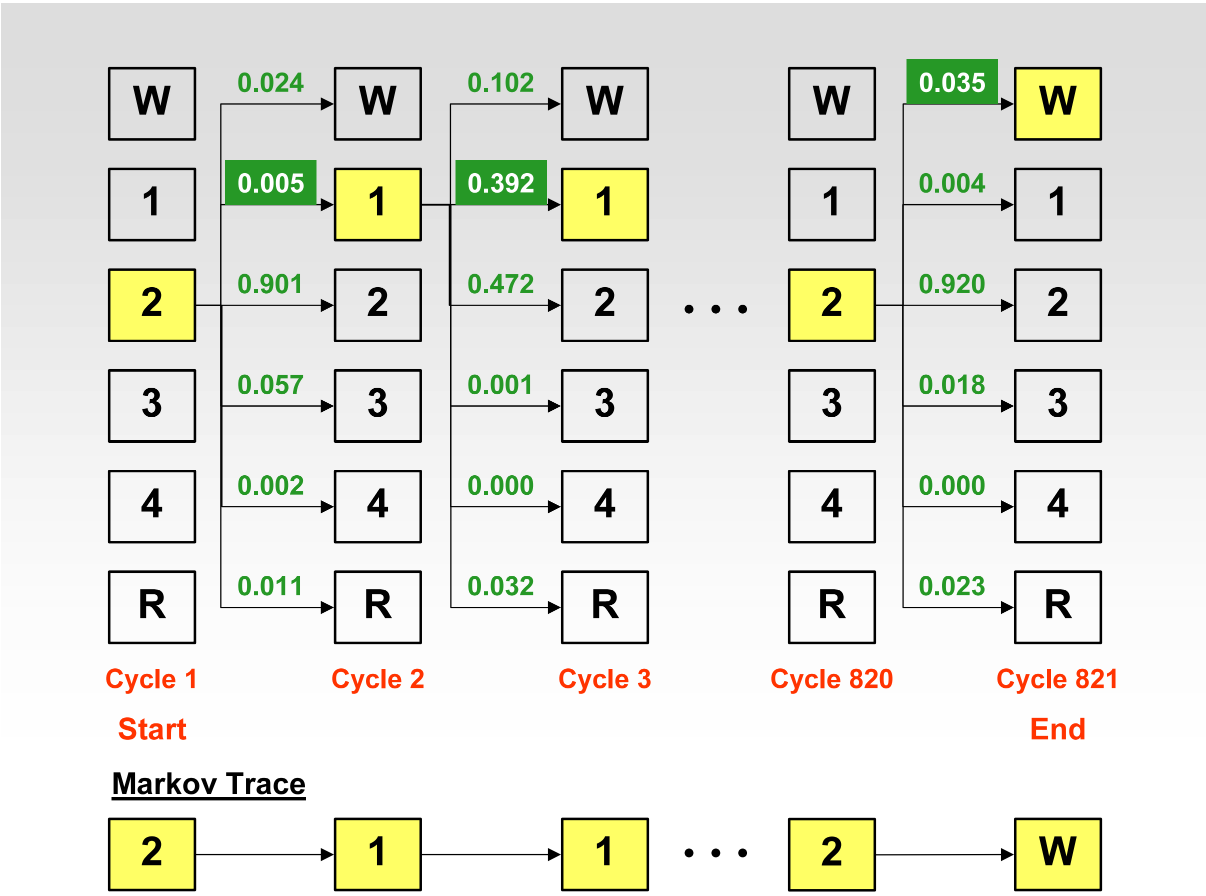


Figure 5.8: Another possible realization of a first-order Monte Carlo trial of the baseline nights.

In the first-order simulation trial shown in Figure 5.8, stage 1 is randomly chosen for cycle #2, which is likely to happen in only 5 out of 1,000 trials. Now, the transition probabilities from cycle #2 to cycle #3 are based on stage 1 as the present sleep stage, and there is an obvious change compared to the transition probabilities from cycle #2 to cycle #3 in the first example shown in Figure 5.7, e.g. the transition probability to stage 2 decreases from 0.901 to 0.472. Based on these transition probabilities, stage #1 is randomly chosen again for cycle #3, which was a likely event with a chance of 392 out of 1,000. This illustrates one of the most important features of Markov models: Realizations in the past may influence realizations in the near and even in the far future.

5.4.2 Noise model

As in the baseline model, Monte Carlo simulation trials were performed in the noise model. Three different noise scenarios, extracted from the timetable of Frankfurt Airport and described in detail below (see Chapter 5.4.2.1), were compared with a noise-free night and with each other according to simulation outcomes. As in the baseline model, a sleep period with a fixed length of 410.5 min (= 6.8 hours) was modeled, which coincides well with today's average adult sleep period time in Germany [37, 39]. For each scenario, 10,000 simulation trials were performed. As the time of falling asleep may substantially influence which part of the night is influenced by aircraft noise and to what extend, separate simulations were run according to different times of falling asleep (see Chapter 5.4.2.2). Figure 5.9 shows one possible realization of a first-order Monte Carlo trial of a night with aircraft noise.

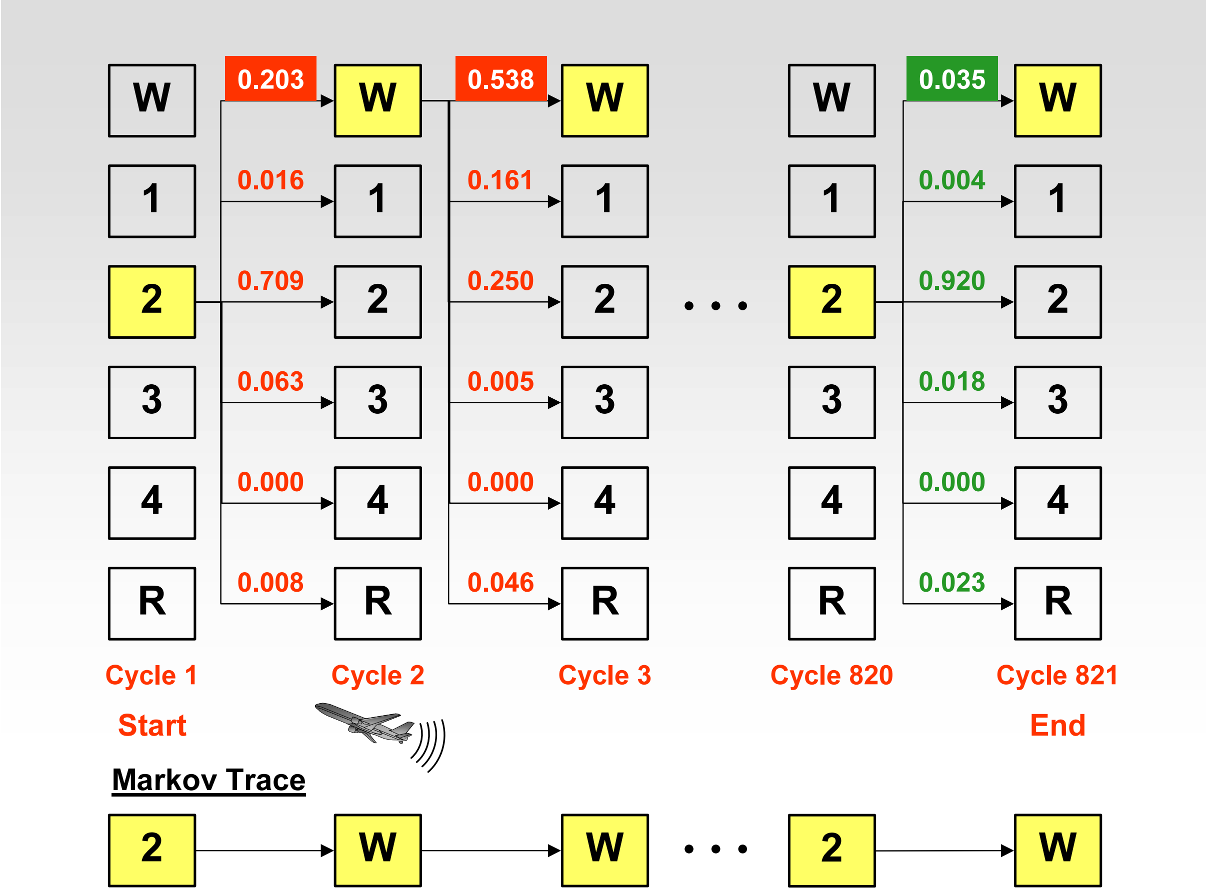


Figure 5.9: One possible realization of a first-order Monte Carlo trial of a night with aircraft noise. Transition probabilities were taken from the actual noise models used for the analyses.

Here, an ANE starts in cycle #2. Hence, the transition probabilities from cycle #1 to cycle #2 are altered compared to the noise-free condition (see Figure 5.7). The change in transition probabilities may persist longer than for one cycle only (see Chapter 5.3.2.2).

5.4.2.1 Description of the three noise scenarios

The three noise scenarios are based on data extracted from the timetable of Frankfurt Airport (valid from 17 July until 29 October, 2005). They are described in detail below.

Noise scenario 1

The first scenario consisted of a traffic pattern currently exercised at Frankfurt Airport, i.e. without restrictions[3] during the night. The scheduled times were extracted from the timetable only for take offs[4] and for the reference date Tuesday, August 16, i.e. in the middle of summer holidays. Therefore, passenger traffic is expected to be a little higher than during off-season. Although there might be fewer freight flights during this period, this should be compensated by tourist flights that take off during the night (a cheap slot) and head towards popular holiday locations. Taking off during the night allows those airlines a so-called "third rotation".

For the simulation, it was assumed that all take offs on August 16 were carried out from one runway (runway 25) and in one direction, and that the modeled residents lived close enough to the runway that they heard each plane before the flight paths diverge into different directions. Hence, the investigated traffic pattern will be found only at special locations around the airport and on days with high to extreme traffic volumes in reality, which is a conservative approach, as noise effects are rather overestimated than underestimated.

The flight schedule for Frankfurt Airport (take offs only) for Tuesday, 16 August 2005 (noise scenario 1) is shown in Figure 5.10.

---

[3] Actually, there are some restrictions, e.g. some older and very noisy planes (so-called Chapter 2 aircrafts) are not allowed to take off during the night and fees for take offs and landings are depending on the noise generated by the aircraft. For detailed information see http://www.boeing.com/commercial/noise/frankfurt.html (last visited 12 Apr 2006).

[4] For a discussion why only times of take offs, and not of approaches, were extracted from the timetable, please refer to Chapter 7.5.1.

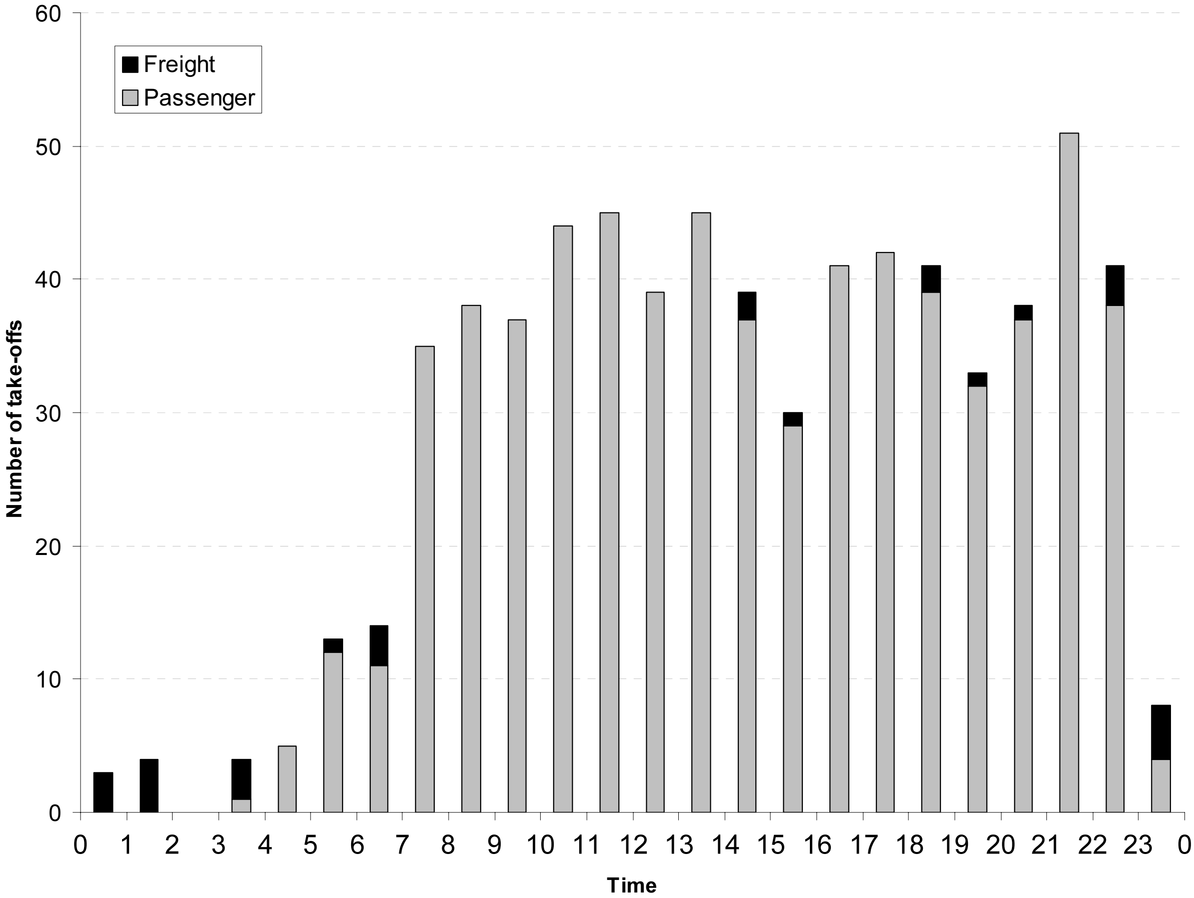


Figure 5.10: Noise scenario 1: Flight schedule for Frankfurt Airport (take offs only) for Tuesday, 16 August 2005. Freight includes mail delivery flights.

Traffic density was very high between 07:00 and 23:00 with more than 30 take offs per hour. During the time period from 23:00 until 05:00, which is planned to be free of air-traffic in the future, traffic density was low with less than 10 take offs per hour (average: four). The time period between 05:00 and 07:00 took an intermediate position with 13 and 14 take offs per hour, respectively. Freight traffic dominantly took place during the night, although there were some exceptions where take offs during the day were necessary to arrive at the destination on time.

There were 662 take offs of passenger flights (95.9%) and 28 take offs of freight flights (4.1%) in total. The night time is officially defined as the time period from 22:00 until 06:00. During this period, there were 60 take offs of passenger flights (76.9%) and 18 take offs of freight flights (23.1%)

altogether. During daytime (06:00 until 22:00), there were 602 take offs of passenger flights (98.4%) and 10 take offs of freight flights (1.6%) in total.

Noise scenario 2

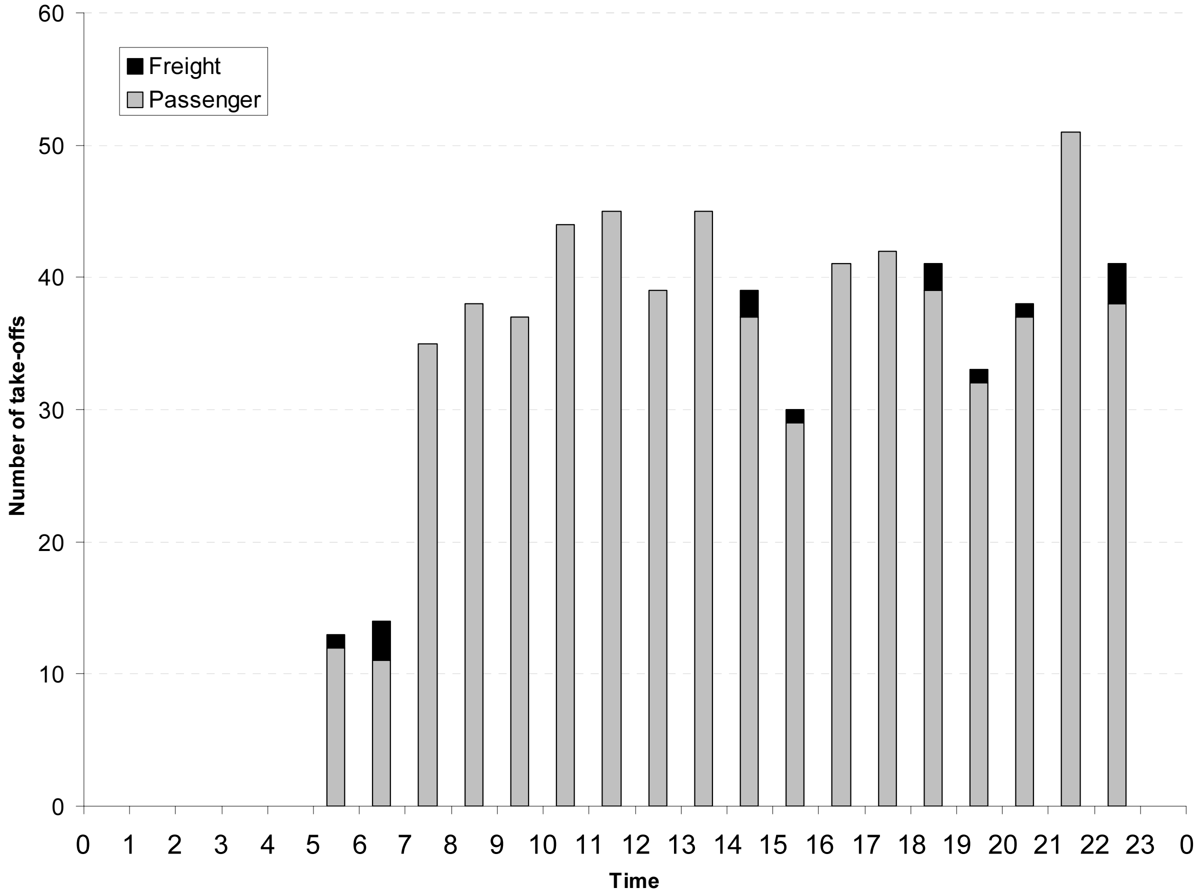


Figure 5.11: Noise scenario 2: Flight schedule for Frankfurt Airport (take offs only) for Tuesday, 16 August 2005. Here, the 24 flights between 23:00 and 05:00 were cancelled.

The second noise scenario is shown in Figure 5.11. The 24 flights between 23:00 and 05:00 were cancelled, otherwise the scenario did not differ from noise scenario 1.

Noise scenario 3

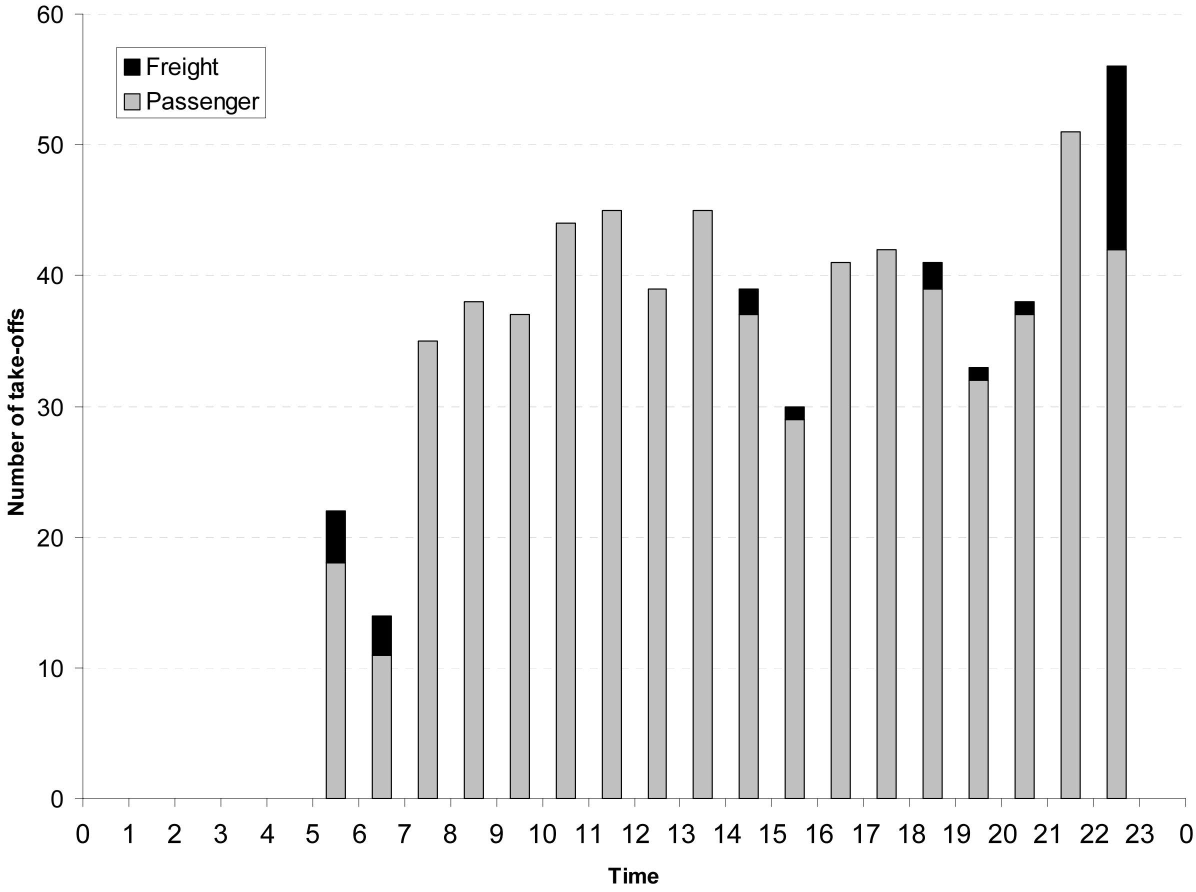


Figure 5.12: Noise scenario 3: Flight schedule for Frankfurt Airport (take offs only) for Tuesday, 16 August 2005. Here, the 24 flights between 23:00 and 05:00 were rescheduled to periods 22:00 to 23:00 (n=15) and 05:00 to 06:00 (n=9), respectively.

The third noise scenario is shown in Figure 5.12. The 15 take offs between 23:00 and 02:00 were rescheduled to the period between 22:00 and 23:00, resulting in a total of 56 take offs in this period[5]. The nine take offs between 03:00 and 05:00 were rescheduled to the period between 05:00 and 06:00, resulting in a total of 22 take offs in this period.

---

[5] Frankfurt's parallel runway system is not independent. Therefore, it may not be possible to handle 56 take offs per hour in reality. For a detailed discussion of this aspect see Chapter 7.5.1.

Comparison of the three noise scenarios

The time period from 22:00 until 06:00, that differs between the three noise scenarios, is shown in Figure 5.13. For simulation purposes, take offs were distributed equidistantly over the hour according to the number of take offs per hour.

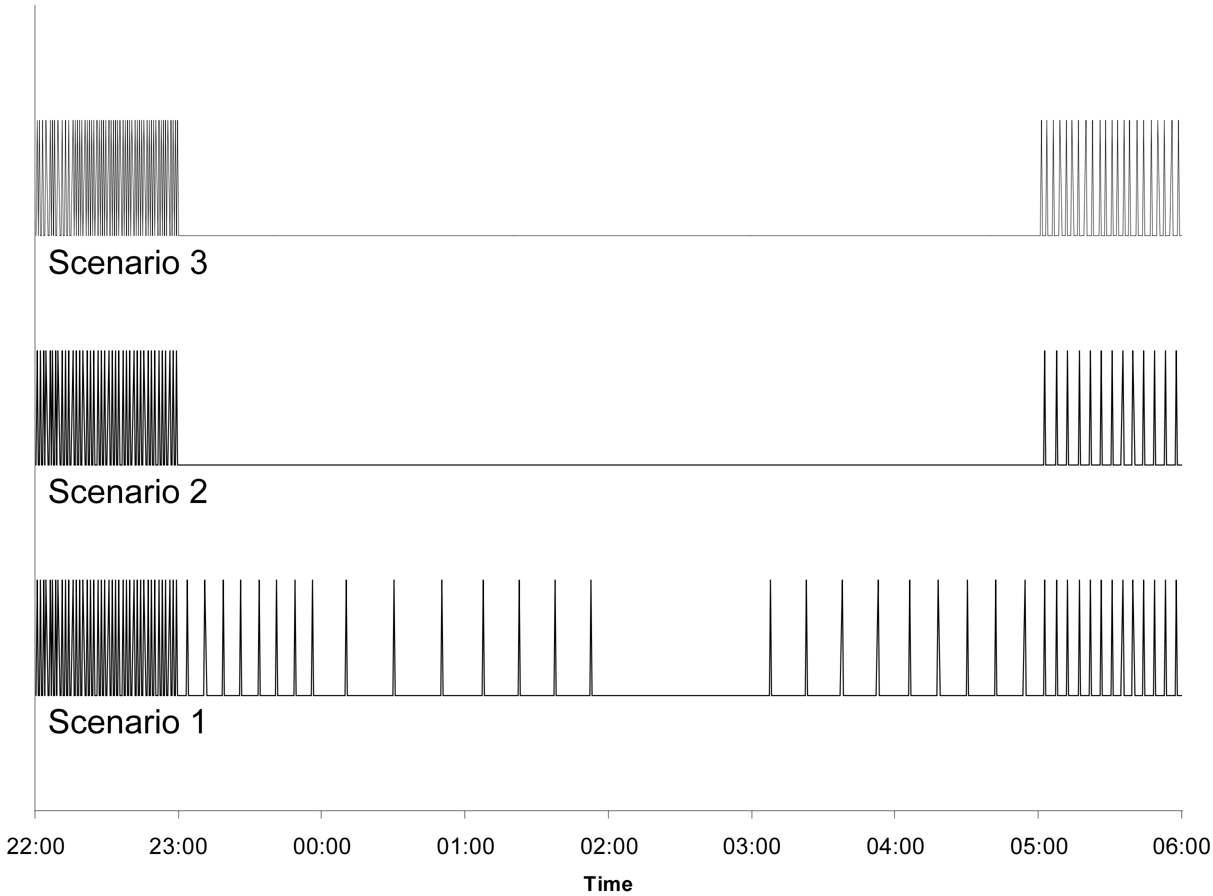


Figure 5.13: Comparison of the three noise scenarios according to the distribution of take offs used for the simulations (each spike marks one take off). Distribution of take offs did not differ between noise scenarios before 22:00 or after 06:00, and is therefore not shown here.

5.4.2.2 Consideration of different times of falling asleep

For the construction of the baseline night model, all subjects were synchronized to time of sleep onset. Naturally, residents living in the vicinity of Frankfurt Airport do not go to bed or get up all at the same time.

Depending on time of sleep onset, the same ANE is experienced at different times in the course of the sleep process: If someone falls asleep at 23:00, an ANE at 02:00 will be experienced three hours after sleep onset. Someone else falling asleep at 01:00 will experience the same ANE only one hour after sleep onset. As the reactions to the ANE are likely to differ depending on elapsed sleep time, different bed times have to be taken into account if the Markov model should validly predict changes in sleep structure for a larger population.

Fortunately, the sleep times of the adult German population have been evaluated using questionnaires [26]. The distribution of times when the adult German population is falling asleep is shown in Figure 5.14. In the investigated sample of adult subjects, 69% of the population was falling asleep between 22:00 and 00:00. 40% of the subjects fell asleep before 23:00.

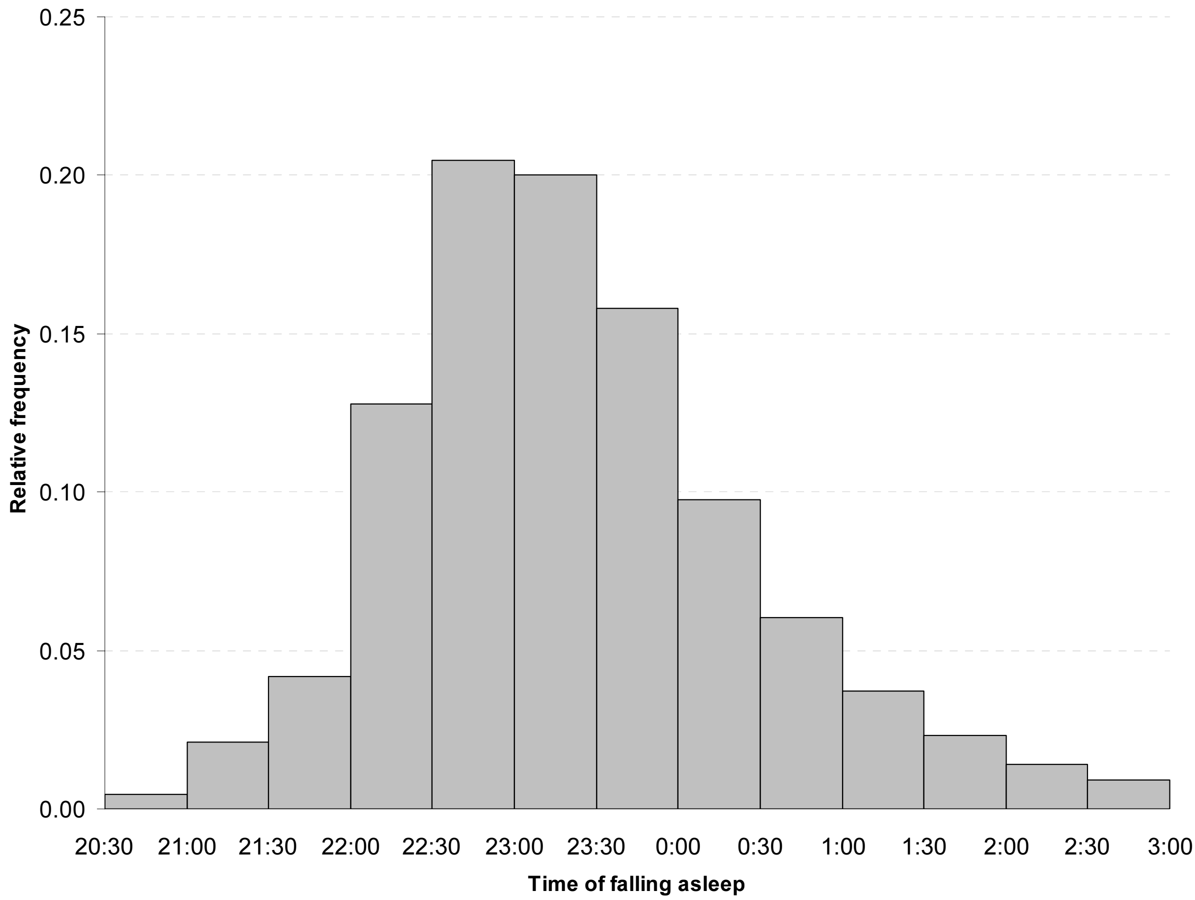


Figure 5.14: Frequency distribution of times of falling asleep for an adult subject sample (n=4357) [26].

The model was run several times with different starting times representing the different times of when people are falling asleep. For simplification and for a better comparison with noise-free baseline nights, the sleep period time was kept constant for everyone (410.5 min). Afterwards, the results were weighted using the relative frequency of people falling asleep at the peculiar point in time as weights, but results will also be given for each of the different groups according to time of falling asleep.

5.5 Outcomes

The outputs and variables described in the next sections can be characterized as typical descriptors of sleep structure and were therefore used as the main outcome variables in this thesis. Regression model selection and validation and comparison of simulation results were based on these outcomes.

5.5.1 Markov trace

If the realizations of the different sleep stages are plotted against cycle number during a simulation trial, a Markov trace is generated, which is the simulation equivalent of the human hypnogram (see Figure 3.3) and therefore allows for a comparison of the results of single simulation trials with human hypnograms.

5.5.2 Time spent in different sleep stages

The sleep stage composition of the night constitutes a very important aspect of sleep structure. As already mentioned above, the different sleep stages differ qualitatively and quantitatively in their contribution to sleep recuperation. Stages 3, 4 and REM are said to be very important for recuperation, whereas stage 1 and especially Wake contribute only little if all [14, 47]. As the main goal of this thesis was to investigate the influence of aircraft noise on sleep, the time spent in the different sleep stages represents an important outcome variable. It was stored accordingly for each simulation trial with the unit "minutes of sleep period time".

5.5.3 Number of sleep stage changes

Sleep stage changes, defined as transitions from one sleep stage to another (e.g. S1→S2), are a physiological part of sleep and necessary to generate the typical periodicities observed in human hypnograms. Environmental noise may lead to sleep fragmentation, which is characterized by frequent intermittent arousal, that are often accompanied by sleep stage changes.

