Index

Introduction

Among neurological diseases, sleep disorders are those commonly considered to be less serious and disabling. For this widespread belief, often people are poorly inclined to report their sleeping problems to their personal health-care provider. However these disturbances affect a large segment of world's population. Their prevalence has been estimated to be approximately 10%.

The causes of sleep problems are different and often are consequences of other clinical conditions. Whatever the causes are, it has been demonstrated that sleep deprivation affects the overall life quality, resulting in impaired memory and cognitive functions and compromising both productivity and wellness. Sleep disturbances constitute therefore a substantial socio-economic burden and, for this reason, the appropriate diagnosis and treatment of these pathologies represent great challenges for clinicians and pharmaceutical companies. The latter have made huge investments in this research field, trying to develop safer and more effective drugs (hypnotics) able to regulate the sleep-wake alternation. Sleep is not a homogeneous state of unconsciousness, but it is characterized by an internal structure, called '*sleep architecture*', described by different sleep stages (awake, stage 1, stage 2, deep sleep, REM) and transitions among them. It has been demonstrated that sleep architecture follows specific patterns during the night and that the maintenance of such structure is important to guarantee a restorative sleep.

Sleep stages are commonly assessed through '*polysomnography*', a multi-channel diagnostic technique which consists in the simultaneous recording of electroencephalogram, electrooculogram, electromyogram and other relevant features. On the basis of the characteristics of these recordings, international regulatory authorities provided indications on the endpoints to be evaluated for assessing the severity of the sleep conditions and the effect of drugs for the relief of these conditions. In this context, aggregated variables characterizing each patient's sleep, such as the *'Wake After Sleep Onset*', (WASO, the amount of time during the night spent in the awake state after being fallen asleep), or the '*Latency to Persistent Sleep*', (LPS, the time spent in the awake state before falling asleep), are typically used. In the clinical practice, sleep disorders and, in turn, the effect of hypnotics are described on the basis of these aggregated parameters which, anyway, reflect only the overall trend of sleep during the night. Transforming the PSG data into these aggregate parameters implies however a substantial loss of information regarding the sleep architecture. On the contrary, the recent research has demonstrated that the maintenance of a physiological sleep architecture is just as important as the total sleeping time, so that this is an important differentiation criteria for new drugs used for the treatment of sleep disturbances.

In the effort of better understanding the sleep-wake cycle, mathematical models should be considered as a natural way to properly and quantitatively integrate complex information obtained from brain and sleep studies. In particular, it would be suitable to describe the sleep architecture through a mathematical model, to correctly describe and/or predict the sleep stages time-course along the night. In this respect, the polysomnograghic data can be described using a finite discrete stochastic process assuming values in a finite discrete set (the sleep stages). One of the most interesting approaches proposed in the literature to describe such data [1] is the '*Markov-chain model*', which assumes the sequence of sleep stages to be governed by the following property: the probability of transition from state *m* to state *k* at a certain time of the night depends only on the state of departure *m* and not on the entire past history of the states.

Recently, this model has been further developed including multinomial random variables for simultaneously characterizing all the possible transitions from a specific stage at a certain time. Multinomial logistic functions have been then introduced for describing the corresponding transition probabilities [2].

The aim of this thesis is to evaluate whether the multinomial model proposed in [2] adequately describes the underlying system. Data from a clinical study conducted in patients with a diagnosis of primary insomnia are used for this analysis.

Different diagnostic methods to investigate the adequacy and the performance of a categorical model are proposed and implemented in this thesis. I mainly focus my attention on those methods based on stochastic simulations considered more appropriate for categorical type of data as suggested in several papers in the literature [3, 4 ,5]. In addition, the model assessment includes also the evaluation of the estimation method adopted to identify multinomial model parameters. For this reason, I introduce in this thesis a diagnostic method derived by a recently proposed [6,7] approach for checking both model and estimation method performance. The key contribution of this work is to translate complex diagnostic methods based on simulations and statistics of re-estimated parameters in a visual evaluation of the relevant model characteristics: in particular, the transition probabilities between sleep stages. This kind of visual comparison – applied for the first time in categorical models - has been called '*Visual Estimation Check'* (VEC).

This thesis is structured as follows: Chapter 1 provides some background on sleep physiology and its regulation mechanisms, including an overview on the instrumentation and the guidelines for assessing sleep stages. Chapter 2 describes different estimation methods, with particular emphasis on the population approach, adopted in our analyses, and its application to categorical data. Chapter 3 presents the Markov-chain model used for describing sleep architecture, including some details of the basic theory. Results are presented in Chapter 4 where the validation of the model and the outcome of the different diagnostic methods applied in this work are described. Finally, a critical discussion, also on possible further developments of this research project, is

provided in the Conclusions. Code implementing models using a non-linear mixed effect (NONMEM) approach and methods for validation analysis (in R) are reported in Appendix.

Chapter 1

Sleep: physiology fundamentals, treatment of insomnia and quantitative methods for its assessment

Sleep is a complex, highly organized physiological process that is fundamental to life. It is a natural state of bodily rest characterized by total or partial unconsciousness and the inactivity of nearly all voluntary muscles. Both sensory and motor activities are suspended implying reduced response to stimuli and minimal movement. However, the responsiveness to endogenous and exogenous stimuli is not completely absent during sleep and the condition of unconsciousness is reversible.

Sleep is not a homogeneous state of unconsciousness but it is characterized by an internal structure, called '*sleep architecture*', defined by different stages and the transition between them. It has been demonstrated that the maintenance of such architecture is fundamental to determine sleep quality and, therefore, the physical and mental well-being.

On the other hand, disorders in the natural pattern of sleep may lead to adverse consequences and may seriously affect patients' health, productivity and life quality [8,9,10].

The prevalence of sleep disturbances indicates that it is a very common problem affecting both men and women, elderly and young population. The causes of sleep problems are different and often are related to other clinical pathologies. For these reasons, the appropriate diagnosis and the treatment of sleep disorders is becoming more and more relevant.

However, the purposes and mechanisms regulating sleep are not completely clear and the deep understanding of sleep architecture and patterns represents a great challenge to clinicians and pharmaceutical companies, that are leading an intense research in this field.

The first attempt to describe human sleep was made in 1930 by Berger, the father of electroencephalography [11]. He obtained the first sleep recording and noted that the alpha rhythm disappeared when his subject fell asleep. A second important achievement was made in 1937 when Loomis et al. [12] published the first continuous overnight EEG sleep recording in humans and proposed a scheme, the so called 'sleep staging', to summarize the EEG recording in

a reduced dataset. They proposed a classification of the EEG activity recorded during sleep into 5 stages: A, B, C, D and E, on the basis of the predominant EEG rhythm in a fixed time domain.

In 1953 Aserinsky and Kleitmain [13] discovered episodic electro-oculagraphic (EOG) activity occurring during sleep stage B every 90-120 minutes. Initially this activity was supposed to be an artifact due to instrumentation, but subsequent studies demonstrated that these episodes were actually occurring. These events were called 'Rapid Eye Movements' (REMs).

The authors tried also to establish the relation existing between REM stage and dreaming. It was noted that dreaming happened in 20 of the 27 instances after the awakening from REM sleep stage. In 1957 Dement and Kleitman [14] suggested a new classification for sleep stages: sleep stages were divided into four Non-REM (NREM) stages and a REM stage.

The next major improvement on sleep architecture description was made in 1959 when Jouvet [15] observed by the electromyography (EMG) technique muscular atonia related to the REM stage. He also introduced the concept that REM stage was a state in which the brain was 'active'.

The staging criteria were standardized in 1968, when Rechtschaffen and Kales developed and published '*A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*', [16], establishing the major rules for classifying sleep stages during a standardized sleep recording. This manual, generally reported as the '*R & K Manual',* received a general consensus and became a gold standard in sleep measurements. In the '*R & K Manual*', NREM sleep was divided into four stages (stages 1, 2, 3 , 4), with slow-wave sleep or deep sleep comprising stages 3 and 4 and, for contrast, light-sleep comprising stages 1 and 2. REM sleep was sometimes reported as stage 5.

In 2004, the American Academy of Sleep Medicine (AASM) proposed several changes in the scoring system indicated by the '*R & K standard'*, the most significant being the combination of stages 3 and 4 into a unique stage, called Stage N3. These proposed changes were published in 2007 in '*The AASM Manual for the Scoring of Sleep and Associated Events*' [17].

1.1. Sleep stages

Sleep stages and the other features related to sleep are commonly assessed by polysomnography in a specialized sleep laboratory. Polysomnography is a worldwide standardized procedure used for the diagnosis of sleep disorders. As a result of such a technique, it is possible (a) to recognize some important features/pathologies accompanying/affecting sleep (apneic episodes, restless leg syndrome, etc), (b) to exactly assess the sequence of sleep stages occurring during the night according to the '*Standard manual's rules*' and (c) to assess some important aggregated parameters that help in quantifying the severity of the pathology under examination.

1.1.1. Instrumentation: Polysomnography

Plysomnography, abbreviated PSG, is a multi-channel tool that allows to simultaneously record relevant activities occurring during sleep. On the basis of these recorded parameters, the rules established by the 'Standard manual' permit to assess for each time interval of the night the corresponding sleep stage.

A polysomnogram typically records a minimum of eleven channels requiring a minimum of 22 wires attached to the patient. A computer system is used for recording, storing and displaying the data obtained by each channel. During sleep the computer monitor can display multiple channels continuously. In addition, sleep laboratories have video camera installed in the patient's room in order to record also sleep-related behavioral features, as sleep-talking, snoring etc.

The recording must be carried out with the least discomfort for the patient, hence it must be conducted in a quiet room similar to comfortable bedroom. Recording tools should be physically separated from the patient and wires and electrodes must disturb the patient as less as possible. Light and noises should be appropriately shielded. After the application of the electrodes and the calibration of the system, the lights should be turned off approximately at patient's usual bed-time. It should be taken note of any episode of snoring, sleep-vocalization or other activities occurring during sleep. In a standardized clinical study aimed to investigate any potential drug effect, the registration of the PSG measurement occurs during a predefined time interval (from the "light off" to the "light on" time that typically lasts 8 hours).

Measurements include many recordings [18] listed below.

- *EEG: Electroencephalography of brain waves*

  The EEG remains the primary variable for staging the sleep. A single central channel is considered sufficient for basic sleep staging. However, it is recommended to provide redundancy to permit adequate assessments of other EEG characteristics. Consequently, six or more channels of cortical activity are typically used, with electrodes preferentially attached to the scalp near the frontal (Fp1, Fp2), central (C3, C4) and occipital (O3, O4) portions of the brain.

- *EOG: Electrooculography of eye movements*

  Electrooculography provides important information regarding sleep onset and REM recognition. At least two electrodes are recommended: one positioned one centimetre above the outer canthus of the right eye and the other placed one centimetre below the outer canthus of the left eye.

- *EMG: Electromyography of skeletal muscle activity*

  Submental EMG activity is used to determine the level of muscle tone, which significantly decreases during REM sleep. This channel is relevant to determine both sleep onset and REM occurrence. This channel also provides information regarding patient movements. A single channel is considered enough: electrodes are placed under the chin, in the submental region.

- *ECG: Electrocardiography of heart functions*

  Electrocardiography allows to control heart functions during sleep. In fact, it can assess the severity of some cardiorespiratory disfunctions, such as in sleep apnea. A single ECG channel is sufficient for PSG recording. Typically, two electrodes are placed one in the sternal area and the other at a lateral chest location.

- *Respiratory Effort and Airflow*

  The monitoring of respiration during PSG is necessary for the detection of apneas and hypopneas. It is important to record at least 2 key parameters: (a) air exchange through both the mouth and nose; (b) expansion and relaxation of the thorax and abdomen indicating respiratory effort.

  Air exchange can be measured by using different types of transducers. The measurement of the expansion and relaxation of thorax and abdomen may be accomplished by measuring intercostal EMG and thoracic/abdominal impedance or using strain gauge devices. Respiratory pauses are considered apneic episodes when persist at least 10 sec. The absence of respiratory effort and, consequently, of airflow indicates the presence of a 'central apnea', while the presence of respiratory effort without air exchange represents an 'obstructive apnea'. Oxygen saturation may drop significantly, usually in a direct relationship with the lengths of apneas.

- *Blood Oxygenation (P_{O_2} or O_2 saturation)*

  The measurement of blood oxygen in the PSG system facilitates detecting apneas episodes and other respiratory difficulties. Monitoring of blood oxygenation indicates the severity of breathing disfunctions. Pulse oxymetry, a noninvasive technique, is recommended. The sensors of this tool can be easily attached to the patient's fingertips or earlobs and can also be easily interfaced with polygraphic recorder. It is also possible to measure transcutaneous P_{O_2} using a Clark oxygen electrode attached directly to the skin.

- *Expired CO_2*

  Quantifying expired amount of CO_2 is another useful tool for monitoring respiratory functions. During expiration, the partial pressure of CO_2 approximates the intra-alveolar P_{CO_2} and this type of information can be considered as a marker of air exchange enabling the evaluation of some respiratory pathologies, as pulmonary disease, that modify the respiratory variables.

  A small diameter tube connecting the instrumentation to patient's mouth or nostrils is often used to measure air flow.

- *Body/Limb Movements*

  Some sleep disturbances are caused by the restless legs syndrome (RLS) or the periodic myoclonus in sleep (PMS). These pathologies imply movements compromising a good sleep. These movements are easily detectable in a sleep laboratory videotaping the patient or measuring muscular activity by means of EMG or accelerometers.

- *Behavioral Observation*

  An integral part of PSG procedure is the observation of patient's behavior, noting any vocalization or snoring during sleep. This is easily implemented through a system composed by a video-camera, a microphone and a monitor carefully observed by the assigned technician.

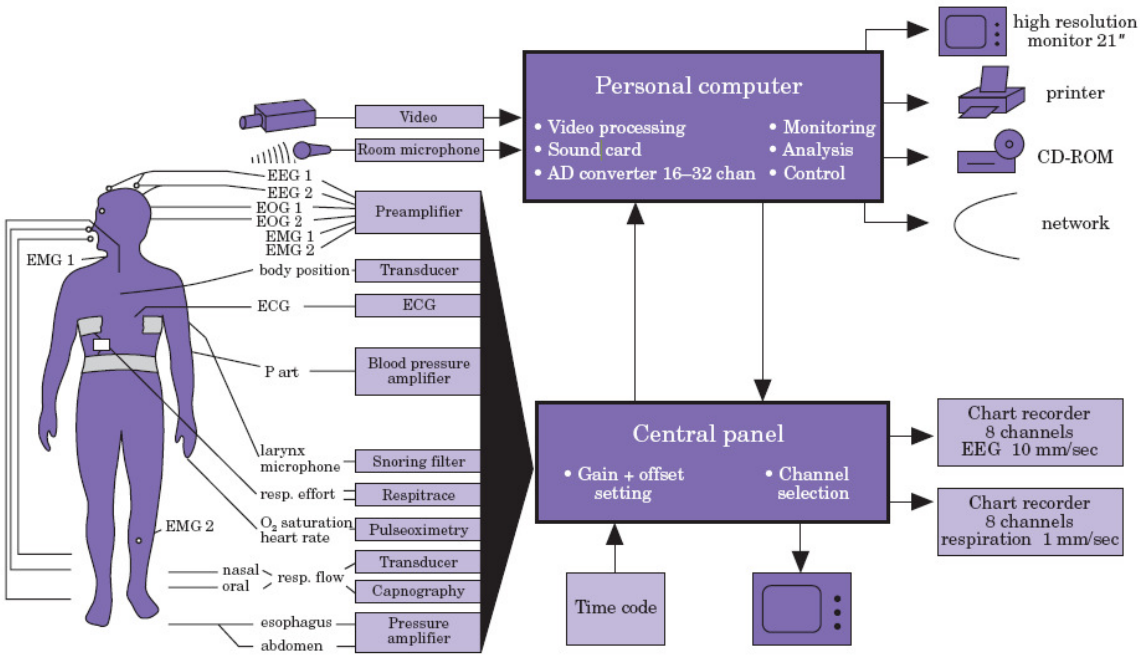


Figure 1.1: Typical polysomnographic installation.

In addition to recorded trace, the results of sleep monitoring procedure are typically reported with few additional information: (a) patient's information for his/her identification; (b) patient's history, including any existing pathology and relevant drugs assumed in the preceding 30 days; (c) recording conditions, including beginning and ending time of recording, list of recorded channels (each being labeled), and, when relevant, anatomic location of sensors; (d) a summary of polygraphic characteristics such as sleep statistics (aggregated parameters as sleep latency, total sleep time, etc), respiratory characteristics (presence/absence of snoring, number of apneic episodes, etc), heart rate values, movements (frequency of occurrence, etc), behavioral observations, EEG characteristics (basic characterization of the awake and sleep patterns and a description of any other feature).

The 'Standard manual' provides the rules for staging the sleep according to the obtained physiological parameters of interest. Night time, that is the recording time, is divided into 30-second intervals and the corresponding sleep stage is assessed for each interval, in an automatic computer-based way according to those rules.

1.1.2. Stage classification

The first step in the automatic sleep analysis is signals pre-processing. This step is necessary to reduce the enormous amount of raw data and remove artefacts affecting signals. This allows automatic statistical tools to analyze the data. As a result of this process, it is possible to clearly distinguish specific features necessary to discriminate sleep stages as follows:

- EEG: it is possible to recognize the amplitude and power of regular waves (beta, alpha, theta and delta waves) and specific patterns overlapped to the basic rhythm: K-complexes, sleep spindles, and vertex sharp waves.

  K-complexes are specific waveforms occurring during sleep and distinguishing a specific sleep stage (stage 2). They are characterized by a brief negative high-voltage peak, usually greater than 100 μV, followed by a slower positive complex at around 350 to 550 ms and a final negative peak at 900 ms. K-complexes occur spontaneously or as a response to exogenous or endogenous stimuli. A sleep spindle is a burst of brain activity visible on an EEG. It consists of 12-16 Hz waves that occur for 0.5 to 1.5 seconds. Vertex sharp waves are particular waveforms occurring during sleep but not characterizing any single stage. Their amplitude is 50-150 μV.

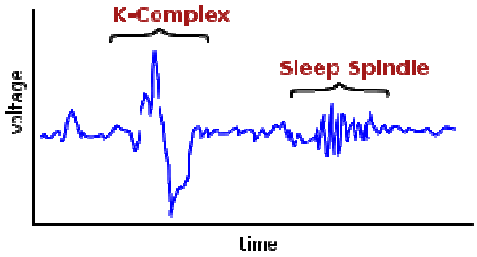


Figure 1.2: EEG specific feature that help in assessing sleep stages: K-complex and Sleep Spindle.

- EOG: the specific patterns are rapid eye movements and slow eye movements which have to be distinguished from each other.

The second step is the combination of these extracted features and waveforms in order to assess the correct sleep stage. The 'Standard manual' establishes that each 30-second interval should be categorized as AWAKE (AW) or sleep stage 1, 2, 3, 4 or REM, according to the following criteria.