An awakening from sleep stage 4 and the process of transitioning back to sleep stage 4 may, e.g., go along with the following transitions: S4→W→W→S1→S2→S2→S3→S4. If there had not bee the aircraft noise event, the hypnogram of the same subject might have looked like this: S4→S4→S4→S4→S4→S4→S4→S4. In this example, several sleep stage changes (S4→W, W→S1, S1→S2, S2→S3 and S3→S4) are induced by the aircraft noise event. In general, an increased number of sleep state changes indicates a higher degree of sleep fragmentation and therefore a relevant alteration of sleep structure.

5.5.4 Sleep quality index

The aim of the noise model was to predict noise induced changes in sleep structure, e.g. changes in the amounts of the different sleep stages. With noise, SWS and REM sleep may be reduced, while the proportions of Wake and sleep stages S1 and S2 may be increased under the influence of aircraft noise. Although the interpretation of these results is relatively straightforward, there are two main disadvantages of the presentation of results in form of changes in the amounts of the different sleep stages alone:

(1) It may be difficult for laypeople like airport residents or politicians, who are not familiar with the physiology of sleep, to interpret the results of the model.

(2) The Markov model of sleep consists of six states. Noise induced changes in sleep structure may lead to alterations in the amounts of all sleep stages simultaneously. Even experts in the field of sleep medicine will have problems in interpreting these multivariable results.

Therefore, it was desirable to express these multivariable results of the analyses in a single value, which should capture the most important aspects of sleep structure and sleep quality. In sleep medicine, the sleep efficiency index (SEI) is used for a similar purpose. Here, the time spent sleeping, i.e. in all sleep stages other than Wake, is divided by sleep period time (SPT). Therefore, SEI is solely dependent on the time spent awake and on sleep period time. The other sleep stages are not differentiated, although there is consensus that some sleep stages (e.g. S3, S4 and REM) are more important for recuperation. Hence, it would be desirable to weigh the different sleep stages according to their assumed restorative power.

Although little is known about the biologic mechanisms of how and how much the different sleep stages contribute to sleep restoration, it may be possible to sort the non-REM sleep stages according to their assumed restorative powers, i.e. in increasing order: S1→S2→S3→S4. The estimation of the position of REM sleep in this chain proofs more difficulty. Its restorative power will most probably be higher than that of S2, but is it higher or lower than S4, or even lower than S3? Work recently published demonstrates that the contributions of REM sleep to recuperation may not even be comparable to those of NREM sleep in a way that REM sleep e.g. contributes more to consolidation of emotional memory, whereas SWS does to consolidation of declarative memory [22, 24].

Despite of these facts, it has been proposed that the higher the arousal threshold of a certain sleep stage, the more this sleep stage will contribute to recuperation. To put it even simpler, the more important a sleep stage is for recuperation, the higher is the interest of the body in protecting those sleep stages from sleep disrupting stimuli like noise.

Table 5.3 shows the probabilities of noise induced awakenings under the influence of aircraft noise depending on the sleep stage prior to the arousing stimulus. Depending on the awakening probability, the different sleep stages can be sorted according to their assumed restorative power, and the distance between the sleep stages can be estimated, too.

Arbitrarily, the value of 0, meaning no restorative power, was attributed to stage S1 (the sleep stage with the highest probability of arousal) and the value of 1, meaning full restorative power, was attributed to S4 (the sleep stage with the lowest probability of arousal). The values of the other sleep stages can be calculated according to Table 5.3. The utility value of stage Wake was also set to zero.

Table 5.3: Attribution of utility values to the different sleep stages based on awakening probability p(0|Sleep Stage) observed during playback of aircraft noise. Each payoff is calculated as $(P_{S1} - P_{sleep\ stage}) / (P_{S1} - P_{S4})$.

Sleep Stage	**p(0|Sleep Stage)**	**Utility Value**
S1	0.445	0.000
S2	0.223	0.657
REM	0.161	0.840
S3	0.148	0.879
S4	0.107	1.000

For each sleep stage, the cumulative duration (in minutes of sleep period time) has to be multiplied by the utility value from Table 5.3. Then, the separate values have to be added to receive a total score. This total score will be called "sleep quality index score" (SQI score). High amounts of SWS and REM sleep lead to high SQI scores, whereas high amounts of S2 and especially of Wake and S1 lead to low SQI scores. Average sleep stage amounts of empirical data of noise-free baseline nights lead to a total SQI-score of 276.6. The contributions of the different sleep stages amounted to: 138.3 (S2), 35.5 (S3), 28.5 (S4) and 74.3 (REM).

The SQI-score of 276.6 was set to an SQI of 100%. SQI-scores of simulated noise nights were likely to be lower than 276.6 due to decreases in SWS and REM sleep and increases in S2, S1 and Wake. Therefore, the impact of aircraft noise on sleep quality may be expressed as reductions in SQI (with expected values smaller than 100%).

5.6 Validation

Predicted probabilities of the autoregressive multinomial logistic regression models had to be validated with empirical data (see Chapter 5.6.1). Additionally, the outcomes of Markov model simulation trials had to be validated with empirical data (see Chapter 5.6.2).

5.6.1 Validation of regression results

Formal tests for goodness of fit for logistic regression models, such as the Hosmer-Lemeshow test [30], may not be valid if there is dependence in the outcome series [20]. Therefore, a graphic method was used to assess goodness of fit, plotting as a function of elapsed sleep time sleep stage amounts empirically observed as well as sleep stage amounts predicted by the regression models. Predictions for T1 were weighted based on the amounts of the different sleep stages at T0. This procedure is described in detail and exemplified in Appendix A (Chapter 9).

For the assessment of goodness of fit, the predicted relative frequencies of all sleep stages and at every point in time ($T_{0.5}$ - $T_{410.5}$) were compared to the observed data. In the regression model building process (see Chapter 5.3.2), figures of these comparisons were produced for all regression models and compared with each other. Differences of predicted and observed relative frequencies were visually inspected for heteroscedasticity. For the final model, mean, standard deviation, 2.5, 25, 50, 75 and 97.5 percentiles were calculated for the difference of predicted transition probabilities and observed transition probabilities. The differences were tested for normality with the Kolmogorov-Smirnov-test (SPSS Inc., Version 11.0). A significant test results indicated non-normal data and therefore a possible systematic bias of transition probability predictions of the regression model.

Noise nights

The methods described for the assessment of goodness-of-fit for the baseline night regression models was identically used for the assessment of goodness-of-fit of the noise regression models. Because of the relatively sparse data compared to the baseline model, the 410.5 min night was divided in six periods of 58.5 min and one period of 59.5 min. Observed and predicted probabilities were averaged over these intervals and compared with each other.

5.6.2 Simulation of baseline nights: Outcome validation

The results of the Monte Carlo simulation trials were aimed at reproducing characteristic features of the original baseline nights.

Markov traces of individual first-order trials should reflect key features of human hypnograms (face validity). Therefore, the Markov traces of the first 100 Monte Carlo simulation trials were stored and compared to human hypnograms by the author according to the following criteria:

(1) Slow wave sleep stages 3 and 4 should dominate the first half, whereas stage 2 and REM should dominate the second half of the night.

(2) The periodic nature of sleep cycles (NREM→REM→NREM→REM→etc.) should be mirrored by the model.

(3) The density of awakenings, i.e. the number of awakenings per hour, should increase towards the end of the night.

(4) There should be no sleep onset REM-episodes, i.e. REM sleep at the very beginning of the night.

For each sleep cycle, averaging over all simulated Markov traces leads to a probability distribution of sleep stages. This probability distribution was compared to the probability distribution of sleep stages observed in the original data, i.e. for every sleep epoch after sleep onset, the proportions of the different sleep stages were calculated based on empirical data of the 125 subjects. Both probability distributions were visually compared in graphs where simulated and observed probabilities were simultaneously plotted against elapsed sleep time. The simulated probability distribution should resemble the probability distribution observed in the original data as closely as possible. Ultradian rhythms observed in the original data should be reproduced by the simulations.

Finally, the outcome variables of Monte Carlo simulation trials (time spent in different sleep stages, number of sleep stage changes and SQI-score) were compared to empirically observed outcomes in baseline nights. For this purpose, mean, relative bias, standard deviation, 2.5, 25, 50, 75 and 97.5 percentiles were calculated for simulated and observed data and compared. Relative bias was defined as (predicted value / observed value – 1) x 100 [%].

The result of statistical tests comparing two independent groups depends on the size of the groups. As the number of simulation trial runs can be chosen at will and without noticeable expense, it was not reasonable to compare the results of simulated and observed data with a statistical test.

5.7 Hypotheses

The following hypotheses about the influence of aircraft noise on the outcome variables are proposed:

On the one hand, the amounts of Wake and S1, classical indicators of sleep fragmentation [47], are expected to increase with aircraft noise. Likewise, the number of sleep stage changes is expected to increase. On the other hand, amounts of S3, S4 and REM are expected to decrease simultaneously. Amounts of S2 are expected to change only marginally, as both the probability of awakenings from S2 and the probability of changes from S3, S4 and REM to S2 should increase under the influence of noise. The SQI is expected to decrease under the influence of noise.

It is hypothesized that the effects of noise on sleep are attenuated, but not eliminated by the introduction of a noise-free period from 23:00 – 05:00, i.e. effects decrease in the order scenario 1 → scenario 2 → noise-free night.

The effect of rescheduling take offs from 23:00 – 05:00 to the periods before and after (scenario 3) was unpredictable and therefore one of the main reasons for the analyses presented in this thesis: Rescheduling of take offs (scenario 3) will definitely go along with more severe changes in sleep structure than cancellation of these take offs (scenario 2). But will the changes in sleep structure be more or less severe than those observed in scenario 1, i.e. in the present condition without a noise-free period? Will the introduction of a noise-free period benefit or harm airport residents, if take offs that formerly took place during the night are rescheduled to periods before 23:00 and after 05:00?

6 Results

6.1 Description of Empirical Data

Key features of the noise-free baseline nights are described in detail in this chapter.

The different sleep stages are not uniformly distributed over the night. As was already mentioned in Chapter 3.1 and shown in the hypnogram in Figure 3.3, SWS dominates the first half of the night, whereas the probability of light sleep and REM sleep is higher in the second half of the night. Additionally, the relative frequencies of the different sleep stages are not equal: The relative frequency of S2 is highest (51.2%), the one of S1 lowest (1.7%). The distribution of the probability of the different sleep stages during the night derived form the first 410.5 min of the sleep period of the noise-free baseline nights of 125 subjects is shown in Figure 6.1.

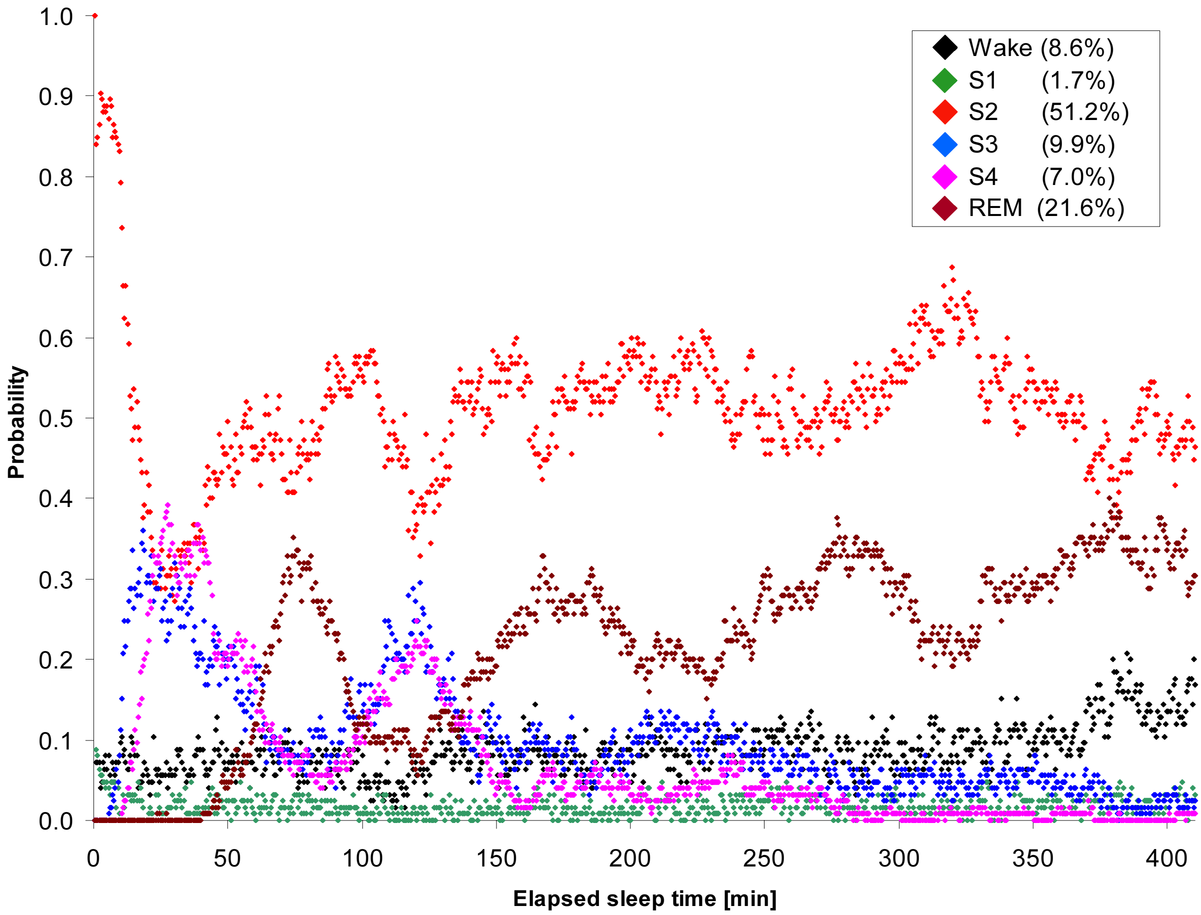


Figure 6.1: Empirical probability distribution of the different sleep stages depending on elapsed sleep time. Proportions of the different sleep stages based on sleep period time are given in parenthesis in the legend.

Wake and S1 are almost uniformly distributed across the night. As sleep pressure decreases with elapsed sleep time, the probability of Wake tends to increase towards the end of the night. S1 is slightly increased during the first minutes of sleep, indicating that the process of falling asleep was not finished in all subjects with the first occurrence of S2 (the definition of sleep onset). Ultradian rhythms are clearly visibly in the probability distribution of sleep stages S2, S3, S4 and REM. The probability of S2 simultaneously decreases as the probabilities of SWS or REM sleep increase. With a latency of circa 45 min after sleep onset, the probability of REM sleep increases steadily during the night superimposed by a periodic ultradian pattern. The probabilities of S3 and S4 increase shortly after sleep onset. At least in younger subjects, S3 can be described as a transitional stage between S2 and S4. Therefore, the probability peak of S3 precedes that of S4 by a few minutes. The probability of SWS decreases during the night. At the end of the night, S3 only rarely and S4 almost never occurs. Ultradian rhythmicity can clearly be seen in the probabilities of S2, S3, S4 and REM at the beginning of the sleep period, as all data were synchronized for sleep onset. This rhythmicity is more and more obscured towards the end of the night.

Average transition probabilities from sleep stage at T0 to sleep stage at T1, unconditional for elapsed sleep time or for the sleep stage distribution of sleep stages further in the past (T-1, T-2, etc.), are shown in Table 6.1. For all sleep stages, the probability to stay in the same sleep stage is highest, except for sleep stage S1, which is known to be a transitional stage between waking and sleeping. Certain transitions occur seldom (e.g. Wake→S3 or S4) or never (e.g. S1 or REM→S4).

Table 6.1: Transition probability matrix of sleep stages at T0 and T1. Transition probabilities are based on 102,500 transitions in total and are given as percentages of sleep stage amounts at T0.

			Sleep Stage at T1						
			Wake	S1	S2	S3	S4	REM	Total
Sleep Stage at T0	Wake	#	6003	810	1452	21	1	552	8839
		% of T0	67.9%	9.2%	16.4%	.2%	.0%	6.2%	
	S1	#	220	627	840	1	0	85	1773
		% of T0	12.4%	35.4%	47.4%	.1%	.0%	4.8%	
	S2	#	1528	249	48102	1761	45	863	52548
		% of T0	2.9%	.5%	91.5%	3.4%	.1%	1.6%	
	S3	#	183	1	1458	7587	888	11	10128
		% of T0	1.8%	.0%	14.4%	74.9%	8.8%	.1%	
	S4	#	109	1	62	760	6222	1	7155
		% of T0	1.5%	.0%	.9%	10.6%	87.0%	.0%	
	REM	#	817	89	567	1	0	20583	22057
		% of T0	3.7%	.4%	2.6%	.0%	.0%	93.3%	
Total		#	8860	1777	52481	10131	7156	22095	102500
		% of T0	8.6%	1.7%	51.2%	9.9%	7.0%	21.6%	100.0%

It was mentioned in Chapter 3.3 that transition probabilities in Markov models depend only on the current state and not on past states (so-called Markovian assumption). In this way, the Markov model has no memory for states preceding the current state. It was therefore important to find out whether transition probabilities derived from empirical data actually depend on states further in the past, i.e. on sleep stage history. Detailed analyses showed that transition probability do depend on sleep stage history. This is exemplified for sleep stage 2 in Figure 6.2.

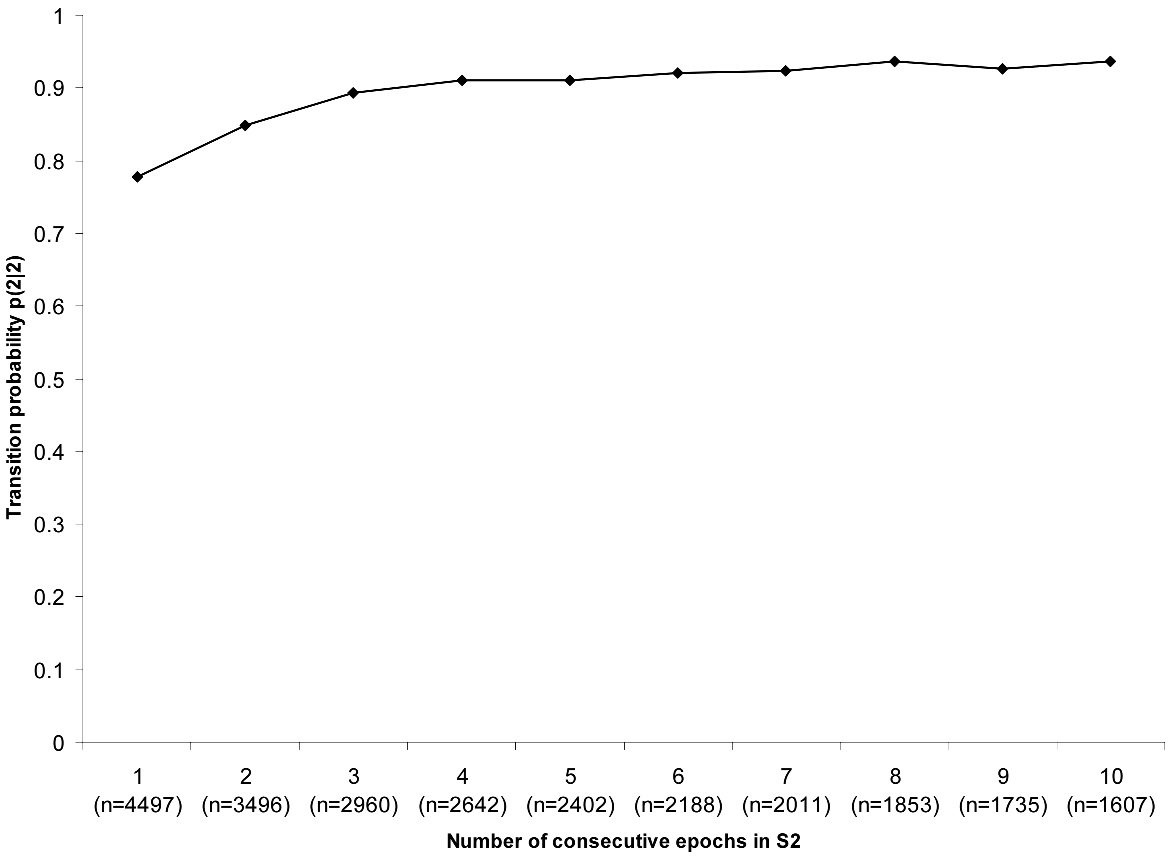


Figure 6.2: Transition probabilities from S2 to S2 depending on the number of epochs already spent in S2 calculated from noise free baseline nights (n = number of epochs with the respective history).

The probability to stay in sleep stage 2 steadily increased with the time already spent in stage 2, especially during epochs 1 to 4. Implications of this finding for the validity of the models described below will be discussed in Chapter 7.3.2.

6.2 Baseline model results

6.2.1 Estimation of transition probabilities: Results of the regression model building process

As none of the models described in Chapter 5.3.2 managed to reproduce ultradian rhythms in Monte-Carlo simulation trials (see Chapter 9.2), a more sophisticated model was built that was aimed at reproducing this rhythmicity:

NREM-episodes and REM-episodes were introduced in Chapter 3.1 and visualized in Figure 3.3. Each sleep cycle consists of one NREM-episode and one REM-episode, which therefore constitute the basic elements of ultradian rhythmicity. The beginning and the end of every REM-episode were marked visually, and the variable "elapsed sleep time" was substituted by the variable "elapsed NREM-episode-time" or "elapsed REM-episode-time". For each NREM- and REM-episode, counting from sleep onset, separate first-order autoregressive multinomial logistic regression models were built, including "elapsed NREM- or REM-episode-time" as the only additional explanatory variable. Altogether, seven NREM- and six REM-models were built.

In the Monte-Carlo simulation, a REM-episode was entered every time the random number generator produced a sleep stage REM during a NREM episode. As stages Wake, S1 and S2 are quite common during a REM-episode, and do not automatically go along with the termination of the REM-episode, it had to be decided whether the REM-episode was terminated whenever the random number generator led to Wake, S1 or S2, or not. Hence, the probability of REM-episode-termination was calculated based on elapsed time of the REM-episode, where the order of the REM-episode was accounted for. If the night ends during a REM-episode, this REM-episode is censored. Therefore, a Cox regression was used to estimate probability of REM sleep termination conditional on having stayed in the REM-episode until that moment (*Baseline* statement in Proc PHREG, SAS Version 8.2). Altogether, six Cox regression models were built, one for each REM-episode. Every time the random number generator produced a sleep stage Wake, S1 or S2 during a REM-episode, probability of REM sleep termination was calculated, and, based on this probability, the result of another run of the random number generator decided whether the REM sleep episode was terminated or not. However, even using this more complex approach, no ultradian rhythmicity was seen averaging over the Markov traces of all Monte Carlo simulations.

Comparing the simple first-order autoregressive model with the other models, none of the more complex models with additional explanatory variables lead to a noticeable improvement of goodness-of-fit or to a decrease in bias of main parameters of interest compared to the simple first-order model. Therefore, it was decided to use the most parsimonious, yet biologically reasonable transition probability model, i.e. the simple first-order autoregressive model with elapsed sleep time as the only additional explanatory variable. This model will be called "the final model" or "AR1" from now on. The regression results of this model are shown in Chapter 10.1.

As the relative frequencies of the different sleep stages change in the course of the night (see Figure 6.1), transition probabilities are not constant during the night. This is illustrated in Figure 6.3 for the transition probability p(3|2), i.e. the probability of a transition from sleep stage S2 to sleep stage S3: Transition probabilities from S2 to S3, as sleep pressure in general, tend to decrease during the night, resulting in SWS being entered less frequently. Additionally, a cyclic pattern in the transition probability can be observed, visualizing ultradian rhythms of sleep. This rhythmicity is more and more obscured towards the end of the night.

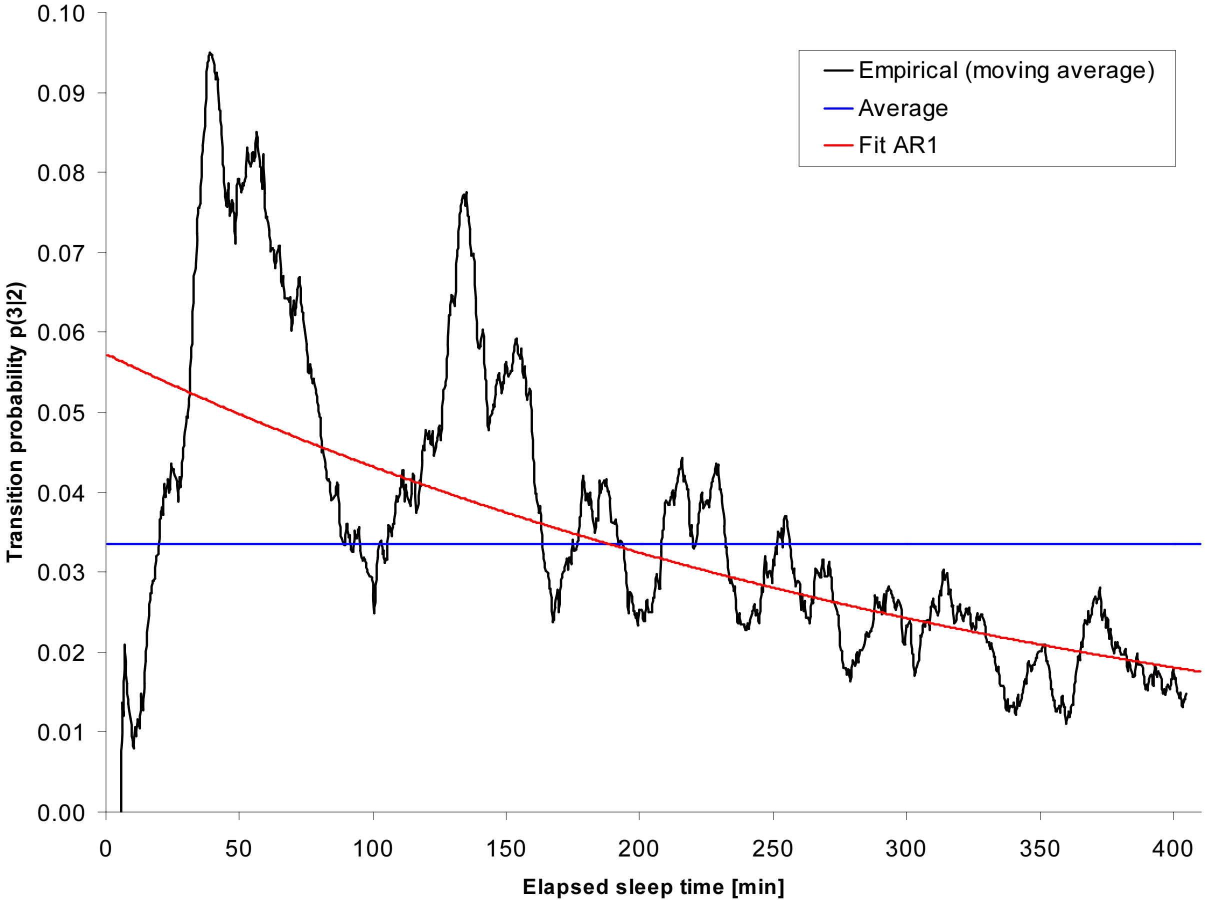


Figure 6.3: Probability p(3|2) of transitions from S2 to S3 depending on elapsed sleep time. A moving average of +/- 10 epochs based on the raw data (Empirical) is compared to the whole-night average transition probability (Average) and to the transition probabilities predicted by a first-order autoregressive model (AR1) containing elapsed sleep time as the only additional explanatory variable (Fit AR1).

The whole-night empirical average transition probability p(3|2) from Table 6.1 (blue line) fits the original probabilities poorly. The decreasing trend in the course of the night is not reproduced. The simple first-order autoregressive model (red line, see Chapter 6.2.1) containing elapsed sleep time as the only additional explanatory variable fits the original data much better, although ultradian rhythms are not reproduced by this model, either.

6.2.2 Estimation of transition probabilities: Validation of the final regression model

The goodness-of-fit of the final model was assessed in several ways. A graphical analysis is shown in the figures in Appendix A (9.1), where, for each sleep stage, two diagrams are shown. In the first diagram, predicted and observed probabilities are plotted simultaneously against elapsed sleep time. In the second diagram, the differences p(predicted) - p(observed) are plotted against elapsed sleep time.

Table 6.2 provides a numerical summary of all goodness-of-fit evaluations. Mean, standard deviation (SD) and 2.5, 25, 50, 75 and 97.5 percentiles are given. Additionally, the results of the Kolmogorov-Smirnov test for normality of the differences p(predicted) - p(observed) are shown for each of the six sleep stages, where p-values < 0.05 indicate a significant deviation from normality. The mean difference p(predicted) - p(observed) was zero for all sleep stages, i.e. there was no systematic bias in the predicted probabilities.

Table 6.2: Descriptive statistics of differences p(predicted) – p(observed) and results of the Kolmogorov-Smirnov-test for normality of the differences p(predicted) – p(observed).

			Percentiles						
Stage	**Mean**	**SD**	**2.5**	**25**	**50**	**75**	**97.5**	**Z(KS)**	**p(KS)**
Wake	0.000	0.019	-0.037	-0.012	-0.000	0.014	0.036	0.810	0.528
S1	0.000	0.011	-0.023	-0.007	0.001	0.008	0.019	1.769	0.004
S2	0.000	0.026	-0.052	-0.018	-0.000	0.017	0.052	0.628	0.826
S3	0.000	0.020	-0.048	-0.010	0.001	0.012	0.041	1.792	0.003
S4	0.000	0.011	-0.026	-0.005	0.001	0.005	0.025	2.764	<.001
REM	0.000	0.015	-0.034	-0.009	0.001	0.010	0.027	1.358	0.050

KS: Kolmogorov-Smirnov-test for normality; Z: test value of KS-test, p: p-value of KS-test

Histograms of the 821 differences are shown in Figure 6.4 for each of the six sleep stages. A normal distribution with mean and standard deviation derived from the observed differences is superimposed onto each distribution.

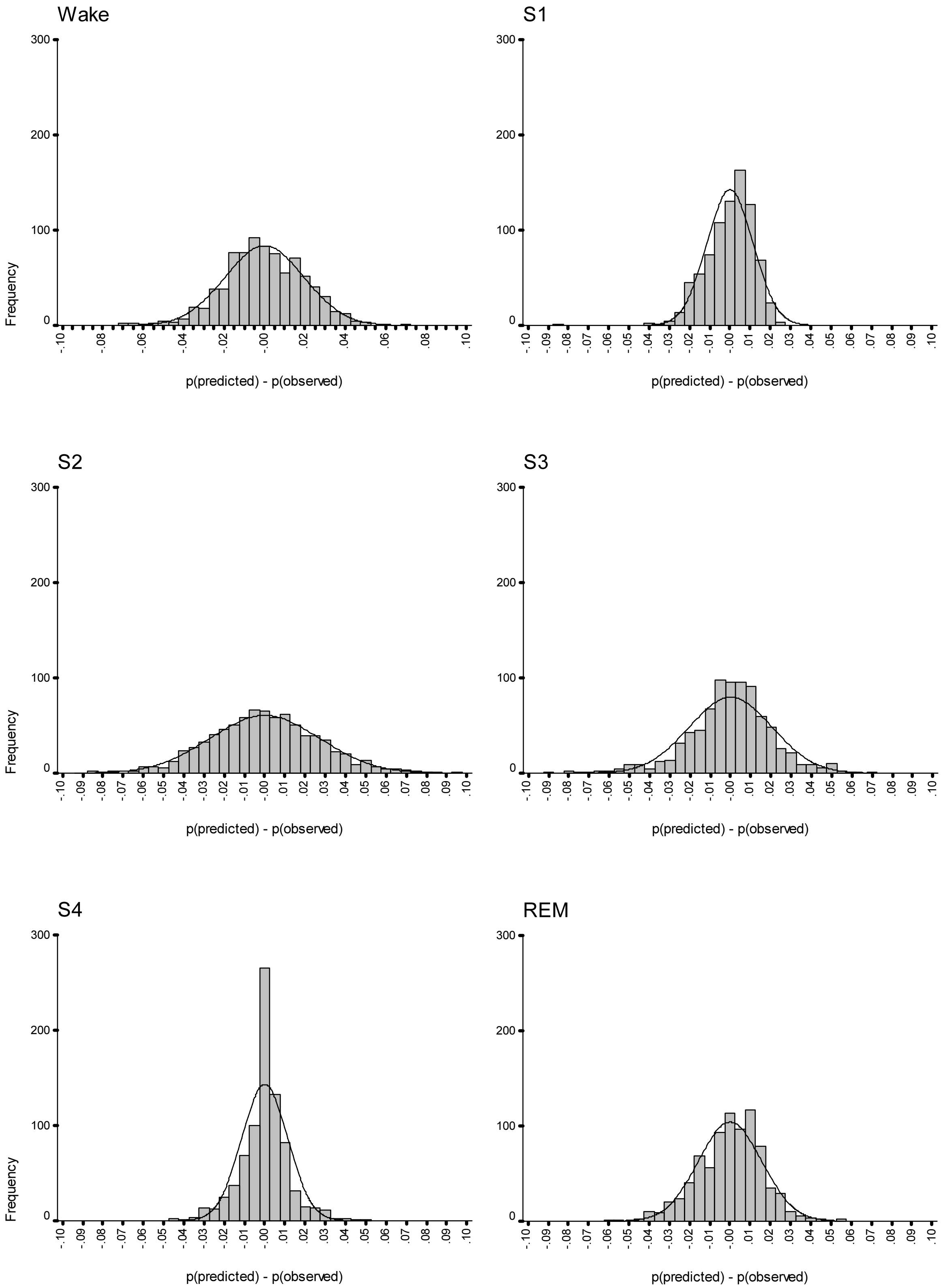


Figure 6.4: Histograms of the differences in predicted and observed probabilities for the six sleep stages. A normal distribution with mean and standard error derived from the observed differences is superimposed onto each distribution.