- AWAKE (AW)
  (a) EEG: Wakefulness with eyes closed is characterized by an EEG rhythm predominantly in the alpha range (from 8 Hz to 12 Hz). Opening the eyes or engaging a significant mental task diminishes or blocks alpha activity. Sleep onset epoch is determined when alpha activity decreases to duration of less than 50% of an epoch or when a sleep spindle, K-complex, vertex wave or theta activity occurs; otherwise wakefulness is scored.
  (b) EMG: Moderately muscle activity can be present
  (c) EOG: Eye movements may occur, both rapid and slow

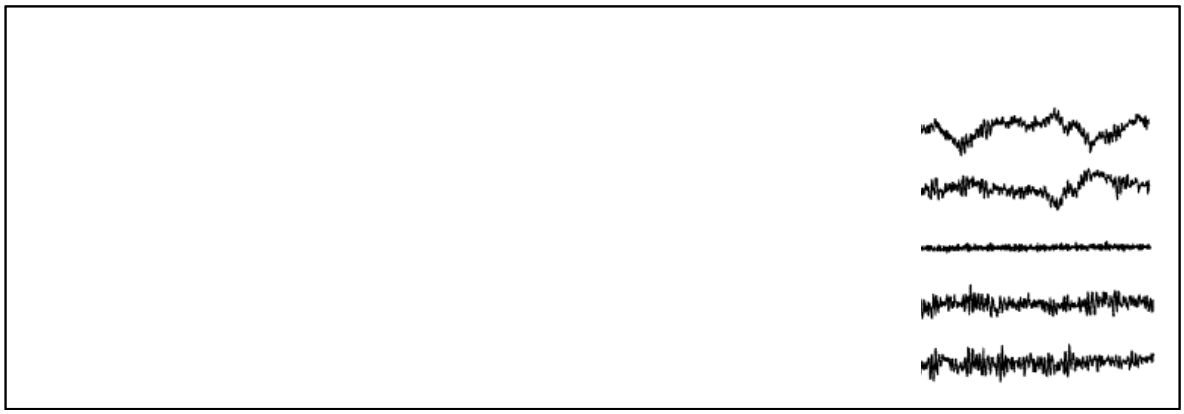


Figure 1.3: Typical EEG, EOG and EMG activities related to the AWAKE state.

- STAGE 1 (ST1)
  (a) EEG: Stage 1 marks the transition from alpha waves to theta waves (ranging from 4 Hz to 7 Hz). This sleep stage is transitional and does not last long. Often this state is not perceived as sleep when a person is asked; it is also referred to as 'somnolence or drowsy sleep'. There are no K-complexes or spindles.
  (b) EMG: Loss of some muscle tone.
  (c) EOG: Slow eye movements.

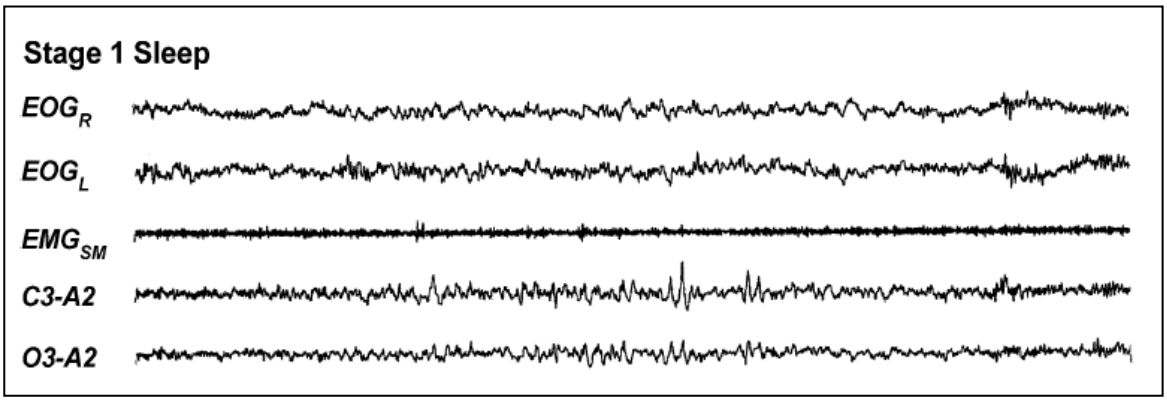


Figure 1.4: Typical EEG, EOG and EMG activities related to STAGE 1 sleep.

- STAGE 2 (ST2)
  (a) EEG: Stage 2 is characterized by the presence of sleep spindles and K-complexes overlapping on a theta range EEG rhythm. Both patterns are clearly recognizable from the background activity.
  (b) EMG: Muscle tone is slightly lower than during stage 1 sleep.
  (c) EOG: Usually there are no more eye movements.

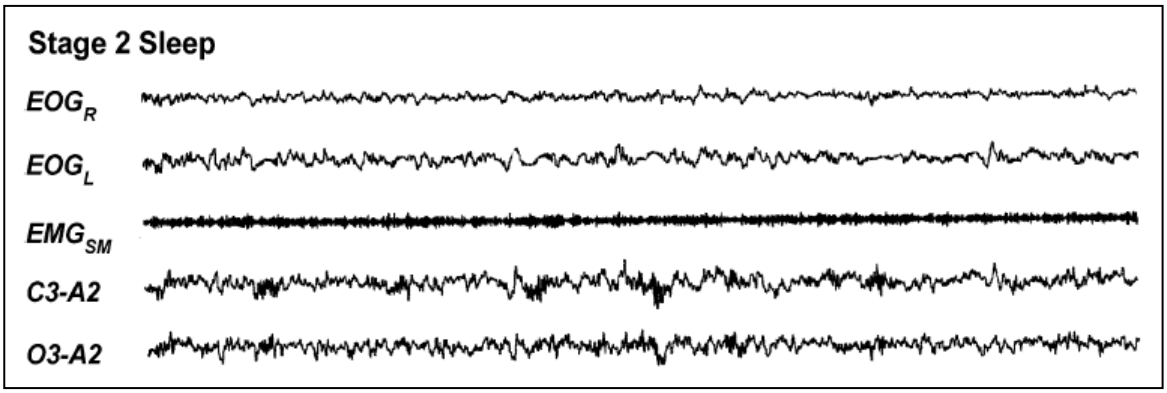


Figure 1.5: Typical EEG, EOG and EMG activities related to STAGE 2 sleep.

- STAGE 3 (ST3)
  (a) EEG: Stage 3 is designated when there is 20% to 50% of high amplitude (more than 75 μV) delta activity in an epoch.
  (b) EMG: Muscle tone is much lower.
  (c) EOG: No eye movements.

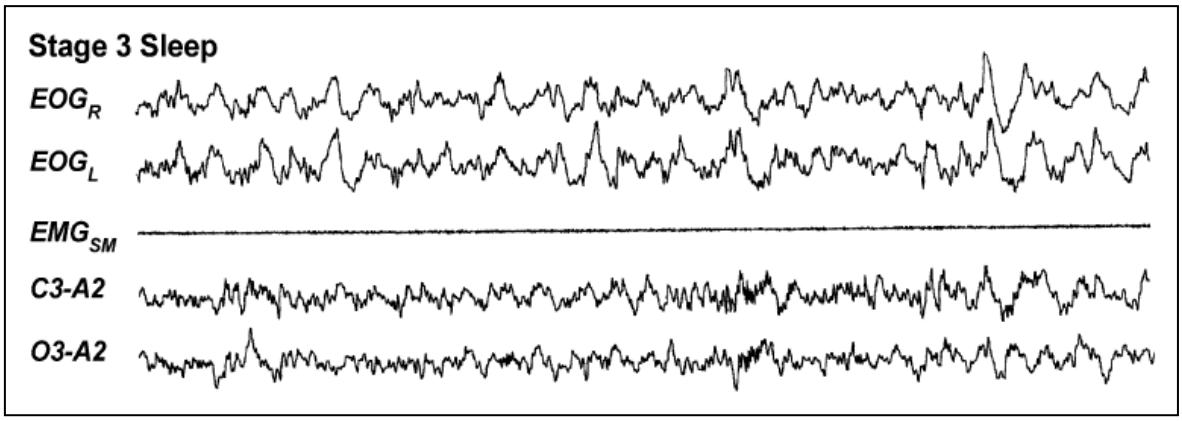


Figure 1.6: Typical EEG, EOG and EMG activities related to STAGE 3 sleep.

- STAGE 4 (ST4)
  (a) EEG: Stage 4 is scored when high amplitude delta activity covers more than 50% of an epoch. To awake a person from stage 3 or stage 4 is very difficult and usually it takes a few minutes to regain full consciousness. They are also called 'slow-wave sleep' or 'deep sleep'. Slow wave sleep has a restorative function, resulting more copious after physical works or sleep deprivation.
  (b) EMG: Muscle tone is much lower.
  (c) EOG: No eye movements.

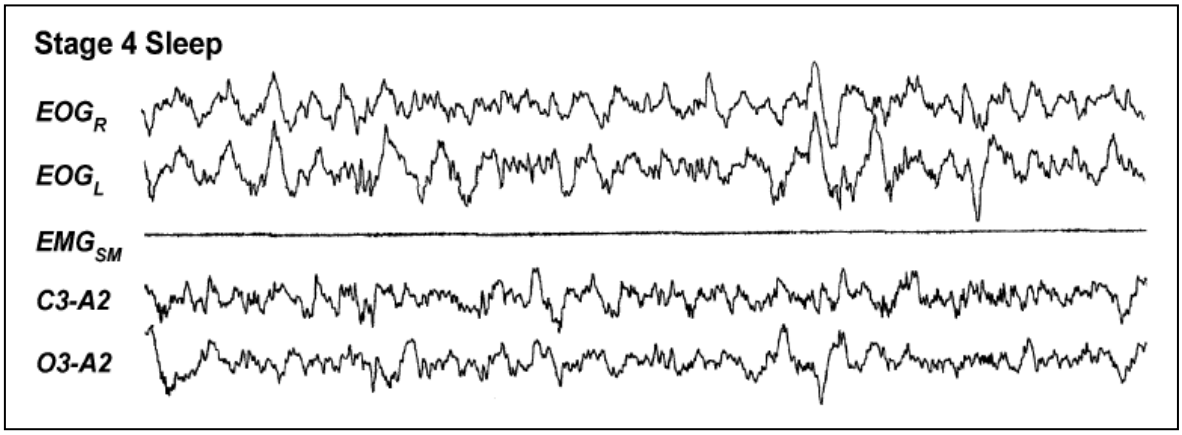


Figure 1.7: Typical EEG, EOG and EMG activities related to STAGE 4 sleep.

- REM (REM)
  (a) EEG: During REM sleep the EEG shows mixed frequencies (theta and alpha range). Saw-tooth theta waves may occur. REM sleep may be split into 2 different phases: periods when eye movement activity is high, and period when REM background activity continues without phasic activity. These two phases are called '*phasic REM activity*' and '*tonic REM activity*'. REM stage is often accompanied by dreaming activity [13]. The particular EEG recorded during REM indicates high cerebral activity. This high cerebral activity has been confirmed by cerebral blood flow during REM activity measured with positron emission tomography technique.
  (b) EMG: Muscle tone is the lowest, presumably to protect from possible self-damage caused as a consequence of dreaming activity.
  (c) EOG: Rapid eye movements

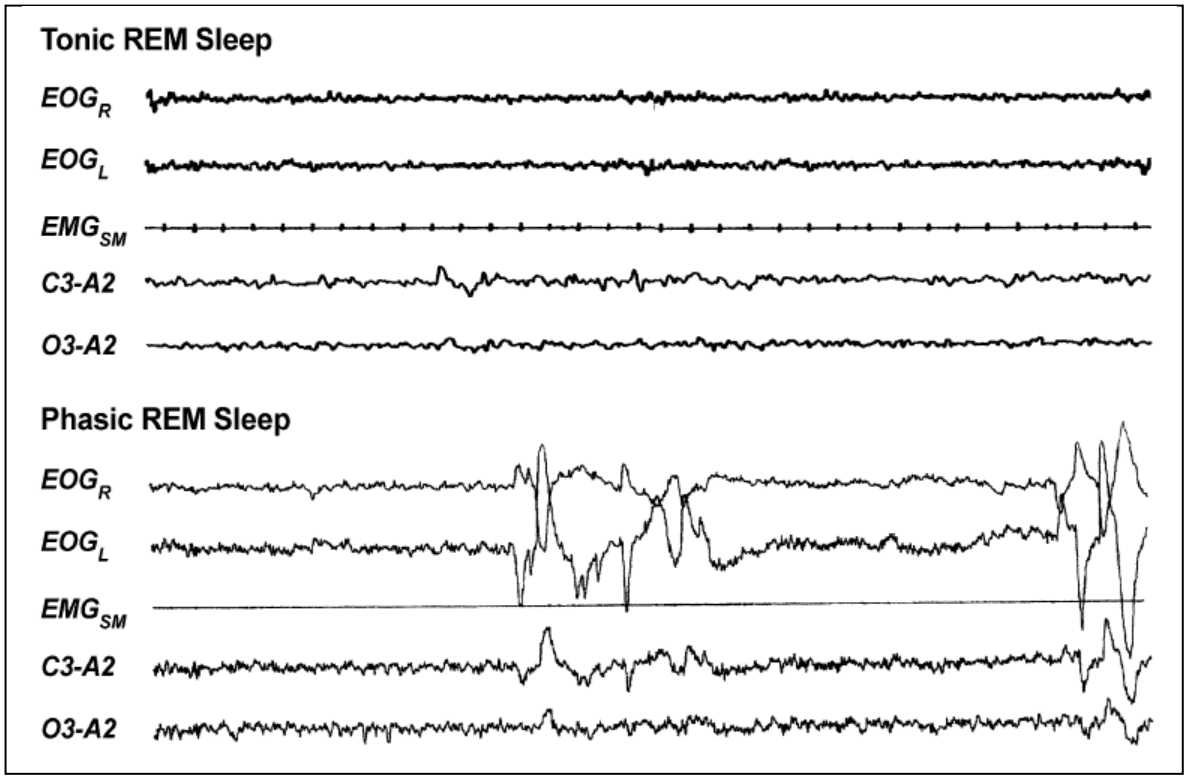


Figure 1.8: Typical EEG, EOG and EMG activities related to Tonic REM sleep and Phasic REM sleep.

1.1.3. Sleep statistics

Some important characteristics of sleep can be derived from plysomnographic data. Such characteristics are in the clinical practice used for quantifying both the severity of sleep disorders and hypnotics efficiency. These parameters are also called '*aggregated parameters*' because they reflect the overall trend of sleep during the night. The following definitions of sleep parameters are listed for reference:

- *Total recording time* (TRT): the duration of time from the start to the end of a recording, usually 8 hours.
- *Time in bed* (TIB): the duration of time from 'light off' to final awakening.
- *Sleep onset* (SO): the first epoch followed by 19 epochs of 'non awake' stages.
- *Sleep period time* (SPT): the duration of time from SO to final awakening.
- *Latency to persistent sleep* (LPS): the duration of time from 'light off' to SO.
- *Wake after sleep onset* (WASO): the total time spent awake during SPT.
- *Total sleep time* (TST): the amount of actual sleep time in a recording.
- *Sleep efficiency* (SE): the ratio of total sleep time to time in bed, i.e. TST/TIB*100.
- *Time spent in each of the sleep stages* (tAW, tST1, tST2, tST3, tST4, tREM).
- *Number of transition to each stage* (nAW, nST1, nST2, nST3, nST4, nREM).
- *Mean extension of each stage* (meanAW, meanST1, meanST2, meanST3, meanST4, meanREM).

1.1.4. Normal Sleep Pattern

The sequence of sleep stages across the night is called '*sleep architecture*' and usually follows a particular structure made by repeated 'sleep cycles' of alternated REM and Non-REM sleep, each lasting approximately 90-120 minutes. These cycles are repeated three to six times each night. Sleep stage 2 is the predominant stage, covering more than half of the night, and REM stage counting for another 20% to 25%. Stage 1 sleep is a transitional state, consequently it covers a minimal part of the night, approximately 1% to 5%, essentially in the phase of sleep onset and in the transition to light sleep. Slow wave sleep predominates the first third of the night, while REM sleep predominates in the last half of the night.

Spontaneous arousals and brief awakenings are common during or after REM sleep. Awakenings are defined as prolonged arousals implying cardiovascular activation [19,20,21].

Total duration of sleep varies from 7.5 and 8.5 hours in the majority of healthy persons. Mental and physical tasks preceding sleep time seriously influence sleep architecture: physical work increases deep-sleep, while mental work increases REM sleep percentages.

As mentioned above, the maintenance of such pattern is fundamental to preserve sleep quality.

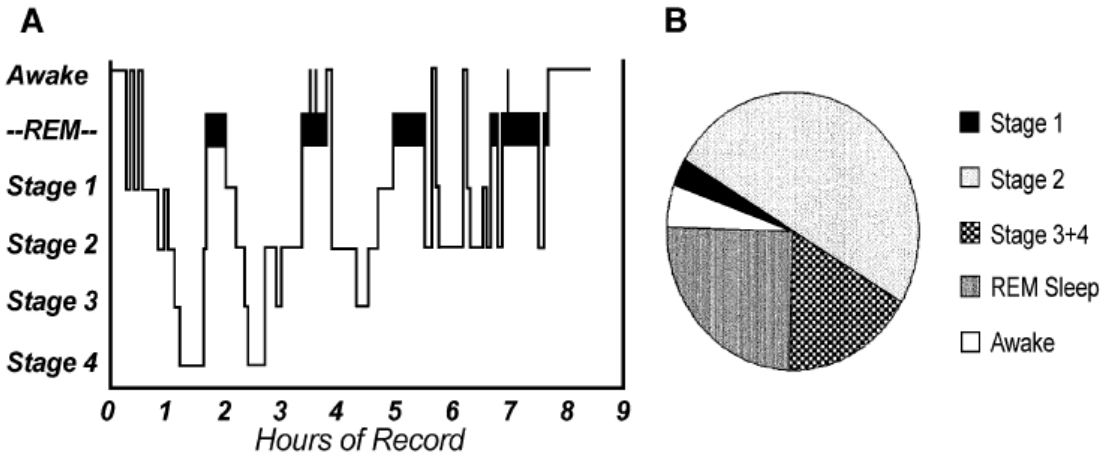


Figure 1.9: Normal sleep pattern. On the left a normal sequence of sleep stages during the night. On the right a diagram showing the total time spent in each stage.

1.2. Mechanisms regulating sleep

Maintaining a dynamic balance between wakefulness and sleep is necessary to preserve the overall wellness and restore mental and physical capabilities. The mechanisms regulating this balance are not completely known but many papers in literature [20,22] suggest that three key drives influence this process: (a) the autonomic nervous system, (b) the homeostatic system, (c) the circadian rhythm. These mechanisms are independent but clearly are involved to coordinate the sleep and wakefulness periodicity, allowing for adaptation to sudden circumstances provoking shifts in the time and duration of sleep.

1.2.1. Autonomic nervous system

In general, a drop in the sympathetic activation accompanied by a simultaneous increase of the parasympathetic activity favourites sleep. Consequently, anything, exogenous or endogenous, that increases sympathetic outflow can disturb sleep. For example, assuming caffeine or nicotine (exogenous) or being in particular states of anxiety (endogenous) are well known factors to disturb sleep.

During stressful periods, increased sympathetic activation results in higher levels of cortisol and adrenocorticotrophic hormone which increase wakefulness and inhibit restorative slow-wave sleep resulting in a ‘unrestorative’ sleep.

The autonomic mechanism is responsible of a 'survival function' in case of necessity during the night: it promotes quick response and sustained alertness. However, when this mechanism does not properly work, it may contribute to sleep disturbances.