In the following, the goodness-of-fit results will be described separately for each of the six sleep stages. They will be interpreted and discussed in detail in Chapter 7.

Wake There was a good fit between predicted and observed probabilities (Figure 9.1). The increasing probability of wakefulness towards the end of the night is reproduced by the model. There is no heteroscedasticity over elapsed sleep time (Figure 9.2). The differences p(predicted) - p(observed) fluctuate with a standard deviation of 0.019 around a zero mean. The 2.5 and 97.5 percentiles are –0.037 and 0.036, respectively. The histogram in Figure 6.4 and the results of the Kolmogorov-Smirnov test do not indicate significant deviations from a normal distribution (p=0.528).

S1 The predicted probabilities matched well with the observed probabilities (Figure 9.3). The steep decrease in probability of S1 at the beginning of the night is reproduced by the model. Both predicted and observed probabilities remain on a low level during the whole night. There is no heteroscedasticity over elapsed sleep time (Figure 9.4). The differences p(predicted) - p(observed) fluctuate with a standard deviation of 0.011 around a zero mean. The 2.5 and 97.5 percentiles are –0.023 and 0.019, respectively. The histogram in Figure 6.4 is slightly skewed to the left and the results of the Kolmogorov-Smirnov test indicate a significant deviation from a normal distribution (p=0.004).

S2 There was a good fit between predicted and observed probabilities (Figure 9.5). Ultradian rhythms are reproduced by the predicted probabilities. There is no heteroscedasticity over elapsed sleep time (Figure 9.6), although predicted probabilities tend to be a little lower at the very beginning of the night. The differences p(predicted) - p(observed) fluctuate with a standard deviation of 0.026 around a zero mean. The 2.5 and 97.5 percentiles are –0.052 and 0.052, respectively. The histogram in Figure 6.4 and the results of the Kolmogorov-Smirnov test do not indicate significant deviations from a normal distribution (p=0.826).

S3 The predicted probabilities matched well with the observed probabilities (Figure 9.7). Ultradian rhythms are reproduced by the predicted probabilities. The predicted probabilities tend to be a little higher at the very beginning of the night. The variance in the differences p(predicted) - p(observed) decreases simultaneously with the probability of being in stage S3 towards the end of the night (Figure 9.8). The differences fluctuate with a standard deviation of 0.020 around a zero mean. The 2.5 and 97.5 percentiles are –0.048 and 0.041, respectively. The histogram in Figure 6.4 and the results of the Kolmogorov-Smirnov test show significant deviations from a normal distribution ($p=0.003$).

S4 There was a good fit between predicted and observed probabilities (Figure 9.9). Ultradian rhythms are reproduced by the predicted probabilities. The variance in the differences p(predicted) - p(observed) decreases simultaneously with the probability of being in stage S4 towards the end of the night (Figure 9.10). The differences fluctuate with a standard deviation of 0.011 around a zero mean. The 2.5 and 97.5 percentiles are –0.026 and 0.025, respectively. The histogram in Figure 6.4 and the results of the Kolmogorov-Smirnov test show significant deviations from a normal distribution ($p<.001$).

REM The predicted probabilities matched well with the observed probabilities (Figure 9.11). Ultradian rhythms are reproduced by the predicted probabilities. The predicted probabilities are higher in the first 40 min of the night, where the observed probability of REM is zero or close to zero. This is also seen in Figure 9.12. After 50 min of sleep time, there is no heteroscedasticity over the rest of elapsed sleep time. The differences fluctuate with a standard deviation of 0.015 around a zero mean. The 2.5 and 97.5 percentiles are –0.034 and 0.027, respectively. The histogram in Figure 6.4 and the results of the Kolmogorov-Smirnov test do not indicate significant deviations from a normal distribution ($p=0.0502$).

Altogether, the probabilities estimated by the simple first-order autoregressive multinomial logistic regression model match excellently with the observed probabilities, although there are some deviations from the observed probabilities in the very beginning of the night for sleep stages S2, S3 and REM.

6.2.3 Simulation of baseline nights: Description and validation of outcomes

Of the 100 first simulation runs, two Markov traces were selected by the author as extreme examples: One reproduced key features of human sleep best (Figure 6.5), the other one did so worst (Figure 6.6).

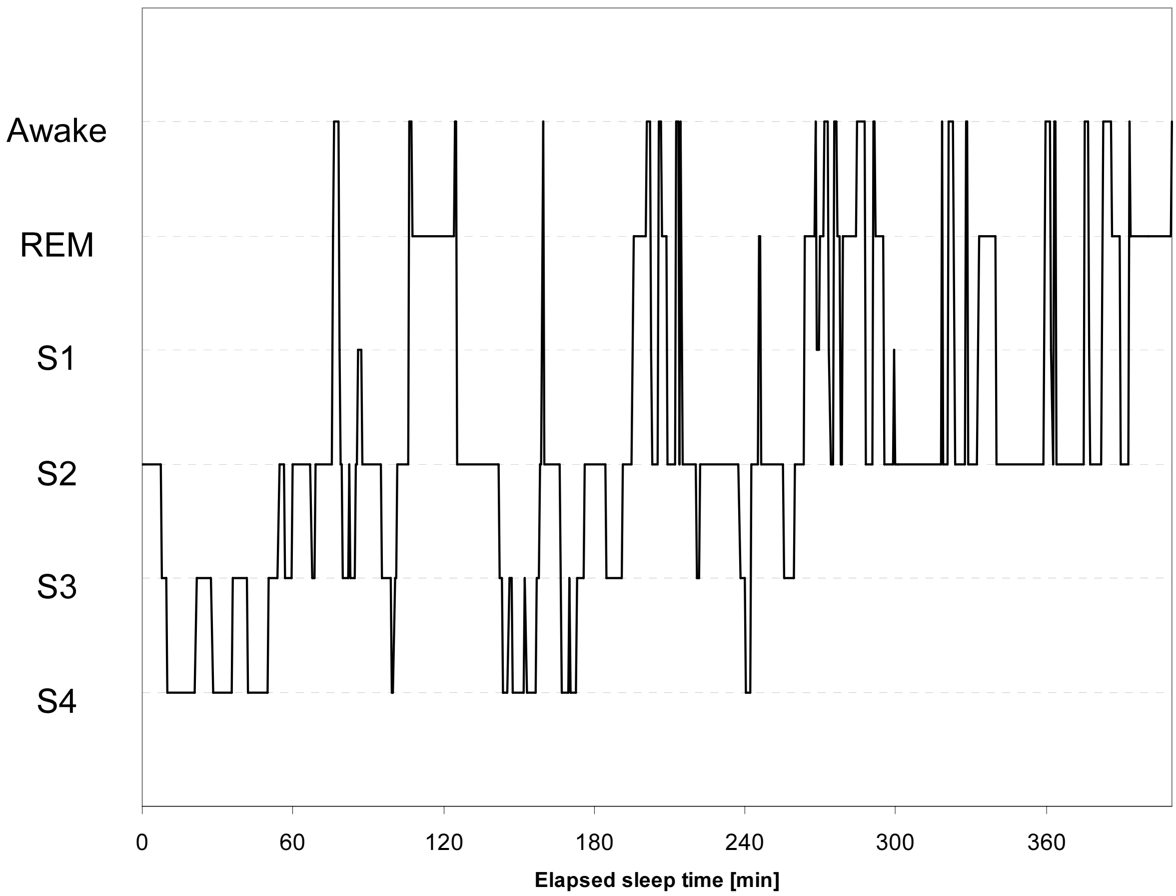


Figure 6.5: "Best" Markov trace of the first 100 simulation trials.

In Figure 6.5, SWS is entered quickly. The first REM-episode starts with a reasonable latency period after sleep onset. Five distinct REM-episodes can be differentiated. Ultradian rhythms are clearly visible. SWS dominates the first half, whereas REM sleep and light sleep dominate the second half of the night. The number of awakenings increases towards the end of the night. Even for experts, it would be hard to distinguish the hypnogram of Figure 6.5 from a human hypnogram.

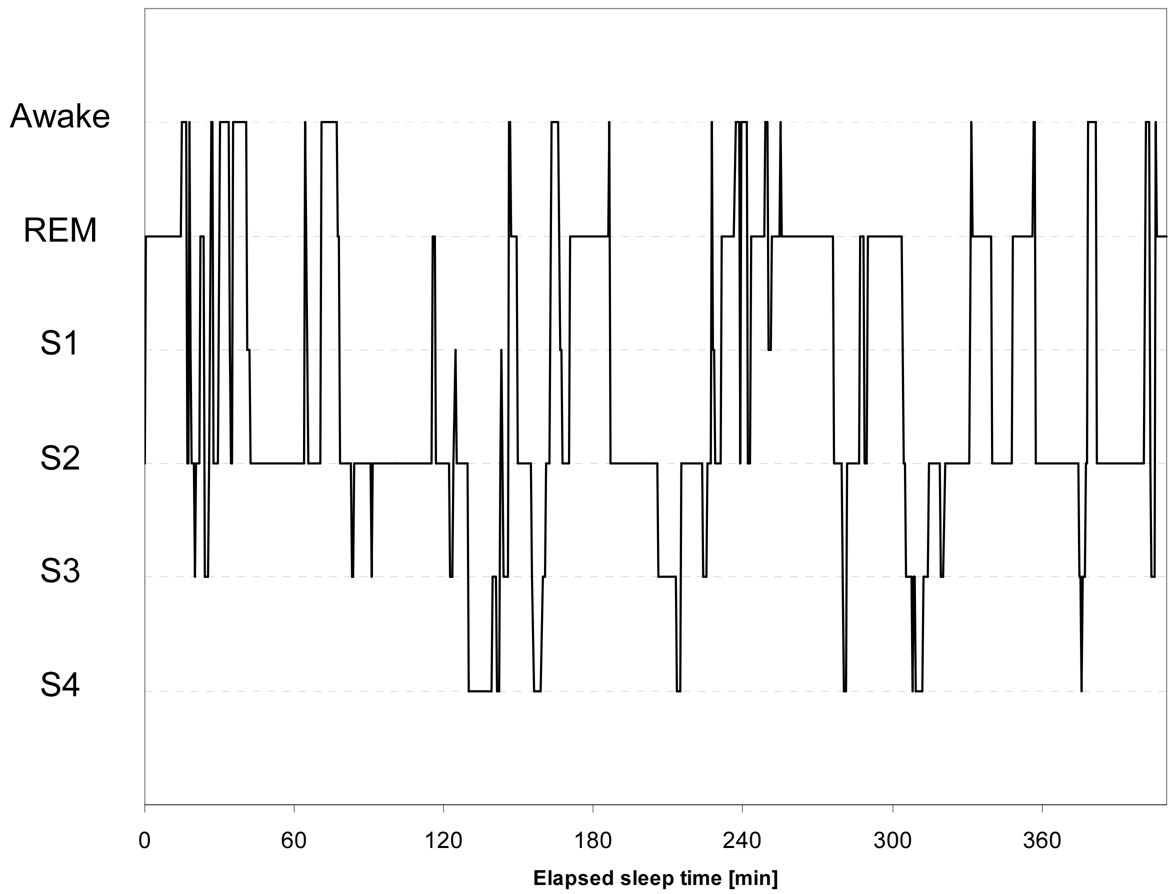


Figure 6.6: "Worst" Markov trace of the first 100 simulation trials.

In Figure 6.6, the night starts with a sleep onset REM, i.e. immediately after sleep onset the first REM-episode is entered. Sleep onset REM is usually not seen in healthy sleepers, although it may be found in jet-lag or in shift workers. It is a typical sign of narcolepsy. Sleep onset REM-episodes occur in simulation trials because transition probabilities to stage REM estimated by the regression model (AR1) at the beginning of the night are close to zero, but not equal to zero (as in the observed data). In Figure 6.6, SWS seems to be focused on the middle, not on the first half of the night. There is no clear tendency of an increasing density of awakenings towards the end of the night. Still, ultradian rhythms can be identified.

Altogether, the simulation was able to reproduce key features of human sleep even on the level of single simulation trials. Most of the hypnograms show the classical features of human sleep, such as the dominance of SWS in the first and of REM sleep in the second half of the night (see also figures in the Appendix). For many of them, it would be hard for an expert to distinguish the Markov trace of the simulation (e.g. Figure 6.5) from a typical human hypnogram.

For each of the 821 sleep epochs after sleep onset, the probability being in one of the different sleep stages was estimated for every sleep stage by averaging over Monte Carlo simulation outcomes. For the assessment of goodness-of-fit, these probabilities were then graphically compared to the empirical data, i.e. probabilities observed in the 125 subjects. For this purpose, simulated and observed probabilities were jointly plotted against elapsed sleep time for each of the six sleep stages. The figures are shown in the Appendix (9.2).

Table 6.3 provides a numerical summary of goodness-of-fit evaluations. Mean, standard deviation (SD), relative bias and the 2.5, 25, 50, 75 and 97.5 percentiles of both simulated (above and bold) and observed (below) data are given. The simulated means matched the observed means very well. Differences (simulated – observed) amounted to (relative bias given in parenthesis): Wake +0.1 min (+0.2%), S1 +0.1 min (+0.7%), S2 +0.5 min (+0.2%), S3 -0.1 min (-0.3%), S4 ±0.0 min, REM -0.5 min (-0.6%), SQI-score -0.2 (-0.1%) and number of sleep stage changes +0.2 (+0.2%).

Table 6.3: Comparison of simulated (above and bold font, n=10,000) and observed (below and normal font, n=125) data.

				Percentiles				
Stage	**Mean**	**Rel. Bias**	**SD**	**2.5**	**25**	**50**	**75**	**97.5**
Wake [min]	**35.5** 35.4	**+0.2%**	**8.9** 25.2	**19.5** 8.2	**29.0** 18.0	**35.0** 25.5	**41.0** 45.0	**54.0** 104.1
S1 [min]	**7.2** 7.1	**+0.7%**	**2.7** 7.0	**2.5** 0.0	**5.0** 2.0	**7.0** 5.5	**9.0** 10.0	**13.0** 31.9
S2 [min]	**210.9** 210.4	**+0.2%**	**27.6** 35.4	**155.5** 145.5	**192.5** 185.0	**211.0** 211.5	**229.5** 230.3	**264.0** 281.3
S3 [min]	**40.4** 40.5	**-0.3%**	**11.8** 19.3	**19.0** 4.6	**32.0** 29.3	**39.5** 40.5	**48.0** 51.0	**65.5** 90.3
S4 [min]	**28.6** 28.6	**±0.0%**	**14.7** 25.4	**4.5** 0.0	**17.5** 2.3	**27.0** 26.5	**38.0** 46.0	**61.5** 81.9
REM [min]	**87.9** 88.4	**-0.6%**	**27.5** 22.2	**38.0** 48.6	**69.0** 73.8	**86.5** 88.0	**105.0** 98.8	**145.5** 131.7
SQI-score	**276.5** 276.7	**-0.1%**	**10.7** 25.3	**255.5** 214.1	**269.1** 265.3	**276.7** 280.4	**283.8** 293.0	**297.3** 320.4
Number of Sleep Stage Changes	**107.2** 107.0	**+0.2%**	**12.2** 26.1	**84.0** 57.3	**99.0** 87.5	**107.0** 105.0	**116.0** 123.5	**131.0** 158.9

Rel. Bias: relative bias; SD: standard deviation; SQI: Sleep Quality Index

Boxplots of the comparison of simulated and observed data are shown in Figure 6.7 for the six sleep stages.

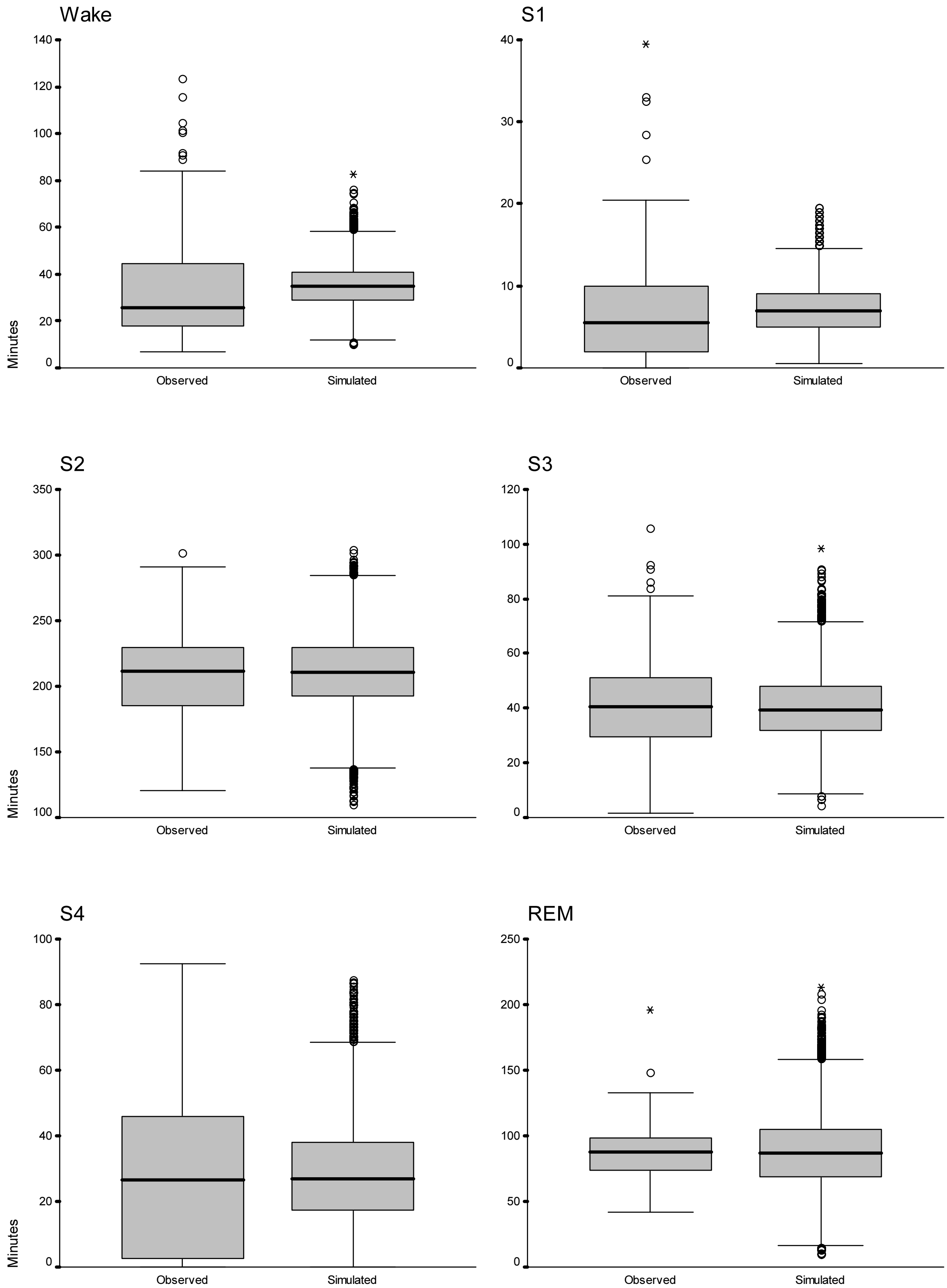


Figure 6.7: Boxplots of simulated (n=10,000) and observed (n=125) data. Whiskers indicate 1.5-fold inter-quartile range. Outliers (>1.5 and <3-fold inter-quartile range) are shown as circles. Extreme values (>3-fold inter-quartile range) are indicated with an asterisk.

In the following, the degree of agreement between simulated and observed probabilities will be described separately for each of the six sleep stages. They will be discussed in detail in Chapter 7.

Wake The simulated fit the observed probabilities well (Figure 9.13). The increasing probability of wakefulness at the end of the night is reproduced by the model. The distribution of observed data is somewhat skewed to the right (see Figure 6.7). This is due to a few subjects with longer periods of Wake during the baseline night which is not reproduced by the model: The median of simulated and observed values differs by 9.5 min, although mean values differ only by 0.1 min.

S1 Except for the first few minutes of the night, observed probabilities are reproduced well by the simulation (Figure 9.14). As in stage Wake, the distribution of observed data is somewhat skewed to the right (see Figure 6.7). The variance of the observed data is larger than that of the simulated data (SD 7.0 min versus 2.7 min). The median (7.0 min versus 5.5 min) and average values (7.2 min versus 7.1 min) match closely between simulated and observed data.

S2 Although not comparable to the fit of the weighted results of the regression model, the simulated probabilities display the trend of observed probabilities well (Figure 9.15). Ultradian rhythms are not reproduced by the simulation. The distribution (Figure 6.7), median and average values match closely between simulated and observed data.

S3 The simulated probabilities display the trend of observed probabilities well (Figure 9.16). Ultradian rhythms are not reproduced by the simulation. The distribution (Figure 6.7), median and average values match closely between simulated and observed data. The variance of simulated data is a little lower than that of observed data (SD 11.8 min versus 19.3 min).

S4 The simulated probabilities display the trend of observed probabilities well (Figure 9.17). Ultradian rhythms are not reproduced by the simulation.

The variance and inter-quartile range of the distribution of observed data are higher than those of the simulation (Figure 6.7). The amounts of sleep stage S4 depend strongly on the age of the subject. Older subjects sometimes show no SWS at all (a fact reproduced by the simulation). Nevertheless, median and average values match closely between simulated and observed data.

REM The simulated probabilities display the trend of observed probabilities well (Figure 9.18). Probabilities are overestimated in the first 50 min of the night by the simulation. This can be attributed to the estimation of transition probabilities to stage REM in the first minutes of the night that are close to but not equal to zero, as in reality. Ultradian rhythms are not reproduced by the simulation. The variance and inter-quartile range of the distribution of observed data are lower than those of the simulation (Figure 6.7). Nevertheless, median and average values match closely between simulated and observed data.

Although ultradian rhythms are not reproduced by the simulation trials, the simulated data fit the observed data reasonably well. Especially, the mean values of the main parameters of interest (time spent in different sleep stages, SQI-scores and number of sleep stage changes) are reproduced excellently.

6.3 Noise model results

6.3.1 Estimation of transition probabilities: Differences between noise and baseline conditions

First, the question had to be answered for how many epochs transition probabilities are influenced by an ANE. Hence, for each of the ten epochs following the start of an ANE, a first-order autoregressive multinomial logistic regression model was built with a noise indicator variable (1 or 0) as the only additional explanatory variable (elapsed sleep time was per definition identical between both conditions). In this case, the regression output of the noise indicator variable indicates direction, magnitude and statistical significance of the difference in transition probabilities between noise-free epochs and epochs with ANEs (see Chapter 5.3.2.2). All ANEs were shorter than 60 s, and therefore always ended before the beginning of noise epoch #4. Regression coefficients (betas) of noise indicator variables and their statistical significance are shown in Figure 6.8 to Figure 6.12 for transitions to stages Wake, S1, S3, S4 and REM (S2 served as reference category).

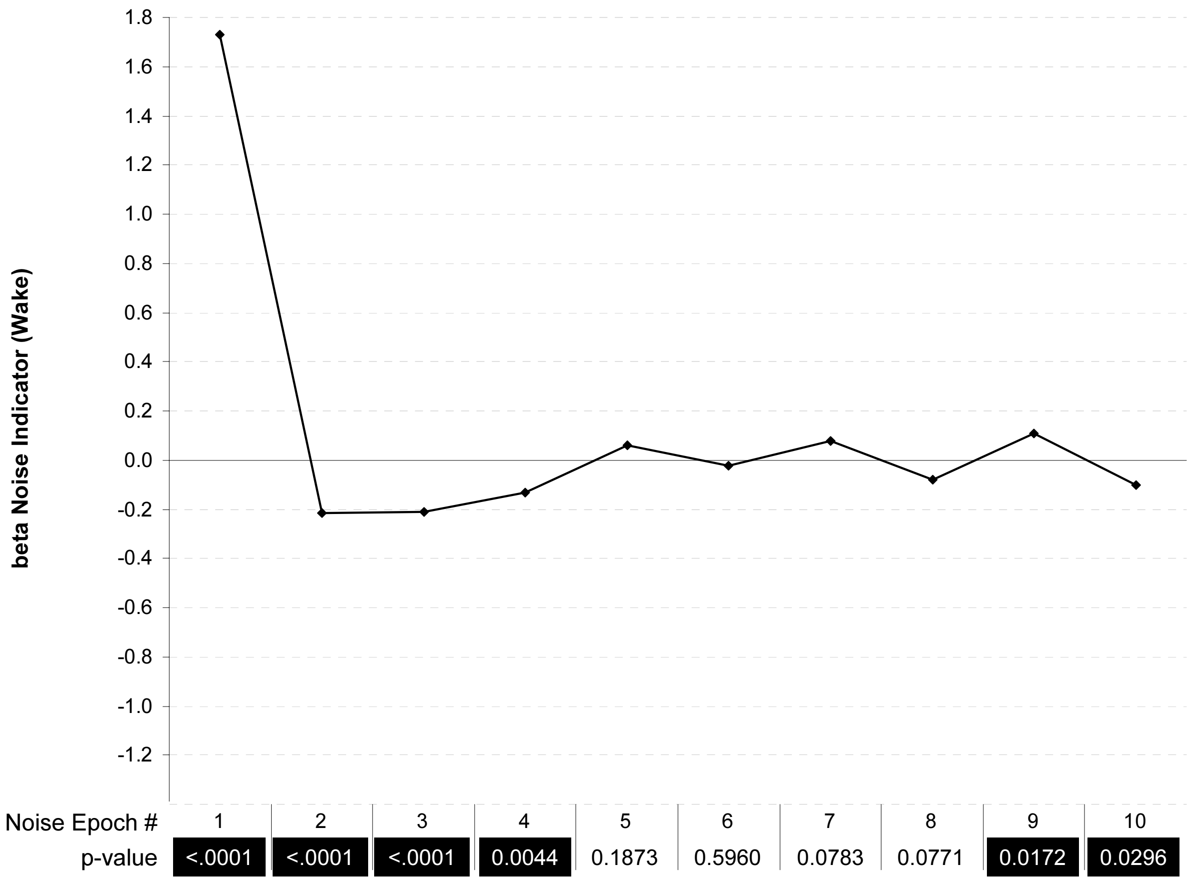


Figure 6.8: Size and statistical significance of noise indicator regression coefficents (beta) for transitions to stage Wake in noise epochs #1 to #10. p-values <0.05 are highlighted.

For stage Wake, the regression coefficient of the noise indicator variable was markedly and significantly increased for transitions to epoch #1, i.e. here transitions to Wake were more likely in noise compared to baseline conditions. Betas were significantly decreased for transitions to epochs #2, #3 and #4, but the size of the decrease was much smaller than the size of the increase for transitions to epoch #1. For transitions to epochs #5, #6, #7 and #8, there was neither a significant nor a relevant difference in betas between noise and baseline conditions. Betas for transitions to epochs #9 and #10 differed significantly between noise and baseline conditions: There was a small increase for transitions to epoch #9 and a small decrease for transitions to epoch #10.

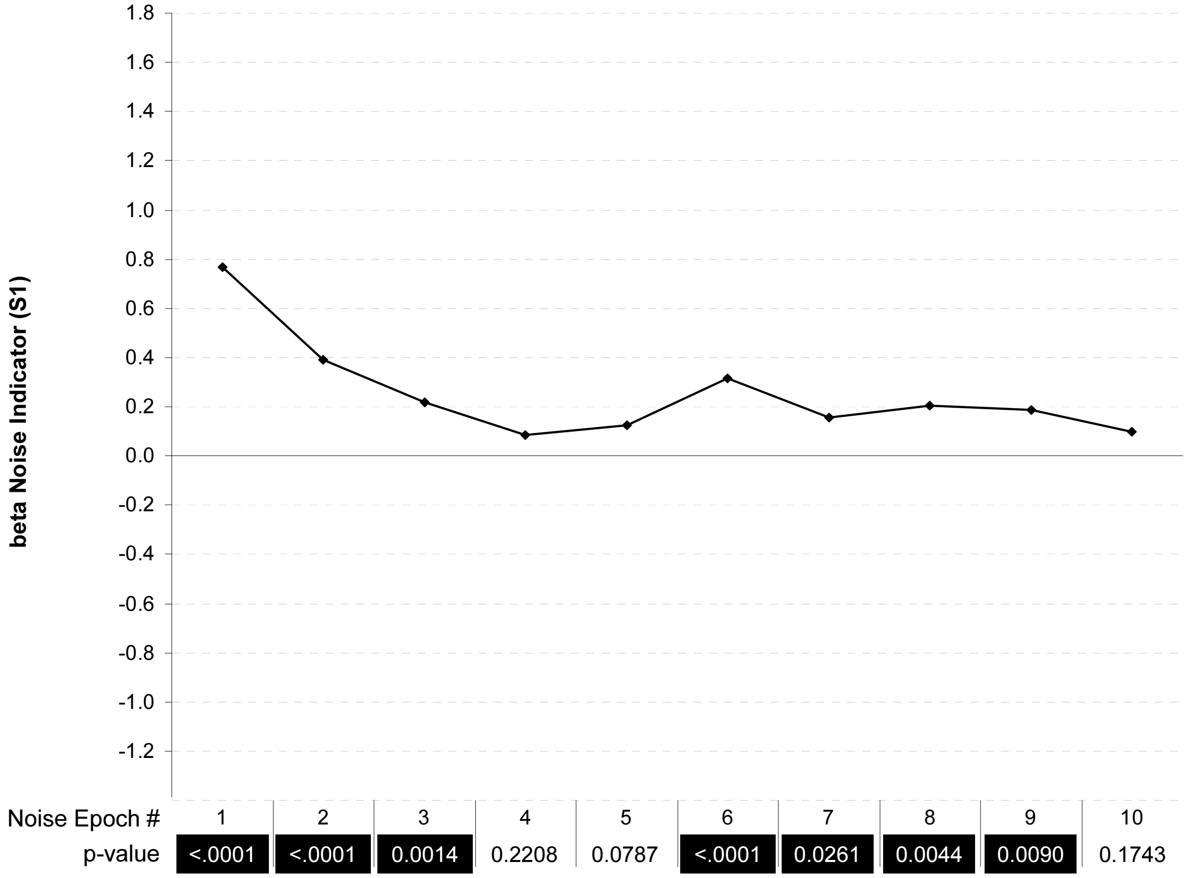
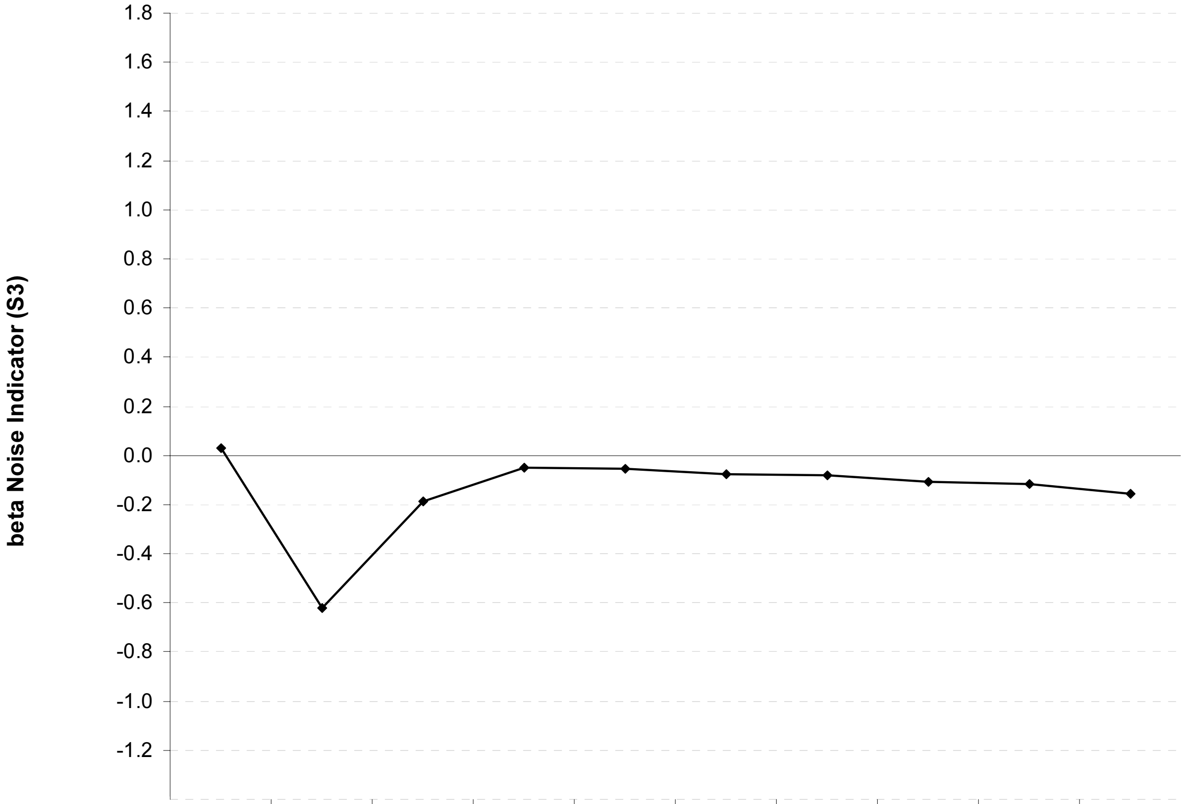


Figure 6.9: Size and statistical significance of noise indicator regression coefficents (beta) for transitions to S1 in noise epochs #1 to #10. p-values <0.05 are highlighted.

For stage S1, the regression coefficients of the noise indicator variable were increased for transitions to all epochs, i.e. transitions to S1 were more likely in noise compared to baseline conditions for all epochs. The increase was markedly for transitions to epochs #1, #2 and #6, and significantly for transitions to all epochs except for epochs #4, #5 and #10.



Noise Epoch #	1	2	3	4	5	6	7	8	9	10
p-value	0.5813	<.0001	0.0004	0.3419	0.2923	0.1289	0.1295	0.0360	0.0271	0.0026

Figure 6.10: Size and statistical significance of noise indicator regression coefficents (beta) for transitions to S3 in noise epochs #1 to #10. p-values <0.05 are highlighted.

For stage S3, the regression coefficients of the noise indicator variable were decreased for transitions to all epochs except for noise epoch #1, i.e. in 9 out of 10 cases transitions to S3 were less likely in noise compared to baseline conditions. The decrease was markedly for transitions to epochs #2 and #3 and less pronounced for transitions to epochs #4 to #10. Transition probabilities differed significantly between noise and baseline conditions for transitions to epochs #2, #3, #8, #9 and #10.

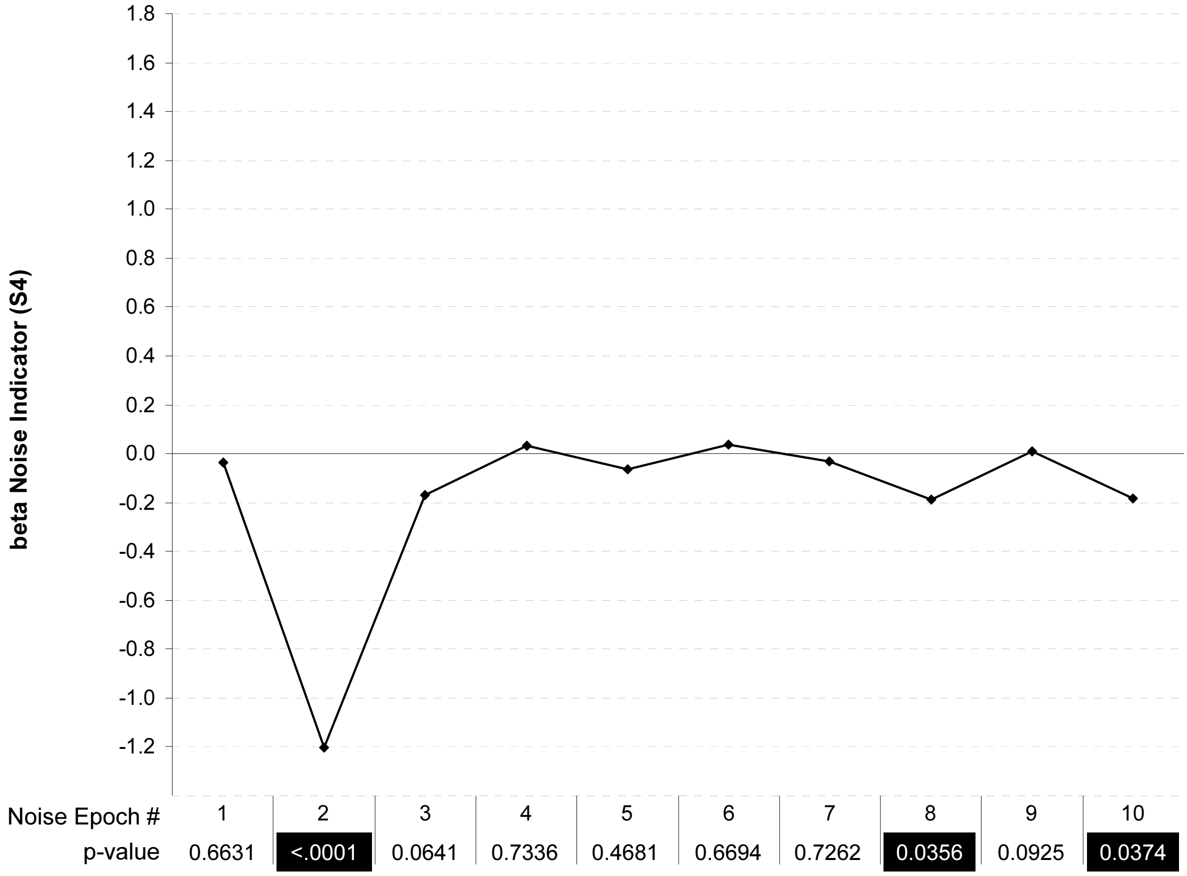


Figure 6.11: Size and statistical significance of noise indicator regression coefficents (beta) for transitions to S4 in noise epochs #1 to #10. p-values <0.05 are highlighted.

For stage S4, the regression coefficients of the noise indicator variable were decreased for transitions to all epochs except for epochs #4, #6 and #9, i.e. in 7 out of 10 cases transitions to S4 were less likely in noise compared to baseline conditions. The decrease was markedly for transitions to epoch #2 only, and less pronounced for transitions to epochs #3, #8 and #10. Betas for transitions to epochs #4 to #7 and to epoch #9 were practically zero. Transition probabilities differed significantly between noise and baseline conditions for transitions to epochs #2, #8, and #10.

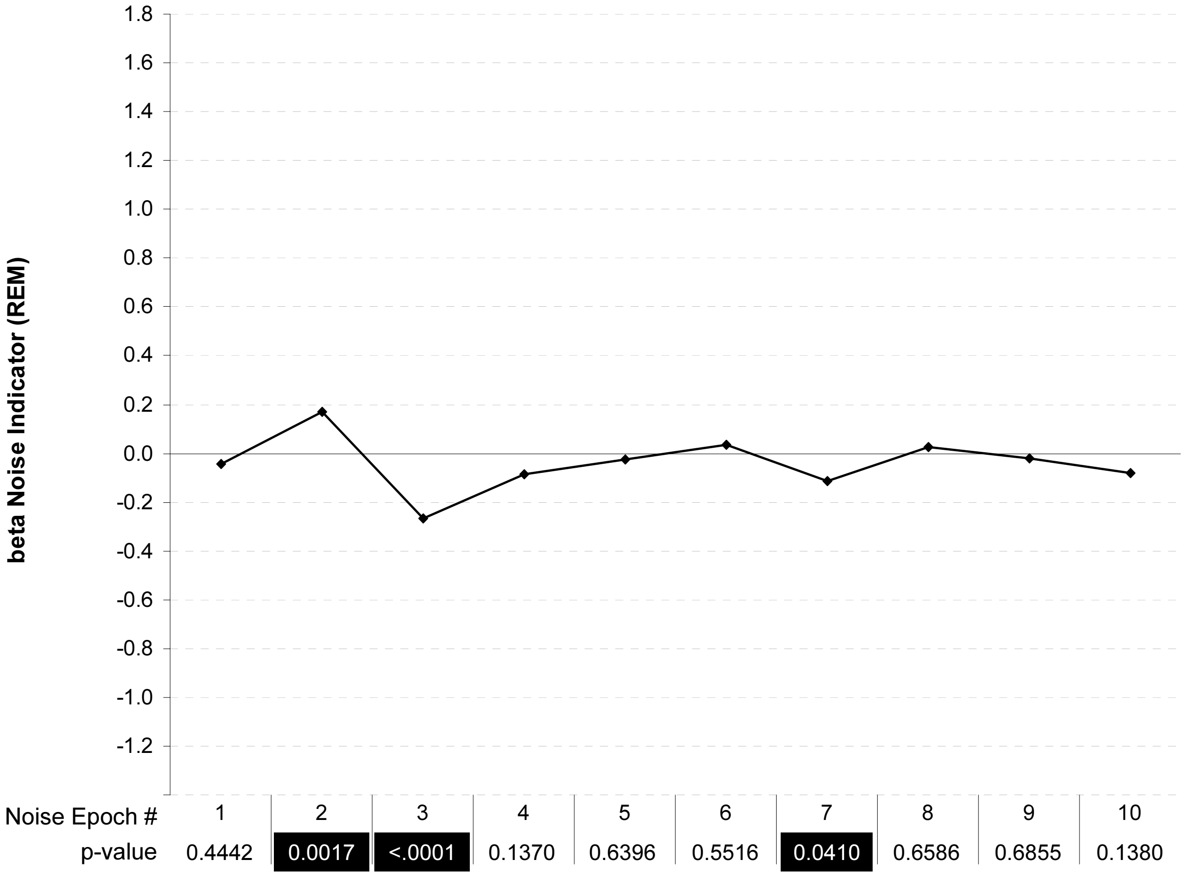


Figure 6.12: Size and statistical significance of noise indicator regression coefficents (beta) for transitions to stage REM in noise epochs #1 to #10. p-values <0.05 are highlighted.

For stage REM, the regression coefficient of the noise indicator variable was markedly and significantly increased for transitions to epoch #2 and was markedly and significantly decreased for transitions to epoch #3. Betas hardly differed between noise and baseline conditions for transitions to epochs #1, #4 to #6 and #8 to #10. There was a small but significant decrease for transitions to epoch #7.

In summary, regression coefficients for the noise indicator variable differed markedly in two cases (Wake, S1) for transitions to epoch #1, in four cases (S1, S3, S4 and REM) for transitions to epoch #2 and again in two cases (S3 and REM) for transitions to epoch #3. From transitions to epoch #4 on, pronounced decreases or increases in regression coefficients for the noise indicator variable were no longer seen, except for transitions to S1 in epoch #6.

The regression coefficients differed statistically significantly from zero in two cases (Wake, S1) for transitions to epoch #1, in all five cases for transitions to epoch #2 and in four cases (Wake, S1, S3 and REM) for transitions to epoch #3. For transitions to epoch #4, betas differed significantly from zero only for stage Wake, and none of the betas differed significantly from zero for transitions to epoch #5. From transitions to epoch #6 on, some of the regression coefficients differed statistically significantly from zero again.

Based on these findings, it was decided that transition probabilities in noise and baseline conditions differed markedly and significantly between noise and baseline conditions for epochs #1, #2 and #3 only. This decision and other aspects of the results shown in the previous figures will be discussed in detail in Chapter 7.4.2.

6.3.2 Estimation of transition probabilities: Results of the regression models

For each noise epoch, a separate first-order multinomial logistic regression model with elapsed sleep time as the only additional explanatory variable was built. The regression results of the three noise epoch models are shown in Appendix B (Chapters 10.2, 10.3 and 10.4).

In Table 6.4, average whole night transition probabilities calculated by the model for noise-free baseline nights and by the models for noise epochs #1, #2 and #3 are shown. Transition probabilities to lighter sleep stages should increase under the influence of noise, while transition probabilities to deeper NREM sleep stages or REM should simultaneously decrease. This was especially seen for noise epoch #1. Sometimes, the effect persisted until noise epoch #2 (e.g. p(0|2)) or even until noise epoch #3 (e.g. p(0|3)). In some cases, after effects of an initial reaction to noise can be seen in later noise epochs: After an initial decrease in p(2|1) in noise epoch 1, p(2|1) increased in both noise epoch 2 and noise epoch 3 compared to results of the model for noise-free baseline nights. For a detailed discussion of these results see Chapter 7.4.2.

Table 6.4: Mean transition probabilities calculated by the model for noise-free baseline nights and by the models for noise epochs #1, #2 and #3 (Wake = 0, REM = 5). Arrows indicate direction of change compared to noise-free nights. p(y|x) = transition from stage x to stage y.

	Noise-free	**Noise Epoch #1**	**Noise Epoch #2**	**Noise Epoch #3**
p(0|0)	0.675	0.819↑	0.529↓	0.615↓
p(0|1)	0.124	0.425↑	0.053↓	0.100↓
p(0|2)	0.029	0.192↑	0.040↑	0.027↓
p(0|3)	0.023	0.123↑	0.039↑	0.027↑
p(0|4)	0.031	0.103↑	0.034↑	0.022↓
p(0|5)	0.038	0.135↑	0.033↓	0.042↑
p(1|0)	0.094	0.053↓	0.152↑	0.153↑
p(1|1)	0.354	0.261↓	0.323↓	0.299↓
p(1|2)	0.005	0.015↑	0.008↑	0.006↑
p(1|3)	0.000	0.002↑	0.002↑	0.001↑
p(1|4)	0.000	0.001↑	0.001↑	0.000
p(1|5)	0.005	0.016↑	0.006↑	0.007↑
p(2|0)	0.167	0.091↓	0.250↑	0.198↑
p(2|1)	0.474	0.289↓	0.500↑	0.573↑
p(2|2)	0.915	0.743↓	0.898↓	0.914↓
p(2|3)	0.170	0.163↓	0.340↑	0.216↑
p(2|4)	0.015	0.017↑	0.047↑	0.012↓
p(2|5)	0.028	0.039↑	0.036↑	0.040↑
p(3|0)	0.003	0.001↓	0.003	0.002↓
p(3|1)	0.001	0.002↑	0.000↓	0.001
p(3|2)	0.034	0.040↑	0.029↓	0.032↓
p(3|3)	0.732	0.638↓	0.581↓	0.688↓
p(3|4)	0.132	0.119↓	0.157↑	0.102↓
p(3|5)	0.000	0.000	0.000	0.000
p(4|0)	0.000	0.000	0.000	0.001↑
p(4|1)	0.000	0.000	0.000	0.000
p(4|2)	0.001	0.000↓	0.000↓	0.000↓
p(4|3)	0.075	0.073↓	0.038↓	0.065↓
p(4|4)	0.822	0.760↓	0.761↓	0.863↑
p(4|5)	0.000	0.000	0.000	0.000
p(5|0)	0.061	0.036↓	0.066↑	0.031↓
p(5|1)	0.048	0.024↓	0.124↑	0.027↓
p(5|2)	0.016	0.010↓	0.026↑	0.021↑
p(5|3)	0.001	0.000↓	0.000↓	0.003↑
p(5|4)	0.000	0.000	0.000	0.000
p(5|5)	0.929	0.809↓	0.925↓	0.911↓

6.3.3 Estimation of transition probabilities: Validation of the regression models

A graphical analysis of goodness-of-fit of the noise models is shown in Chapter 9.3. The 410.5 min night was divided in six periods of 58.5 min and one period of 59.5 min. Observed and predicted probabilities were averaged over these intervals and compared with each other (see Chapter 7.4.2).

In the following, goodness-of-fit results will be described separately for each of the six sleep stages. They will be discussed in detail in Chapter 7.

Wake The probability of being in stage Wake is higher for all noise epochs and all sleep episodes compared to the noise-free baseline night, i.e. transition probabilities to stage Wake increase under the influence of noise and in the course of the night (Figure 9.19). The probability of being in stage Wake decreases from noise epoch #1 (N1) to noise epoch #3 (N3), and almost reaches baseline night levels in N3. The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N1 in sleep episode 1 (-18.6%) and in N2 in sleep episode 1 (+16.4%).

S1 The probability of being in stage S1 tends to be higher for all noise epochs and all sleep episodes compared to the noise-free baseline night, i.e. transition probabilities to stage S1 increase under the influence of noise (Figure 9.20). The probability of being in stage S1 increases from noise epoch #1 to noise epoch #3: After a noise induced awakening subjects often return to sleep via stage S1. The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N2 in sleep episode 1 (-19.3%) and in N2 in sleep episode 4 (+27.6%).

S2 The probability of being in stage S2 tends to be lower for all noise epochs and all sleep episodes compared to the noise-free baseline night (Figure 9.21), which may be explained by increased transition probabilities from stage S2 to stage Wake (see above). The probability of being in stage S2 increases from noise epoch #1 to noise epoch #3: After a noise induced awakening subjects often quickly return to sleep stage S2. The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N1 in sleep episode 2 (-6.3%) and in N1 in sleep episode 1 (+9.5%).

S3 The probability of being in stage S3 is lower for noise epochs #2 and #3, while it is almost unchanged for noise epoch #1 (Figure 9.22). The latter may be interpreted as a lagged response to noise. The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N1 in sleep episode 3 (-6.1%) and in N3 in sleep episode 2 (+6.4%).

S4 The probability of being in stage S4 tends to be lower for all noise epochs and all sleep episodes compared to the noise-free baseline night (Figure 9.23), which may be explained by increased transition probabilities from stage S4 to lighter sleep stages and to stage Wake. The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N1 in sleep episode 2 (-4.0%) and in N1 in sleep episode 1 (+3.3%).

REM The probability of being in REM sleep is lower for all noise epochs and all sleep episodes compared to the noise-free baseline night (Figure 9.24). The probabilities predicted by the models fit the observed data well. The most extreme differences can be found in N2 in sleep episode 2 (-5.8%) and in N3 in sleep episode 1 (+30.7%).

6.4 Results of comparing three noise scenarios

The three noise scenarios were (for details, see Chapter 5.4.2.1):

- Scenario 1: Present Frankfurt Airport noise pattern. Take off times for 16 August 2005 (Tuesday) were extracted from the time schedule of the airport.
- Scenario 2: As scenario 1, but take offs between 23:00 and 05:00 cancelled.
- Scenario 3: As scenario 2, but with redistribution of take offs that formerly took place between 23:00 and 05:00 to periods 22:00 to 23:00 and 05:00 to 06:00, respectively.

Table 6.5 summarizes Monte Carlo simulation results for the noise-free night and the three noise scenario nights. For noise scenario nights, separate simulations were run and weighted depending on time of falling asleep (for details see Chapter 5.4.2.2).

Table 6.5: Comparison of weighted averages of each of 10,000 first-order simulation trials for the four scenarios. 2.5 and 97.5 percentiles are given in parenthesis.

	No Noise	**Scenario 1** Present	**Scenario 2** NoNoise 23-5h	**Scenario 3** Redistribution
Wake [min]	**35.5** (19.5, 54)	**43.1** (26.2, 62.8)	**40.0** (23.1, 59.5)	**41.7** (24.9, 61.1)
S1 [min]	**7.2** (2.5, 13)	**9.2** (3.8, 15.7)	**8.3** (3.3, 14.7)	**8.7** (3.7, 15.2)
S2 [min]	**210.9** (155.5, 264)	**212.8** (161.8, 262.2)	**211.2** (158.9, 261.8)	**210.8** (158.7, 261.5)
S3 [min]	**40.4** (19, 65.5)	**37.2** (17.7, 60.6)	**38.9** (18.4, 63.4)	**38.3** (18, 62.3)
S4 [min]	**28.6** (5.5, 61.5)	**23.5** (2.2, 53)	**26.3** (5.1, 57.9)	**25.7** (2.9, 57)
REM [min]	**87.9** (38, 145.5)	**84.7** (38, 139.1)	**85.8** (37.9, 142.2)	**85.3** (37.6, 141.1)
SQI-Score	**276.5** (255.5, 297.3)	**267.2** (246.1, 288.2)	**271.3** (250.1, 292.5)	**269.4** (248.3, 290.8)
SQI [% of No Noise]	**100%**	**96.6%**	**98.1%**	**97.4%**
Number of Sleep Stage Changes	**107.2** (84, 131)	**121.3** (96.4, 146.9)	**115.7** (91.3, 140.8)	**118.3** (93.5, 143.8)

In the following, the simulation results will be described separately for each of the six sleep stages.

Wake Time spent awake increased in all noise scenarios compared to the noise-free night. Time spent awake increased in noise scenario 1 by 7.6 min

(+21.4%)[6]. In scenario 2, time spent awake was lower compared to scenario 1, but it still increased by 4.5 min (+12.7%) compared to a noise-free night. Redistribution of noise events (scenario 3) diminished this relative decrease in time spent awake, which increased by 6.2 min (+17.5%) compared to the noise-free night.

S1 Time spent in stage S1 increased in all noise scenarios compared to the noise-free night. It increased in noise scenario 1 by 2 min (+27.8%). In scenario 2, time spent in stage S1 was lower compared to scenario 1, but it still increased by 1.1 min (+15.3%) compared to a noise-free night. Redistribution of noise events (scenario 3) diminished this relative decrease in time spent in stage S1, which increased by 1.5 min (+20.8%) compared to the noise-free night.

S2 Time spent in stage S2 was only little changed, if all. Relative changes were smaller than 1% (212.8 min compared to 210.9 min).

S3 Time spent in stage S3 decreased in all noise scenarios compared to the noise-free night. It decreased in noise scenario 1 by 3.2 min (-7.9%). In scenario 2, time spent in stage S3 was higher compared to scenario 1, but it still decreased by 1.5 min (-3.7%) compared to a noise-free night. Redistribution of noise events (scenario 3) diminished this relative increase in time spent in stage S3, which decreased by 2.1 min (-5.2%) compared to the noise-free night.

S4 Results for stage S4 are very similar to those of stage S3. As expected, time spent in stage S4 decreased in all noise scenarios compared to the

[6] At this point, the following should be noticed and kept in mind: It is likely that the effects of noise on sleep structure are rather overestimated than underestimated by the models and the simulations. This is due to the fact that data are based on laboratory studies. The sensitivity of subjects to aircraft noise induced changes in sleep structure is known to be greater in the laboratory compared to the field, i.e. the familiar environment (see Chapter 7).

noise-free night. It decreased in noise scenario 1 by 5.1 min (-17.8%). In scenario 2, time spent in stage S4 was higher compared to scenario 1, but it still decreased by 2.3 min (-8%) compared to a noise-free night. Redistribution of noise events (scenario 3) diminished this relative increase in time spent in stage S4, which decreased by 2.9 min (-11.1%) compared to the noise-free night.

REM Time spent in REM sleep decreased in all noise scenarios compared to the noise-free night. It decreased in noise scenario 1 by 3.2 min (-3.6%). In scenario 2, time spent in REM sleep was higher compared to scenario 1, but it still decreased by 2.1 min (-2.4%) compared to a noise-free night. Redistribution of noise events (scenario 3) diminished this relative increase in time spent in REM sleep, which decreased by 2.6 min (-3%) compared to the noise-free night.

SQI Sleep quality, as measured by the sleep quality index SQI-score (for a description see Chapter 5.5.4), diminished in all three noise scenarios compared to noise-free nights, with a maximum decrease of 4.4% in scenario 1. The cancellation of take offs between 23:00 and 05:00 lead to a relative increase of SQI compared to scenario 1, but it still decreased by 1.9% compared to noise-free nights. Redistribution of take offs in scenario 3 diminished this relative increase (97.4% compared to noise-free nights).

Sleep Stage Changes The number of sleep stage changes also increased in all noise scenarios compared to noise-free nights (maximum: +13.2% in scenario 1), indicating a higher degree of sleep fragmentation under the influence of noise. This goes well along with the findings for stage Wake and especially for stage S1.

A graphical analysis of weighted simulation results is shown in Figure 6.13 for the six sleep stages.

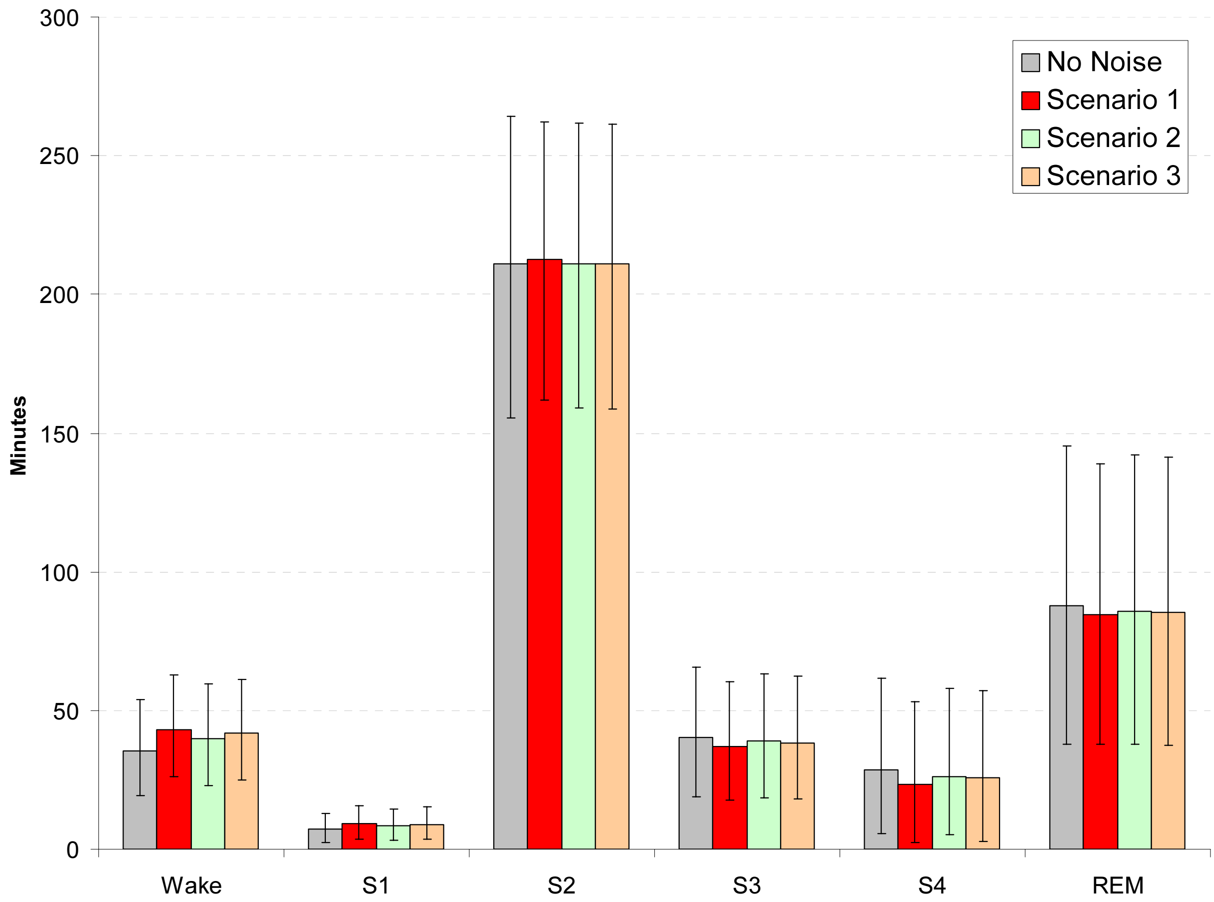


Figure 6.13: Comparison of weighted simulation results of the four scenarios. Whiskers indicate 2.5 and 97.5 percentiles for each of 10,000 first-order simulation trials. Scenario 1: Present; Scenario 2: No Noise 23:00-05:00; Scenario 3: Redistribution

Although data of a laboratory study were used to estimate transition probabilities for the noise scenarios, the differences between the noise conditions and the noise-free nights were not profound. Additionally, there is considerable overlap of the 2.5 and 97.5 percentile range of the simulation results. But nevertheless, aircraft noise lead to obvious changes in the average duration of the different sleep stages, that go well along with the hypotheses (i.e. increase in amounts of Wake and S1, decrease in amounts of SWS and REM sleep).

The results presented in Table 6.5 and Figure 6.13 are weighted averages based on separate simulations for different times of falling asleep. Figure 6.14 shows a detailed analyses of changes in the amounts of the different sleep stages for the three noise scenarios depending on the time of falling asleep. For the interpretation of these figures, the typical pattern of the noise scenarios has to be borne in mind. A period with a low traffic density between 23:00 and 07:00 is preceded and followed by periods with relatively high traffic densities. This is typical for all three scenarios (see Figure 5.13).

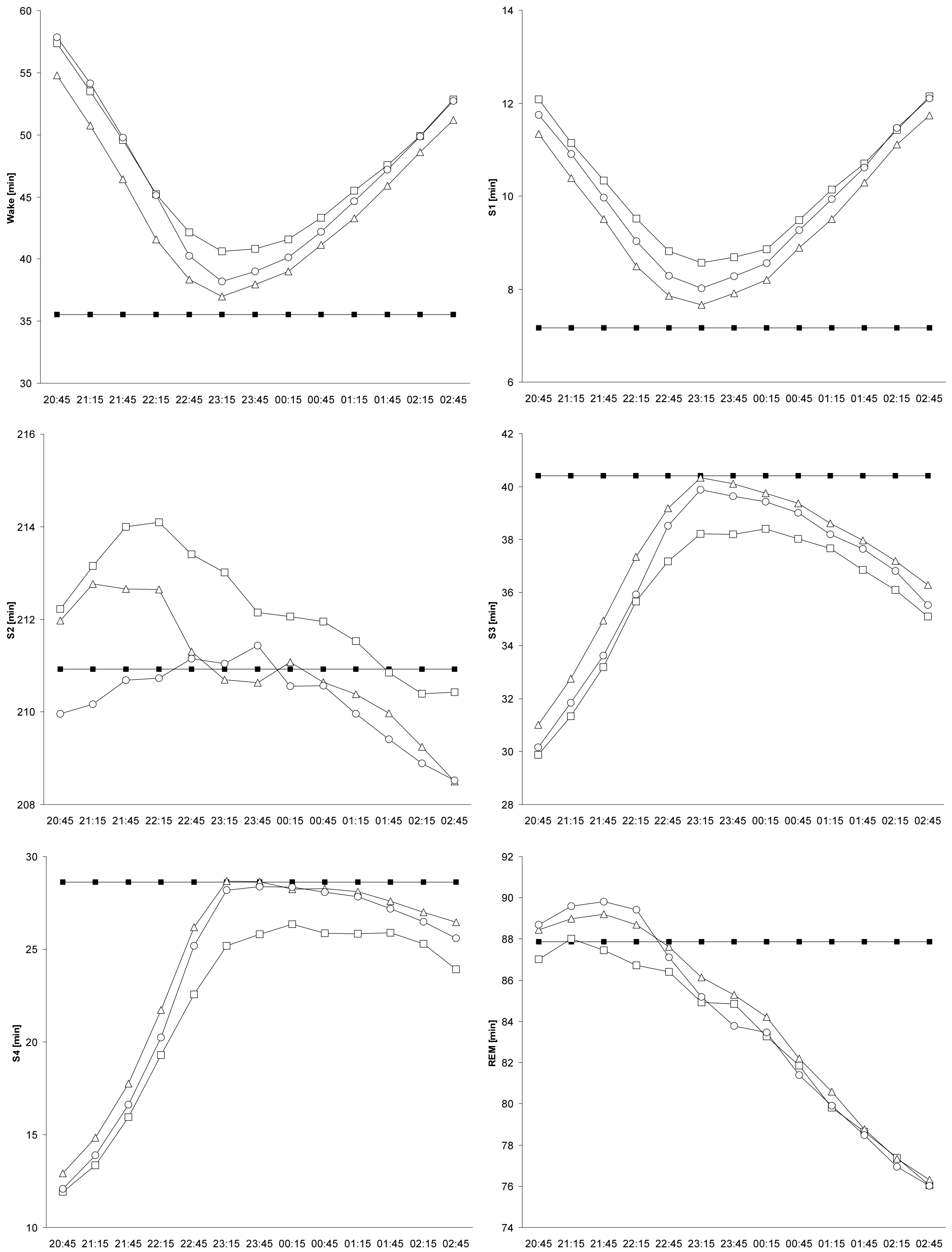


Figure 6.14: Simulation results of the four scenarios no noise (black squares, independent of time of falling asleep), scenario 1 (white squares), scenario 2 (triangles) and scenario 3 (circles) depending on time of falling asleep (ordinate) for the six sleep stages. Graduation of abcissa differs between plots.

Minimal changes in sleep structure compared to noise-free nights were observed in subjects falling asleep at 23:15 (exception: stage REM). Earlier or later bed times were accompanied by increased amounts of stage Wake and S1. S3 and S4 especially decreased if subjects fell asleep earlier than 23:15. REM sleep steadily decreased with times of falling asleep later than 22:45. For times of falling asleep before 22:45, REM sleep was slightly increased in noise scenarios 2 and 3, i.e. in the scenarios with noise-free periods from 23:00 to 05:00, compared to noise-free nights.

Compared to the average duration of 210.9 min in noise-free nights, amounts of S2 were only moderately altered by aircraft noise, if all.