1.2.2. Homeostatic control

The homeostatic process is determined by the amount of sleep and wakefulness. It has been demonstrated that the longer the sleep deprivation lasts the stronger the need to sleep becomes. Many sleep-deprivation studies have been conducted to determine the effects produced by sleep loss: sleep deprivation negatively affects attention, cognitive abilities and behavioural attitudes, entailing irritability and more easily (potential) demoralization.

Reduction of sleep time the previous night linearly increases the speed of falling asleep the following day: this can be quantified using a standardized measure of daytime sleepiness: the Multiple Sleep Latency Test (MSLT) [23].

Since the slow-wave sleep has a fundamental restorative function, partial or total sleep deprivation increases the amount of slow wave sleep rate during recovery sleep [19,22].

Physiologically, sleep deprivation results in increased cortisol levels, thyroid activity and catecholamine turnover [24,25].

1.2.3. Circadian rhythm

Independently of the homeostatic process, a circadian process influences the sleep-wake alternation according to the 24-hours light-dark cycle. The Central Nervous System CNS area involved with this mechanism is the suprachiasmatic nucleus (SCN), which is considered to be a biological clock. SCN controls a variety of biological functions, including the core body temperature. For this reason, the core body temperature is often used as a monitoring marker for circadian clock's period, phase and amplitude.

In general, in humans, the body temperature is characterized by:

- a nadir early in the morning (between 3 AM and 5 AM)
- a peak in the early evening (between 5 PM and 8 PM)
- a decline after midday (between 1 PM and 3 PM)

It has been demonstrated [26,27] that, according to this rhythm:

- maximum sleepiness occurs when body temperature reaches nadir
- maximum alertness occurs at body temperature peak
- drowsiness appears when temperature starts to fall

A variety of hormonal and metabolic functions are strictly related to the circadian process: blood levels of cortisol, prolactin, growth hormone, thyroid stimulating hormone and melatonin are driven by this daily clock.

1.3. Functions related to sleep

In addition to obvious physical and mental restorative functions, it is possible to attribute to sleep two further roles: (a) the capability to restore/maintain the immunologic system and (b) memory consolidation.

It has been demonstrated, in fact, that a low quality sleep or sleep deprivation seriously affects the immune system. For example, in [28] the authors proved that a group of 24-hour sleep-deprived rats had a 20% decrease in white blood cell count, compared with a control group.

It is also well demonstrated that sleep deprivation negatively affect memory functions, compromising cognitive tasks as decision making, reasoning, etc [29].

1.3.1. Memory and learning

A variety of studies have tried to clarify the relationship existing between sleep and memory consolidation. The mechanisms involved in the memory process consist of brain structure's reorganization into new neuronal connections on the basis of nerve cell dendrites' information/stimuli. This process needs to be fulfilled during the absence of continuous stimuli and this is the likely reason for which memory consolidation is accomplished during sleep.

There are essentially two types of learning and associated memory.

- The '*declarative memory*' which is tested through a series of pairs of related terms, such as hand-glove, leaf-tree etc. This type of memory is involved in learning foreign languages.

- The '*procedural memory*' which is tested through the learning of mirror-writing. This test consists on writing some sentences while the writing-hand is hidden being helped by a mirror only. Procedural memory is necessary for all types of task involving movements learning, such as car driving, skiing, etc.

Several studies have shown that not only REM sleep, but also NREM sleep is important for memory functions. In particular, in [30] it has been demonstrated that memory is influenced

differentially by certain sleep stages: subjects tested in the early night, in which slow wave sleep activity is prevalent, demonstrated a better declarative memory with respect to a control group, while subjects tested at the end of the night, when REM sleep is prevalent, performed better in the procedural memory test. This implies that improvements of declarative memory can be attributed to NREM sleep (SWS in particular) while those of procedural memory to REM sleep.

1.4. Sleep disorders

As mentioned above, disruptions in the correct maintenance of sleep architecture and sleep-wake balance may lead to serious consequences for individuals and society in general, compromising both productivity and wellness.

Sleep disorders are very common complaints affecting a large segment of world's population. The prevalence of such disorders is estimated to be approximately the 10% [31,32].

Sleep disorders consequences represent a substantial economic burden and, for this reason, it is a major objective for clinicians properly diagnosing and treating this kind of pathology.

The guidelines for diagnosing sleep disorders are listed in the '*Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition-Text Revision*' (DSM-IV-TR) [33].

Pharmaceutical companies have invested a lot of resources to develop new hypnotic drugs with a more safe and effective profile for the treatment of insomnia.

1.4.1. Classification/Definitions

According to the '*Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition-Text Revision*' (DSM-IV-TR) [33], sleep disorders can be divided into two main categories.

- *'Insomnia'*: defined as difficulty in falling asleep (sleep onset), difficulty in staying asleep (sleep maintenance) or low quality of sleep (non-restorative sleep). These symptoms may occur simultaneously or affect differentially some individuals, causing for example exclusively early morning awakenings.

  Duration of insomnia can be transient (from 1 to several nights), short term (from several days to a month) or chronic (lasting for months or years). Transient and short term insomnia often may be caused by changes in the sleep environment or by acute stressful experiences. Chronic insomnia, instead, may be of primary nature or secondary to other conditions/pathologies. When insomnia is secondary to another pathology, often it is not

possible to establish whether insomnia is the cause or the effect, but any specific treatment should take into account the coexistence of these pathologies.

There are six major categories in which chronic insomnia can be classified: medical (related to medical diseases), psychiatric (following or preceding psychiatric disorders), circadian (disordered circadian rhythm), behavioural, pharmacologic (due to drugs' undesired effects) and primary (not conditioned, independent disorder).

Transient and chronic insomnia obviously entail different therapeutic approaches. One of the most important objective in treating transient insomnia is to prevent its evolution to chronic insomnia.

Furthermore, it emerges from several studies [31,32] that risk factors exist for chronic insomnia, such as age, gender, medical disease, psychiatric disease and shift work.

- *'Disorders associated with excessive sleepiness'*, for example *narcolepsy*: these are chronic sleep disorders characterized by excessive daytime sleepiness (EDS) which involves extreme daytime fatigue and that may lead to fall asleep at inappropriate times. Narcoleptics usually experience disturbed nocturnal sleep and an abnormal daytime sleep pattern, which is often confused with insomnia.

1.4.2. Consequences of insomnia

Consequences associated to transient or short term insomnia are similar to those already highlighted about sleep deprivation studies: there is an increase in daytime sleepiness and an impairment of psychomotor function. Daytime sleepiness can be quantified using the MSL test, while psychomotor impairment can be assessed by performing different psychomotor tests, such as reaction time test, vigilance test etc. Furthermore, transient/short term insomnia affect significantly the ability of sustain attention, resulting in increased daytime lapses, that are periods of lack of responsivity (most likely effects of the so called 'microsleep').

Chronic insomnia is more complex and its consequences vary significantly from case to case. Some of the consequences attributable to chronic insomnia may be daytime fatigue and sleepiness, as in the case of rheumatoid arthritis and other medical diseases provoking fragmentation of sleep. However, in most cases people suffering of chronic insomnia do not show daytime sleepiness, but rather a greater alertness as shown by the higher mean latency in MSL test when compared to a control group without insomnia.

Many papers in literature [34,35] have also shown that often insomnia is a precursor for depression, reducing subject's quality of life.

It has been demonstrated that consequences of insomnia for the society in general are economically relevant. People affected by this pathology are more likely to absenteeism and have

higher rates of accidents, decreased productivity and quality of life [32]. Direct costs, those of medical care, have been estimated to be around $14 billion [36]. There are also relevant indirect costs due to decreased economic output attributable to insomnia; these have been quantified in around $100 billion [37].

1.4.3. Treatment of insomnia

In our culture, there is the tendency to not consider insomnia as a true medical disease, but rather as a minor problem, and people often do not report their sleeping difficulties to the general practitioner. There is also a general concern about long term use of hypnotic drugs, regarding in particular potential addiction and abuse liability. Some physicians, in fact, are reluctant to prescribe hypnotic drugs because they are considered symptomatic drugs without solving the real causes of insomnia.

However, physicians are now more experienced in treating such problems and are now available new hypnotic agents more safety and effective.

The purpose of insomnia treatment is obviously to improve patient's quality of life. In doing so, it is necessary to identify and remove any existing problem that may cause insomnia. This goal may be achieved through an appropriate treatment consisting in pharmacologic therapy combined to educational and behavioural approaches.

In the past, bromides, barbiturates, paraldehyde and methaqualone have been used as hypnotics, but although they have proved sedating properties, they also have significant toxicity problems. For this reason, their use is no more recommended. Current hypnotic drugs indicated in treating insomnia include traditional benzodiazepines and non-benzodiazepines. The non-benzodiazepines are positive allosteric modulators of the GABA-A receptor. Like the benzodiazepines, they exert their effects by binding to and activating the benzodiazepine site of the receptor complex.

- *Traditional benzodiazepines* have been available since the 1960s. These types of medications vary significantly in their elimination half-lives and subsequent duration of action. The half-life times range from few hours to few days, and this is the major concern about their use. The longer half-life time, in fact, may have daytime undesired effects after drug's intake, compromising daytime impairment and increasing the risk of accidents and falls. Among the benzodiazepines approved for the treatment of insomnia, the fast acting ones with short half-lives such as estazolam, triazolam, and temazepam are recommended. Longer-acting benzodiazepines such as nitrazepam and diazepam have residual effects that may persist into the next day and are, in general, not recommended.

- *Newer nonbenzodiazepines,* (zolpidem and zaleplon, zopiclone, eszopiclone) are available since 1990s. These medications have a shorter half-life time (1-2 hours), so there is a very little risk of morning residual effect. Their action implies a rapid sleep onset, so the patient can take them just before going to bed. Serious adverse reactions are infrequent. The short term safety and efficacy of these new hypnotics have been well proved [38,39] and clinical experience supports safety in long-term intermittent use. Continuous long term use is not recommended as tolerance, dependence and addiction can occur.

Hypnotics' efficacy, as already mentioned, is generally determined on the basis of the effect they exert on the aggregated parameters. Hypnotic medications currently in use are demonstrated to positively influence the *Latency to Persistent Sleep* (LPS), inducing a rapid sleep onset [38,40].
Integrated to the pharmacologic approach, it is of relevant importance the evaluation of specific behaviours and circumstances that favourite insomnia. For this reason it is indispensable the continuous monitoring and the constant support of the patient.

1.5. Aim of this work

Sleep disorders are commonly assessed through 'polysomnography', a quantitative diagnostic technique that consists in the simultaneous recording of the principal electrophysiological activities involved in the sleep process.
Currently, the clinical use of polysomnography is mainly related to assess sleep aggregated characteristics (WASO, LPS, etc.). However, averaging PSG data into these parameters implies a substantial loss of information regarding the sleep architecture considered as a time-course pattern among specific sleep stages. The maintenance of this internal structure has been demonstrated to be of significant importance.
A Multinomial Mixed-Effect Markov-chain model has been recently proposed to properly describe the sleep architecture.
The aim of the work presented in this thesis is to assess the appropriateness of this Markov-chain model, evaluating whether the proposed model adequately describes the underlying system, provides precise predictions for the parameters of interest and avoids biased estimates.

Chapter 2

Model identification methods

2.1. Introduction

As described in the previous chapter, the main objective of this thesis is to evaluate a recent multinomial proposed for analyzing polisomnography data [2]. However, bias and imprecision in the model estimations and predictions may depend not only on possible model misspecifications but also on weaknesses in the estimation methods. When dealing with data from clinical studies the key objectives of any modeling analysis is to establish both population mean and individual responses, since the drug under development has to be effective and safe for the whole population.

Let the gathered data arise from *M* subjects participating in a certain clinical study and let *X* be the data matrix, where each row contains individual data on *N* different times:

$$X = \begin{bmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_M^T \end{bmatrix} = \begin{bmatrix} x_{11}\, x_{12} \cdots x_{1N} \\ x_{21}\, x_{22} \cdots x_{2N} \\ \vdots \quad \vdots \quad \cdots \quad \vdots \\ x_{M1} x_{M2} \cdots x_{MN} \end{bmatrix}.$$

Suppose such data need to be described through a mathematical model. Classical approaches, e.g. Least Squares (LS), Weighted Least Squares (WLS), Maximum Likelihood (ML) estimate, Bayesian approach, etc., are generally designed for individual fitting:

$$X_i = f_i(z, \varphi_i) + v_i,$$

where X_i are the individual data, f_i the corresponding model prediction, z the independent variable, φ_i the individual parameters and v_i the vector of the random error affecting individual data. The random error is generally due to measurement errors and noise and it is supposed to be drawn from a Gaussian distribution with zero mean and covariance matrix equal to Σ, $v_i \sim N(0, \Sigma)$.

These approaches present some identification problems in case of noisy or sparse individual data: individual estimates in certain cases may be not accurate or impossible to obtain. Furthermore, information at the mean and individual levels are often needed. Consequently, when working in similar contexts, it is suitable to build models and identify the corresponding parameters using the so called '*population approaches*'.

Let the individual parameters of the population under examination belong to a certain distribution (gaussian, lognormal, etc.), characterized by mean θ and covariance Ω:

$$\varphi_i \sim Distr(\theta, \Omega).$$

It is then possible to think that individual data are realizations of such distribution. The population approach allows to investigate, with varying levels of precision, both the mean parameter and the variability in the population, i.e. the first two moments of the parameters distribution, $\Theta = (\theta, \Omega)$.

In this chapter, population approaches are illustrated in detail with particular emphasis on the '*mixed-effect modelling approach*' and its application in the case of categorical data.

2.2. Naïve average data approach

When clinical studies are performed with the same drug administration and sampling schedules for all the subjects in a population, the simplest method for analyzing data is to use a '*Naïve Average Data*' (NAD) approach, which allows to obtain an estimate, $\hat{\theta}$, of the mean population parameter, θ.

This approach consists in:

(a) computing the average value of the data for each sampling time:

$$\bar{x}_t = \frac{1}{M}\sum_{i=1}^{M} x_{it}\,, \qquad t = 1, \ldots, N$$

where *t* indicates the sample number, *i* the subject number, *N* the total number of sampling times and *M* the total number of subjects (the resulting mean vector is $\bar{X}^T = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N]$);

(b) fitting a model to the mean data vector: $\bar{X} = f(z, \hat{\theta}) + v$.

This approach is very easily implementable, but it is rather poor in terms of performance and it can be very misleading (i.e., leading to wrong models). Imagine, for example, to have gathered data which can be described with a mono-exponential decay law:

$$X_i = A\, exp(-K_i z) + v_i.$$

The mean curve used by NAD approach seems to exhibit a bi-exponential decay, as shown in Figure 2.1. That is, data averaging produces a distorted picture of the underlying model.

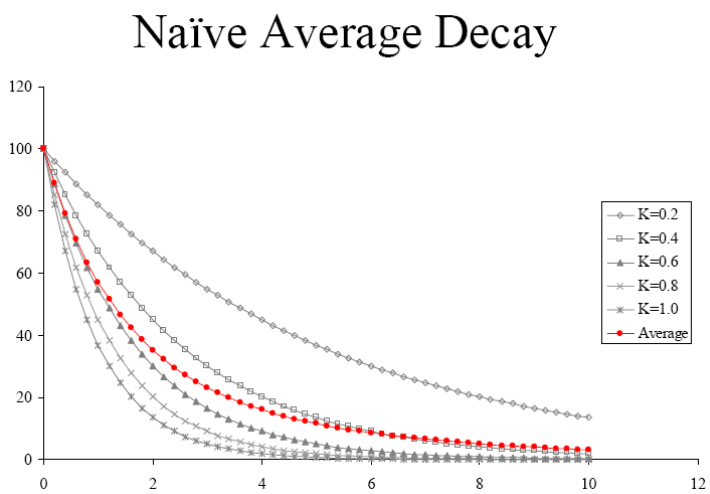


Figure 2.1: Naïve average data approach applied to mono-exponential data. Figure obtained from 'Models and control of Biological Systems 2' course 2007-2008.

Furthermore, such approach cannot provide any information on the individual behaviour: all sources of variability disappear with the computation of the mean response.

2.3. Naïve pooled data approach

The '*Naïve Pooled Data*' (NPD) method was proposed by Sheiner and Beal in [41]. Unlike the NAD approach, the NPD approach is far more general: it considers all data as belonging to one unique individual. The subscript *i* disappears and the subscript *t* indexes all the available (*z*, *x*) pairs:

$$x_t = f(z_t, \varphi_{NPD}) + v_t,$$

where *f* is the model prediction for the '*reference subject*'. The latter is characterized by a set of parameters, φ_{NPD} , which can be estimated using, for example, the LS method, where the following global '*objective function*' (OF) is minimized:

$$OF(\varphi_{NPD}) = \sum_{i=1}^{M}\sum_{t=1}^{N}(x_{it} - f(z_t, \varphi_{NPD})))^2.$$

The resulting value is therefore an estimate of the mean population parameters, $\hat{\theta}$.
This approach performs well in case of small variations between subjects but cannot provide any information on the individual behaviour: all sources of variability are confounded together.

2.4. Standard two-stage approach

The '*Standard Two-Stage*' (STS) approach consists in:

(a) estimating individual parameters, φ_i, by individually fitting each subject data (through for example a classical ML approach or a WLS method as in the following formula):

$$\hat{\varphi}_i = argmin_{\varphi_i}[X_i - f_i(z, \varphi_i))]^T \Sigma^{-1}[X_i - f_i(z, \varphi_i)],$$

where Σ is the covariance matrix of the residual errors;

(b) computing population mean and covariance as empirical mean and covariance on the $\hat{\varphi}_i$, as follows:

$$\hat{\theta} = \frac{1}{M}\sum_{i=1}^{M}\hat{\varphi}_i,$$

$$\hat{\Omega} = \frac{1}{M}\sum_{i=1}^{M}(\hat{\varphi}_i - \hat{\theta})(\hat{\varphi}_i - \hat{\theta})^T,$$

where $\hat{\theta}$ and $\hat{\Omega}$ are the estimated population mean and covariance, respectively.
The advantage of the STS approach is its simplicity, but the STS estimator is demonstrated to be polarized. Ideally, it would be suitable to estimate the true population variability only, by using the true individual residuals, r_i:

$$\Omega = \boldsymbol{E}[r_i r_i^T] = \frac{1}{M}\sum_{i=1}^{M}(\varphi_i{}^* - \theta^*)(\varphi_i{}^* - \theta^*)^T ,$$

where $\varphi_i{}^*, \theta^*$ are the true individual and mean parameters. But using the STS approach the individual residuals are instead:

$$r_i = (\hat{\varphi}_i - \hat{\theta}) = \hat{\varphi}_i - \varphi_i{}^* + \varphi_i{}^* - \theta^* + \theta^* - \hat{\theta}$$

Error on the individual estimate — *Population variability* — *Error on the mean estimation*

Moreover, population parameters are obtained in a second step, non influencing individual estimates. For this reason, individual estimates are not improved with respect to the estimates obtained with the traditional approaches. Furthermore, individual fitting in the first step needs 'rich data' to be correctly performed, otherwise it tends to over-estimate parameters dispersion.

2.5. Iterative two-stage approach

The '*Iterative Two-Stage*' (ITS) approach is an iterative method which consists in the following steps.

(a) Individual parameters, φ_i, are estimated by individually fitting each subject data.

(b) Population mean and covariance are computed as empirical mean and covariance, as in the STS approach, $\hat{\theta}_{STS}$ and $\hat{\Omega}_{STS}$.

(c) $\hat{\theta}_{STS}$ and $\hat{\Omega}_{STS}$ are used as initial estimates for $\hat{\theta}$ and $\hat{\Omega}$: $\hat{\theta}^{(0)} = \hat{\theta}_{STS}$, $\hat{\Omega}^{(0)} = \hat{\Omega}_{STS}$.

(d) Individual parameters at *k*-th iteration, $\hat{\varphi}_i^{(k)}$, are estimated with a Bayesian approach, for example a MAP approach, using as prior for the estimands the parameters obtained at the previous iteration:

$$\hat{\varphi}_i^{(k)} = argmin_{\varphi_i}[X_i - f_i(z,\varphi_i)]^T \Sigma^{-1}[X_i - f_i(z,\varphi_i)] + \\ +[\varphi_i - \hat{\theta}^{(k-1)}]^T \{\hat{\Omega}^{(k-1)}\}^{-1}[\varphi_i - \hat{\theta}^{(k-1)}].$$

(e) Individual parameters obtained at the previous step are therefore used to re-estimate population parameters, $\hat{\theta}^{(k)}$, $\hat{\Omega}^{(k)}$, where k indicates the number of iterations:

$$\hat{\theta}^{(k)} = \frac{1}{M}\sum_{i=1}^{M}\hat{\varphi}_i^{(k)},$$

$$\hat{\Omega}^{(k)} = \frac{1}{M}\sum_{i=1}^{M}(\hat{\varphi}_i^{(k)} - \hat{\theta}^{(k)})(\hat{\varphi}_i^{(k)} - \hat{\theta}^{(k)})^T.$$

(f) Steps (d) and (e) are repeated until convergence.

This method allows the estimation of the population parameters, θ and Ω, to be estimated. It also provides individual parameter values, estimated using information derived not only from individual data but also from population parameters distribution (priors). However, ITS method, being an iterative method, involves a certain computational burden and may produce polarized estimates when the number of iterations is excessive.

2.6. Non-linear mixed-effect approach: theory

The '*Non-linear Mixed-Effect*' (NLME) approach is one of the most interesting for population analysis, and it is particularly well suited for biological and medical data, which display heterogeneity of responses to stimuli and treatments. This approach is based on the assumption that the (unknown) process to be described is characterized by a typical behaviour which is common to the whole population and by some sources of variability that make the individual behaviours differ from the typical one. The latter is determined by the so called '*fixed effects*', while the identified sources of variability are of two different types and are called '*random effects*'. The first source of variability is the intrinsic difference that exists among subjects: one individual is obviously different from another one. This is called '*inter-individual*' or '*between subject*' variability. Mostly in the medical field, to understand the variability between subjects is as important as to understand the characteristics of the typical individual. The second source of variability is the '*residual error*' (also called '*noise* or '*intra-individual error*'): this is the difference between the prediction of the model for the individual and individual measured observations. It is also called '*intra-individual*' or '*within subject*' variability. For taking into account all of these assumptions, the mixed-effect approach specifies the model in a hierarchical fashion, integrating an '*individual*' model and a '*population*' one. In this way, it allows to estimate both the vector of population characteristics, $\Theta = (\theta,\ \Omega)$, and the individual parameters, φ_i.

Individual model

The individual model is aimed to describe individual data specifying the relationship between the dependent variables, independent variables and individual parameters.

Let

- z be the independent variable, for example '*time*' in a time series;
- z_t be the *t*-th value of the independent variable, $t = 1, \ldots, N$;
- x_{it} be the *t*-th observation of the *i*-th individual, $t = 1, \ldots, N$ and $i = 1, \ldots, M$;
- φ_i be the vector of model parameters of subject *i*.

Each individual measure, x_{it}, can be described by the individual model in this way:

$$x_{it} = f_i(z_t, \varphi_i) + v_{it}, \quad \forall t = 1, \ldots, N,$$

where $f_i(z_t, \varphi_i)$ is the individual model prediction and v_{it} is the residual error.

Using a vector notation:

$$\begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iN} \end{bmatrix} = X_i = f_i(z, \varphi_i) + v_i = \begin{bmatrix} f_i(z_1, \varphi_i) \\ f_i(z_2, \varphi_i) \\ \vdots \\ f_i(z_N, \varphi_i) \end{bmatrix} + \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iN} \end{bmatrix}.$$

Classical assumptions for the residual errors v_{it}, or are that they

(a) have zero mean,

(b) are uncorrelated,

(c) are normally distributed.

Therefore, v_{it} are independently normally distributed with

$$E[v_i|\varphi_i] = 0,$$
$$Cov(v_i|\varphi_i) = R_i(\varphi_i, \xi),$$

where R_i is a diagonal matrix depending on ξ (a constant characteristic across individuals) and possibly on the individual parameters (according to the error model structure). In practice,

$$v_i \sim N\big(0, R_i(\varphi_i, \xi)\big).$$

Population model

A model for φ_i is also needed in order to account for inter-individual variability among the φ_i. In particular, the population model relates the individual parameters to the covariate vector, the fixed effects and the inter-individual random effects.

Let

a_i be the covariate vector, i.e. the set of individual values for weight, age, etc.,

η_i be the vector of inter-individual random effects associated with the subject *i*,

θ be the vector of fixed effects.

A general population model is given by

$$\varphi_i = d(\theta, \eta_i, a_i),$$

where d is a multi-dimensional function, and η_i are supposed to be drawn from a normal distribution having zero mean and Ω covariance matrix, i.e.

$$\eta_i \sim N(0, \Omega).$$

Eventually, the model can be expressed as follows:

$$X_i = f_i(z, d(\theta, \eta_i, a_i)) + v_i, \quad \eta_i \sim N(0, \Omega), \quad v_i \sim N\big(0, R_i(\varphi_i, \xi)\big).$$

Most of the non-linear mixed-effect modelling methods estimate the parameters using a ML approach: the data probability is given by a function of the model parameters and parameter estimates are chosen to maximize this probability.

The overall likelihood is the product of all individual likelihoods L_i and since the likelihood must account for the random effects on the individual level, the individual likelihood is expressed as the integral over all possible values of η_i:

$$L = \prod_{i=1}^{M} L_i = \prod_{i=1}^{M} \int l(X_i|\theta, \eta_i, a_i) h(\eta_i|\Omega) d\eta_i \qquad [2.1]$$

where h is a multivariate normal density function with zero mean and covariance matrix Ω.

2.7. Non-linear mixed-effect approach: NONMEM as software platform

NONMEM version VI (Icon Development Solutions) [42] is a software platform that allows to perform population analysis with mixed-effect models and whose name stands for 'NON linear Mixed Effect Modeling'.
The basic steps for running NONMEM are:

(a) to organize data input,
(b) to write the control file, which specifies the mixed-effect model,
(c) to run the model and obtain model parameter estimates.

(a) NONMEM needs '*data input*' files organized into records with some pre-defined items as follows. The subject number, called 'ID', is the first item in each record. The records appear in subject order and, within a subject, they are organized by time if a time series is being analyzed. Time specification, called 'TIME', is usually the second item in each record. The dependent variable, called 'DV', is the third one. Additional items can be added if the model requires so, for example subject covariates, such as weight, height, gender, etc.

(b) Once data are organized, the model is specified in the '*control file*', or '*control stream*', which for an estimation problem typically contains the following control elements.

$PROB	States the problem being solved.
$DATA	Specifies the name of the data file (created at step (a)).
$INPUT	List the names of the data records in the input file, in the exact order of data file columns.
$PRED	Describes ta routine that predicts the observations. It is the main part of the control stream in which the mixed effect model is specified.
$THETA	List the initial estimates of the fixed effects parameters.
$OMEGA	List the initial estimates of the variance of the inter-individual random effects (called ETA's). Note that if they are all fixed to zero the estimation method

	becomes NAD.
$SIGMA	List the initial estimates of the variance of the intra-individual random effects (called EPS's).
$EST	Provides the parameters that control the estimation process. Typically includes: METHOD (indicating which estimation method has to be applied), MAX (maximum number of iterations), LIKELIHOOD (indicating whether the likelihood is defined by the modeller; this option is necessary in case of categorical data modelling).
$COVARIANCE	Implies the estimation of the full variance-covariance matrix of the parameter estimates. This step is useful to get standard errors of the estimated parameters.
$TABLE	Produces an output table of the results.

NONMEM needs specific keywords for each feature of a mixed-effect model, in particular:

(i) THETA is a fixed effect parameter, i.e. an element of θ.;

(ii) ETA is an inter-individual random effects, i.e. an element of η_i;

(iii) EPS is an intra-individual random errors, i.e. an element of v_i;

(iv) F is the individual model prediction, $f_i(z_t, \varphi_i)$;

(v) Y is an observation, x_{it} ., or the correspondent user-defined likelihood (if LIKELIHOOD is specified in $EST).

NONMEM only deals with normally distributed random variables (ETA's and EPS's). However, in the $PRED field it is possible to implement three different individual parameter distributions (see VAL) and error structures (see Y):

(i) additive model:

VAL=THETA+ETA

Y=F+EPS

(ii) constant coefficient of variation model:

VAL=THETA* (1+ETA)

Y=F*(1+EPS)

(iii) log-normal model:

VAL=THETA*EXP(ETA)

Y=F*EXP(EPS)

where VAL stands for the value for the specific individual parameter of interest.

(c) NONMEM estimates model parameters with a ML approach. However, the likelihood of non-linear mixed-effect models, Equation 2.1, is often difficult to compute in a close form because of non-linearity of the random effects. To deal with these problems, three methods have been implemented in NONMEM, all of them based on likelihood approximations: the '*First Order*', the '*First Order Conditional Estimation*' and the '*Laplace*' methods.

2.7.1. First order

The First Order (FO) method is the simplest approximation method that can be applied when dealing with ML estimation in non-linear mixed-effect modelling. It consists in linearizing the mixed-effect model,

$$X_i = f_i(z, d(\theta, \eta_i, a_i)) + v_i, \quad \eta_i \sim N(0, \Omega), \quad v_i \sim N\big(0, R_i(\varphi_i, \xi)\big).$$

(in which the dependency from z has been omitted for simplicity), through a first order Taylor expansion around the mean random effects of the population, $\bar{\eta}_i = 0$:

$$X_i = f_i\{d(\theta, \eta_i, a_i)\} + v_i$$

$$X_i \simeq f_i\{d(\theta, 0, a_i)\} + \underbrace{\frac{\partial f_i\{d(\theta, 0, a_i)\}}{\partial \varphi_i} \frac{\partial d(\theta, 0, a_i)}{\partial \eta_i}}_{Z_i(\theta, 0)} (\eta_i - 0) + v_i$$

$$X_i \simeq f_i\{d(\theta, 0, a_i)\} + Z_i(\theta, 0)\eta_i + v_i.$$

In this way, the random effects v_i and the individual errors η_i, (assumed to be independent) enter the approximation in a linear way. Since they are normally distributed, also the marginal distribution of X_i is normally distributed, with the following first two moments:

$$E[X_i] = f_i\{d(\theta, 0, a_i)\},$$

$$Cov(X_i) \simeq E\big[X_i {X_i}^T\big] = E[(Z_i(\theta, 0)\eta_i (Z_i(\theta, 0)\eta_i)^T] + E[v_i {v_i}^T] =$$

$$= Z_i(\theta,0)E[\eta_i{\eta_i}^T]Z_i(\theta,0)^T + E[v_i{v_i}^T] =$$
$$= Z_i(\theta,0)\Omega Z_i(\theta,0)^T + R_i(d(\theta,0,a_i),\xi) = V_i(\theta,0,\Omega).$$

Therefore, for each subject the likelihood is

$$L_i(\theta,\Omega) = \frac{1}{\left((2\pi)^N \det\left(Cov(X_i)\right)\right)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left[X_i - E[X_i]\right]^T Cov(X_i)^{-1}\left[X_i - E[X_i]\right]\right) =$$
$$= \frac{1}{\left((2\pi)^N \det\left(V_i(\theta,0,\Omega)\right)\right)^{1/2}} \exp\left(-\frac{1}{2}[X_i - f_i\{d(\theta,0,a_i)\}]^T V_i(\theta,0,\Omega)^{-1}[X_i - f_i\{d(\theta,0,a_i)\}]\right)$$

and the overall likelihood is given by the product of all the individual likelihoods:

$$L(\theta,\Omega) = \prod_{i=1}^{M} L_i(\theta,\Omega).$$

The population parameters $\Theta = (\theta,\Omega)$ can be estimated by maximizing $L(\theta,\Omega)$ or, similarly, minimizing the following objective function:

$$OF = -2\log\left(L(\theta,\Omega)\right).$$

For simplifying the notation, let's call $f_i\{d(\theta,0,a_i)\} = f_i$. The objective function becomes:

$$OF = -2\log\left(\prod_{i=1}^{M} \frac{1}{\left((2\pi)^N \det\left(V_i(\theta,0,\Omega)\right)\right)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}[X_i - f_i]^T V_i(\theta,0,\Omega)^{-1}[X_i - f_i]\right)\right)$$

Individual parameters φ_i are estimated in a second step. In particular, the inter-individual random effects are estimated with a '*Maximum A Posteriori*' (MAP) approach using as prior for the their distribution

$$\eta_i \sim N\left(0,\widehat{\Omega}\right),$$

where $\widehat{\Omega}$ is the previously estimated population variability.
In particular, calling $R_i(\varphi_i,\xi) = R_i$, η_i estimates are obtained by minimizing:

$$\hat{\eta}_i = argmin_{\eta_i}\left[X_i - f_i\left\{d\left(\hat{\theta},\eta_i,a_i\right)\right\}\right]^T {R_i}^{-1}\left[X_i - f_i\left\{d\left(\hat{\theta},\eta_i,a_i\right)\right\}\right] + {\eta_i}^T \widehat{\Omega}^{-1}\eta_i$$

Once the inter-individual random effects are computed, the individual parameters can be obtained according to the population model:

$$\hat{\varphi}_i = d\left(\hat{\theta}, \hat{\eta}_i, a_i\right)$$

This step is called the '*Posthoc Step*' and the estimates obtained through it are also called '*Posthoc estimates*' or '*Empirical Bayes Estimates*' (EBE's).

The First Order approximation method guarantees quick computational times. However, especially for models that are nonlinear in the parameters, the approximation about the expectation of the random effects, $\bar{\eta}_i = 0$, might be rather poor, resulting in inconsistent estimates of the fixed effects:

$$\hat{\theta}_{FO} \rightarrow \theta_{true} + bias.$$

2.7.2. First order conditional estimation

The '*First Order Conditional Estimation*' (FOCE) method is a more accurate algorithm. The FOCE algorithm approximates η_i by an individualized estimate, ${\eta_i}^*$, defined by a second optimization problem:

$$X_i \simeq f_i\{d(\theta, {\eta_i}^*, a_i)\} + \underbrace{\frac{\partial f_i\{d(\theta, {\eta_i}^*, a_i)\}}{\partial \varphi_i}\frac{\partial d(\theta, {\eta_i}^*, a_i)}{\partial \eta_i}}_{Z_i(\theta, {\eta_i}^*)}(\eta_i - {\eta_i}^*) + {v_i}^*$$

$$X_i \simeq f_i\{d(\theta, {\eta_i}^*, a_i)\} + Z_i(\theta, {\eta_i}^*)(\eta_i - {\eta_i}^*) + {v_i}^*$$

$$X_i \simeq f_i\{d(\theta, {\eta_i}^*, a_i)\} + Z_i(\theta, {\eta_i}^*)\eta_i - Z_i(\theta, {\eta_i}^*){\eta_i}^* + {v_i}^*,$$

where ${v_i}^* \sim N(0, R_i(d(\theta, {\eta_i}^*, a_i), \xi))$.

The approximate marginal distribution of X_i is therefore characterized by the following mean and covariance:

$$E[X_i] = f_i\{d(\theta, {\eta_i}^*, a_i)\} - Z_i(\theta, {\eta_i}^*){\eta_i}^*$$
$$Cov(X_i) \simeq Z_i(\theta, {\eta_i}^*)\Omega Z_i(\theta, {\eta_i}^*)^T + R_i(d(\theta, {\eta_i}^*, a_i), \xi) = V_i(\theta, {\eta_i}^*, \Omega)$$

The overall likelihood, obtained as in the FO approximation method, can be used for estimating (θ, Ω). These population estimates can be subsequently used to update the estimates for η_i.

2.7.3. Laplace approximation

Laplace method is based on the Laplacian approximation of the exact marginal likelihood specified by the hierarchical nonlinear model:

$$L = \prod_{i=1}^{M} L_i = \prod_{i=1}^{M} \int l(X_i|\theta, \eta_i, a_i) h(\eta_i|\Omega) d\eta_i.$$

Given a complex integral, $\int b(x)dx$, $b(x)$ can be re-expressed as $e^{logb(x)} = e^{g(x)}$ and $g(x)$ can be approximated by a second-order Taylor expansion around a point x_0 as:

$$g(x) \simeq g(x_0) + (x - x_0)g'(x_0) + \frac{(x - x_0)^2}{2!} g''(x_0).$$

The approximated integration is the first order Laplacian approximation to the true integration:

$$\int b(x)dx = \int e^{g(x)}dx = \int e^{g(x_0)+(x-x_0)g'(x_0)+\frac{(x-x_0)^2}{2!}g''(x_0)}dx. \qquad [2.2]$$

The Laplacian approximation given by Equation 2.2, considering η_i as x and therefore $l(X_i|\theta, \eta_i, a_i)h(\eta_i|\Omega)$ as $b(x)$, allows to compute the approximated likelihood in close form. The approximation implemented in NONMEM is around the conditional estimates of η_i for the non-linear model. The conditional estimates or the empirical Bayes estimates (EBE) of η_i are the best predictors of the random effects η_i.

2.8. Categorical data likelihood

Data can be considered as continuous or categorical. Continuous data consist of variables with an infinite number of values, while categorical data consist of variables taking values in a finite set and can therefore be placed into mutually exclusive categories.

Categorical data can be further divided into nominal or ordinal. The nominal ones are unordered and can be classified by '*names*', while the ordinal ones have a natural order and can be organized into '*levels*', even if the exact distance between the levels is generally unknown. Examples of

nominal categorical variables are race or sex, while examples of ordinal categorical variables are age or pain intensity (e.g., absent, mild, moderate, severe).

Sleep data can be considered as categorical data: sleep data derived from PSG recordings can be 'categorized' into 6 mutually exclusive categories (AW, ST1, ST2, ST3, ST4, REM), according to the criteria illustrated in Chapter 1. Such categories do not have a recognized natural order, therefore sleep data can be considered as nominal categorical data.

Modeling categorical data is slightly more complex than modeling continuous ones. There are two major differences between models for continuous and categorical data analysis:

(a) the probability of a certain outcome, rather than the values of the outcome itself, is modeled;

(b) consequently, no residual error is defined in the model.

The model for categorical data becomes:

$$p_{it} = f_i(z_t, d(\theta, \eta_i, a_i)), \qquad \eta_i \sim N(0, \Omega),$$

where p_{it} is the probability of a certain outcome for the *i*-th subject at time z_t.

Consequently, each observation's likelihood is the probability of the observation itself as it is modeled, and the likelihood of Equation 2.1, $l(x_i|\theta, \eta_i)$, can be expressed by the product of the probabilities of the outcomes. Such likelihood is said 'user-defined', because it is explicitly defined in the model. In this case the only estimation method which can be adopted with NONMEM VI is the Laplace one. A more detailed explanation on likelihood expression in case of categorical data and in the specific Markov-chain model context is provided in Chapter 3.2.8.