Overall, Figure 6.14 reveals that the effects of the different times of falling asleep were much stronger than the differences between the three noise scenarios. E.g., the largest difference in amounts of stage Wake between the noise scenarios was 3.8 min between scenario 1 and 2 when falling asleep at 22:45. Compared to that, the largest difference in amounts of stage Wake within one of the noise scenarios was 19.7 min (+51.5%) for scenario 3 when falling asleep at 20:45 compared to 23:15.

7 Discussion

The major findings of this thesis are summarized and interpreted in Chapter 7.1. The data basis, derived from the DLR laboratory study STRAIN, is discussed in Chapter 7.2. In the following chapters, methods and results are discussed separately for the baseline model (Chapter 7.3), the noise models (Chapter 7.4) and the comparison of three noise scenarios with each other and with a noise-free night (Chapter 7.5). Limitations of the analyses are summarized in Chapter 7.6 .This discussion always includes a detailed analysis of the limitations of methods and results.

7.1 Summary and interpretation of major findings

In this thesis, Markov state transition models were used for the simulation of noise-free baseline nights. Autoregressive multinomial logistic regression was used to estimate the probability of transitions between different sleep stages. The calculations were based on polysomnograms of undisturbed baseline nights of 125 subjects who participated in an experimental laboratory study on the influence of aircraft noise on sleep. Graphical analyses of goodness-of-fit showed that the final regression model, with present sleep stage (at T0) and elapsed sleep time as the only explanatory variables, fit the original data well. First-order Monte Carlo simulations were also able to correctly display the trend of the changing probabilities of the different sleep stages during the night, although ultradian rhythms were not reproduced. Differences in average values of the main outcome variables (time spent in different sleep stages, number of sleep stage changes and SQI-score) between observed data and data predicted by the model were negligible. Additionally, typical features of human hypnograms were reproduced by Markov traces of Monte Carlo simulation trials.

The influence of aircraft noise on transition probabilities was assessed in three noise models, each for one of three epochs following the onset of the

ANE. The models were also based on autoregressive multinomial logistic regression and – as the baseline model - contained present sleep stage and elapsed sleep time as the only explanatory variables. The estimations were based on 26,135 ANEs administered during the DLR laboratory study. Inspection of transition probabilities and a graphical analysis of goodness-of-fit showed that the models fit the original data well and produced plausible results.

The influence of aircraft noise on sleep structure was estimated for three noise scenarios that were compared with each other and with a noise-free night: Scenario 1 was based on the official time schedule for Frankfurt Airport for 16 August 2005 for take offs only. Scenario 2 was identical to scenario 1, but all take offs between 23:00 and 05:00 were cancelled. Scenario 3 was identical to scenario 2, but take offs that took place between 23:00 and 05:00 in scenario 1 were rescheduled to the time periods 22:00 to 23:00 and 05:00 to 06:00. The primary question of this comparison was, whether a noise-free period remained a benefit for the affected population if take offs that formerly took place during the night were rescheduled to periods before and after the noise-free period.

The results of the different noise scenarios corroborated the proposed hypotheses: Under the influence of noise, amounts of Wake and S1 increased, amounts of S3, S4 and REM decreased while amounts of S2 remained almost unchanged. Cancellation of take offs between 23:00 and 05:00 (scenario 2) resulted in a benefit for the simulated population in terms of sleep structure, but the absolute changes in the amounts of the different sleep stages were rather small. This may be attributed to relatively low traffic densities at Frankfurt Airport between 23:00 and 05:00 (on average four take offs per hour) compared to other airports (e.g. Cologne-Bonn during the week).

Even in the case of rescheduling all take offs that took place between 23:00 and 05:00 to the time periods 22:00 to 23:00 and 05:00 to 06:00, a worst-case scenario, there was still a benefit of the noise-free period for affected residents. With absolute differences between scenario 3 and scenario 1 of Wake -1.4 min, S1 -0.5 min, S2 -2.0 min, S3 +1.1 min, S4 +2.2 min and REM +0.6 min (see Table 6.5 and Figure 6.13), the resulting net benefit was rather small.

Different times of falling asleep were accounted for in the analyses. Separate simulations were performed and a weighted average was calculated, with the relative frequencies of the population falling asleep at a peculiar point in time as weights. Overall, the effects of the different times of falling asleep were much stronger than the differences between the three noise scenarios. For example, the largest difference in amounts of stage Wake between the noise scenarios was 3.8 min between scenario 1 and 2 when falling asleep at 22:45. Compared to that, the largest difference in amounts of stage Wake within one of the noise scenarios was 19.7 min (+51.5%) for scenario 3 when falling asleep at 20:45 compared to 23:15. This emphasizes the importance of traffic density during shoulder, i.e. late evening and early morning, hours for sleep structure.

Therefore, if the models and the results of the simulations are valid, affected residents will benefit from the introduction of a noise-free period from 23:00 to 05:00 even if the traffic that took place between 23:00 and 05:00 will be completely rescheduled to the periods before 23:00 and after 05:00 (a worst case scenario). At the same time, the observed net benefit will probably be very small, which may be attributed to relatively low traffic densities at Frankfurt Airport between 23:00 and 05:00 at present. The results also indicate that periods with high air traffic densities, i.e. before 23:00 and after 07:00, may disturb sleep much stronger than the traffic presently taking place during the night. Especially the sleep of subjects who choose to go to bed or have to go to bed (e.g. children, shift workers) earlier than 22:30 or later than 01:00 may substantially be disturbed by noise emitted from air traffic during shoulder hours, in particular as traffic densities will increase during shoulder hours after the extension of Frankfurt Airport.

The introduction of a noise-free period between 23:00 and 05:00 will most probably lead to a shrinkage of noise protection areas around the airport, because the night is officially defined as the period from 22:00 to 06:00 and only some of the flights that formerly took place between 23:00 and 05:00 will be rescheduled to the period between 22:00 and 23:00 and between 05:00 and 06:00. Paradoxically, for people who live just outside of this noise protection zone and who usually fall asleep earlier than 22:30 or later than 01:00, the introduction of a noise-free period between 23:00 and 05:00 will most probably lead to more severe changes in sleep structure compared to the status quo with no prohibition of air traffic during the night.

Markov state transition models have previously been used by other groups in order to model human sleep and the impact of drugs or other external influences on sleep, e.g. [33, 34, 52, 53]. To the knowledge of the author, this is the first time that the influence of aircraft noise on sleep structure was modeled with a Markov state transition model. Additionally, it is the first time that the effects of unprecedented situations on sleep structure were investigated and reported.

A variety of mathematical procedures have been applied for the calculation of transition probabilities by other authors in the past. These ranged from simply calculating average transition probabilities over constant periods of the night [53] to complex and sophisticated mixed effects models [33]. In this thesis, autoregressive multinomial logistic regression [20] was used for the calculation of transition probabilities for the first time. All data were integrated into a single model, as opposed to some approaches of the past where several models were built for this purpose (e.g. [33]). Additionally, it was not necessary to divide the sleep period time into segments with constant transition probabilities (e.g. [34, 53]). In autoregressive models, realizations of past states are explicitly used to predict future states. In the context of modeling human sleep, this allowed for the implementation of past sleep stage realizations as covariates in the autoregressive regression models in order to predict transition probabilities to future sleep stages.

7.2 Discussion of empirical data

Empirical data, on which both the model for noise-free baseline nights and the noise models are based on, were generated in an experimental laboratory study at the DLR Institute of Aerospace Medicine between 1999 and 2003, with 128 subjects of both genders aged 19 to 65. Selection criteria were applied in order to assure the internal validity of the study [9]. Therefore, the results of the model should only be directly transferred to subjects eligible for study participation.

Reactions to aircraft noise observed in laboratory studies are usually much stronger than those observed in field studies conducted in the subjects' homes [6, 40]. Possible reasons are the unfamiliar environment of the laboratory and the presentation of noise events that may not typically occur at the subjects' homes. Additionally, the application of electrodes and sensors disturbs sleep per se. Hence, the ecologic validity of the model results is limited. In the future, models based on data of field studies should be built and the results should be compared to the results of the models presented in this thesis.

Nevertheless, it should be stressed that, as the data are derived from laboratory studies, the results presented in this thesis rather overestimate than underestimate the influence of aircraft noise on sleep structure. Likewise, the effects of the introduction of a noise-free period from 23:00 to 05:00 are also rather overestimated than underestimated by the models presented here.

7.3 Discussion of the baseline model

7.3.1 Methods

In order to prevent censored data and to simplify the comparison of empirical and predicted data, the model of the noise-free baseline nights was restricted to the first 410.5 min of sleep period time. Although this coincides well with the average sleep time of the adult German population [37, 39], shorter and longer sleep period times were not accounted for. As all subjects were synchronized to sleep onset, no information was available about the process of falling asleep. For noise-free nights, average times of falling asleep can be assumed, but this may not apply once aircraft noise is involved (see Chapter 7.5.1)

As was shown in Chapter 5.3, autoregressive logistic regression is a powerful tool for modeling data with repeated measurements in the same subjects: The underlying theory is easily understood and standard statistical software can be used. A potential drawback, especially of autoregressive models of higher order, is that there is no information on past events for the prediction of the very first transitions. Nevertheless, this posed no problem in this thesis: The simulation started with sleep onset and therefore information on sleep stage history prior to sleep onset was available. In contrast to other model types, where the correlation structure of dependent data is modeled indirectly, the conditional distribution of each response is expressed as an explicit function of the past response (the history) and covariates. This is a fairly strong assumption, and perhaps this is one of the reasons why autoregressive logistic regression has only sparsely been applied in the past [20].

7.3.2 Results

In the model building process, several models were tested for their validity in reproducing the original data of noise-free baseline nights. As the graphical analysis of goodness-of-fit showed, even the simple first-order autoregressive model with elapsed sleep time as the only additional variable fit the original data excellently. The fit and the precision of the prediction of empirical data did not improve relevantly for models containing additional explanatory variables. Therefore, it was decided to use the most parsimonious, yet biologically reasonable model, i.e. the simple first-order autoregressive model with elapsed sleep time as the only additional explanatory variable as the final model. One has to bear in mind that this simple model already contains seven explanatory variables in total (intercept, five indicator variables indicating present sleep stage (at T0), the continuous variable "elapsed sleep time"), and therefore generates several degrees of freedom.

Transition probabilities do not solely depend on the current sleep stage, but also on sleep stage history (see Chapter 6.1). The final regression model was a first-order autoregressive model, where transition probabilities are per definition calculated depending on the current sleep stage only. Nevertheless, serious implications for the validity of the transition probabilities calculated by the first-order regression model are unlikely, as transition probabilities depending on the current sleep stage only may also be interpreted as the weighted average of several different transition probabilities, that each depend on their individual history of past states. In this case, the weights are given by the relative frequencies of the individual histories. Put differently, the history of past states will always implicitly be accounted for even if no memory of past states is explicitly incorporated into the model.

Distributions of the differences between predicted and observed data were always symmetrical and unbiased, although the distributions of S1, S3 and

S4 deviated significantly from the normal distribution (see Figure 6.4 and Table 6.2).

The main parameters of interest, generated by averaging over 10,000 first-order Monte Carlo simulation trials, fit the original data excellently: Predictions deviated between -0.6% (REM) and +1.4% (S1) from the observed averages. Average SQI-scores and number of sleep stage changes were reproduced precisely. Because of the central limit theorem, the distribution of simulation results tended to be symmetrical, whereas some distributions of the original data were skewed (Wake, S1 and S4, see Figure 6.7).

Typical features of human sleep, like the dominance of SWS in the first half of the night or the ability to clearly distinguish between single NREM- and REM-episodes, were generally reproduced in Markov traces of individual trials, although in some cases features were observed that are usually not seen in healthy human sleep (e.g. sleep onset REM-episodes, see below).

Although the final regression model reproduced the original data well, there were some potential limitations that could not be eliminated even by more complex models:

(1) There was a systematic bias in the differences of predicted and observed probabilities of sleep stages S2, S3 and REM at the very beginning of the night (see Figure 9.5, Figure 9.7 and Figure 9.11). As REM sleep occurs for the first time only after a certain latency after sleep onset in the healthy sleeper, the overestimation of transition probabilities to stage REM in the beginning of the sleep period also lead to a systematic overestimation of REM sleep amounts in the first 50 minutes of the sleep period in the first-order simulation trials (see Figure 9.18) and to sleep onset REM episodes in Markov traces of individual trials (see Figure 6.6), that are usually only seen in narcoleptic patients or under extreme conditions (e.g. jet-lag or shift work). On the one hand, it would have been possible to divide the

night into different sections and to built different regression models for each of the sections. Actually, this method was already implemented by others [34, 53]. On the other hand, the attractive approach of one relatively simple model explaining many typical features of human sleep would have been discarded.

(2) The analysis of goodness-of-fit showed that the sleep stage distribution across the night predicted by the model reproduced the observed data well. Therefore, ultradian rhythms of the original data were also reproduced by the model, if the sleep stage distribution of the original data at T0 was given. This did not apply for the first-order simulation trials shown in Figures 8.13 to 8.18, where the simulation starts with sleep stage S2 and otherwise solely depends on transition probabilities calculated by the regression model. Here, the trend of the probability distribution of the different sleep stages was reproduced well across the night, but no ultradian rhythmicity was replicated. Several explanations are possible: The model used for the prediction of transition probabilities may simply be too basic for the replication of ultradian rhythms. On the other hand, even more complex models were not able to reproduce ultradian rhythmicity. One model was specially designed for the reproduction of ultradian rhythms (Cox model, see Chapter 6.2.1), but it also failed. Kemp and Kamphuisen [34] already mentioned in their 1986 paper that the NREM-REM-periodicity may be obscured by averaging over different subjects with varying individual periodicities. This may well be the case here, as the regression models were based on 125 different subjects. Superimposing many periodicities with varying period lengths should lead to a decrease in the amplitude of the overall periodicity, but the general trend across the night should be preserved unbiased. This is exactly what can be observed both in the empirical data at the end of the night and after averaging over several first-order trials that are

based on regression results of several subjects with different periodicities.

Nevertheless, both time trends across the night (see Figures 8.13 to 8.18) and whole night averages of the main parameters of interest (time spent in different sleep stages, SQI-score and number of sleep stage changes, see Table 6.2) were reproduced excellently by first-order Monte Carlo simulation trials of the final model. This was a prerequisite for using autoregressive multinomial logistic regression in conjunction with first-order simulation trials to predict noise induced changes in sleep structure.

7.4 Discussion of the noise models

7.4.1 Methods

In this first analysis, different maximum SPLs of ANEs were not differentiated. Rather, all noise events were treated as if they consisted of the same maximum SPL, and this SPL can be thought of as an "average" maximum SPL of all ANEs played back during the first 410.5 min of sleep period time (see Table 5.1). With maximum SPLs between 45 dB and 80 dB, some levels applied in the laboratory study were higher than those usually seen in the field [6]. Again, this rather leads to an overestimation of the effects of aircraft noise on sleep structure and of the effects of the introduction of a noise-free period (see Chapter 7.1). With a model containing the maximum SPL of an ANE as an explanatory variable, the effects of the maximum SPL of single ANEs on changes in sleep structure could be investigated more precisely, but this was not the primary aim of this first analysis.

In sleep medicine, the so-called sleep efficiency index (SEI) is usually used as an indicator of sleep quality. The SEI is calculated as the percentage of all sleep stages except Wake and based on sleep period time or time in bed. In the noise-free baseline nights, an average SEI of (410.5-35.4)/410.5 = 0.914 was observed. It would have been possible to compare SEIs of the noise scenarios with this value, but as sleep period times were the same under all conditions (410.5 min), this would have been identical to comparing the amounts of stage Wake between conditions. Therefore, by using the SEI as a measure of sleep quality, the concept "sleep quality" would have been reduced to the differentiation between stage Wake and the other five sleep stages.

As was explained in detail in Chapter 3.4, the effects of noise on sleep have often been investigated only in the context of awakenings in the past, but it is very desirable to predict not only noise induced awakenings, but also the consequences for the structure of sleep in total, and this was one of the main objectives of this thesis. Therefore, a new sleep quality index score was introduced, where the sleep stages were weighted for their assumed contribution to sleep recuperation. The weights were calculated from the awakening probabilities of the different sleep stages relative to that of stage S1, whose recuperative value was arbitrarily set to zero (see Chapter 5.5.4).

It should be stressed that the concept of the SQI presented here is new. It has neither been introduced to nor discussed by the scientific community so far. Yet, the method that generates the SQI is plausible. It manages to accumulate some of the aspects of sleep structure into a single value, where sleep stages that are thought to be very important for the restorative power of sleep are accordingly provided with higher weights. Of course, it is questionable whether it is reasonable to treat sleep stages, whose mechanisms and functions may differ on a fundamental level, on an equal footage by integrating them into one index.

However, taking into account the different vulnerabilities of the sleep stages for external stimuli, a weighted score for sleep quality as introduced by the SQI may be superior to the SEI.

7.4.2 Results

Transition probabilities differed markedly and statistically significantly between noise and noise-free conditions only for the first three epochs after the start of a noise event (see Chapter 6.3.1). Hence, three noise models were built, one for each of the three epochs.

Transition probabilities to lighter sleep stages should increase under the influence of noise, while transition probabilities to deeper NREM sleep stages or REM should simultaneously decrease. This was especially seen for noise epoch #1. Sometimes, the effect persisted until noise epoch #2 (e.g. $p(0|2)$) or even until noise epoch #3 (e.g. $p(0|3)$). In some cases, the initial effect was reversed in epoch #2 or epoch #3 or both: E.g., after an initial decrease in $p(2|1)$ in noise epoch #1, $p(2|1)$ increased in both epochs #2 and epoch #3 compared to results of the model for noise-free baseline nights (see Table 6.4). Two reasons may explain this phenomenon:

(1) The effects of transitions to lighter sleep stages that occur in the first noise epoch are often very short. Basner et al. showed that, after an awakening induced by ANEs with maximum SPLs up to 65 dB, about 50% of the subjects had fallen asleep again after just one epoch of stage Wake [8]. Therefore, transitions back to sleep in general and to SWS and REM sleep in particular may become more likely compared to noise-free conditions in epochs #2 and #3.

(2) The cohort of subjects is split into two populations by an ANE: Those who react (e.g. S2→Wake) and those who do not (e.g. S2→S2). Subjects in the first group are likely to return to sleep quickly (see above). In the example, transitions from Wake to the other sleep

stages may become more likely in the following epochs. Simultaneously, the sleep pressure of those who do not react is likely to be higher compared to those who do react. In the example, transitions to lighter sleep stages (S2→S1, Wake) may become less likely whereas transitions to stages S2, S3, S4 and REM may become more likely in the following epochs.

Both mechanisms, i.e. the initial effect of aircraft noise on transition probabilities and secondary consequences of this initial effect, should be captured by a noise model. Hence, three noise models were built, considering both primary and secondary consequences on sleep.

For S1 and S3, regression coefficients for the noise indicator variable remained positive (S1) or negative (S3) for all or most of the 10 epochs that were compared between noise and baseline conditions. This indicates that transition probabilities between noise and baseline conditions may differ not only while or shortly after, but also between ANEs, although the differences were small and mostly non-significant. This may be attributed to partial sleep deprivation, accumulated over the nine consecutive experimental nights of the laboratory study, as was explicitly shown by Basner and Samel in 2005 [9]. It emphasizes the fact that subtle and more complex interactions and countermeasures of the human body exist, and some of them may not have been replicated by the comparatively simple models chosen for the analyses [9].

Some of the regression coefficients for the noise indicator variable differ significantly from zero from epoch #6 or #7 on. These effects are most likely caused by data derived from nights in which 128 ANEs were played back. Here, the minimum interval between two ANEs was three minutes or six epochs, i.e. the next ANE started in epoch #7 of the analyses.

The graphical analysis of goodness-of-fit of the noise models, which is shown in Chapter 9.3, indicated that the sleep stage fractions observed under the influence of noise were reproduced well by the noise models for noise epochs #1, #2 and #3. Altogether, the observed fractions of the different sleep stages under the influence of noise all corroborated the hypotheses about the influence of noise on sleep structure (i.e. increases in Wake and S1, decreases in S3, S4 and REM).

7.5 Discussion of the comparison of three noise scenarios

7.5.1 Methods

The following paragraphs discuss whether the three noise scenarios assumed for the simulations were realistic according to present knowledge about air traffic at Frankfurt Airport nowadays and in the future.

If an airport is planned to be relevantly expanded in Germany, this expansion has to be agreed upon in an official approval process. Part of this approval process is the description of the status quo of air traffic and projections for the future both in case of an expansion (*Planungsfall*) and in case of no expansion (*Prognosenullfall*). These numbers were taken from the documentation of the official approval process [31] and are shown in Table 7.1.

Table 7.1: Air traffic movements (take offs and approaches) during the night and during the day for different traffic scenarios based on the six months with the highest air traffic volume of the year (here: May – October). [1], [4] and [5] taken from [31].
* calculated as shown in the text below.

Air traffic	Total	Day 06:00 – 22:00	Night 22:00 – 06:00	Percentage Night
[1] Status quo 2000	240,217	215,082	25,135	10.5%
[2] Status quo 2005* (noise scenarios 1 and 3)	253,920	225,216	28,704	11.3%
[3] Status quo 2005* (noise scenario 2)	245,088	225,216	19,872	8.1%
[4] No expansion 2015	261,976	234,619	27,357	10.4%
[5] Expansion 2015 with noise-free period	344,926	317,449	27,477	8.0%
[6] Expansion 2015* without noise-free period	344,926	308,907	36,019	10.4%

During the period May until October 2000, which corresponds to the six months with the highest air traffic volume in 2000, a total of 240,217 movements was observed (see [1] in Table 7.1) at Frankfurt Airport. 10.5% of these movements took place during the night, which is officially defined as the period from 22:00 until 06:00. If noise scenario 1, i.e. the number of planes taking off on 16 August 2005, is projected to a period of six months by multiplying by 184 days and then again by two (to account for approaches), 253,920 aircraft movements would be expected (see [2] in Table 7.1). This number falls between the status quo 2000 (240,217 movements) and the projection for 2015 if there is no expansion (261,976 movements, see [4] in Table 7.1), which seems reasonable.

The percentage of the total traffic that takes place during the night was 11.3% in noise scenarios 1 and 3, and therefore somewhat higher than in

[1] and [4] of Table 7.1. This is most probably caused by the fact that 16 August 2005 lies in the middle of summer holidays, with additional air traffic caused by charter airlines using cheap slots during the night. Therefore, the chosen approach, again, rather overestimated the air traffic volume and its effects during the night.

If the airport will be expanded, expected aircraft movements go up by more than 80,000 and reach a total of 344,926 movements in 2015 (see [5] in Table 7.1). Of course, the additional noise burden associated with this increase of air traffic will also be spread among airport residents living in several locations around the airport, once the new runway is opened. Projections predict that, once a noise-free period is implemented during the night, the percentage of night traffic decreases from 10.4% (no expansion 2015) to 8.0%. This percentage is similar to the one of scenario 2, where all movements between 23:00 and 05:00 were cancelled (see [3] in Table 7.1). If the percentage of night traffic remained at 10.4% (see [6] in Table 7.1), 36,019 movements would be expected during the night in contrast to 27,477 with a night traffic percentage of 8.0% (see [5] in Table 7.1). The decrease of night traffic percentage to 8.0% after the implementation of a noise-free period could either be caused by flight cancellations (some carriers may shift to other airports without restrictions during the night) or by rescheduling to slots before 22:00 or after 06:00. Probably both will happen at the same time (see below). Nevertheless, the 28,705 aircraft movements during the night assumed in scenario 1 go well along with the projected numbers for 2015, both with and without an expansion of Frankfurt Airport.

It was not accounted for military machines in noise scenarios 1, 2 and 3. The number of military aircraft movements highly depends on the momentary political situation. It is expected that the U.S. Army will have retreated from Frankfurt Airport in 2015 [23]. Therefore, military machines were also not accounted for in these scenarios (see [4] to [5] in Table 7.1).

For the simulation, it was assumed that all take offs on 16 August 2005 were carried out from one runway (runway 25), and that the modeled residents live close enough to the runway that they hear each plane before the flight paths diverge into different directions. As was already stated in Chapter 5.4.2, the investigated traffic pattern will be found only at special locations around the airport and only on days with high to extreme traffic volumes in reality, which is a conservative approach.

The direction of take offs and approaches depends on wind direction: Both take offs and approaches are performed "against the wind" (with headwind). If wind direction remains stable during the night, people will either experience aircraft noise from starting or from landing planes, but never simultaneously from starting and landing planes. In this thesis, only take off times were extracted from the time schedule for two reasons:

(1) Because of noise emitted from the engines, SPLs emitted from starting planes are usually higher than those emitted from landing planes. As SPLs applied in the DLR laboratory studies were on average louder than those observed in the field, selecting times of starting planes seemed to be the favorable choice.

(2) The distribution of take offs and approaches over 24 hours of the day differs for starting and landing planes. Compared to starting planes, the frequency of landing planes is higher during the late evening and lower during early morning hours. Sleep debt decreases with elapsed sleep time. Simultaneously, the vulnerability of sleep for noise induced sleep disturbances increases towards the end of the night. Therefore, concentrating on starting planes, that are expected more often in early morning hours, should lead to more severe impacts of noise on sleep. Choosing take offs, therefore, represents a conservative approach for the affected population. Nevertheless, the effect of different distributions of flight movements over 24 hours of the day (e.g.

take offs vs. landings) should be investigated more closely in the future.

In noise scenario 3, all take offs scheduled for the period 23:00 to 05:00 on 16 August 2005 were redistributed to the two periods 22:00 to 23:00 and 05:00 to 06:00. An expert opinion [23] on the possible consequences of a period without traffic from 23:00 to 05:00 forecasts for carriers that:

- 36.5% of air traffic movements will be rescheduled to the periods 22:00 to 23:00 and 05:00 to 06:00
- 17% will be rescheduled to periods before 22:00 and after 06:00
- 35.7% will be relocated to other airports and
- 10.8% will be cancelled.

Therefore, the assumption made in this thesis corresponds to a *worst case scenario*, possibly leading to an overestimation of the effects of scenario 3 on sleep structure.

The 15 take offs between 23:00 and 02:00 on 16 August 2005 were rescheduled to the period between 22:00 and 23:00, resulting in a total of 56 take offs in this period. On independent parallel runway systems, 70 or even more take offs per hour can easily be handled. Frankfurt has a parallel runway system, but it is not independent as the distance between the two runways is too short. Therefore, in reality it may not be possible to handle 56 take offs per hour. It would have been possible to reschedule the take offs between 23:00 and 02:00 to a longer time period, e.g. from 19:00 to 23:00. This would have resulted in lower peak traffic densities, that, on the one hand, could be handled by Frankfurt's parallel runway system with a higher probability. On the other hand, only 17% of air traffic that formerly took place between 23:00 and 05:00 are expected to be rescheduled to periods before 22:00 and after 06:00 (see above).

Additionally, results are not believed to differ substantially between both possible scenarios (rescheduling to 22:00 – 23:00 or to 19:00 – 23:00).

As was already mentioned in Chapter 7.3.1, the data of all subjects were synchronized to sleep onset. Therefore, a possible influence of aircraft noise on sleep latency, i.e. the time needed to fall asleep, could not be taken into account. If there is an influence of aircraft noise on sleep latency, it is expected that subjects that go to bed early (before 23:00) will be especially affected by aircraft noise because of the high traffic densities before 23:00. On the other hand, an influence of aircraft noise on sleep latencies of people that regularly go to bed after 23:00 is expected to be negligible.

7.5.2 Results

A comparison of the main outcome parameters between noise-free baseline nights and the weighted results of noise scenarios 1, 2 and 3 was given in Table 6.5 of Chapter 6.4.

The hypotheses proposed in Chapter 5.7 were all corroborated by the weighted results of the simulation: Amounts of stage Wake increased under the influence of noise. Also, amounts of stage S1 and the number of sleep stage changes, classical indicators of sleep fragmentation [47], increased in noise compared to noise-free nights. Amounts of stage S2 were only slightly increased. Amounts of S3, S4 and REM all decreased under the influence of noise. The SQI-score accordingly decreased in all noise scenarios.

The absolute differences between noise scenario 1 compared to a noise-free night amounted to: Wake +7.6 min, S1 +2.0 min, S2 +1.9 min, S3 –3.2 min, S4 –5.1 min and REM -3.2 min (see Table 6.5 and Figure 6.13). These effects are certainly not negligible, but at the same time they are not alarming. It has to be kept in mind that, because of the conservative approach and the use of laboratory instead of field data, absolute changes in the amounts of the different sleep stages are likely to be overestimated.

The relatively moderate changes in sleep structure may be attributed to the fact that at present traffic densities between 23:00 and 05:00 at Frankfurt Airport are rather low (four take offs per hour in scenario 1) and that about 70% of the adult German population fall asleep between 22:15 and 23:45 [26]. Accordingly, their sleep will only be little affected by aircraft noise in the first hours of sleep.

As hypothesized, the cancellation of air traffic between 23:00 and 05:00 without rescheduling to periods before 23:00 and after 05:00 lead to improvements of sleep structure and would therefore be beneficial for airport residents. On the other hand, the absolute improvements were rather low, which may, again, be attributed to the relatively low density of air traffic during the night (23:00 - 05:00) at present (scenario 1). The absolute differences scenario 2 - scenario 1 amounted to: Wake -3.1 min, S1 –0.9 min, S2 –1.6 min, S3 +1.7 min, S4 +2.8 min and REM +1.1 min (see Table 6.5 and Figure 6.13).

At least some of the traffic that formerly took place between 23:00 and 05:00 will be rescheduled to periods before 23:00 and after 05:00 once a noise-free period is established [23]. The expansion of the airport will also lead to increments in traffic densities before 23:00 and after 05:00 per se (see [5] in Table 7.1). At the same time, the new runway may be used to spread air traffic over different locations. Therefore, the development of air traffic before 23:00 and after 05:00 is rather unpredictable once the airport is expanded, but it is still very likely that it will increase at most locations around the airport. Hence, it is expected that some of the benefits of the noise-free period (scenario 2) will be undone by the process of rescheduling (scenario 3). One of the main questions of this thesis was whether a noise-free period with more traffic before and after will be better or worse for the sleep of airport residents. Although all traffic that formerly took place between 23:00 and 05:00 was rescheduled to periods 22:00 to 23:00 and 05:00 to 06:00, a worst case scenario, there was still a benefit of scenario 3 compared to scenario 1 in terms of sleep structure. On the other hand, the observed absolute differences between scenario 3 and scenario 1, i.e. the net benefit for the affected population, were relatively small: Wake –1.4 min, S1 –0.5 min, S2 –2.0 min, S3 +1.1 min, S4 +2.2 min and REM +0.6 min (see Table 6.5 and Figure 6.13)

Figure 6.14 reveals the influence of time of falling asleep on noise induced changes in sleep structure. The optimal time of falling asleep according to sleep structure was 23:15. Compared to noise-free nights, the increase of Wake and S1 and the decrease of S3, S4 and REM were minimal. Earlier or later bed times went along with an overlap of the beginning or the end of the sleep period with the periods of high traffic density either from 20:00 to 23:00 or from 05:00 to 10:00. Therefore, Wake and S1 increased in situations were time of falling asleep both preceded or followed 23:15. S3 and S4 especially decreased if subjects fell asleep earlier than 23:15: In this case, the first half of the sleep period, which is usually dominated by SWS, overlapped with the high traffic density period from 20:00 to 23:00.

REM sleep steadily decreased with times of falling asleep later than 22:45. In that case, the last half of the sleep period, which is usually dominated by REM sleep, more and more overlapped with the high traffic density period from 05:00 to 10:00. Surprisingly though, for times of falling asleep before 22:45, REM sleep slightly increased in noise scenarios 2 and 3, i.e. in the scenarios with noise-free periods from 23:00 to 05:00, compared to noise-free nights. This may be caused by noise-induced shortenings of NREM-cycles for the benefit of longer REM-cycles, as noise obviously lead to a shortening of SWS in these situations.

Compared to the average duration of 210.9 min in noise-free nights, amounts of S2 were only moderately altered by aircraft noise, if all. This may be explained by the fact that noise may lead simultaneously to an increase in transition probabilities from S2 to Wake and S1 and from S3, S4 and REM to S2.

Overall, Figure 6.14 reveals that the effects of the different times of falling asleep were much stronger than the differences in effects between the three noise scenarios. For example, the largest difference in amounts of stage Wake between the noise scenarios is 3.8 min between scenario 1 and 2 when falling asleep at 22:45. Compared to that, the largest difference in amounts of stage Wake between one of the noise scenarios and the noise-free night is 22.3 min for scenario 3 when falling asleep at 20:45.

Therefore, the remote benefits of the introduction of a noise-free period between 23:00 and 05:00 could be easily exceeded, at least by some people, by small changes of sleep habits, i.e. by advancing or postponing the time of falling asleep by 30 min to one hour. This is shown exemplarily for the SQI-score in Figure 7.1.