2.9. Conclusion

In this chapter we have provided an overview of different estimation methods with a primary focus on the mixed-effect approach in a context of categorical data. In the next chapter a full description of the key features of the Multinomial Markov-Chain model for the analysis of polisomography data will be presented together with relevant information about data collected in a recent clinical study in primary insomnia.

.

Chapter 3

A Markov-chain model for the study of polysomnographic signals

3.1. Transitions between sleep stages and Markov chains: state of art

As mentioned in the first chapter, polysomnography allows to evaluate some relevant clinical parameters (LPS, WASO, etc) characterizing patients' sleep quality. The clinical efficacy of an hypnotic compound is usually evaluated on the basis of its effects on these aggregated parameters. However, this kind of approach brings mainly to the evaluation of the overall effect of the drug under examination and provides only few insights on the internal structure of sleep.

Besides supplying values for the aggregated parameters, polysomnography enables the assessment of the sleep stages occurring during the 8-hour night time (Figure 3.1). Their sequence is determined on a 30-second temporal grid and is therefore made of 960 'samples', one for each 'epoch' (30 second interval). Each sequence can be considered as a realization of a finite discrete stochastic process assuming values in a finite discrete set, i.e. the sleep stages. Therefore appropriate mathematical models can describe sleep structure allowing to predict feasible sleep stages time-courses over the night.

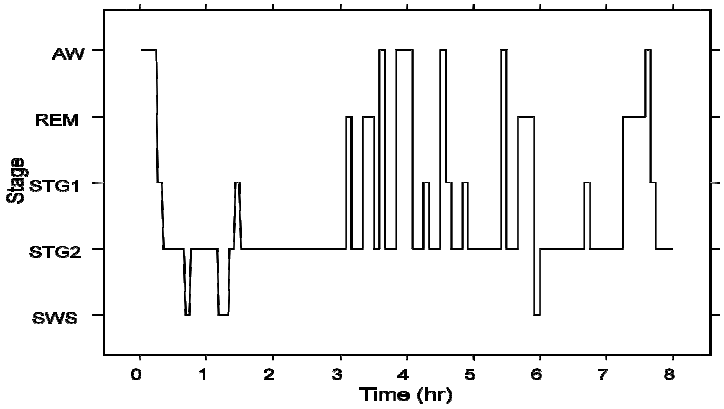


Figure 3.1. A realization of sleep stochastic process.

Since, as illustrated in paragraph 1.1.4, there are specific sleep patterns with some preferential transitions, sleep stages can be considered as a sequence of stochastic variables with some dependency among each other.

A mathematical model able to describe this feature is the '*Markov chain model*', whose basic concepts are those of *state* and *state transition*. The '*Markov-chain model*' determines the probabilities of transitioning between different states based on the recent and the present states or, in the less general case, on the present state only. Such model is adaptable to a large variety of situations: for example, it can be applied to the classic stochastic problem of the Drunkard's Walk, in which the position reached in the next step only depends on the present position and not on the way this present position was reached. It can be also used to predict weather conditions given the weather on the previous days or to describe all the board games played with dice (Monopoly for example), in which the future position on the board depends only on the current state and the next roll of the dice.

Some attempts to apply a Markov-chain model for describing sleep data are reported in the literature: in 2002 Gregory and Cabeza [43] used this approach for describing the internal architecture of sleep in rats. However, they modeled sleep as a two-state process, considering only REM and non-REM sleep. Kemp et al. [44] modeled rates of transition among the various sleep stages, but either assumed constant rates throughout the night or estimated their dynamics by smoothing observed transition frequencies by hand. More recently, Karlsson et al. [1] and Kjellsson et al. [3] proposed a Markov-chain model using a mixed-effect approach and modeling the transition probabilities among different stages as binary logistic functions. Finally, the model I validate here [2] is a refinement of the latter model, which considers the transition probabilities as multinomial logistic functions.

This chapter provides some mathematical background on Markov-chain models, the description of the sleep data used for its implementation and the presentation of specific features adopted in order to maximize the data likelihood.

3.2. Markov-chains: mathematical definitions

Let a random process be a finite sequence $X = \{X_t\}_{t\geq 1} = \{X_1, X_2, \dots\}$ of random variables taking values in a discrete set $\mathcal{S}$. The elements of $\mathcal{S}$ are called '*states*' of the system and thus $\mathcal{S}$ the '*state space*'.

The index t of X_t is usually thought as a '*time index*', even though it is not necessarily related to the concept of time, but rather expressing the ordered evolution of the process. Consequently, X_t represents the state of the process at '*time*' z_t, where z is the '*time*' vector.

The process $X = \{X_1, X_2, \dots\}$ is called *'Markov chain'* if it the conditional probability between the outcomes at different times satisfies the '*Markov property*'.

Definition 3.1. The sequence $\{X_1, X_2, \dots\}$ of random variables taking values in $\mathcal{S}$ is said to have the *Markov property* if

$$P(X_{t+1} = x_{t+1} \mid X_t = x_t,\ X_{t-1} = x_{t-1}, \dots,\ X_1 = x_1) = P(X_{t+1} = x_{t+1} \mid X_t = x_t)$$

for every sequence $x_1, \dots,\ x_t,\ x_{t+1}$ of elements of $\mathcal{S}$ and for every $t \geq 1$.

This property states that the probability of an event one step into the future conditioned on the entire past up to the present time *t* is equal to the conditional probability of the future event given just the present one. In particular, a sequence of such random variables is said to be a '*First-order Markov chain process*' because each outcome depends exclusively on the previous state.

Definition 3.2. The process *X* is an '*Nth-order Markov chain process*' if the dependency between the random variables constituting the process involves *N* successive steps in the sequence, that is if the probability of the future outcome is conditioned on the *N* previous states:

$$P(X_{t+1} = x_{t+1} \mid X_t = x_t, \dots,\ X_1 = x_1) = P(X_{t+1} = x_{t+1} \mid X_t = x_t, \dots,\ X_{t-N} = x_{t-N}).$$

Definition 3.3. Given a First-order Markov chain process $X = \{X_1, X_2, \dots\}$ with k and m in its state space $\mathcal{S}$, the conditional probability

$$p_{km}(t) = P(X_{t+1} = m \mid X_t = k) \geq 0$$

is called the '*transition probability*' from *k* to *m* at time z_t. If the transition probabilities do not depend on time, i.e. $p_{km}(t) = p_{km}(t+h) = p_{km},\ \forall t \in \mathbb{N}^+, \forall h \in \mathbb{N}^+$, the Markov chain is said to be '*time-homogeneous*'.

Assuming a finite state space $\mathcal{S}$, that is $\mathcal{S} = \{0, 1, \dots, S\}$;, it is useful to collect the transition probabilities from/to the states of $\mathcal{S}$ in a matrix:

$$P(t) = \begin{bmatrix} p_{00}(t) & p_{01}(t) & \cdots & p_{0S}(t) \\ p_{10}(t) & p_{11}(t) & \cdots & p_{1S}(t) \\ \vdots & \vdots & \vdots & \vdots \\ p_{S0}(t) & p_{S1}(t) & \cdots & p_{SS}(t) \end{bmatrix},,$$

which is called '*transition probability matrix*'. Each row of $P(t)$ represents all the transition probabilities from a single state of $\mathcal{S}$: therefore, the probabilities in each row must sum up to 1:

$$\sum_{m=0}^{S} p_{km}(t) = 1\,, \qquad \forall\, i \in \mathcal{S}.$$

Assuming to observe M independent realizations of the process of the same length *N*:

$$X^{(1)} = \{x_{11}, x_{12}, x_{13}, \dots, x_{1N}\}$$
$$X^{(2)} = \{x_{21}, x_{22}, x_{23}, \dots, x_{2N}\}$$
$$\dots\dots$$
$$X^{(M)} = \{x_{M1}, x_{M2}, x_{M3}, \dots, x_{MN}\},$$

it is possible to define some statistics depending on the data:

- $N_{km} \triangleq$ number of transitions from state *k* to state *m* in all the realizations,
- $N_k \triangleq$ number of transition starting from state *k,*
- $TR \triangleq$ total number of transitions in the data,
- $SO_k \triangleq$ number of occurrences of state *k* in the data.

Once these statistics are available, the frequency of occurrence of each stage is computed as:

$$\hat{\rho}_k = \frac{SO_k}{\sum_{k=0}^{S} SO_k},\ \forall\, k\ \in \mathcal{S}, \qquad [3.1]$$

,

and the '*transition frequencies'* between stages are calculated as:

$$\hat{f}_{km} = \frac{N_{km}}{N_k},\ \forall\, k, m \in \mathcal{S} \qquad [3.2]$$

Every single realization of the process can be considered as a path in time through the state space, and its probability

$$P((X_1, \dots, X_t) = (x_1, \dots, x_t))$$

is just the joint probability of $(X_1, \dots, X_t)$. It is possible to demonstrate that, according to the Markov property characterizing the process, such joint probability can be expressed in terms of the transition probabilities and the probability mass function of X_1:

Theorem 3.1. For a Markov chain X and for any path $\{x_1, x_2, x_3, \dots, x_t\}$, the conditional probability of the path conditioned on the first value is the product of the transition probabilities between successive states of the path:

$$P\big((X_2, \dots, X_t) = (x_2, \dots, x_t) \big| X_1 = x_1\big) = p_{x_1 x_2} p_{x_2 x_3} \cdots p_{x_{t-1} x_t} \quad [3.3]$$

and consequently the probability (not-conditioned) of the path is:

$$P((X_1, \dots, X_t) = (x_1, \dots, x_t)) = P(X_1 = x_1)\, p_{x_1 x_2} p_{x_2 x_3} \cdots p_{x_{t-1} x_t} \quad [3.4]$$

3.3. Implementation of Markov-chains is sleep data modeling

As already mentioned, sleep data can be considered as realizations of a Markov-chain process with S = (AW, ST1, ST2, ST3, ST4, REM).

The model developed in [2] is a time-non-homogeneous Markov-chain model which uses multinomial logistic functions for describing transition probabilities between states and uses a mixed-effect approach for describing their parameters.

The following paragraphs present the clinical data obtained from PSG together with the specific model features applied to the Markov-chain model as proposed by Bizzotto et al. [2].

3.3.1. Clinical study

Data were obtained from a polysomnograghic multi-centre, randomized, double-blind, placebo-controlled, parallel three-arm study designed to investigate a new candidate drug [Fig 3.2]. Male and female subjects (18-64 years of age) diagnosed with primary insomnia were chosen as feasible candidates for the study. The eligibility of the subjects was determined on the basis of specific PSG variables (e.g., LPS, WASO) obtained after a screening period consisting of a first clinical screening visit followed by a 2-night PSG recording in a sleep laboratory. After a week of daily placebo administration, subjects were randomized in the study, each arm assuming placebo or two different doses of the drug for 28 days before bedtime. PSG was recorded in three occasions in two consecutive days (1-2, 13-14 and 27-28) for each arm of the study.

Subjects taking part in the selection for this study had a diagnosis of primary insomnia and insomnia symptoms for at least three months, according to the *'Diagnostic and Statistical Manual*

of Mental Disorders - Fourth Edition - Text Revision (DSM - IV - TR)', [33] criteria 307.42. For being included in the study, the mean of PSG variables obtained after the two screening nights had to fall within the following ranges:

- mean TST: between 240 and 390 minutes,
- mean LPS: more than 30 minutes and not less than 20 minutes on either night,
- mean WASO: more than 60 minutes and neither night less than 45 minutes.

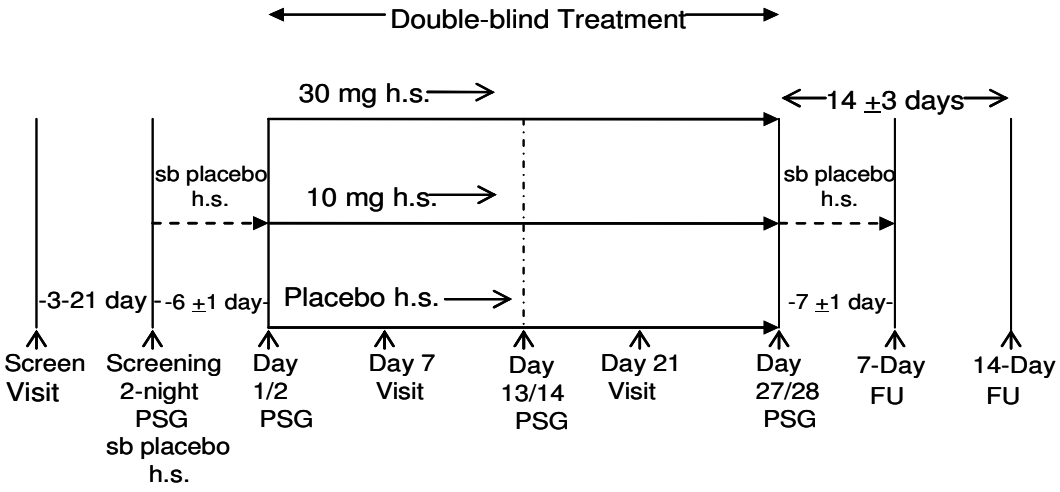


Figure 3.2: Clinical study protocol.

The model was developed based only on the first night of the double-blind treatment nights, from M=116 patients treated with placebo. Since few epochs of stages 3 and 4 were reported, they were merged in a single stage called 'slow wave sleep' stage.

Therefore the stages considered are the awake stage (AW), stage 1 sleep (ST1), stage 2 sleep (ST2), slow wave sleep (SWS) and REM sleep (REM) and the state space is $\mathcal{S}$ = {AW, ST1, ST2, SWS, REM}.

The sequence of sleep stages can be thought as:

$$X^{(1)} = \{x_{11}, x_{12}, x_{13}, \dots, x_{1N}\}$$
$$X^{(2)} = \{x_{21}, x_{22}, x_{23}, \dots, x_{2N}\}$$
$$\dots\dots$$
$$X^{(M)} = \{x_{M1}, x_{M2}, x_{M3}, \dots, x_{MN}\},$$

where $N = 960$ is the number of samples for each subject and $M = 116$ is the number of subjects.

3.3.2. *Multinomial logistic function*

Let x_{it} represent the state (i.e. the sleep stage) of the *i*-th patient at epoch *t* and let each single realization of the process (i.e. each patient's sequence) obey to a first-order Markov-chain, according to Definition 3.1.

Let then the binary variable y_{ikmt} represent the transition of the *i*-th individual from state *k* at epoch (*t*-1) to state *m* at epoch *t*, that is:

$$y_{ikmt} = \begin{cases} 1 \quad \text{if } x_{i(t-1)} = k \text{ and } x_{it} = m, \text{ with } k, m \in \mathcal{S} \\ 0 \quad \text{otherwise} \end{cases}$$

Then, for given values of $k \in \mathcal{S}$ (i.e. the starting state of a transition) and for a given time $t \in \{1, 2, \dots, N\}$ the vector

$$\bar{y}_{ikt} = [y_{ikAWt}, y_{ikST1t}, \dots, y_{ikREMt}]$$

is a multinomial random variable representing all the possible transitions from state *k* at time *t*. This multinomial random variable is characterized by its probability vector:

$$\bar{p}_{ik}(t) = [p_{ikAW}(t), p_{ikST1}(t), \dots, p_{ikREM}(t)] \qquad [3.5]$$

where $p_{ikm}(t) = P(x_{it} = m | x_{i(t-1)} = k) \geq 0$ is the probability of moving from *k* to *m* at time *t*. Since $\bar{p}_{ik}(t)$ is the *k*-th row of the transition probability matrix characterizing the process at time *t*,

$$\sum_{m \in \mathcal{S}} p_{ikm}(t) = 1.$$

The model for the transition from the state *k* of the chain is therefore

$$(\bar{y}_{ikt} | \bar{p}_{ik}(t)) \sim Multinomial(\bar{p}_{ik}(t)). \qquad [3.6]$$

In our context, the transition probabilities represent the model parameters to be estimated according to the sleep data. To avoid estimates constrained between 0 and 1, it is often useful to describe such parameters as logistic functions [3].

The logit of a probability p describing a random variable is defined as the logarithm of its odds, that is the logarithm of the probability of achieving a favorable outcome divided by the probability of failing:

$$G = log \frac{p}{(1-p)}$$

The logit function allows therefore to transform a probability into a not constrained variable, as shown in figure 3.3.

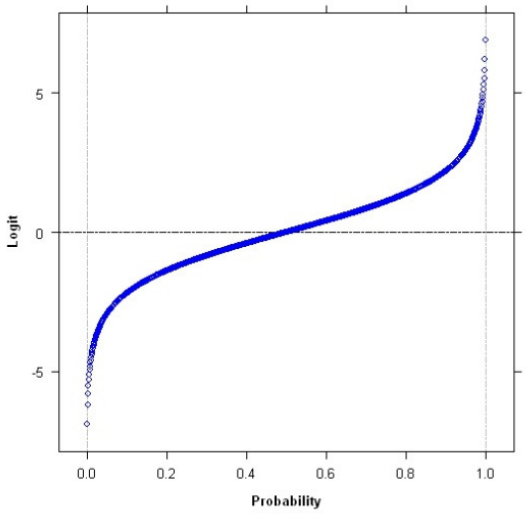


Figure 3.3: The binary logit function.

Since the transitions from a single state of $\mathcal{S}$ come from a multinomial distribution (Equation 3.6), they are transformed into multinomial logit functions. For each subject *i*, each starting state *k* and each epoch *t* the logits for the model are defined as:

$$g_{ikm}(t) = log \frac{p_{ikm}(t)}{p_{ikk}(t)},$$

where *m* takes all the values in $\mathcal{S}$.

In such a way, it is possible to define for each triple (*i*, *k*, *t*) 5 different logit functions, one of which is equal to zero. Taking, for example, the transitions from the AW state at time *t*, it is possible to define the following logits:

$$\begin{cases} g_{iAWAW}(t) = log\frac{p_{iAWAW}(t)}{p_{iAWAW}(t)} \\ g_{iAWST1}(t) = log\frac{p_{iAWST1}(t)}{p_{iAWAW}(t)} \\ g_{iAWST2}(t) = log\frac{p_{iAWST2}(t)}{p_{iAWAW}(t)} \\ g_{iAWSWS}(t) = log\frac{p_{iAWSWS}(t)}{p_{iAWAW}(t)} \\ g_{iAWREM}(t) = log\frac{p_{iAWREM}(t)}{p_{iAWAW}(t)} \end{cases}$$

The corresponding transition probabilities can be therefore obtained as multinomial logistic functions:

$$\begin{cases} exp(g_{iAWAW}(t)) = \frac{p_{iAWAW}(t)}{p_{iAWAW}(t)} \\ exp(g_{iAWST1}(t)) = \frac{p_{iAWST1}(t)}{p_{iAWAW}(t)} \\ \cdots \\ exp(g_{iAWREM}(t)) = \frac{p_{iAWREM(t)}}{p_{iAWAW(t)}} \end{cases}$$

$$\begin{cases} p_{iAWAW}(t) = p_{iAWAW}(t) \\ p_{iAWST1}(t) = exp(g_{iAWST1}(t)) \cdot p_{iAWAW}(t) \\ \cdots \\ p_{iAWREM}(t) = exp(g_{iAWREM}(t)) \cdot p_{iAWAW}(t) \end{cases} \quad [3.7]$$

Recalling then that their sum must be equal to 1:

$$p_{iAWAW}(t) + p_{iAWST1}(t) + p_{iAWST2}(t) + p_{iAWSWS}(t) + p_{iAWREM}(t) = 1$$
$$p_{iAWAW}(t) + exp(g_{iAWST1}(t)) \cdot p_{iAWAW}(t) + \cdots + exp(g_{iAWREM}(t)) \cdot p_{iAWAW}(t) = 1$$

$$p_{iAWAW}(t) = 1/(1 + exp(g_{iAWST1}(t)) + exp(g_{iAWST2}(t)) + exp(g_{iAWSWS}(t)) + exp(g_{iAWREM}(t)))$$

it is possible to compute all the transition probabilities from AW substituting p_{iAWAWt} in Equations 3.7.