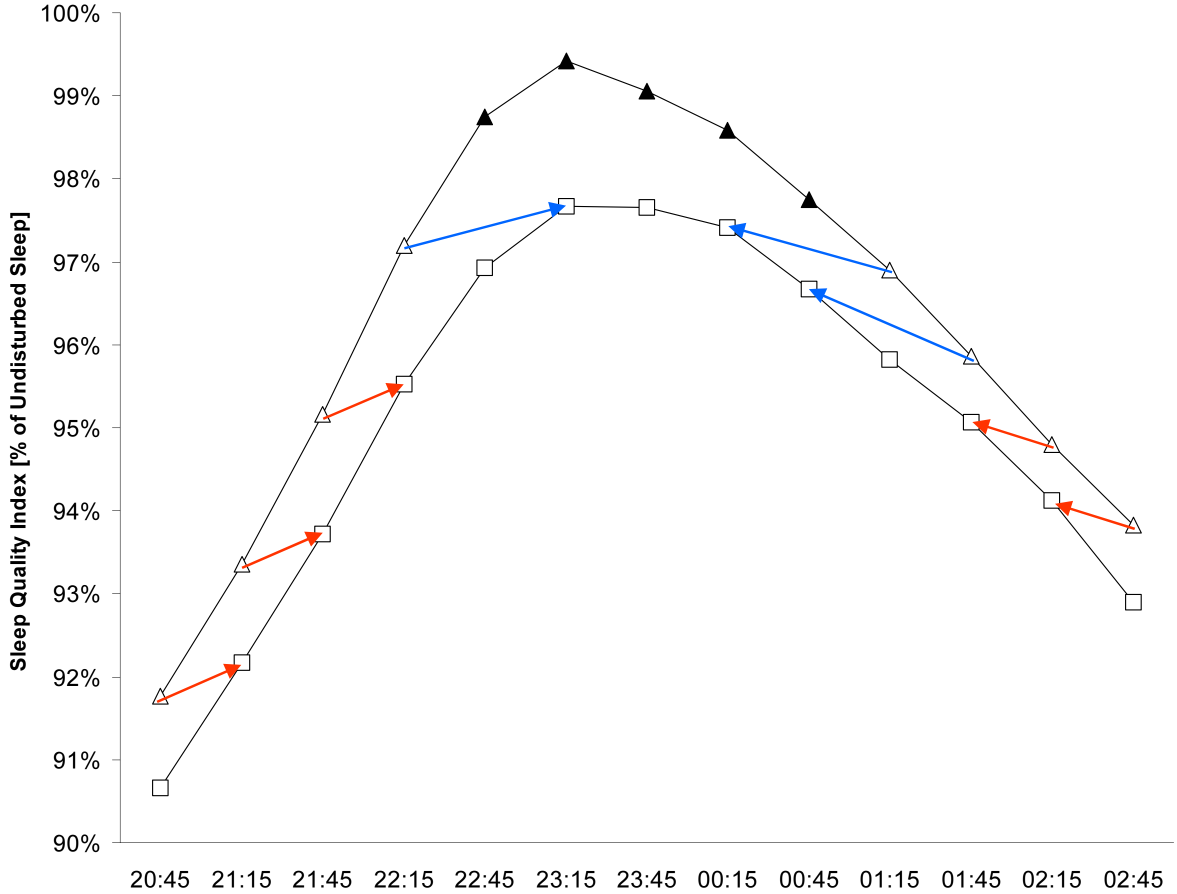


Figure 7.1: Sleep Quality Index (SQI) in % of undisturbed sleep depending on time of falling asleep. Noise scenario 1 (current noise at Frankfurt Airport) is shown as squares while noise scenario 2 (take offs cancelled between 23:00 and 05:00) is shown as triangles.

The beneficial effect of the introduction of a noise-free period form 23:00 to 05:00 could be easily exceeded by postponing sleep period time by 30 min (20:45, 21:15 and 21:45, red arrows) or by one hour (22:15, blue arrow), respectively. It would also be exceeded by advancing sleep period time by 30 min (2:15 and 2:45, red arrows) or by one hour (1:15 and 1:45, blue arrows), respectively. For times of falling asleep between 22:45 and 0:45, there is a true benefit of a noise-free period between 23:00 and 5:00, that cannot be reached or exceeded by advancing or postponing sleep period time. As long as the air traffic pattern presently exercised at Frankfurt Airport and a sleep period time of 410.5 min are concerned, a time of falling asleep of 23:15 can be described as optimal for sleep structure. Earlier or later times of falling asleep go along with more pronounced changes in sleep structure, that outweigh possible beneficial effects of the introduction of a noise-free period by far.

7.6 Limitations and future research needs

Several limitations should be borne in mind when interpreting the results presented in this thesis:

(1) The models were based on data derived from experimental laboratory studies with healthy subjects aged 19 to 65 years. Subjects had to meet inclusion criteria to be eligible for study participation. Although internal validity was maximized by this procedure, external validity of the results may be restricted.

(2) Data of the laboratory study were primarily chosen as data basis because of the controlled conditions (e.g. constant time in bed of 8 hours for all subjects) and the existence of a noise-free baseline night. Nevertheless, reactions of the sleeper to aircraft noise are usually more frequent and more severe in laboratory compared to real life conditions [6, 40]. Additionally, SPLs applied in the DLR laboratory study were on average higher than those that are typically observed in the field, also leading to a possible overestimation of effects. Hence, the ecologic validity of the analyses presented here is limited, although the effects of noise on sleep structure were rather overestimated than underestimated; a conservative approach in favor of the affected population. In the future, data derived from field studies should be used for estimating the effects of aircraft noise on changes in transition probabilities between sleep stages. The results of these analyses should then be compared to the results found in this thesis.

(3) Naturally, the length of sleep period times differs inter-individually. No inferences on the effects of different sleep period lengths on sleep structure could be made in this thesis, as a fixed sleep period time of

410.5 min was used for analyses and comparisons for practical reasons. Hence, the influence of different sleep period lengths on the effects of aircraft noise on sleep should be investigated more closely in the future.

(4) Subjects were synchronized to sleep onset, i.e. the process of falling asleep was no part of the analyses presented here. Of course, aircraft noise may interfere with the process of falling asleep, depending on traffic densities and maximum SPLs of ANEs during evening hours. Hence, the models presented here should be extended in the future to the period before sleep onset in order to incorporate effects of aircraft noise on the process of falling asleep.

(5) Only three specific noise patterns, based on the time schedule for Frankfurt Airport on 16 August 2005 and considering take offs only, were compared in these analyses. Several assumptions had to be made,. e.g. concerning how much and to what periods traffic that formerly took place between 23:00 and 05:00 was rescheduled to periods before 23:00 and after 05:00. These assumptions should be borne in mind for the interpretation of the results presented here. The models developed in this thesis could be used to investigate the influence of various other traffic patterns. It could and should be used to search for optimal traffic strategies in terms of sleep quality in the future.

(6) Traffic densities assumed for the simulations were very high and may not or only very seldom be observed in reality. Rescheduling of all take-offs that formerly took place between 23:00 and 05:00 to shoulder hours has to be called a worst case scenario [23]. Overall, the assumptions made in this thesis rather lead to an overestimation of the effects of aircraft noise on sleep.

(7) In this first approach, different maximum SPLs of ANEs were not differentiated. Rather, all noise events were treated as if they consisted

of the same maximum SPL, and this maximum SPL can be thought of as an "average" maximum SPL of all 26,135 ANEs played back during the first 410.5 min of sleep period time of the subjects of the experimental group (see also Table 5.1). In the future, maximum SPLs of ANEs should be incorporated into the model in order to improve predictions of noise effects on sleep structure.

(8) Regression models based on data of noise-free baseline nights and on nights with noise exposure were thoroughly validated using the original data. Additionally, results of Monte Carlo simulations of noise-free baseline nights were validated with original data. The simulation of noise nights combined data of noise-free baseline nights and noise exposure nights in order to estimate the effects of three different noise scenarios on sleep structure. These scenarios were themselves not applied in the DLR laboratory study, and therefore, it was not possible to validate the results of the noise models with original data. Actually, it was the fundamental idea of this thesis to facilitate the information gathered on single ANEs to study the influence of complex traffic patterns, that may even lie in the future. Although the results of the simulations of noise exposure nights are plausible and go well along with the hypotheses defined in Chapter 5.7., it would nevertheless be very valuable to validate the results of the noise models in the future. This goal could be achieved, e.g., using a specific traffic pattern applied during the DLR laboratory studies.

(9) The analysis of goodness-of-fit showed that the sleep stage distribution across the night predicted by the model reproduced the observed data well. Therefore, ultradian rhythms of the original data were also reproduced by the model, if the sleep stage distribution of the original data at T0 was given. This did not apply for the first-order simulation trials shown in Figures 8.13 to 8.18, where the simulation starts with sleep stage S2 and otherwise solely depends on transition probabilities calculated by the regression model. Here, the trend of the

probability distribution of the different sleep stages was reproduced well across the night, but no ultradian rhythmicity was replicated. Possible explanations were presented in Chapter 7.3.2. The models presented in this thesis should be modified and extended in the future, in order to be able to replicate ultradian rhythms in Monte Carlo simulation trials, too.

(10) Longer or more complex counter measures of the body could not be taken into account by this relatively simple model of human sleep, where the influence of noise was restricted to three epochs after the start of an ANE. Even if the models presented here validly reproduce and extrapolate the empirical data, there may still be restrictions of the original data that have to be borne in mind. Again, an external validation of the model results would be very desirable.

(11) Noise induced changes in sleep structure may lead to short- and long-term consequences for quality of life and health. Although a causal link between noise induced changes in sleep structure and health related long-term consequences is biologically plausible, a scientific corroboration proves difficult [4, 8, 38, 45]. Therefore, epidemiological studies on health related long-term consequences of nocturnal aircraft noise are also encouraged here.

(12) Finally, it is emphasized that the analyses presented in this thesis concentrate on changes in sleep structure as one fundamental effect of aircraft noise on humans with exceptional importance for recuperation and well-being. Nevertheless, there are other important dimensions that may be influenced by aircraft noise and that were not investigated in this thesis. E.g., a noise-free period from 23:00 until 05:00 may be psychologically beneficial for airport residents, irrespective of the results of this thesis, as residents may rely on a six hour period without aircraft noise.

8 Conclusions and recommendations

In this thesis, a mathematical model of human sleep was developed. It was based on noise-free baseline nights of 125 subjects and was able to reproduce key features of undisturbed human sleep without major bias. Additionally, the characteristic sequence of sleep stages seen in hypnograms simulated by the model resembled the natural sequence observed in human hypnograms closely.

The model was extended in order to allow for the prediction of changes in sleep structure in the presence of aircraft noise, based on the data collected in a polysomnographic laboratory study on the effects of aircraft noise on sleep. Three different noise scenarios, derived from the current and from possible future constellations at Frankfurt Airport, were compared with each other and with a noise-free night. The results of this comparison should be used to aid the decision making process of policy makers and legislative bodies. The following implications can be derived from the findings of this thesis:

(1) Affected residents will benefit from the introduction of a noise-free period, even if the traffic that took place between 23:00 and 05:00 will be completely rescheduled to the periods before 23:00 and after 05:00. At the same time, the observed net benefit will probably be very small, which may be attributed to relatively low traffic densities at Frankfurt Airport between 23:00 and 05:00 at present. The results also indicate that periods with high air traffic densities, i.e. before 23:00 and after 07:00, may disturb sleep much stronger than the traffic presently taking place during the night. Especially the sleep of subjects who choose to go to bed or have to go to bed (e.g. children, shift workers) earlier than 22:30 or later than 01:00 may substantially be disturbed by noise emitted from air traffic during shoulder hours, in particular as traffic densities will increase during shoulder hours after the extension of Frankfurt Airport.

(2) In order to protect sleep of the people living in the vicinity of Frankfurt Airport in the case of an extension, measures additional or alternative to a noise-free period from 23:00 to 05:00 should be taken into account. Possible administrative measures are:

- extension of the noise-free period before 23:00 and/or after 05:00,
- partial restrictions of air traffic during shoulder hours or
- optimal distribution of air traffic over different regions, which would be facilitated by a new runway.

(3) Considering the fact that a prohibition of air traffic between 23:00 and 05:00 will lead to legal conflicts and severe competitive disadvantages for some branches of the economy, alternative strategies to a ban of air traffic between 23:00 and 05:00 should also be developed and discussed. These may include improved noise protection of residents by the enhancement of sound insulation and the extension of noise protection areas around the airport [7, 21]. Additionally, advanced operational procedures (noise optimized take offs and landings) and decreases in noise emissions from single aircrafts may also serve to further reduce sleep disturbances induced by aircraft noise.

(4) The models developed in this thesis could be used to search for optimal strategies in order to minimize adverse effects of aircraft noise on sleep structure.

9 Appendix A: Graphical validation of regression and simulation results

In Chapter 9.1, goodness-of-fit tests for the results of the final autoregressive multinomial logistic regression model for noise-free baseline nights (see Appendix B) are shown. Two figures are shown for each sleep stage. In the first figure, predicted and observed probabilities are plotted simultaneously for each of the six sleep stages and for the whole night. In the second figure, differences (predicted - observed) are plotted for each of the six sleep stages for the whole night.

The predicted probabilities were calculated as follows: For each sleep stage and for each of the 820 transitions, six probabilities were calculated: The probability to be in the specific sleep stage at T1 if the present sleep stage (at T0) is Wake, the probability to be in the specific sleep stage at T1 if the present sleep stage (at T0) is S1, etc.. The six predictions for the specific sleep stage at T1 were then weighted based on the probability of the six different sleep stages at T0, which was calculated based on the original data.

This procedure is exemplified in Table 9.1 for the final baseline regression model with elapsed sleep time as the only additional explanatory variable (see Appendix B) for T1 = 200 min (corresponding to 400 epochs) and for sleep stage 2.

Table 9.1: Estimation of probability of S2 at T_{200} based on the final first-order autoregressive baseline regression model with elapsed sleep time as the only explanatory variable. Predicted probabilities after 200 min ($p(S2|T_{200}, PSS)$) are weighted according to the empirically observed relative frequencies of the different sleep stages in the preceding epoch ($p(PSS|T_{199.5})$).

Prior Sleep Stage (PSS)	$p(PSS\|T_{199.5})$ (A)	$p(S2\|T_{200}, PSS)$ (B)	(A) · (B)
Wake	0.088	0.167	0.015
S1	0.000	0.475	0.000
S2	0.584	0.917	0.536
S3	0.048	0.162	0.008
S4	0.048	0.013	0.001
REM	0.232	0.028	0.006
			Σ 0.565

Based on the empirical frequency distribution of sleep stages at $T_{199.5}$, the model predicts a relative frequency of stage S2 after an elapsed sleep time of 200 min of 0.565. The observed relative frequency at T_{200} was 0.560, indicating an excellent fit for this particular sleep stage and at this particular point in time.

In Chapter 9.2, goodness-of-fit tests for the results of 10,000 first-order Monte Carlo simulation trials of noise-free baseline nights are shown. Simulated and observed probabilities are plotted simultaneously for each of the six sleep stages and for the whole night. A single simulation trial refers to one individual randomly traversing a single path through the model. The state of the individual is stored for each cycle in the Markov trace.

Averaging over all simulated Markov traces leads to a probability distribution depending on elapsed sleep time.

In Chapter 9.3, goodness-of-fit tests for the results of the final autoregressive multinomial logistic regression models for epochs #1, #2 and #3 in nights with noise exposure are shown for each of the six sleep stages. Because of the relatively sparse data compared to the baseline model, the 410.5 min night was divided into six periods of 58.5 min and one period of 59.5 min. Observed and predicted probabilities were averaged over these intervals and compared with each other. Average probabilities derived from noise-free baseline nights are also given as black bars.

9.1 Regression results baseline model

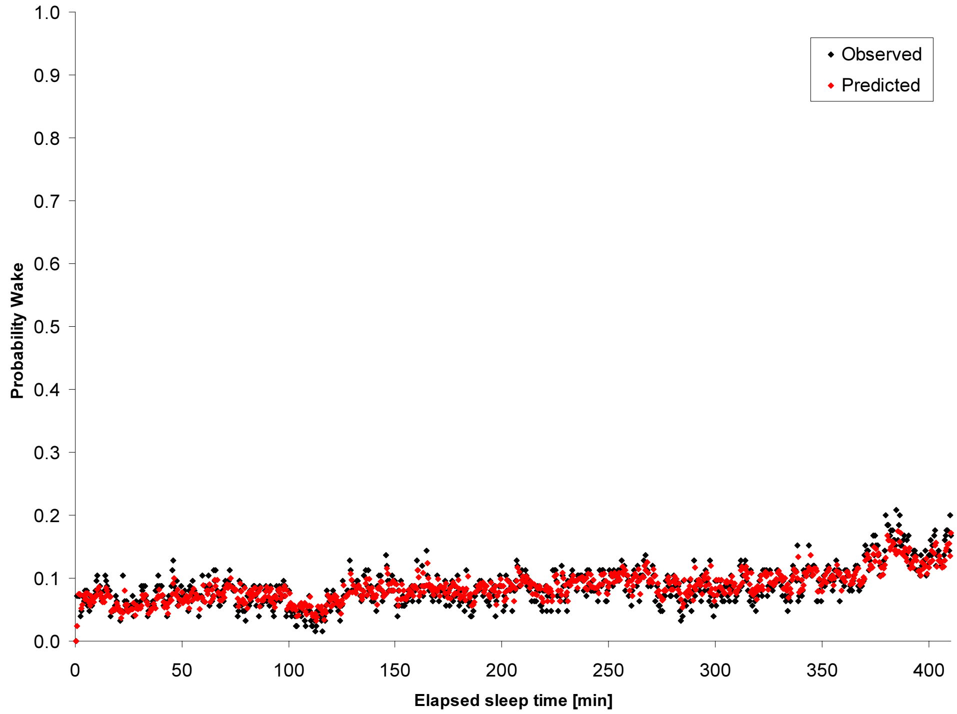


Figure 9.1: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage Wake.

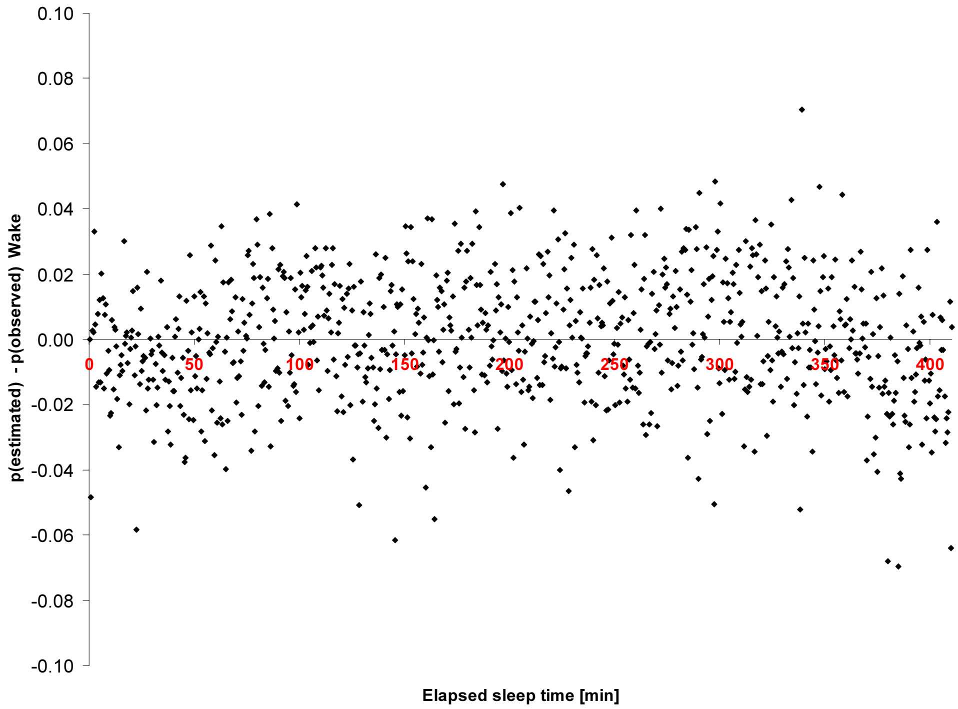


Figure 9.2: Differences p(predicted) – p(observed) for stage Wake.

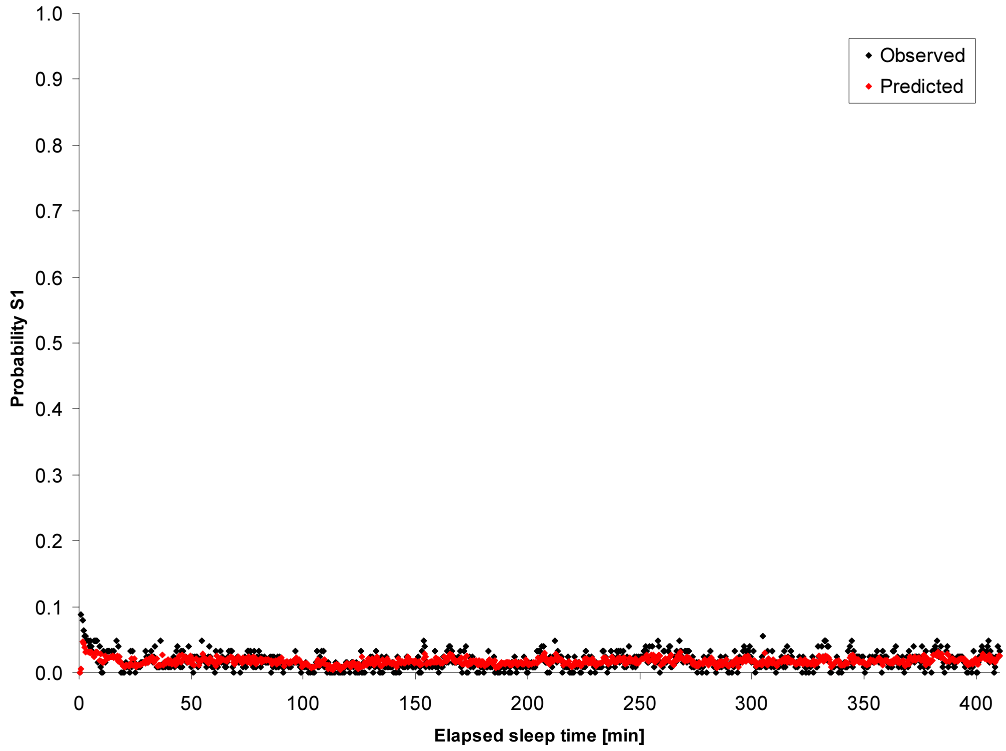


Figure 9.3: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage S1.

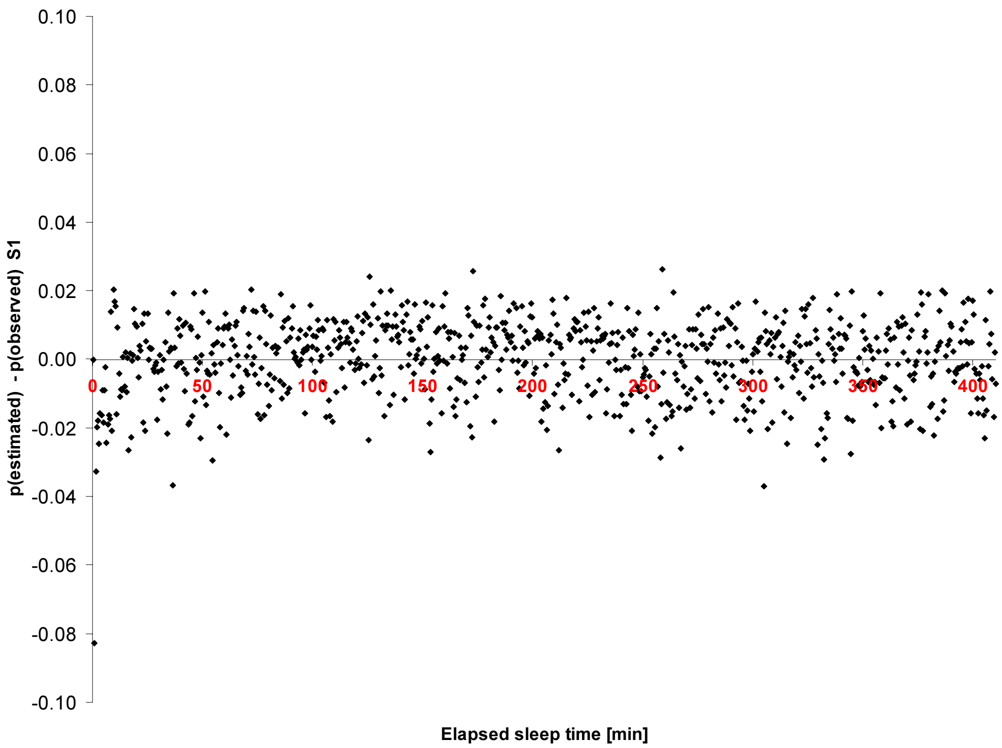


Figure 9.4: Differences p(predicted) – p(observed) for stage S1.

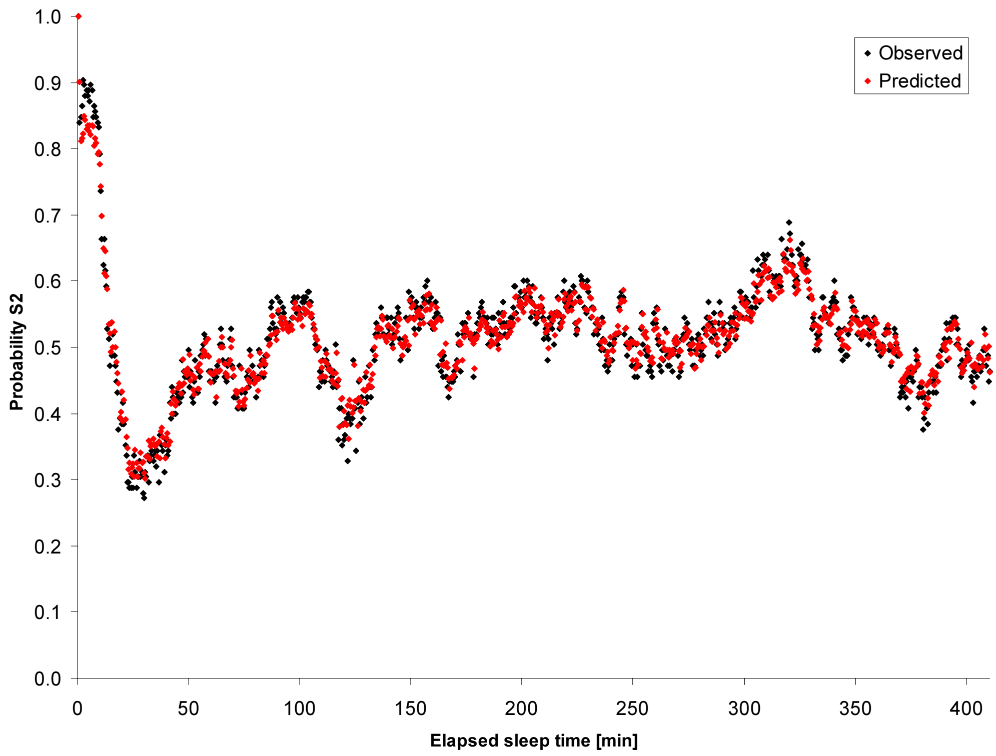


Figure 9.5: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage S2.

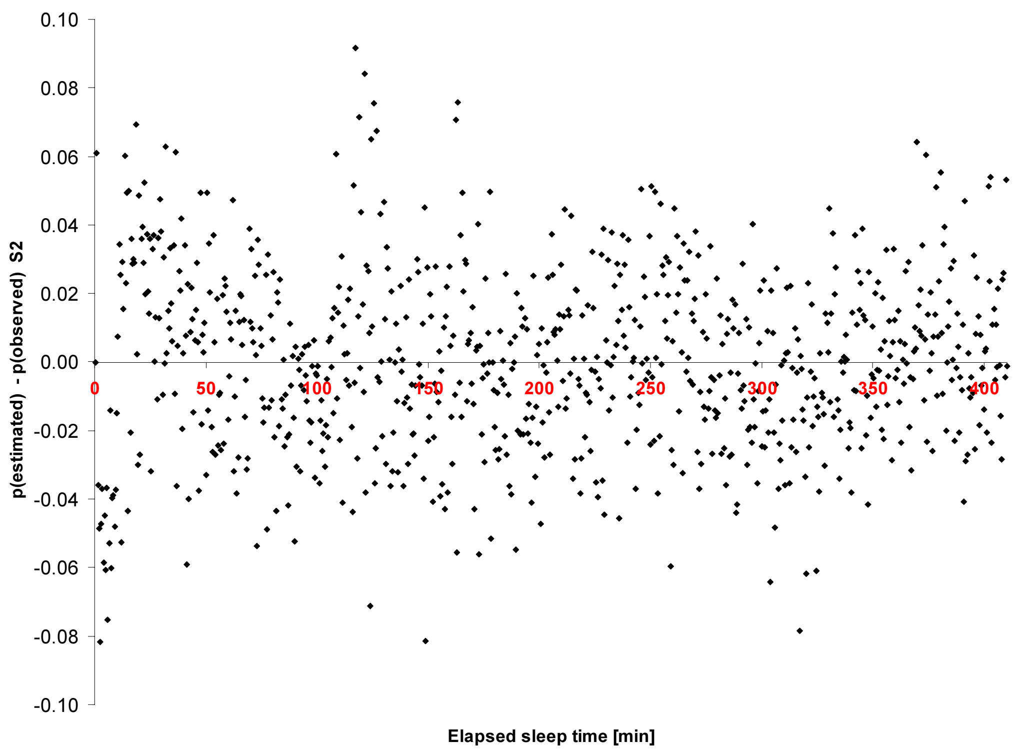


Figure 9.6: Differences p(predicted) – p(observed) for stage S2.

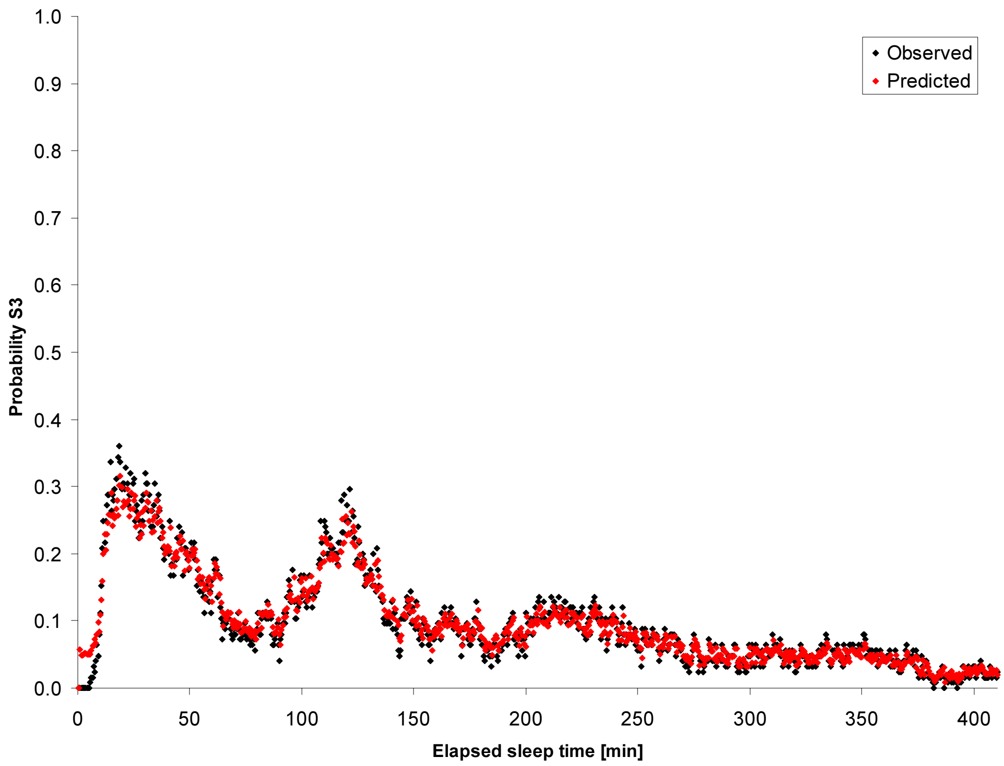


Figure 9.7: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage S3.