Therefore the probability of the transition from *k* to *m* at time t is given by:

$$p_{ikm}(t) = \frac{exp(g_{ikm}(t))}{\sum_{m \in \mathcal{S}} exp(g_{ikm}(t))} \quad [3.8]$$

Hence, instead of the probability vectors $\bar{p}_{ik}(t)$ defined in equation 3.5 the parameters of the model are the corresponding logit vectors:

$$\bar{g}_{ik}(t) = [g_{ikAW}(t), g_{ikST1}(t), \dots, g_{ikREM}(t)],\ k \in \mathcal{S}, \mathrm{i} \in \{1, \dots, M\}, t \in \{1, \dots, N\} \quad [3.9]$$

which fully characterize the model. Note that if no correlation is assumed between logits with different values for *k*, i.e. for different stages of departure, the model can be divided into five different smaller models. Each sub-model, referred to as *'sub-model k'*, describes the transitions from a specific sleep stage and its parameters can be identified separately from the others.
Each sub-model is estimated using a non-linear mixed-effect approach for taking the variability of the population into consideration. That is, each individual logit $g_{ikm}(t)$ is thought to be normally distributed around its typical value:

$$g_{ikm}(t) \sim N(g_{km}(t), \omega^2{}_{km}(t)),$$

where $g_{km}(t)$ is the typical value for the logit and $\omega^2{}_{km}(t)$ is the variance of the inter-individual distribution. Considering the vector-matrix notation

$$\bar{g}_{ik}(t) \sim N(\bar{g}_k(t), \Omega_k(t)) \quad [3.10]$$

where $\bar{g}_{ik}(t)$ is the vector of Equation 3.9 and is assumed normally distributed around the vector of population values for the logit functions $\bar{g}_k(t)$ with covariance matrix $\Omega_k(t)$.

3.3.3. Time dependence of model parameters

As introduced in the first chapter, sleep physiology varies over night time. Consequently, it is logical to assume that the transition probabilities of our model depend on time, this means that the Markov-chain model is non-homogeneous. The temporal dependence of logit functions $\bar{g}_{ik}(t)$ with respect to night time is modeled as a piecewise linear function with three break-points BP = (BPA, BPB, BPC). BPA and BPC are selected at the beginning and at the end of the night window, while BPB is estimated according to the maximum likelihood principle and is assumed to be common to the whole population (no inter-individual variability on it).

The parameters of the model defined so far are the logit population values at the 3 break-points, the associated inter-individual variability and BPB. The population logits are expressed by the following vectors:

$$\bar{g}_{kA} = [g_{kAWA}, g_{kST1A}, \cdots, g_{kREMA}]$$
$$\bar{g}_{kB} = [g_{kAWB}, g_{kST1B}, \cdots, g_{kREMB}]$$
$$\bar{g}_{kC} = [g_{kAWC}, g_{kST1C}, \cdots, g_{kREMC}],$$

in which the letters A, B, C refer to the night time break-points BPA, BPB, BPC. The individual deviation of a specific logit function from its typical value is constrained to be constant at the different break-points, meaning that the individual logits in BP can be expressed by:

$$\begin{aligned} \bar{g}_{ikm} &= [g_{ikmA}, g_{ikmB}, g_{ikmC}] \\ &= [(g_{kmA} + \hat{g}_{ikm}), (g_{kmB} + \hat{g}_{ikm}), (g_{kmC} + \hat{g}_{ikm})] \end{aligned} \quad [3.11]$$

where $g_{kmA}, g_{kmB}, g_{kmC}$ are the typical population values of the logit characterizing the transition from *k* to *m* at times BP = [BPA, BPB, BPC] and $\hat{g}_{ikm}$ is the individual deviation from this logit.

The individual deviation $\hat{g}_{ikm}$ is assumed to be drawn from a normal distribution having zero mean and variance ${\omega^2}_{km}$, that is: $\hat{g}_{ikm} \sim N(0, {\omega^2}_{km})$.

Once $\bar{g}_{ikm}$ are known, $\forall i \in \{1, ..., M\}$, $k \in \mathcal{S}$ and $m \in \mathcal{S}\text{-}\{k\}$, each logit time-course over the night, $g_{ikm}(t)$, is given by linear interpolation of the logit values at two adjacent break-points.

Going back to the vector-matrix notation, Equation 3.10 becomes

$$\bar{g}_{ik}(t) \sim N(\bar{g}_k(t), \Omega_k) \quad [3.12]$$

where the covariance matrix Ω_k is supposed to be not time-depending.

Finally, the transition probabilities profiles are calculated as anti-logit, as described in Equation 3.8. The typical probability profiles estimated from the available dataset for the transitions from ST1 are shown in Figure 3.4. This plot highlights the property of the Markov-chain model built on multinomial logistic functions: by construction, probabilities sum up to one for each time of the night. Figure 3.5 shows the probability profile for the specific transition from ST1 to ST2 with the corresponding 90% prediction interval for the population distribution.

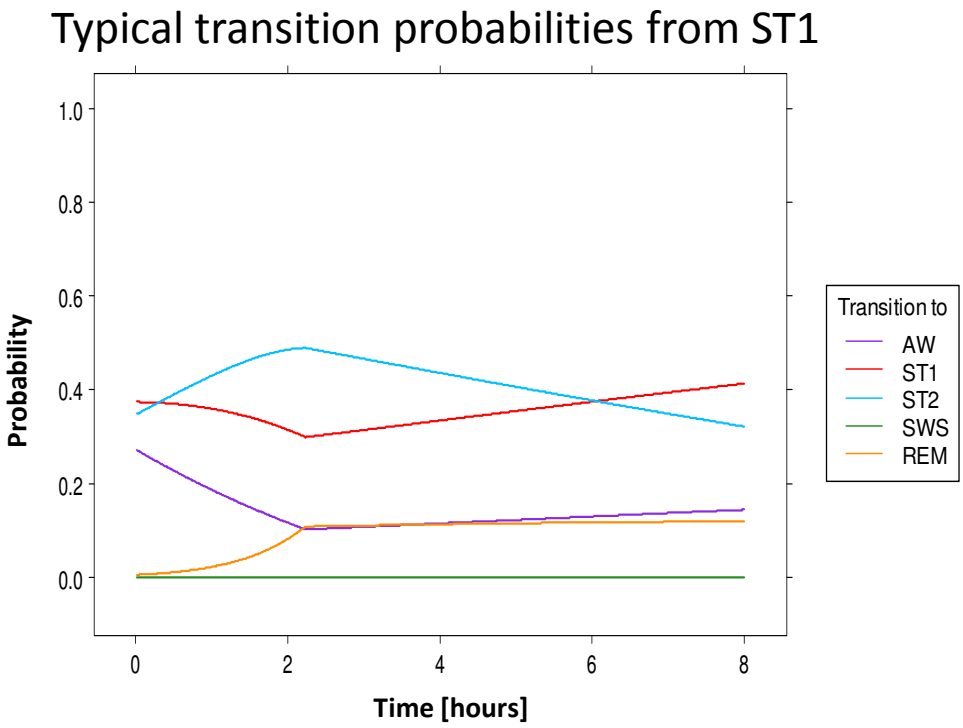


Figure 3.4: All the typical transition probabilities from ST1 estimated from the available dataset

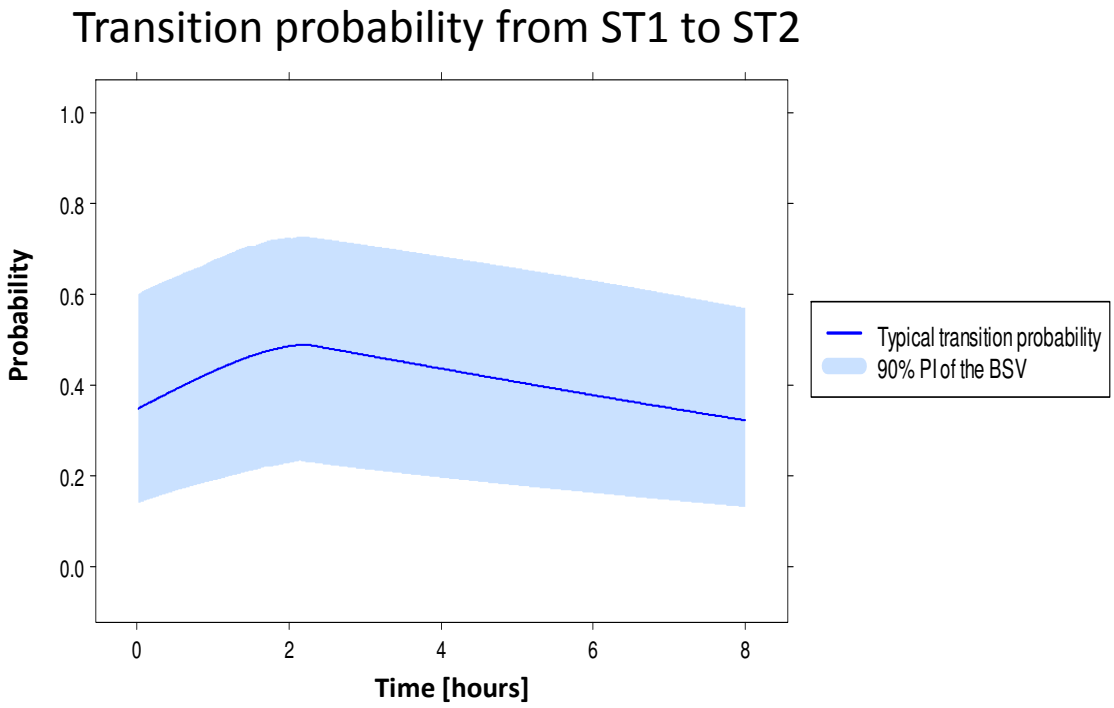


Figure 3.5: Typical transition probability from ST1 to ST2 with the corresponding 90% prediction interval on the between subject variability.

3.3.4. *Stage time effect*

According to the sleep physiology, as presented in Chapter 1, the probability of moving from or staying in a specific stage *k* is influenced by the duration of contiguous time spent in stage *k*. This duration, also defined as the time elapsed since the last change in sleep stage, is called '*stage time*', (t_s or STT). The stage time has been introduced in the model as a predictor for the parameter values and is supposed to modify the logits at the night-time break-points according to an additive law. A temporal grid is defined on the stage time and three break-points, BPs = [BPsa, BPsb, BPsc] are defined on it:

- BPsa, the first break-point, at t = 1 epoch, that is the minimum stage time that can be observed
- BPsc, the last break-point, at the maximum stage time observed in the data with respect to the particular stage of departure
- BPsb, the central one, to be estimated as a parameter of the model.

The temporal grid is therefore divided into two segments on which the '*stage time effect*' (STE) is modeled as a piecewise linear function of stage time. Considering a generic stage $k \in \mathcal{S}$ of departure new vectors are introduced $\forall m \in \mathcal{S}\text{-}\{k\}$:

$$\bar{s}_{km} = [s_{kma}, s_{kmb}, s_{kmc}] \qquad [3.13]$$

where s_{kma}, s_{kmb}, s_{kmc} are the additive effects that stage time has on the logit defined on the transition from k to m for t_s = BPsa, t_s = BPsb and t_s = BPsc, respectively. Obviously, no effect is assumed for t_s = BPsa, ($s_{kma} = 0, \ \forall k, m \in \mathcal{S}$). In each segment of the temporal grid the deviation from the logit predicted by the model is therefore calculated by linear interpolation as:

if $BPsa \leq t_s \leq BPsb$

$$s_{km}(t_s) = s_{kma} \cdot (BPsb - t_s)/(BPsb - BPsa) + s_{kmb} \cdot (t_s - BPsa)/(BPsb - BPsa)$$
$$= s_{kmb} \cdot (t_s - BPsa)/(BPsb - BPsa)$$

endif

if $BPsb < t_s \leq BPsc$

$$s_{km}(t_s) = s_{kmb} \cdot (BPsc - t_s)/(BPsc - BPsb) + s_{kmc} \cdot (t_s - BPsb)/(BPsc - BPsb)$$

endif

being $s_{km}(t_s)$ the stage time effect for stage time equal to t_s.

No inter-individual variability is assumed, so that the stage time influences the whole population in the same manner. Moreover, each stage time effect $s_{km}(t_s)$ is supposed to equally modify the

corresponding logit values $\bar{g}_{ikm}$ at the three night time break-points BP = [BPA, BPB, BPC]. New vectors including the stage time effect are therefore defined in place of those of Equation 3.1:

$$\bar{G}_{ikm}(t_s) = [G_{ikmA}(t_s), G_{ikmB}(t_s), G_{ikmC}(t_s)\,] \qquad [3.14]$$
$$= [(g_{kmA} + \hat{g}_{ikm} + s_{km}(t_s)), (g_{kmB} + \hat{g}_{ikm} + s_{km}(t_s)), (\,g_{kmC} + \hat{g}_{ikm} + s_{km}(t_s))]$$

Once $\bar{G}_{ikm}(t_s)$ are known, $\forall i \in \{1, ..., M\}$, $k \in \mathcal{S}$ and $m \in \mathcal{S}$-$\{k\}$, each logit time-course over the night ($G_{ikm}(t, t_s)$, function of t_s and t) is given by linear interpolation of the logit values at two adjacent break-points.

Figure 3.6 exemplifies how the stage time effect (red line) modifies an individual logit (blue line).

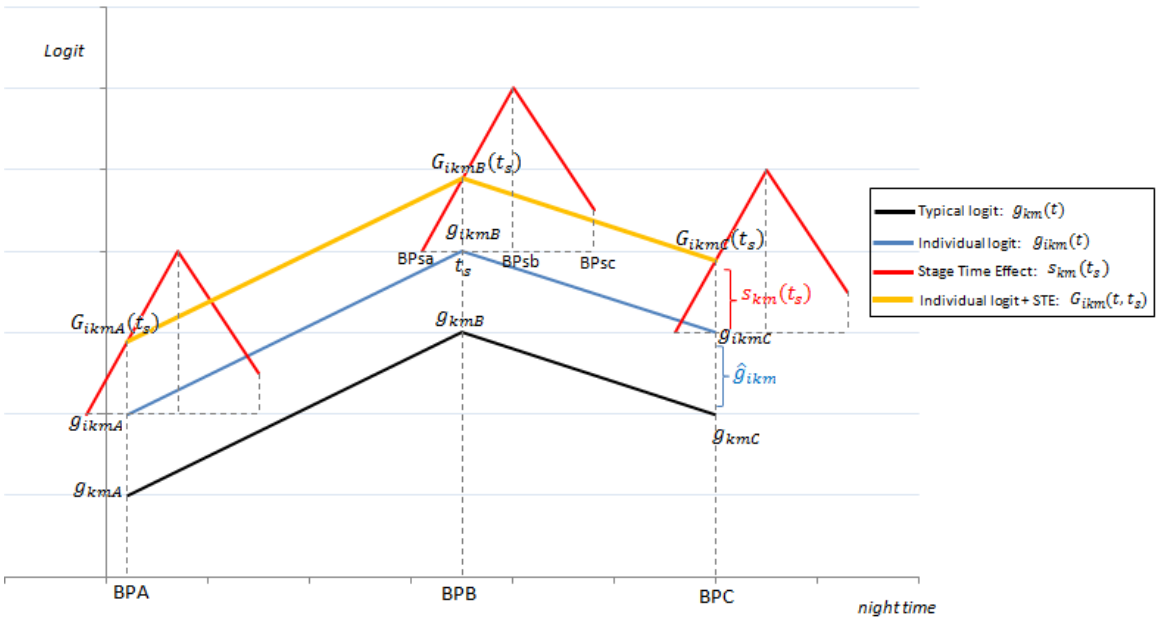


Figure 3.6: An illustrative example on how the individual deviation from the typical behavior and the stage time effect (STE) can affect a logit profile over the night time.

The stage time effects estimated on the available data over the stage time are shown as exponential terms in Figure 3.7. The choice to report an exponential value is related to the logit transformation properties. In this way the additive effect of the STE on the logits is equivalent to a multiplicative effect on probabilities ratios. For example, 'AW / REM' in the 'REM' plot indicates the multiplicative STE on $\frac{p_{iREMAW}(t)}{p_{iREMREM}(t)}$ obtained as the exponential for the additive STE on $g_{iREMAW}(t) = log\,\frac{p_{iREMAW}(t)}{p_{iREMREM}(t)}$.

The night time profiles of the transition probabilities are again calculated as anti-logit from $G_{ikm}(t, t_s)$, as in Equation 3.8.

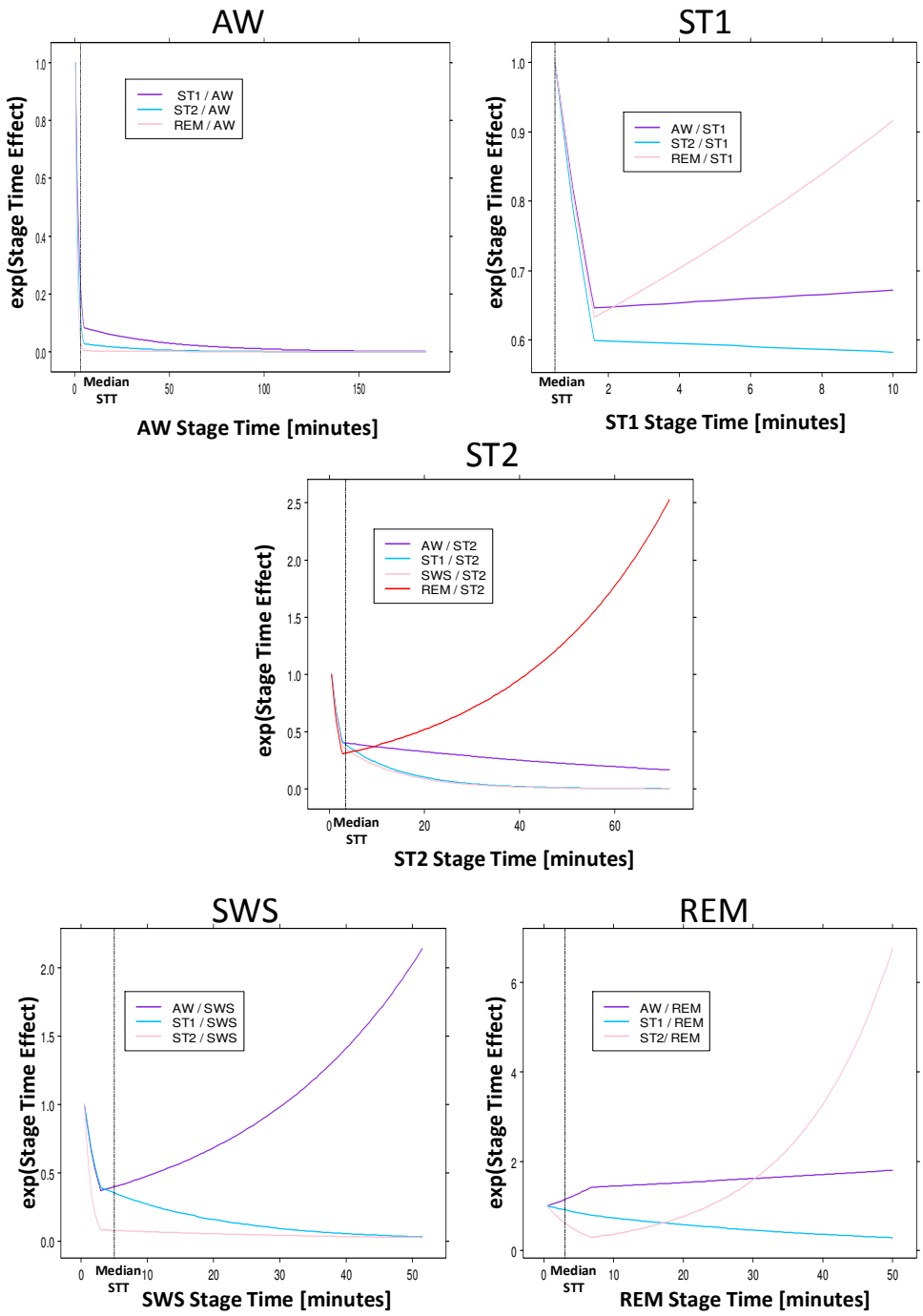


Figure 3.7: Exponential of the stage time effects estimated on the available dataset for all the logits.

Figure 3.8 presents the estimated probability profiles for the transitions from ST1, ST2, SWS and REM at different stage times. In particular, the left column shows the probability profiles estimated for t_s eqaul to 0.5 minutes (i.e. 1 epoch, when no stage time effect is considered), while the right one shows those obtained for t_s equal to the median stage time over the whole night time. The left and right plots for transitions from ST1 are equal, since the median stage time in this case is 1 epoch only (reflecting the fact that stage 1 sleep is a rapid transitional state in sleep). Differently from the plots that illustrates the transition probabilities along the night time, the stage

time appears to be an informative predictor as shown by the dynamic relation between it and the STE (Figure 3.7). Plots for transitions from awake are shown in Paragraph 3.2.7, after the inclusion of an additional model feature.

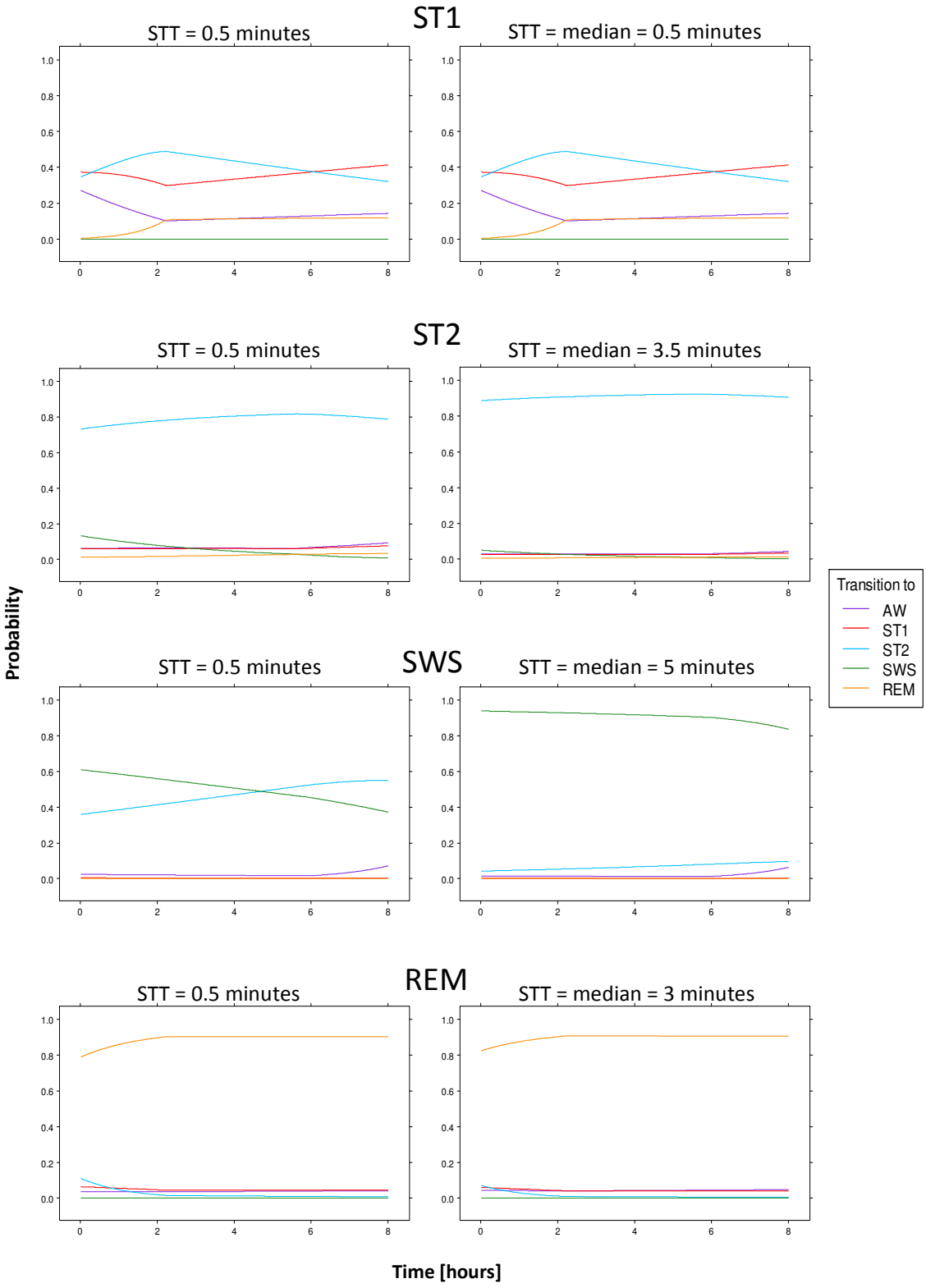


Figure 3.8: Typical transition probabilities for the *stage1*, *stage2*, *sws* and *rem* sub-models. In the left column the profiles are computed for stage time = 0.5 minutes, while in the right one they are computed for the median stage time observed in the data.

3.3.5. *Transition probabilities fixed to zero*

It is known that some of the 25 hypothetically possible transitions between sleep stages occur very hardly. The available dataset confirms this statement, as shown by the *'transition frequencies'*, computed as in Equation 3.1, and here reported:

FROM\TO	AW	ST1	ST2	SWS	REM
AW	0.8303	0.1246	0.0321	0.0006	0.0125
ST1	0.1350	0.4183	0.3471	0.0006	0.0990
ST2	0.0399	0.0381	0.8737	0.0362	0.0121
SWS	0.0207	0.0046	0.1418	0.8324	0.0005
REM	0.0522	0.0540	0.0132	0	0.8806

Table 3.1: Transition frequencies computed from data.

Some transition frequencies are very low and are therefore fixed to zero in order to simplify model identification. The chosen frequency threshold is 0.1% so that the transition frequencies highlighted in red are assumed equal to zero.

This implies that the sub-models *awake*, *stage1*, *sws* and *rem* need one less logit each. Summarizing, the logits considered for each sub-model are shown in table 3.2.

logit\sub-model	*awake*	*stage1*	*stage2*	*sws*	*rem*
1	$G_{iAWST1}(t,t_s)$	$G_{iST1AW}(t,t_s)$	$G_{iST2AW}(t,t_s)$	$G_{iSWSAW}(t,t_s)$	$G_{iREMAW}(t,t_s)$
2	$G_{iAWST2}(t,t_s)$	$G_{iST1ST2}(t,t_s)$	$G_{iST2ST1}(t,t_s)$	$G_{iSWSST1}(t,t_s)$	$G_{iREMST1}(t,t_s)$
3	$G_{iAWREM}(t,t_s)$	$G_{iST1REM}(t,t_s)$	$G_{iST2SWS}(t,t_s)$	$G_{iSWSST2}(t,t_s)$	$G_{iREMST2}(t,t_s)$
4			$G_{iST2REM}(t,t_s)$		

Table 3.2: Logits for the different sub-models.

3.3.6. *Inter-individual variability*

The assumption explained in the previous paragraph implies that the covariance matrices Ω_k of *awake*, *stage1*, *sws* and *rem* sub-models have dimension 3x3, while the covariance matrix for *stage2* sub-model has dimension 4x4. Full covariance matrices were initially assumed.

The statistical significance of each element of these matrices was assessed according to '*log likelihood ratio test*'. Only correlations that achieved statistical significance were included in the final model and are here reported:

- *awake*, using an inter-individual covariance between:

(i) $G_{iAWST1}(t, t_s)$ and $G_{iAWST2}(t, t_s)$ when $t_s = 0$,

- *stage1*, using an inter-individual covariance between:

(i) $G_{iST1AW}(t, t_s)$ and $G_{iST1ST2}(t, t_s)$, $(t_s = 0)$,

(ii) $G_{iST1ST2}(t, t_s)$ and $G_{iST1REM}(t, t_s)$, $(t_s = 0)$,

(iii) $G_{iST1AW}(t, t_s)$ and $G_{iST1REM}(t, t_s)$, $(t_s = 0)$.

Using the following notation for expressing the covariance elements of Ω_k:

$$\omega^2{}_{kmn} = cov\big(G_{ikm}(t, t_s), G_{ikn}(t, t_s)\big) \quad (t_s = 0),$$

the corresponding covariance matrices are:

$$\Omega_{AW} = \begin{bmatrix} \omega^2{}_{AWST1} & \omega^2{}_{AWST1ST2} & 0 \\ \omega^2{}_{AWST2ST1} & \omega^2{}_{AWST2} & 0 \\ 0 & 0 & \omega^2{}_{AWREM} \end{bmatrix},$$

$$\Omega_{ST1} = \begin{bmatrix} \omega^2{}_{ST1AW} & \omega^2{}_{ST1AWST2} & \omega^2{}_{ST1AWREM} \\ \omega^2{}_{ST1ST2AW} & \omega^2{}_{ST1ST2} & \omega^2{}_{ST1ST2AW} \\ \omega^2{}_{ST1REMAW} & \omega^2{}_{ST1REMST2} & \omega^2{}_{ST1REM} \end{bmatrix}.$$

Furthermore, it was found that also the between subject variability on the logit $\bar{G}_{iREMAWt_r}$ of the *rem* sub-model can be fixed to zero, resulting in the following covariance matrix:

$$\Omega_{REM} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \omega^2{}_{REMST1} & 0 \\ 0 & 0 & \omega^2{}_{REMST2} \end{bmatrix}.$$

3.3.7. Initial sleeplessness

One additional feature has been introduced in the model to differentiate the sleep physiology time-course between initial sleeplessness and rest of the night. Consequently, the 8-hour night time is divided into 2 parts: the first ranging from *t* = 2 epochs to *t* = IS, where IS (*'Initial Sleeplessness'*) is the first epoch in which a non-awake state is observed in a specific subject, and the second one filling the remaining part of the night. In the second time interval the logits are

modeled as previously described changing only the position of the first break point of the night time: BPA=IS.

The logits in the first part of the night are modeled again as piecewise linear functions, but no inter-individual variability or stage-time effects are considered. Therefore, the logits in this time interval are simply estimated as population parameters. In particular, three additional break-points are defined, Bpi = [BP1, BP2, BP3], where:

- BP1 is the first one at t = 2 epochs
- BP3 is the last one at the maximum IS observed in the data
- BP2 is the central one and it is considered as an additional parameter to be estimated.

Figure 3.9 shows the estimated typical transition probabilities for the awake sub-model: in the left plot no stage time effect is considered, while in the right one stage time effect is considered and fixed to the value estimated for the median awake time observed in the data. The discontinuity in the profiles for t = BPA is due to the introduction of the Initial Sleeplessness feature.

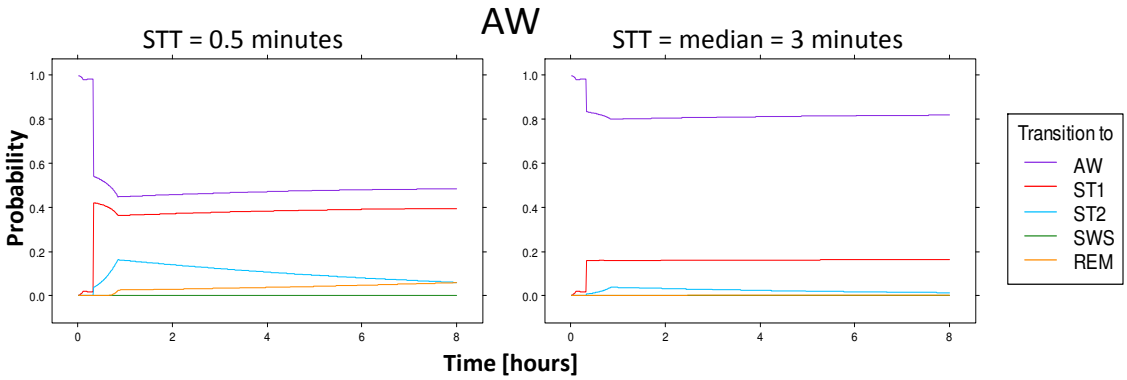


Figure 3.9: Typical transition probabilities for the *awake* sub-model. In the left column the profiles are computed for stage time = 0.5 minutes, while in the right one for the median stage time.

In conclusion, Figure 3.10 shows all the typical transition probabilities with their 90% prediction interval on the between subject variability according to the final model. Note that the profiles for the transitions from AW show inter-individual variability in the second part of the night only, according to model assumptions.

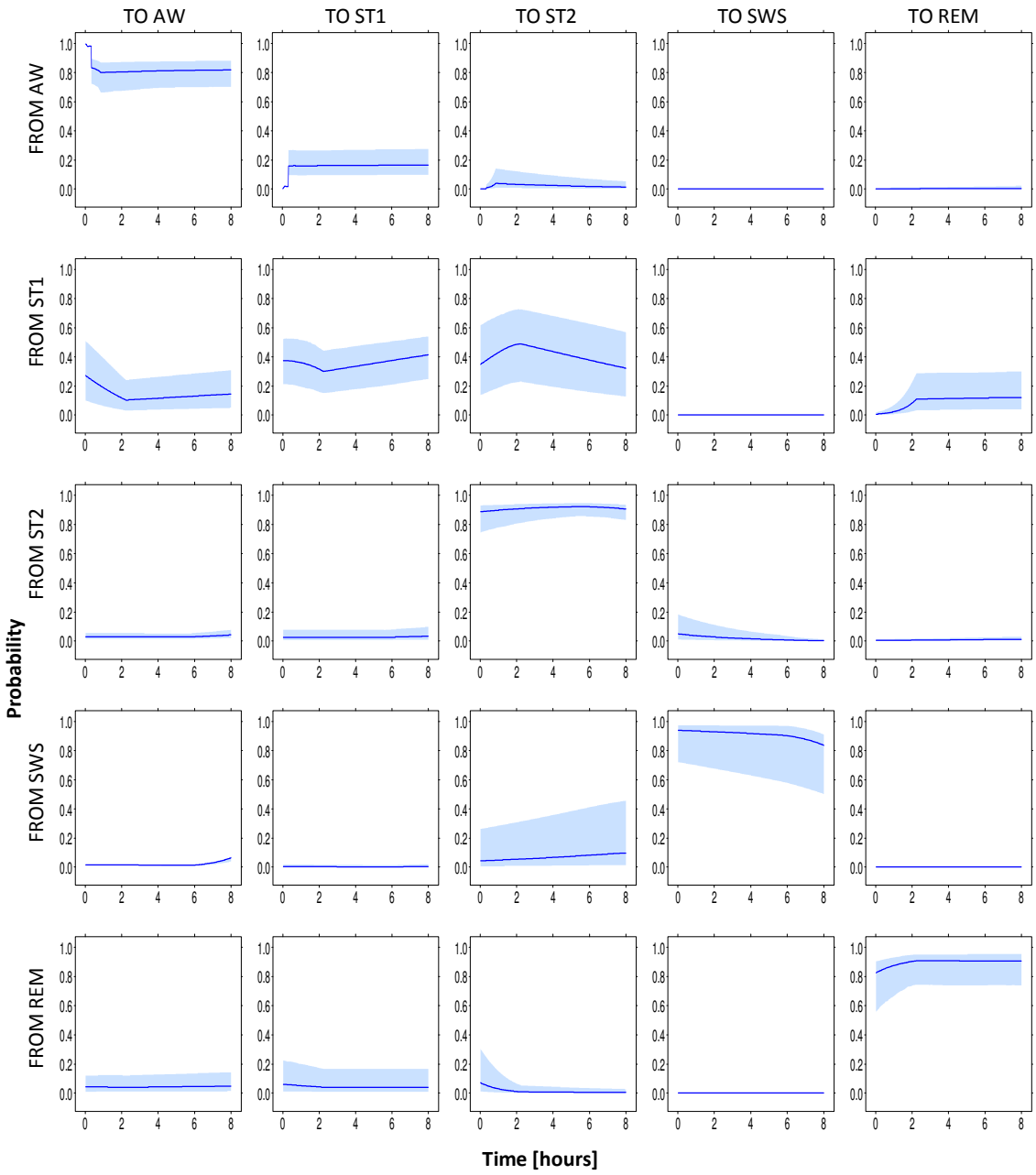


Figure 3.10: All the typical transition probabilities with the corresponding 90% prediction interval on the between subject variability.

3.3.8. Likelihood

Each sub-model was identified using a maximum likelihood approach. As introduced in Chapter 2.5, the individual likelihood is expressed as the integral over all possible values of η_i:

$$L_i = \int l(x_i|\theta, \eta_i) h(\eta_i|\Omega) d\,\eta_i$$

where x_i = data from individual i, θ = population parameters, η_i = random effects on i and h is a multivariate normal density function with zero mean and covariance matrix Ω(covariates are not considered in this model). In our case, i.e. categorical data, the Markov-chain model describes transition probabilities between observations, instead of actual observation values as with continuous data. The probability of a whole realization for the individual i is therefore given by the product of the probabilities of the transitions occurred along the night, as stated in Theorem 3.1, considering as first state of the chain the AW state ($P(X_1 = AW) = 1$). Therefore, $l(x_i|\theta, \eta_i)$ can be expressed as a product of transition probabilities, the ones which actually occurred, so that maximizing the likelihood is equivalent to maximizing the probability of the actually occurred path. Each transition probability is defined in the model as function of various logits, that is

$$p_{ikm}(t) = p_{ikm(t)}\left(\bar{G}_{ik}(t, t_s)\right), \qquad \forall \text{ triple } (i, k, t).$$

Here, *m(t)* indicates that *m* depends on *t* and is equal to x_{it}, and

$$\bar{G}_{ik}(t, t_s) = [G_{ikAW}(t, t_s), G_{ikST1}(t, t_s), G_{ikST2}(t, t_s), G_{ikSWS}(t, t_s), G_{ikREM}(t, t_s)],$$

with t_s equal to the stage time observed at night time t for the subject i ($\bar{G}_{ik}(t, t_s)$ depends on θ and η_i according to the model structure).
Considering sub-model k, the likelihood contribution from individual i is, therefore, defined by

$$L_{ik} = \int \left[\prod_{t: x_{i(t-1)}=k} p_{ikm(t)}\left(\bar{G}_{ik}(t, t_s)\right) \right] h(\eta_i|\Omega) d\eta_i.$$

The overall likelihood function L_k for sub-model k is then the product of the contributions from all the individuals. Each likelihood L_k is maximized using the Laplace method in NONMEM version VI (ICON Development Solutions.) [42].