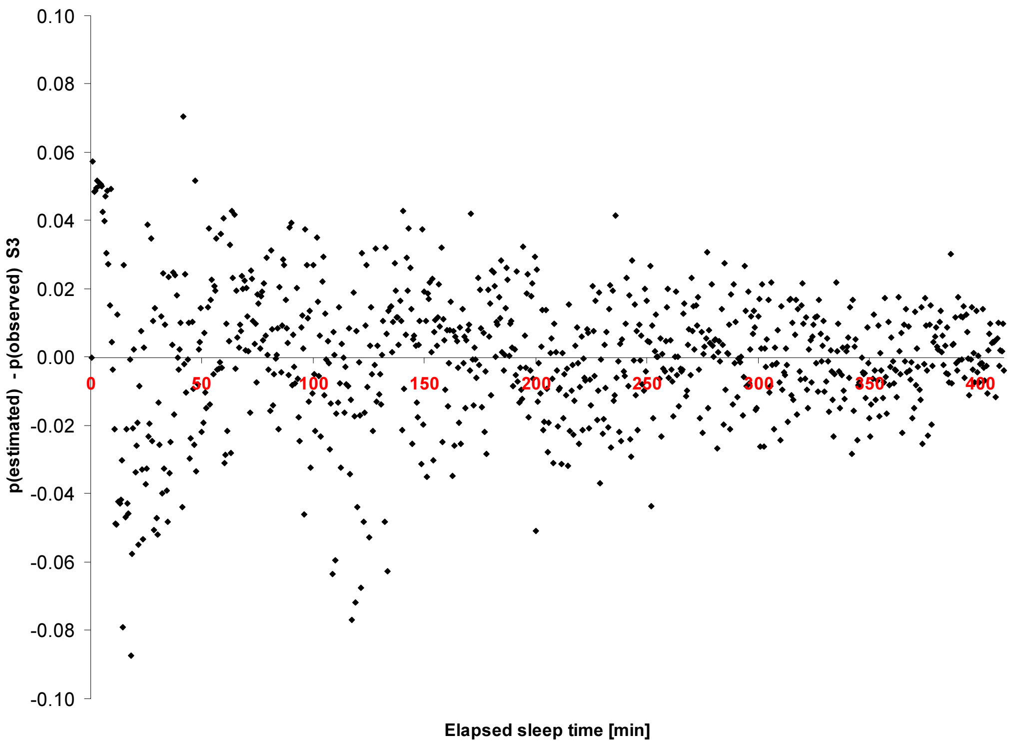


Figure 9.8: Differences p(predicted) – p(observed) for stage S3.

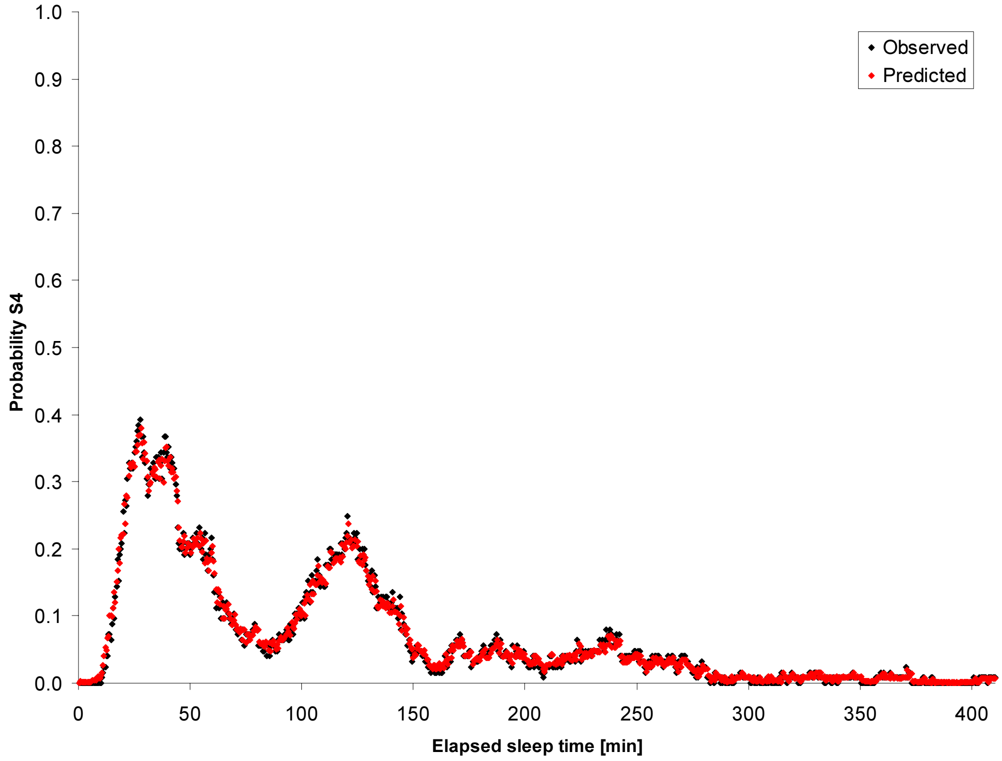


Figure 9.9: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage S4.

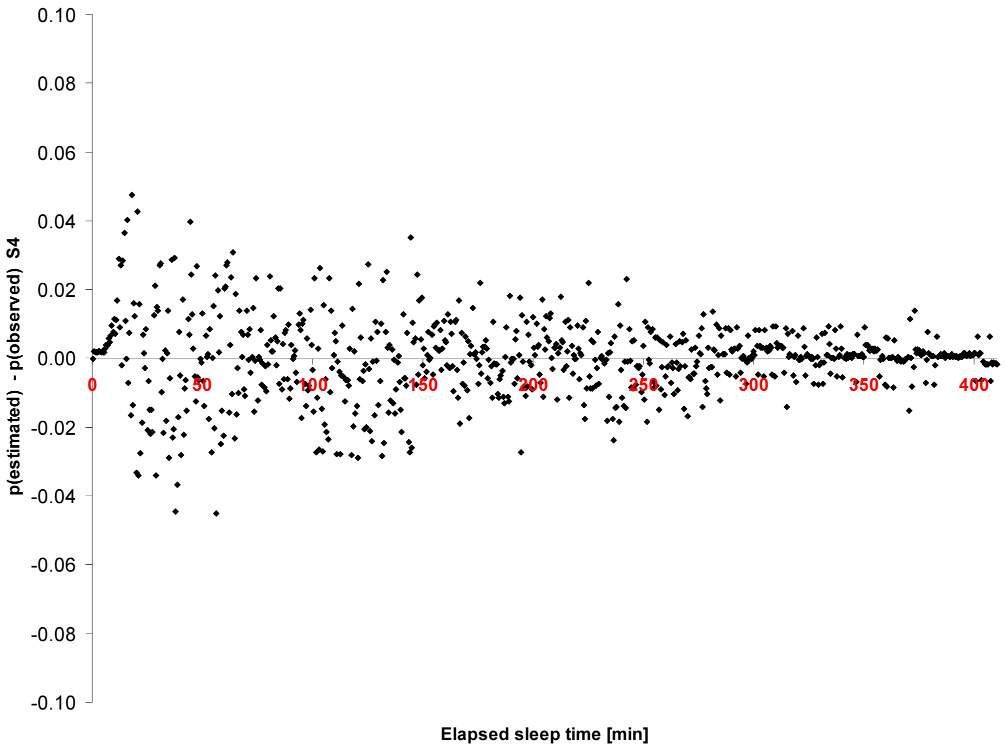


Figure 9.10: Differences p(predicted) – p(observed) for stage S4.

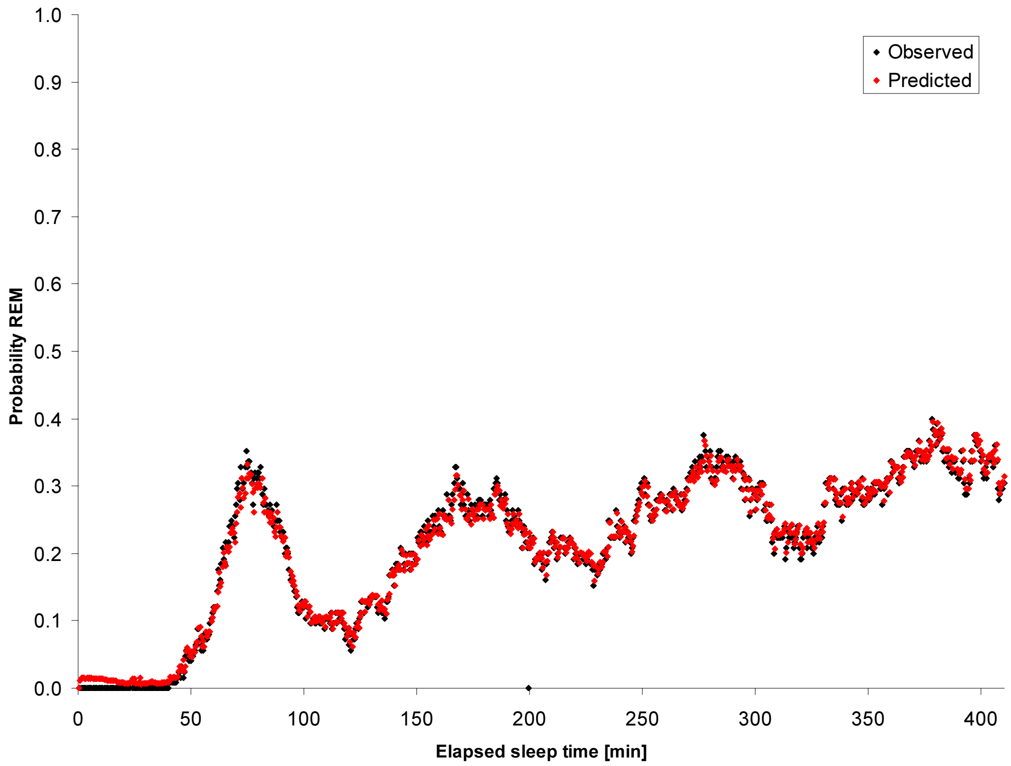


Figure 9.11: Comparison of observed probabilities (black dots) and probabilities predicted by the final regression model (red dots) for stage REM.

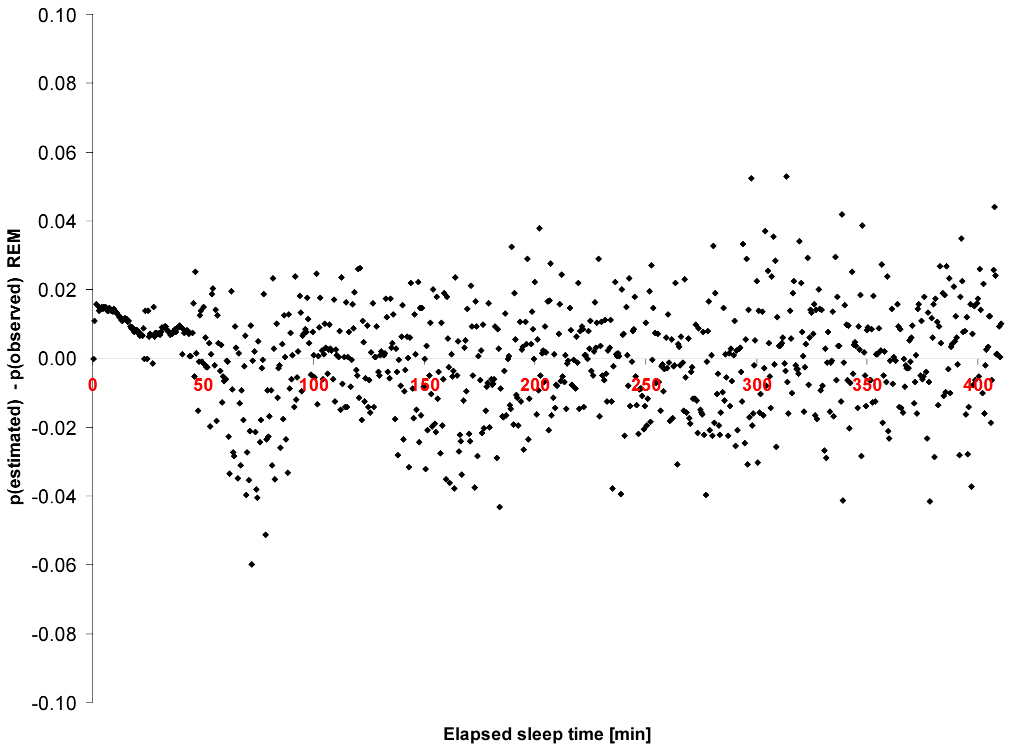


Figure 9.12: Differences p(predicted) – p(observed) for stage REM.

9.2 Simulation results baseline model

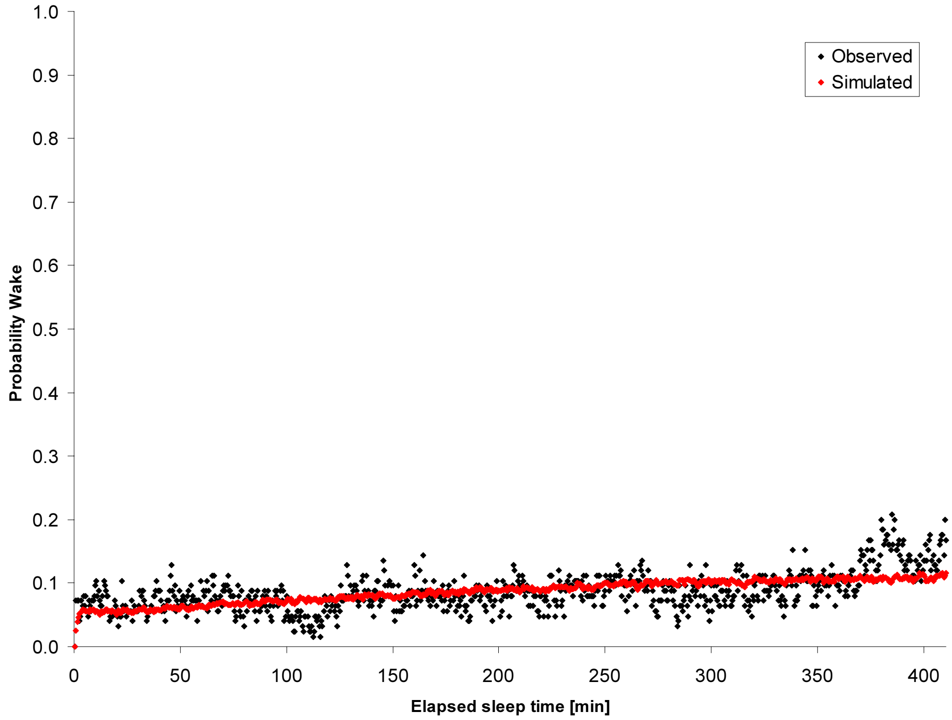


Figure 9.13: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage Wake.

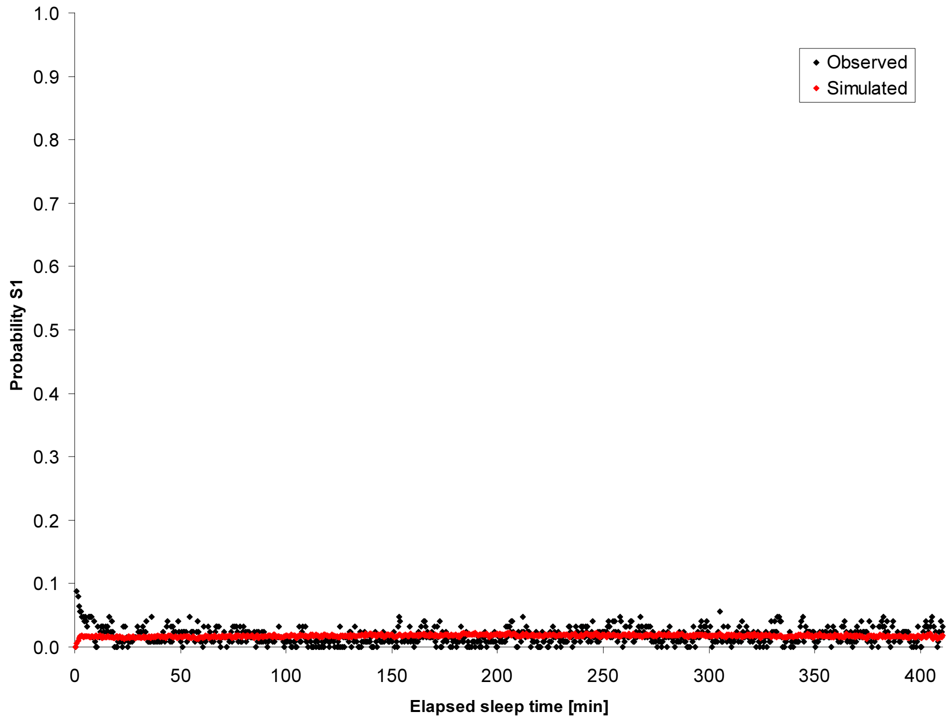


Figure 9.14: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage S1.

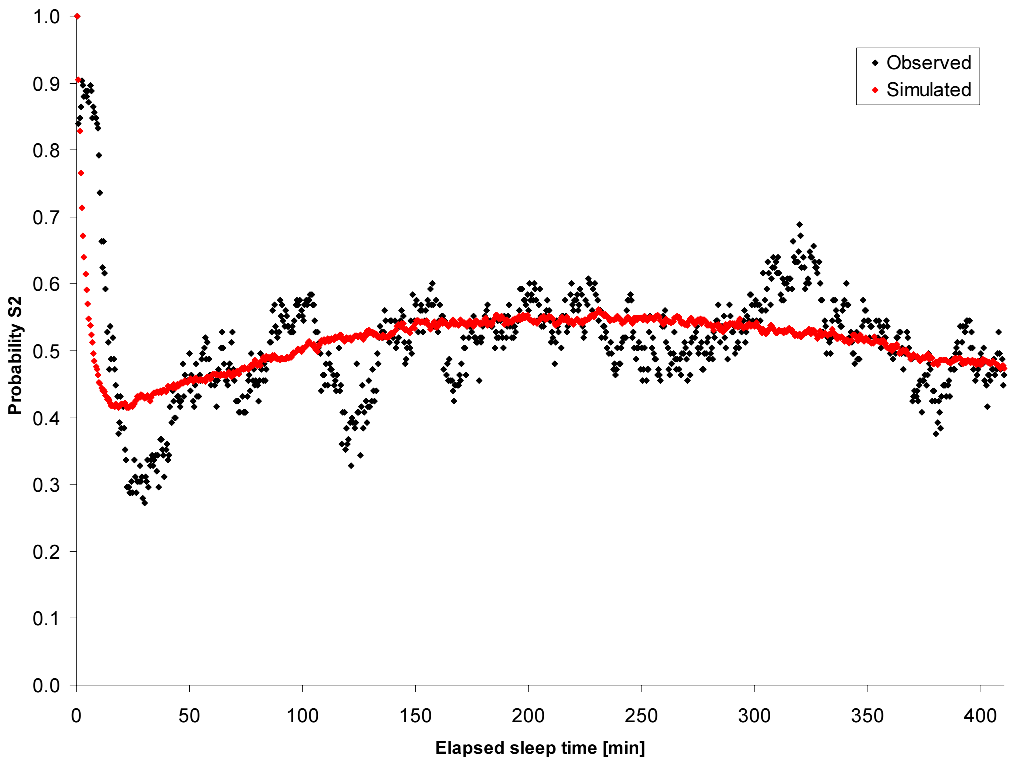


Figure 9.15: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage S2.

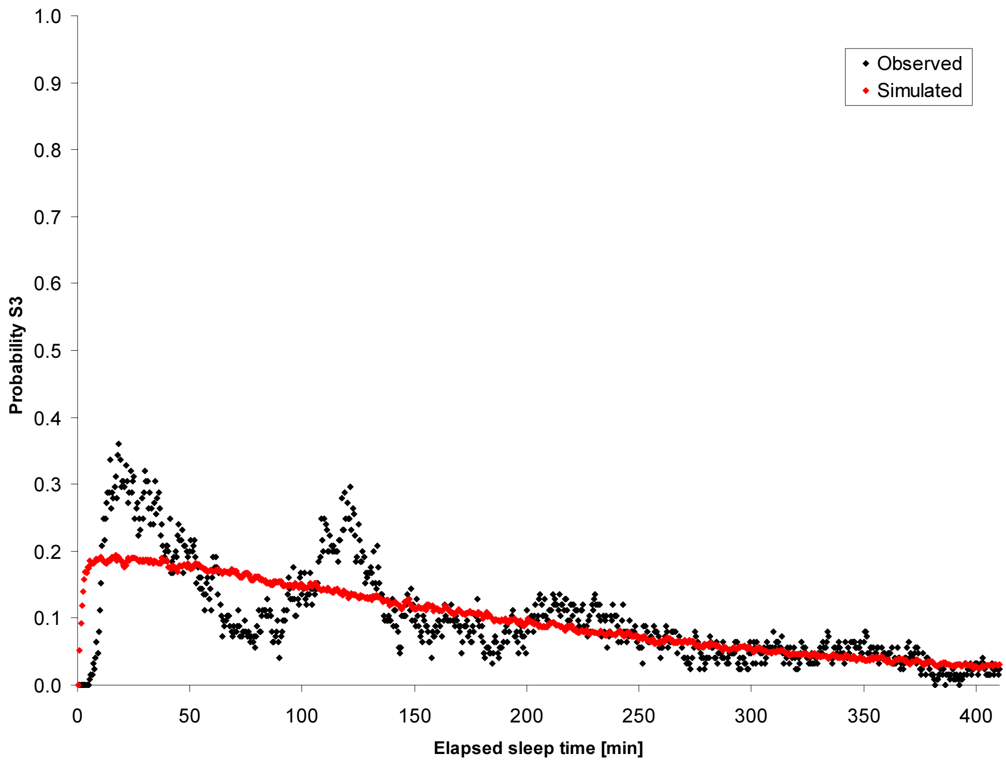


Figure 9.16: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage S3.

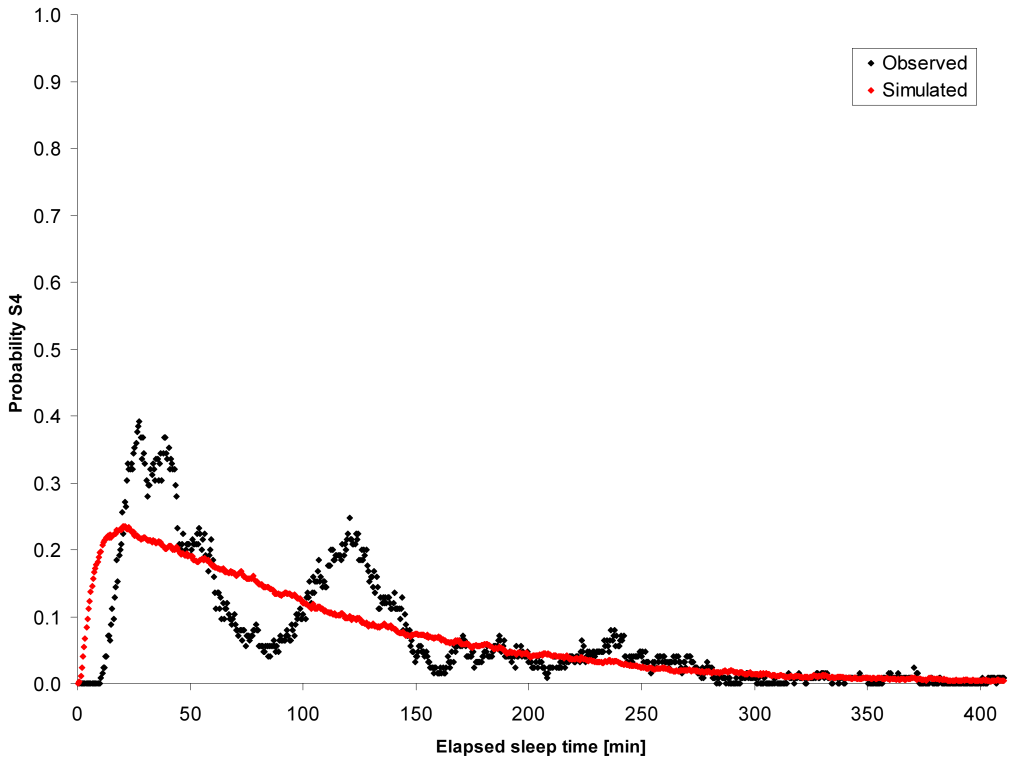


Figure 9.17: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage S4.

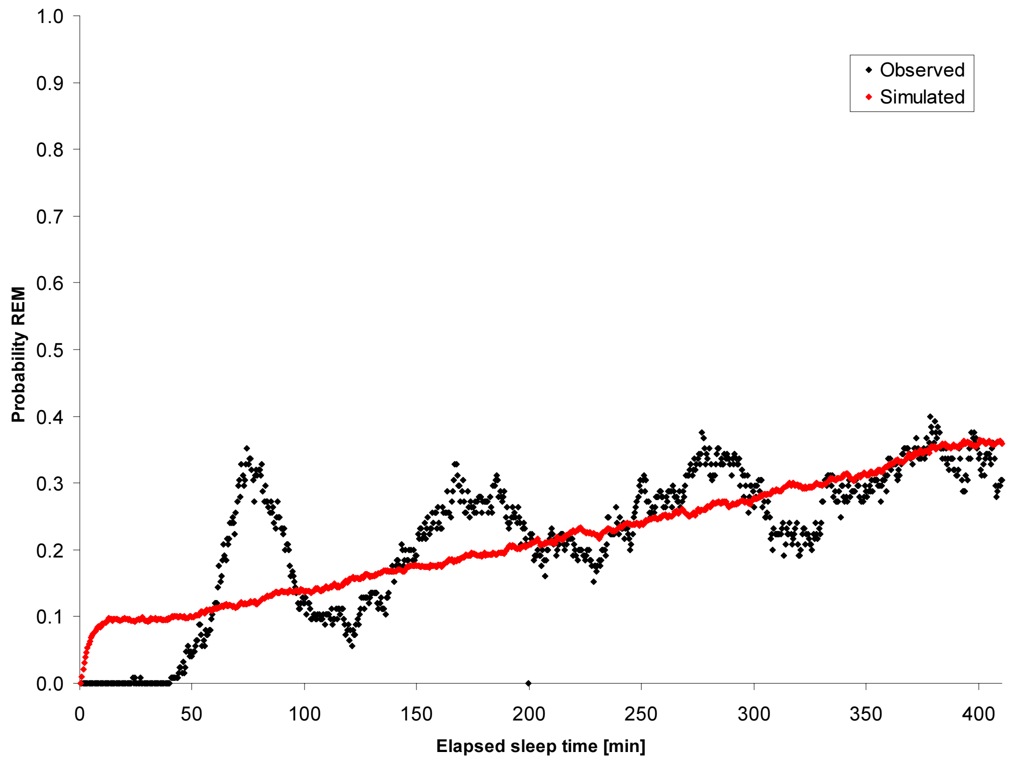


Figure 9.18: Comparison of observed probabilities (black dots) with average probabilities derived from 10,000 Monte Carlo simulations (red dots) for stage REM.

9.3 Regression results noise models

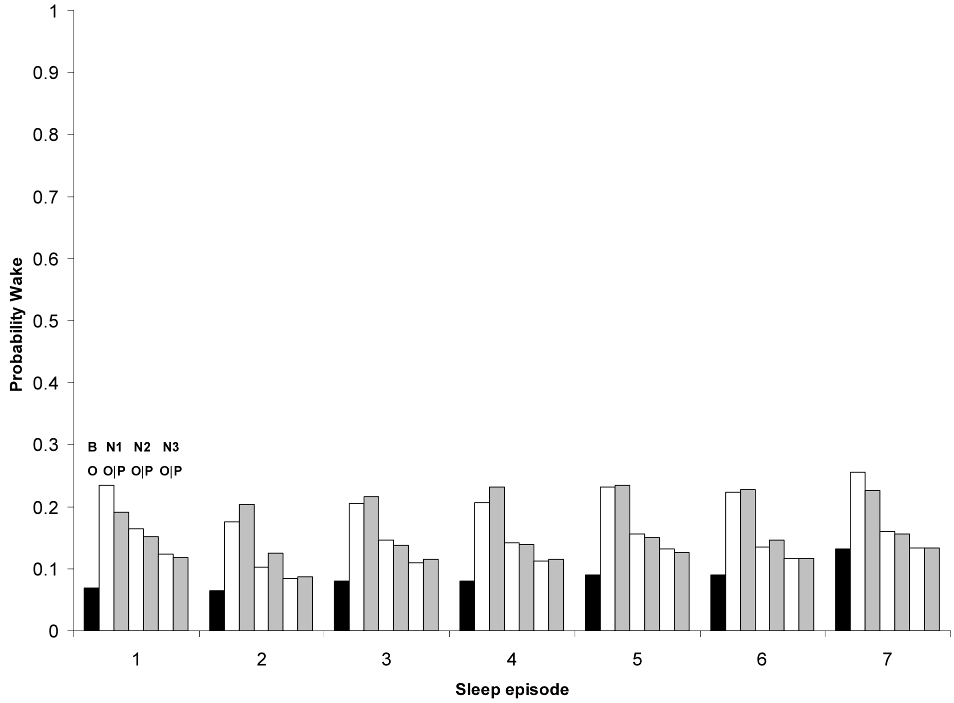


Figure 9.19: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage Wake.

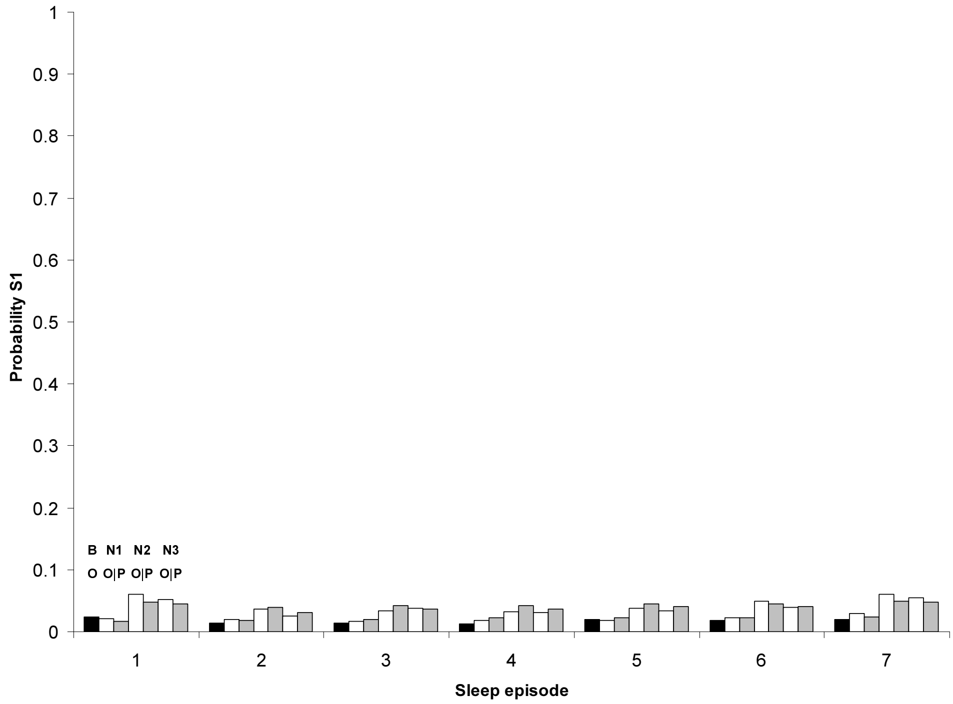


Figure 9.20: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage S1.

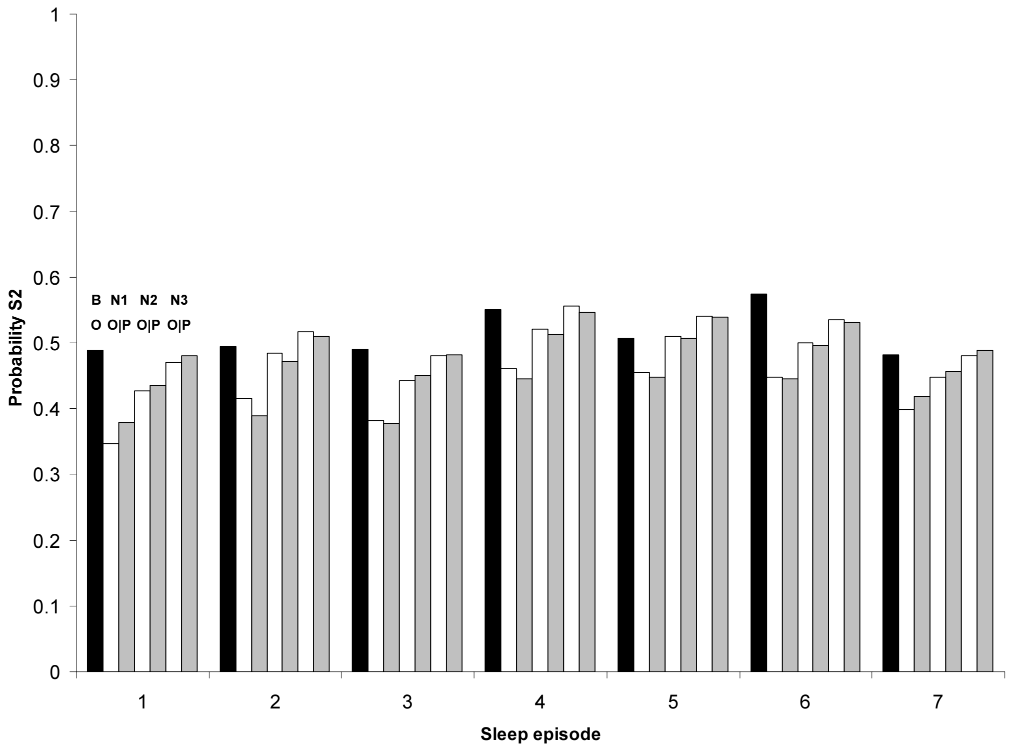


Figure 9.21: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage S2.

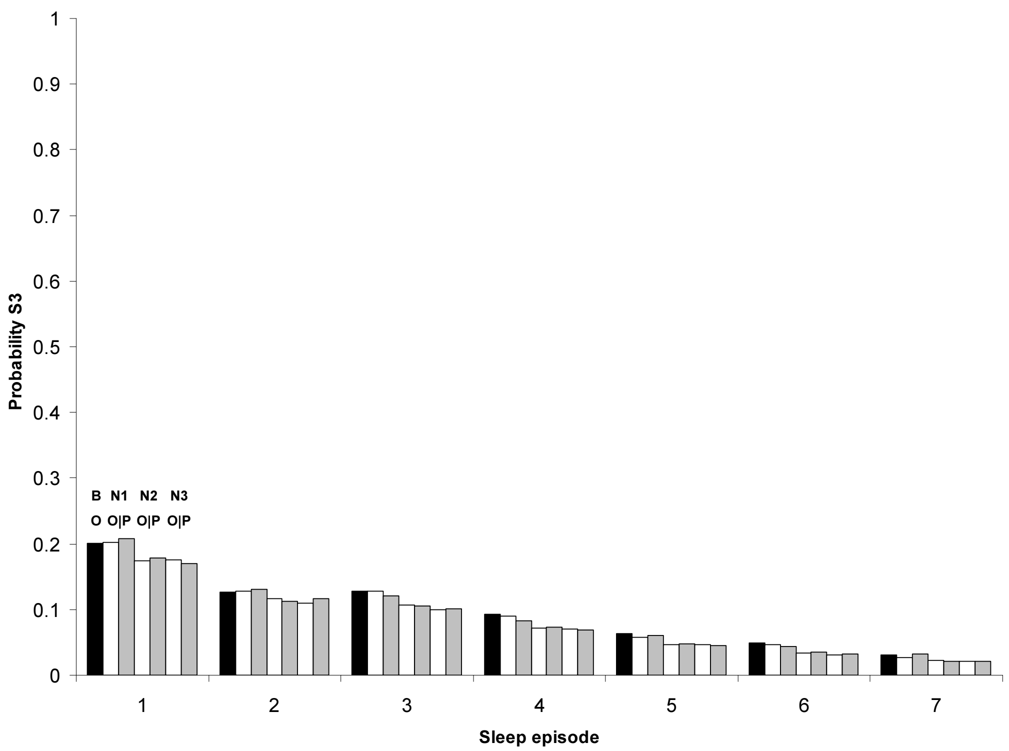


Figure 9.22: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage S3.

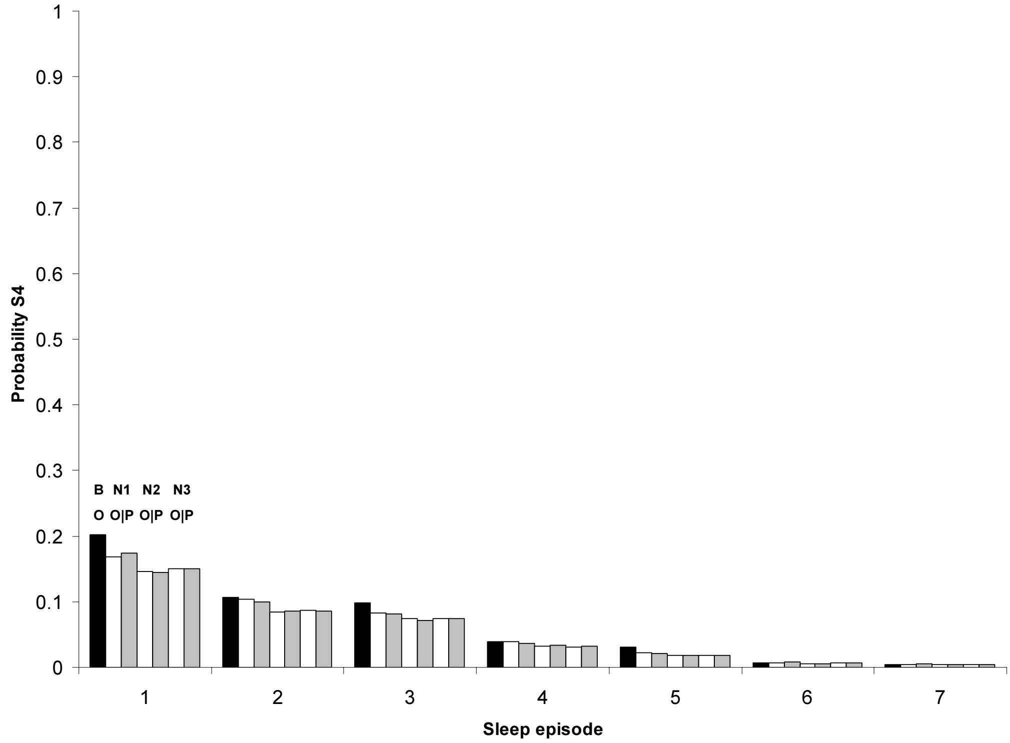


Figure 9.23: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage S4.

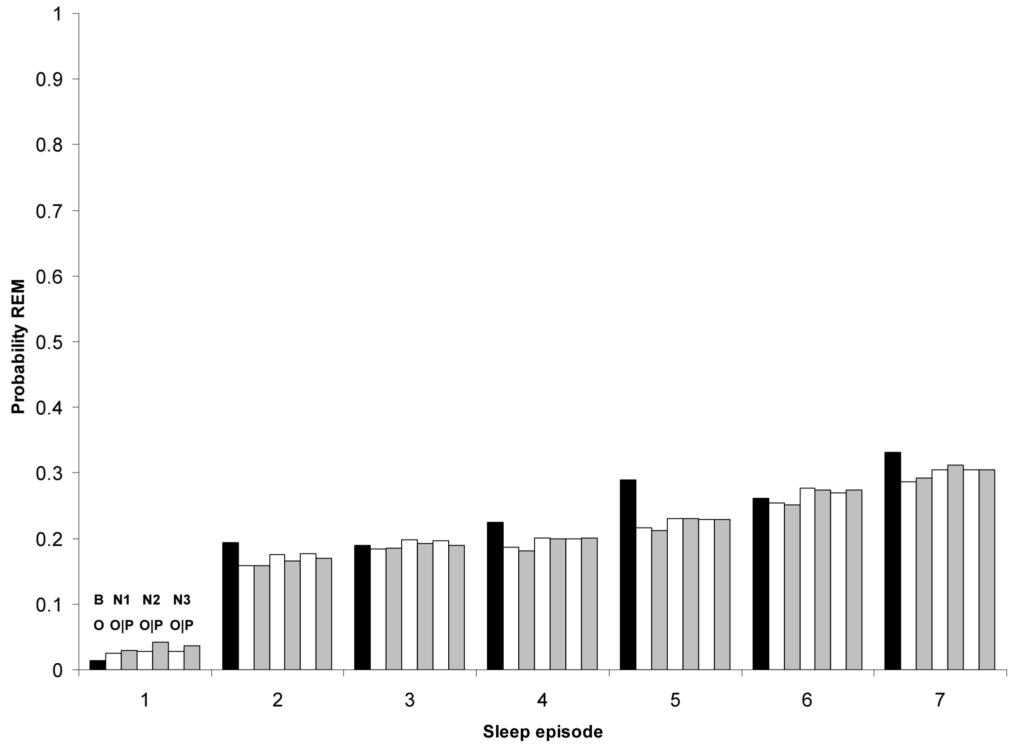


Figure 9.24: Comparison of predicted probabilities (P) for noise epochs N1 to N3 (gray) with observed (O) probabilities (white) and baseline (B) data (black) for stage REM.

10 Appendix B: Regression results

The autoregressive multinomial logistic regression models were calculated with the CATMOD procedure of the SAS System, Version 8.2. The purpose of the models was to predict the next sleep stage (i.e. the sleep stage at T1). For the regression, the explanatory variable "present sleep stage" (i.e. at T0) was converted into five indicator variables S1, S2, S3, S4 and REM. Stage Wake served as the reference category (reference coding).

The continuous explanatory variable "elapsed sleep time since sleep onset" was termed "Transition" (see also Table 5.2). Altogether, 820 transitions between sleep stages are needed to describe a night with 821 epochs or 410.5 min.

Notice that there are five parameter estimates (termed "function number" in the regression outputs) for each independent variable, as there should be for this model with a six-category outcome variable. The function numbers are referenced as follows: 1 = Wake, 2 = S1, 3 = S3, 4 = S4 and 5 = REM. Here, sleep stage S2 served as the reference category, as S2 represents the largest group and changes in transition probabilities to S2 were expected to be small under the influence of noise.

The probability for sleep stage S1 at T1 is calculated as follows (LP = linear predictor, numbers refer to function numbers of the regression outputs):

$$p(S1) = \frac{e^{LP2}}{1 + e^{LP1} + e^{LP2} + e^{LP3} + e^{LP4} + e^{LP5}}$$

Please notice that function number 2 refers to sleep stage S1, and not to S2 (see above). For the calculation of the probabilities of the other sleep stages, LP2 in the exponent of the numerator has to be exchanged by the respective linear predictor (LP1 for Wake, LP3 for S3, LP4 for S4 and LP5 for

REM). The probability for the reference sleep stage S2 at T1 is calculated as follows:

$$p(S2) = \frac{1}{1 + e^{LP1} + e^{LP2} + e^{LP3} + e^{LP4} + e^{LP5}}$$

10.1 Baseline model

The CATMOD Procedure

Data Summary

Response	Y	Response Levels	6
Weight Variable	None	Populations	4589
Data Set	ARNEU	Total Frequency	102500
Frequency Missing	0	Observations	102500

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	5	2651.82	<.0001
S1	4*	.	.
S2	5	16807.26	<.0001
S3	5	2746.09	<.0001
S4	5	1072.97	<.0001
REM	4*	.	.
Transition	5	641.50	<.0001
Likelihood Ratio	2E4	15064.00	1.0000

NOTE: Effects marked with '*' contain one or more redundant or restricted parameters.

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	1.2144	0.0426	811.67	<.0001
	2	-0.4702	0.0651	52.21	<.0001
	3	-3.6542	0.2217	271.69	<.0001
	4	-6.2984	1.0007	39.61	<.0001
	5	-1.3717	0.0661	430.22	<.0001
S1	1	-2.7472	0.0813	1143.24	<.0001
	2	0.2838	0.0687	17.06	<.0001
	3	-2.5433	1.0245	6.16	0.0130
	4	-6.9807#	.	.	.
	5	-1.3017	0.1245	109.37	<.0001
S2	1	-4.8576	0.0391	15398.33	<.0001
	2	-4.6860	0.0773	3678.44	<.0001
	3	0.8986	0.2212	16.50	<.0001
	4	0.2586	1.0108	0.07	0.7981
	5	-3.0316	0.0608	2488.73	<.0001
S3	1	-3.4514	0.0839	1691.85	<.0001
	2	-6.7253	1.0014	45.11	<.0001
	3	5.7615	0.2218	675.03	<.0001
	4	6.5807	1.0007	43.25	<.0001
	5	-3.8353	0.3069	156.18	<.0001
S4	1	-0.7784	0.1621	23.05	<.0001
	2	-3.5858	1.0091	12.63	0.0004
	3	6.5302	0.2567	647.28	<.0001
	4	11.5460	1.0079	131.23	<.0001
	5	-3.0085	1.0094	8.88	0.0029
REM	1	-1.0655	0.0621	294.72	<.0001
	2	-1.2599	0.1222	106.28	<.0001
	3	-2.0445	1.0248	3.98	0.0460
	4	-6.1652#	.	.	.
	5	4.5398	0.0658	4765.56	<.0001
Transition	1	0.000452	0.000069	42.46	<.0001
	2	-0.00026	0.000112	5.46	0.0194
	3	-0.00147	0.000079	344.25	<.0001
	4	-0.00273	0.000151	326.71	<.0001
	5	0.000869	0.000090	92.57	<.0001

NOTE: Parameters marked with '#' are regarded to be infinite.

10.2 Model for noise epoch #1

The CATMOD Procedure
Data Summary

Response	Y	Response Levels	6
Weight Variable	None	Populations	3877
Data Set	ARNEU	Total Frequency	26135
Frequency Missing	0	Observations	26135

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	5	1774.81	<.0001
S1	5	353.47	<.0001
S2	5	2143.02	<.0001
S3	5	844.38	<.0001
S4	4*	.	.
REM	3*	.	.
Transition	5	122.18	<.0001
Likelihood Ratio	2E4	10616.73	1.0000

NOTE: Effects marked with '*' contain one or more redundant or restricted parameters.

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	2.3013	0.0836	757.42	<.0001
	2	-0.4125	0.1465	7.93	0.0049
	3	-4.0295	0.7128	31.96	<.0001
	4	-12.6277	1.0084	156.81	<.0001
	5	-1.0815	0.1513	51.07	<.0001
S1	1	-1.8124	0.1281	200.15	<.0001
	2	0.4352	0.1654	6.93	0.0085
	3	-0.4452	1.2295	0.13	0.7173
	4	-1.1678	70.5303	0.00	0.9868
	5	-1.5643	0.3179	24.22	<.0001
S2	1	-3.5524	0.0781	2068.89	<.0001
	2	-3.3554	0.1374	596.71	<.0001
	3	1.6156	0.7121	5.15	0.0233
	4	4.8135	1.2269	15.39	<.0001
	5	-3.3679	0.1592	447.68	<.0001
S3	1	-2.4651	0.1078	523.02	<.0001
	2	-3.6566	0.4286	72.78	<.0001
	3	5.9772	0.7130	70.28	<.0001
	4	12.6037	1.0067	156.74	<.0001
	5	-4.9858	1.0103	24.35	<.0001
S4	1	-0.4093	0.2473	2.74	0.0979
	2	-2.5576	1.0310	6.15	0.0131
	3	6.5707	0.7472	77.32	<.0001
	4	17.2381	1.0272	281.64	<.0001
	5	-10.3476#	.	.	.
REM	1	-0.9722	0.1072	82.22	<.0001
	2	-0.3825	0.1730	4.89	0.0270
	3	-9.1235#	.	.	.
	4	8.0627#	.	.	.
	5	3.9654	0.1498	700.30	<.0001
Transition	1	-0.00025	0.000082	9.12	0.0025
	2	-0.00030	0.000198	2.32	0.1273
	3	-0.00135	0.000141	91.78	<.0001
	4	-0.00187	0.000273	46.69	<.0001
	5	0.000337	0.000153	4.87	0.0273

NOTE: Parameters marked with '#' are regarded to be infinite.

10.3 Model for noise epoch #2

The CATMOD Procedure
Data Summary

Response	Y	Response Levels	6
Weight Variable	None	Populations	3902
Data Set	ARNEU	Total Frequency	26098
Frequency Missing	0	Observations	26098

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	5	1970.36	<.0001
S1	4*	.	.
S2	4*	.	.
S3	5	1868.85	<.0001
S4	4*	.	.
REM	4*	.	.
Transition	5	109.60	<.0001
Likelihood Ratio	2E4	10583.37	1.0000

NOTE: Effects marked with '*' contain one or more redundant or restricted parameters.

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	0.7674	0.0534	206.77	<.0001
	2	-0.4415	0.0733	36.29	<.0001
	3	-3.8388	0.2373	261.80	<.0001
	4	-14.1409	0.5121	762.60	<.0001
	5	-1.6914	0.0883	366.89	<.0001
S1	1	-2.9919	0.1979	228.47	<.0001
	2	0.0584	0.1059	0.30	0.5809
	3	-9.0011#	.	.	.
	4	-1.6087	128.5	0.00	0.9900
	5	-0.0583	0.1470	0.16	0.6913
S2	1	-3.8725	0.0589	4317.59	<.0001
	2	-4.2785	0.1190	1293.32	<.0001
	3	0.8650	0.2382	13.18	0.0003
	4	6.9155#	.	.	.
	5	-2.1935	0.0834	691.97	<.0001
S3	1	-2.9214	0.1170	623.89	<.0001
	2	-4.7870	0.5034	90.44	<.0001
	3	4.9008	0.2354	433.54	<.0001
	4	12.5687	0.5114	603.91	<.0001
	5	-5.2357	1.0025	27.28	<.0001
S4	1	-1.0691	0.2088	26.22	<.0001
	2	-3.5520	1.0101	12.37	0.0004
	3	5.6143	0.2751	416.62	<.0001
	4	17.5938	0.5196	1146.71	<.0001
	5	-10.8294#	.	.	.
REM	1	-0.8380	0.1179	50.52	<.0001
	2	-1.2592	0.2092	36.24	<.0001
	3	-8.4853#	.	.	.
	4	-1.0821	132.7	0.00	0.9935
	5	4.6170	0.0986	2191.44	<.0001
Transition	1	-0.00004	0.000102	0.16	0.6850
	2	-0.00013	0.000143	0.84	0.3585
	3	-0.00125	0.000155	64.90	<.0001
	4	-0.00150	0.000322	21.60	<.0001
	5	0.000822	0.000145	32.02	<.0001

NOTE: Parameters marked with '#' are regarded to be infinite.

10.4 Model for noise epoch #3

The CATMOD Procedure
Data Summary

Response	Y	Response Levels	6
Weight Variable	None	Populations	4053
Data Set	ARNEU	Total Frequency	26066
Frequency Missing	0	Observations	26066

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	5	1063.35	<.0001
S1	4*	.	.
S2	5	5230.04	<.0001
S3	5	816.33	<.0001
S4	3*	.	.
REM	3*	.	.
Transition	5	181.18	<.0001
Likelihood Ratio	2E4	10029.95	1.0000

NOTE: Effects marked with '*' contain one or more redundant or restricted parameters.

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	0.9691	0.0637	231.56	<.0001
	2	-0.3739	0.0831	20.25	<.0001
	3	-3.8809	0.3592	116.75	<.0001
	4	-5.0150	0.7138	49.36	<.0001
	5	-2.2264	0.1232	326.81	<.0001
S1	1	-2.8770	0.1092	693.54	<.0001
	2	-0.3877	0.0867	20.02	<.0001
	3	-1.9712	1.0622	3.44	0.0635
	4	-8.7818#	.	.	.
	5	-1.2233	0.2092	34.21	<.0001
S2	1	-4.6710	0.0700	4453.33	<.0001
	2	-4.7750	0.1293	1364.17	<.0001
	3	1.2007	0.3594	11.16	0.0008
	4	-3.3496	1.2255	7.47	0.0063
	5	-1.9309	0.1176	269.60	<.0001
S3	1	-3.2425	0.1651	385.89	<.0001
	2	-4.9466	0.7112	48.37	<.0001
	3	5.8730	0.3604	265.51	<.0001
	4	4.9879	0.7146	48.72	<.0001
	5	-2.6269	0.5126	26.26	<.0001
S4	1	-0.5785	0.4239	1.86	0.1724
	2	-9.9189#	.	.	.
	3	6.9644	0.4964	196.86	<.0001
	4	10.6345	0.7835	184.22	<.0001
	5	-8.2248#	.	.	.
REM	1	-1.0694	0.1078	98.34	<.0001
	2	-1.5366	0.1965	61.17	<.0001
	3	-8.2782#	.	.	.
	4	-6.7936#	.	.	.
	5	4.9946	0.1235	1635.28	<.0001
Transition	1	0.000401	0.000118	11.56	0.0007
	2	0.000277	0.000151	3.34	0.0677
	3	-0.00190	0.000172	122.64	<.0001
	4	-0.00285	0.000365	60.92	<.0001
	5	0.000896	0.000164	29.70	<.0001

NOTE: Parameters marked with '#' are regarded to be infinite.

11 Acknowledgements

I would like to especially thank Uwe Siebert, Professor of Public Health, University of Health Sciences, Medical Informatics and Technology (UMIT), Hall in Tirol, Austria, for his tremendous support of my work. Without the time he invested, his open-mindedness and his impressive and broad methodological background, it would have been hard to accomplish what has been done in this master thesis.

I would also like to thank Scott Gazelle, Director of the Institute for Technology Assessment, Massachusetts General Hospital (MGH), Harvard Medical School, Boston, U.S.A. for having me around at his Institute in April 2004 and for providing me with a two year appointment to MGH professional staff.

Thanks to Alexander Samel, Head of the Flight Physiology Division at the DLR-Institute of Aerospace Medicine in Cologne, for his constant support of my career and for his valued advice.

Thanks also to DLR for supporting my participation in the "Master of Science in Epidemiology" program at the University of Bielefeld.

Thanks to all my colleagues at the Flight Physiology Division of the DLR-Institute of Aerospace Medicine, who worked nights and weekends for sampling the data the analyses presented in this thesis are based on.

I would like to thank my parents for their continuous support. Finally, I would like to thank my wife and my in-laws for looking after Bruno and thus providing me with the freedom and time to write this thesis.

12 Abbreviations

Abbreviation	Meaning
ANE	Aircraft Noise Event
EEG	Electroencephalogram/Electroencephalography
EMG	Electromyogram
EOG	Electrooculogram
NREM sleep	all sleep stages except REM, i.e. S1, S2, S3, and S4
OSAS	Obstructive Sleep Apnea Syndrome
p(y|x)	transition probability from stage x to stage y
REM sleep	Rapid Eye Movement sleep
S1, S2, S3, S4	sleep stages 1 to 4
SEI	Sleep Efficiency Index, defined as the time spent in all sleep stages but Wake divided by sleep period time (sometimes also based on TIB)
SPL	Sound Pressure Level
SPT	Sleep Period Time, i.e. the time from sleep onset, defined as the first appearance of sleep stage 2, until the last appearance of any sleep stage but Wake
SQI	Sleep Quality Index, as defined in this thesis (see Chapter 4)
SWS	Slow Wave Sleep, i.e. sleep stages S3 and S4
T0	current sleep stage
T-1	sleep stage preceding the current sleep stage
T-2, T-3, etc.	sleep stage preceding T-1, T-2, etc.
T1	sleep stage following the current sleep stage (that is to be predicted)
TIB	Time in Bed, i.e. time from lights out to lights on
TST	Total Sleep Time, defined as SPT minus the time spent awake

13 References

[1] Ergebnisse der Arbeiten Ö2/Ö3 - Stand der Erkenntnisse von Fluglärmwirkungen - Beschluss der Mediationsgruppe vom 26. November 1999

[2] Bericht Mediation Flughafen Frankfurt/Main. 2000

[3] Ausbau Flughafen Frankfurt Main - Unterlagen zum Planfeststellungsverfahren - Antragsbegründung vom 02.11.2004. Fraport AG, Band A, Antragsteil A2, 2004

[4] Babisch, W: Traffic noise and cardiovascular disease: epidemiological review and synthesis. Noise and Health 2(8): 9-32, 2000

[5] Babisch, W., Beule, B., Schust, M., Kersten, N., and Ising, H.: Traffic noise and risk of myocardial infarction. Epidemiology 16(1): 33-40, 2005

[6] Basner, M., Buess, H., Elmenhorst, D., Gerlich, A., Luks, N., Maaß, H., Mawet, L., Müller, E. W., Müller, U., Plath, G, Quehl, J., Samel, A., Schulze, M, Vejvoda, M., and Wenzel, J.: Effects of nocturnal aircraft noise (Volume 1): Executive summary. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Cologne, Germany, FB2004-07/E, ISSN 1434-8454, 2004

[7] Basner, M., Isermann, U. and Samel, A.: Die Umsetzung der DLR-Studie in einer lärmmedizinischen Beurteilung für ein Nachtschutzkonzept. Zeitschrift für Lärmbekämpfung 52(4): 109-123, 2005

[8] Basner, M., Isermann, U. and Samel, A.: Aircraft noise effects on sleep: Application of the results of a large polysomnographic field study. J.Acoust.Soc.Am. 119(5): 2772-2784, 2006

[9] Basner, M. and Samel, A.: Effects of nocturnal aircraft noise on sleep structure. Somnologie 9(2): 84-95, 2005

[10] Beersma, D. G., Daan, S., and Van den Hoofdakker, R. H.: Distribution of REM latencies and other sleep phenomena in depression as explained by a single ultradian rhythm disturbance. Sleep 7(2): 126-136, 1984

[11] Berger, M: Lehrbuch des normalen und gestörten Schlafes, 1992

[12] Bonnet, M., Carley, D. W., Carskadon, M. A., Easton, P., Guilleminault, C., Harper, R., Hayes, B., Hirshkowitz, M., Ktonas, P. Y., Keenan, S., Pressman, M., Roehrs, T., Smith, J., Walsh, J. K., Weber, S., and Westbrook, P. R.: EEG arousals: Scoring rules and examples. A preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep 15(2): 173-184, 1992

[13] Bonnet, M. H.: Performance and sleepiness following moderate sleep disruption and slow wave sleep deprivation. Physiol.Behav. 37(6): 915-918, 1986

[14] Bonnet, M. H.: Differentiating sleep continuity effects from sleep stage effects (Letter to the editor). J.Sleep Res. 9: 403-406, 2001

[15] Bonney, G. E.: Logistic regression for dependent binary observations. Biometrics 43(4): 951-973, 1987

[16] Borbely, A. A.: A two process model of sleep regulation. Hum.Neurobiol. 1(3): 195-204, 1982

[17] Bowe, T. R. and Anders, T. F.: The use of the semi-Markov model in the study of the development of sleep-wake states in infants. Psychophysiology 16(1): 41-48, 1979

[18] Brooks, D., Horner, R. L., Kozar, L. F., Render-Teixeira, C. L., and Phillipson, E. A.: Obstructive sleep apnea as a cause of systemic hypertension. Evidence from a canine model. J.Clin.Invest 99(1): 106-109, 1997

[19] Daan, S., Beersma, D. G., and Borbely, A. A.: Timing of human sleep: recovery process gated by a circadian pacemaker. Am.J.Physiol 246(2 Pt 2): R161-R183, 1984

[20] de Vries, S. O., Fidler, V., Kuipers, W. D., and Hunink, M. G.: Fitting multistate transition models with autoregressive logistic regression: supervised exercise in intermittent claudication. Med.Decis.Making 18(1): 52-60, 1998

[21] de Witt, S. and Gärtner, K.: Das Schutzkonzept gegen nächtlichen Fluglärm aufgrund der DLR-Studie in der Planfeststellung von Flughäfen und Flugplätzen. Zeitschrift für Lärmbekämpfung 52(4): 124-130, 2005

[22] Drosopoulos, S., Wagner, U., and Born, J.: Sleep enhances explicit recollection in recognition memory. Learn.Mem. 12(1): 44-51, 2005

[23] Fränkle, A. and Flunkert, J.: Praxisorientiertes Umsetzungskonzept zur Verlagerung der Flugbewegungen in den Zeiten des geplanten Nachtflugverbotes am Flughafen Frankfurt. 2004

[24] Gais, S. and Born, J.: Declarative memory consolidation: Mechanisms acting during human sleep. Learn.Mem. 11(6): 679-685, 2004

[25] Griefahn, B.: Präventivmedizinische Vorschläge für den nächtlichen Schallschutz. Zeitschrift für Lärmbekämpfung 37: 7-14, 1990

[26] Griefahn, B.: Gestörte Schlafprofile. In: Bartels, K. H. and Ising, H.: Nachtfluglärmproblematik. 1, Eigenverlag Verein WaBoLu, Berlin, pp 70-74, 2001

[27] Gronefeld, V.: Rechtliche Stellungnahme über die Voraussetzung und die Zulässigkeit der Anordnung eines Nachtflugverbotes nach Maßgabe der Empfehlungen der Mediationsgruppe für die Zukunft des Verkehrsflughafens Frankfurt am Main. 2001

[28] Hale, C. S.: Obstructive sleep apnea and cardiovascular disease and mortality: the argument for causality. J Insur.Med. 37(4): 272-282, 2005

[29] Hobson, J. A., McCarley, R. W., and Wyzinski, P. W.: Sleep cycle oscillation: reciprocal discharge by two brainstem neuronal groups. Science 189(4196): 55-58, 4-6-1975

[30] Hosmer, D. H. and Lemeshow, S.: Applied Logistic Regression, 2nd, Wiley, New York, 2000

[31] Isermann, U., Kowalski, T., and Schmid, R.: Flugbetriebsbedingte und sonstige Geräuschimmissionen ausgehend vom Gelände des Flughafens - Teil A Fluglärm. Band C, Gutachten G10.1, 2004

[32] Jünemann, R.: Verkehrliche Auswirkungen und mögliche Maßnahmen bei Einführung eines Nachtflugverbotes am Flughafen Frankfurt/Main. 2001

[33] Karlsson, M. O., Schoemaker, R. C., Kemp, B., Cohen, A. F., van Gerven, J. M., Tuk, B., Peck, C. C., and Danhof, M.: A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. Clin.Pharmacol.Ther. 68(2): 175-188, 2000

[34] Kemp, B. and Kamphuisen, H. A.: Simulation of human hypnograms using a Markov chain model. Sleep 9(3): 405-414, 1986

[35] Lawder, R. E.: A proposed mathematical model for sleep patterning. J.Biomed.Eng 6(1): 63-69, 1984

[36] McCarley, R. W. and Hobson, J. A.: Neuronal excitability modulation over the sleep cycle: a structural and mathematical model. Science 189(4196): 58-60, 1975

[37] Meier, U.: Das Schlafverhalten der deutschen Bevölkerung - eine repräsentative Studie. Somnologie 8(3): 87-94, 2004

[38] Morrell, S., Taylor, R., and Lyle, D.: A review of health effects of aircraft noise. Australian and New Zealand Journal of Public Health 21(2): 221-236, 1997

[39] Ohayon, M. M. and Zulley, J.: Correlates of global sleep dissatisfaction in the German population. Sleep 24(7): 780-787, 2001

[40] Pearsons, K., Barber, D., Tabachnick, B. G., and Fidell, S.: Predicting noise-induced sleep disturbance. J.Acoust.Soc.Am. 97(1): 331-338, 1995

[41] Plihal, W and Born, J.: Effects of early and late nocturnal sleep on declarative and procedural memory. J.Cognitive Neurosci. 9(4): 534-547, 1997

[42] Rechtschaffen, A., Kales, A., Berger, R. J., Dement, W. C., Jacobsen, A., Johnson, L. C., Jouvet, M., Monroe, L. J., Oswald, I., Roffwarg, H. P., Roth, B., and Walter, R. D.: A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Public Health Service, U.S. Government, Printing Office, Washington, D.C., 1968

[43] Reppert, S. M. and Weaver, D. R.: Coordination of circadian timing in mammals. Nature 418(6901): 935-941, 2002

[44] Riemann, D.: Schlaf und Gedächtnis. In: Schulz, H.: Kompendium Schlafmedizin. 7. Ergänzungslieferung 02/2004, ecomed, Landsberg/Lech, pp 1-2, 1997

[45] Samel, A. and Basner, M.: Extrinsische Schlafstörungen und Lärmwirkungen. Somnologie 9(2): 58-67, 2005

[46] Ursin, R., Moses, J., Naitoh, P., and Johnson, L. C.: REM-NREM cycle in the cat may be sleep-dependent. Sleep 6(1): 1-9, 1983

[47] Wesensten, N. J., Balkin, T. J., and Belenky, G.: Does sleep fragmentation impact recuperation? A review and reanalysis. J.Sleep Res. 8(4): 237-245, 1999

[48] Williams, H. L., Hammack, J. T., Daly, R. L., Dement, W. C., and Lubin, A.: Responses to auditory stimulation, sleep loss and the EEG stages of sleep. Electroencephalography and Clinical Neurophysiology 16: 269-279, 1964

[49] Willich, S. N., Wegscheider, K., Stallmann, M., and Keil, T.: Noise burden and the risk of myocardial infarction. Eur.Heart J. 2005

[50] Winfree, A. T.: Human body clocks and the timing of sleep. Nature 297(5861): 23-27, 1982

[51] Yaggi, H. K., Concato, J., Kernan, W. N., Lichtman, J. H., Brass, L. M., and Mohsenin, V.: Obstructive sleep apnea as a risk factor for stroke and death. N.Engl.J Med. 353(19): 2034-2041, 2005

[52] Yang, M. C. and Hursch, C. J.: The use of a semi-Markov model for describing sleep patterns. Biometrics 29(4): 667-676, 1973

[53] Zung, W. W., Naylor, T. H., Gianturco, D. T., and Wilson, W. P.: Computer simulation of sleep EEG patterns with a Markov chain model. Recent Adv.Biol.Psychiatry 8: 335-355, 1965