3.4. Conclusions

The Multinomial Mixed-Effect Markov-chain model presented in detail in this Chapter has been developed [2] with the precise intent of providing a useful tool for analyzing sleep data in their dynamic behaviour along the night. This model therefore can be extensively applied in the analysis of clinical studies that investigate the efficacy of sleep promoting drugs. A proper

evaluation of model performances is therefore of particular importance. Different diagnostic methods to evaluate model adequacy are discussed in Chapter 4 with particular emphasis on those based on stochastic simulation.

Chapter 4

Assessment of the model against data

4.1. Introduction to diagnostic methods

Mixed-effect models are increasingly used in the pharmaceutical field to analyze and interpret clinical data in a context of population framework and to perform clinical trial simulations. The latter, when properly performed, constitute a valid alternative to extensive studies, since they allow the evaluation of the individual responses to a specific treatment, and consequently the drug efficacy and safety in the whole population.

The Markov-model built for describing sleep architecture and characterized in the previous chapter has been mainly developed with the intent of providing a useful tool for clinical trial simulations. The aim of this thesis is therefore to investigate its performance, evaluating model robustness and potential model misspecifications. In practice, we are interest to understand if the proposed model adequately interprets the underlying process, providing precise predictions for the parameters of interest and avoiding biased estimates.

Statistical theory indicates several diagnostic methods to check model adequacy, most of them involving graphics. In 2007, Karlsson and Savic [4] published a detailed overview of the different model diagnostics to be used in the context of mixed-effect modeling, highlighting for each of them pros and cons. The authors divided such methods into five categories:

(a) typical individual prediction-based methods,
(b) individual parameter estimates-based methods,
(c) residual-based diagnostics,
(d) numerical diagnostics,
(e) simulation-based diagnostics.

The (a) and (b) methods are essentially based on graphics: their aim is to visually evaluate whether there is agreement between the dependent variable (DV), i.e. the observations we want to describe, and population or individual model predictions, respectively. The first method can be misleading when applied to non-linear mixed-effect models and often indicates misspecified

models even when they are adequate. The second one suffers from the opposite drawback: in case of sparse information in the individual data, it becomes difficult to recognize misspecified models since excellent predictions are usually produced (‘perfect fit’ phenomenon). The residual based diagnostics (c), consisting in checking population or individual residual, are strictly connected to the latter methods and substantially have the same flaws.

However, these methods (a, b and c), often referred to as *'goodness of fit inspection methods',* are very easily implementable but not applicable in the categorical data context, in which only probability of observations is estimated and, consequently, residuals are not considered. For this reason, they are not used in this thesis.

Several types of numerical diagnostics (d) can be implemented for checking models adequacy in case of categorical data. These methods are essential importance for model comparison (Loglikelihood ratio test), for the evaluation of model robustness (bootstrap) and for the detection of possible overfit (standard errors for the parameters). Most of them were used during model building procedure, but are not part of this thesis.

Finally, the simulation based methods (e) are considered the most interesting diagnostics in the current use: when dealing with categorical data. In fact, these methods appear very attractive because they preserve the same pros of the classical *'goodness of fit inspection methods'* without being affected by their cons. Such diagnostics consist in comparing a desired statistic derived from raw data with a reference distribution obtained trough stochastic simulation.

A stochastic simulation is a methodology which aims to re-produce a desired set of data from one model through random sampling from probability distributions defined by the model. When repeated many times, this procedure becomes a useful tool to show potential model misspecifications. In fact, when the model is adequate the simulations should mimic real data behavior.

In this Chapter such methods together with their application to the final model are discussed in details. However, it is important to highlight that their application, in addition to that of numerical diagnostics, played an essential role also during the whole model building process, allowing to achieve significant model improvements.

In particular, our aim in applying stochastic simulation-based methods to the final model is not only to check its capabilities in predicting the aggregated parameters (WASO, LPS, etc.) used in the clinical practice for quantifying the severity of sleep disorders, but mainly to evaluate its predictability of sleep physiological pattern whose maintenance has been demonstrated to be relevant. The two inspections were performed through the implementation of ’*Posterior Predictive Check*’ (PPC) on the aggregated parameters and ’*Visual Predictive Check*‘ (VPC) on specific statistics derived from data, as the transition frequencies and the stage frequencies.

However, bias and imprecision in the desired predictions may depend not only on possible model misspecifications but also on non robust estimation methods. In the case of the Markov chain

model, which is complex, highly non-linear and dealing with categorical data, this issue could become very relevant.

For this reason, a recent approach [6,7] for checking both model adequacy and estimation method performance is implemented in this thesis: the 'Stochastic Simulations followed by Estimation' (SSE) based diagnostics. SSE is a two-stage method which consists in the following steps:

1. to simulate a data set from a model with the parameter values estimated from the observed data,
2. to re-estimate the same model on the simulated data,
3. to repeat steps 1 and 2 many times.

In addition to the evaluation of model robustness and misspecifications, simulation followed by re-estimation has been proposed for the assessment of the performance of the estimation method. This is particular critical in our specific case where the estimation method (ML approach) is based on the likelihood as defined by the user and its approximation according to the Laplacian method, as implemented in NONMEM VI. These complexities related to the estimation method should be clearly considered in the validation procedure. SSE is used here for the visual comparison between the transition probabilities estimated from raw data and the confidence intervals on the correspondent transition probabilities estimated during SSE procedure. At the best of my knowledge, this kind of visual comparison was never implemented before. We propose to call it '*Visual Estimation Check*' (VEC), that reflects the similarity with 'VPC' relative to the evaluation of the 'estimation' procedure.

4.2. Simulation-based diagnostics: theory

The theory of Bayesian inference is now recalled in order to give an adequate theoretical basis to the simulation-based validation methods implemented here. In general, statistical inference consists in making conclusions from numerical data about quantities that are not directly observable or quantities potentially observable but not yet observed. In the Bayesian framework, the model for describing such quantities is a *'full probability model*' which combines the information derived from data and a priori knowledge on parameters.

Let X be the matrix of the observed data, in which each row contains, for example, scores belonging to a single subject at different times ($x_{it}, i \in [1, \dots, M], t = 1, \dots, N$). $\tilde{X}$ is the unknown but potentially observable matrix of future outcomes. Let Θ be the vector of unobserved parameters of interest, where $\Theta = (\theta, \Omega)$, with θ = population parameters and Ω = variance-covariance matrix for between-subject variability.

In order to make inferences about Θ given a realization of the process, X, a model providing a joint probability distribution for Θ and X is needed:

$$p(\Theta, X) = p(\Theta)p(X|\Theta) \quad [4.1]$$

where the joint probability density function is given by the '*Bayes rule*': it is the product of two densities referred to as '*prior distribution*', $p(\Theta)$, and '*likelihood distribution*', $p(X|\Theta)$. The prior distribution reflects the a priori expectations about the parameters before seeing the data), while the likelihood quantifies the probability that the data are generated by that set of parameters.
The inference on parameters is provided by their '*posterior distribution'*, that is the probability distribution of the unknown parameters conditioned on the observed realization:

$$p(\Theta|X) = \frac{p(\Theta,X)}{p(X)} = \frac{p(\Theta)p(X|\Theta)}{p(X)} \quad [4.2]$$

Here, it is possible to omit $p(X)$ (which does not depend on Θ) and rewrite the posterior distribution as:

$$p(\Theta|X) \propto p(\Theta)p(X|\Theta) \quad [4.3]$$

A similar objective is to make inferences about future potentially observable outcomes $\tilde{X}$, after a single realization of the process, X, is available. Before the data X are considered, the distribution of the unknown but observable X is:

$$p(X) = \int p(X, \Theta)d\Theta = \int p(\Theta)p(X|\Theta)d\Theta \quad [4.4]$$

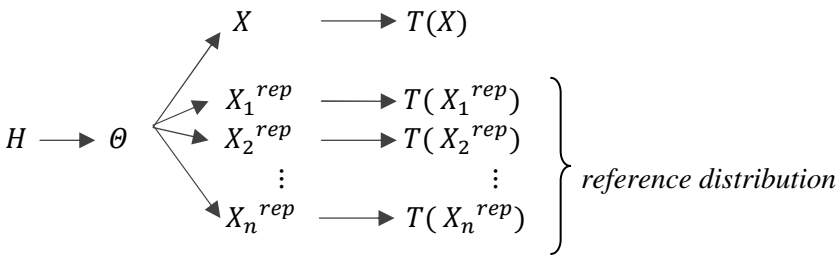
which is the marginal distribution of X ,also referred to as the '*prior predictive distribution*', reflecting the a priori expectation on the outcome. Once the data have been observed, the unknown observable outcome $\tilde{X}$ can be predicted from the same process. The distribution of $\tilde{X}$ given X is called the '*posterior predictive distribution*':

$$p(\tilde{X}|X) = \int p(\tilde{X}, \Theta|X)d\Theta = \int p(\tilde{X}|\Theta, X)p(\Theta|X)d\Theta = \int p(\tilde{X}|\Theta)p(\Theta|X)d\Theta \quad [4.5]$$

and is given by an average of conditional predictions, $p(\tilde{X}|\Theta)$, over the posterior distribution of Θ.

The inference just introduced on $\tilde{X}$ (in the form of [4.5]), can be used to assess model adequacy. In fact, let consider T a specific statistic on the data, i.e. a function from data space to real numbers. As suggested by Gelman [45], the '*posterior predictive check*' consists in the comparison between the observed value of $T(X)$ and the distribution of $T(\tilde{X})$ induced by the posterior predictive distribution [4.5]. This comparison provides a measure of the plausibility of the realization X.

Let consider now to repeat n times the experiment that produced the current data X using the same model, H, and the same unknown value of Θ. Let ${X_j}^{rep}, j = 1, \dots, n$ be the set of replicated data. The desired posterior distribution [4.5] can be therefore approximated with the distribution obtained by ${X_j}^{rep}, j = 1, \dots, n$.



In practice, the posterior distribution of $\tilde{X}$ is obtained through simulation of new data using the model, H. Because Θ is unknown, but assumed to have the same value that generated X, simulations can be done using draws from its posterior distribution given X, $p(\Theta|X)$. Finally, the observed statistic, $T(X)$, can be compared to the reference distribution obtained computing T for each replicated dataset.

In a non Bayesian context, the posterior distribution for the parameters, $p(\Theta|X)$, is not provided by the estimation method: in such cases an approximation to this distribution is needed. In particular, since a maximum likelihood (ML) approach is adopted for the estimation step (see Chapter 2), the approximation is based on the ML estimate $\hat{\theta}$. Then the '*degenerate distribution*' $\Theta = \hat{\theta}$ with probability equal to 1 is used to simulate new datasets.

4.3. Stochastic simulations

Once each different sub-model (*awake*, *stage1*, *stage2*, *sws* and *rem*) is identified using the maximum likelihood approach implemented in NONMEM VI, a set of population parameters, including both typical parameters (θ) and inter-individual variance-covariance (Ω), is available for the simulation step, to be repeated *n* times (*n* is chosen equal to 100).

The aim of this procedure is to generate from the developed model and its estimated parameters a new dataset composed by the polysomnogram outcome for M = 116 (as in the observed dataset) new patients. A model that includes all the 5 sub-models is therefore needed in order to produce a new sequence of sleep stages for each potential patient. This model is written for NONMEM VI and its code is presented in Appendix A. The code can be summarized through the following procedure:

1. The AW state is assumed at the beginning of the night for subject i: x_{it}= AW for t = 1 epoch, and the starting state k is equal to AW.
2. t is increased by 1 and a k-sub-model, *sm*, is chosen according to k value (*sm* = *awake* if k = AW, *stage1* if k = ST1, *stage2* if k = ST2, *sws* if k = SWS, *rem* if k = REM,)
3. The transition probabilities at time t are simulated through the sub-model *sm*: the ML estimates obtained from the estimation step on sub-model *sm* are used to sample and re-construct the individual logits at time t and thereafter the corresponding individual transition probabilities (*pAW*, *pST1*, *pST2*, *pSWS* and *pREM*).
4. The 5 transition probabilities are placed side by side in a probability scale ranging from 0 to 1, as shown in Figure 4.1 on the x-axis.
5. A random variable, called R, is drawn from a uniform distribution in [0, 1].
6. On the basis of the position of R in the probability scale of Figure 4.1, one of the 5 possible transitions is supposed to occur: if, for example, R falls into the *pST1* interval as shown in Figure 4.1, the transition to ST1 is then assumed: $y_{ikmt} = 1$, m = ST1, and, having consequently x_{it} = ST1.
7. Variables providing stage times for each sleep stage and initial sleeplessness length (if a non-awake state has already occurred) are updated according to the value of x_{it}.
8. k assumes x_{it} value.
9. Steps 2 to 8 are repeated until t = 960 epochs.
10. Steps 1 to 9 are repeated for i in [1, …, M].

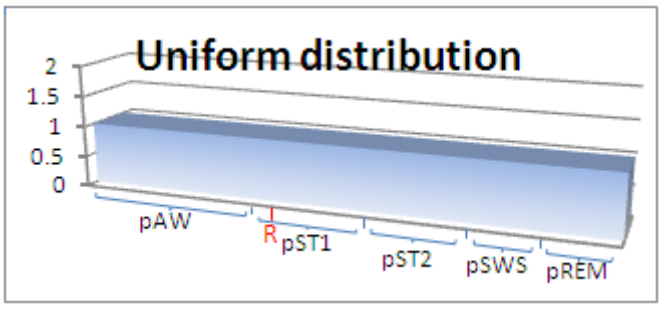


Figure 4.1: Sampling for a multinomial variable with probability vector [pAW, pST1, pST2, pSWS, pREM].

4.3.1. Dataset for simulation

A brief explanation of the datasets used as input files for the simulation model is provided below. Each dataset (real and simulated) is initially formed by four columns indicating the subject identification number (ID), the PSG recording visit (Visit), the night time expressed in epochs (Time) and the corresponding sleep stage (STAGE), taking values in {0, 1, 2, 3, 5} (0 for AW, 1 for ST1, 2 for ST2, 3 for SWS, 5 for REM). Each column length is therefore M x N = 116 x 960.

ID	Visit	Time	STAGE
1	3	1	0
1	3	2	0
1	3	3	1
⋮	⋮	⋮	⋮
1	3	960	3
2	3	1	0
⋮	⋮	⋮	⋮

Five more columns called MDV0, MDV1, MDV2, MDV3 and MDV5 ('MDV' stands for 'missing data value'), indicating whether the previous stage is 0, 1, 2, 3, 5 or not, are added to the dataset. For example, MDV0 is equal to 0 if the previous state is AW, and 1 otherwise. Furthermore, five binary columns called AW, ST1, ST2, ST3 and REM are added for indicating if the current state is 0, 1, 2, 3 or 5, respectively. For example, AW is equal to 1 if the current state is AW, and 0 otherwise.

ID	Time	STAGE	MDV0	MDV1	...	MDV5	AW	ST1	...	REM	STT	SL	IS
1	1	0	1	1		1	1	0		0	0	0	4
1	2	0	0	1	...	1	1	0	...	0	1	0	4
1	3	1	0	1		1	0	1		0	2	1	4
1	4	0	1	0		1	1	0		0	1	1	4
⋮		⋮						⋮					⋮
2	1	0	1	1		1	1	0		0	0	0	30

These columns are useful during the estimation step for extracting relevant data for each single sub-model, and during the validation process for easily computing the statistics of interest. Furthermore, a column counting for the individual stage time (the number of epochs since the last transition) is added and called STT. Finally, due to the introduction of the Initial Sleeplessness feature for the *awake* sub-model, two more columns are needed: a binary column called SL

('sleep'), which is 1 from the epoch of first non-awake state up to the last epoch for each subject, and one called IS ('Initial Sleeplessness'), that reports the number of awake epochs before the first episode of non-awake (IS including also the first non-awake epoch).

4.4. Posterior predictive check (PPC)

The model performance in the prediction of the aggregated parameters (WASO, LPS, etc) was tested through the implementation of the '*posterior predictive check*' (PPC). Since these parameters (or at least some of them) are usually considered as the 'efficacy endpoints' of clinical studies aimed to test efficacy of hypnotic drugs, this term will also be used as synonymous of 'aggregated parameters'.

The individual values for each of these parameters, derived from the observed data, are compared to the corresponding values computed from the simulated data. In particular, considering a specific endpoint, the median of the individual values is computed in each dataset (observed or simulated) and the relative deviations of medians is calculated as follows:

$$RelativeDeviation = \frac{MedianEndpoint_{simulated} - MedianEndpoint_{observed}}{MedianEndpoint_{observed}}.$$

For each endpoint, the distribution of relative deviations is computed and plotted in a box-whisker plot (Figure 4.2) using the R package (R 2.10.0 from the R Development Core, 2009). If the model is adequate, relative deviations need to be distributed around the zero line.

The parameters of interest represented in Figure 4.2 are the following: Latency to Persistent Sleep (LPS), Wake After Sleep Onset (WASO), Total Sleep Time (TST), time spent in each stage (tAW, tST1, tST2, tSWS, tREM), time spent in non-REM sleep (tNREM), sleep efficiency in 0-2 hours of bed time (SE1), 2-4 hours of bed time (SE2), 4-6 hours of bed time (SE3), 6-8 hours of bed time (SE4), mean extension of each sleep stage (meanAW, meanST1, meanST2, meanSWS, meanREM), number of transitions to each stage (nAW, nST1, nST2, nSWS, nREM).

4.4.1. PPC results

Posterior predictive check indicates a good agreement between simulated and observed efficacy endpoints in most cases (Fig. 4.2). Only 1 out of 23 median aggregated parameters computed from the real study fall outside the range of the median values computed from the simulated studies: this parameter is the time spent in SWS (tSWS), which is under-predicted. Figure 4.3 shows the histogram of the relative deviations for the simulated tSWS medians, together with the

corresponding density function (darkgreen line) and confidence interval (red dashed lines) computed as in the box-whisker plots, i.e.:

$$[Q1 - 1.5 * (Q3-Q1), Q3 + 1.5 * (Q3-Q1)],$$

where Q1 and Q3 are the first and the third quartile of the distribution, respectively. It is notable to observe that the zero relative deviation (vertical black line) falls outside the latter confidence interval. This bias is probably due to the low frequency of appearance of SWS during the night, and the consequent low number of observations available for the transitions from this sleep stage.
Other PPC plots were produced considering statistics different from the median: they are not reported in this thesis but they were all considered relevant in judging the model predictive performance, in terms of both typical outcomes and variability extent in the population.
The PPC outcome presented here can be compared with the one presented in [2], where the Markov model with multinomial logistic functions is introduced but not fully implemented with all the new features presented in Chapter 3. Although, the aim of this thesis is not related to the model development, it is worth to mention that the PPC for the revised model shows a general improvement in the predictive performance on the aggregated parameters. In particular, a significant improvement in LPS and nAW predictions is obtained, which can be associated to the introduction of the '*Initial Sleeplessness*' feature in the awake sub-model.
In conclusion, an example of the use of PPC has been provided to show how the proposed model is suitable to predict aggregated characteristics of PSG data in a patient population.

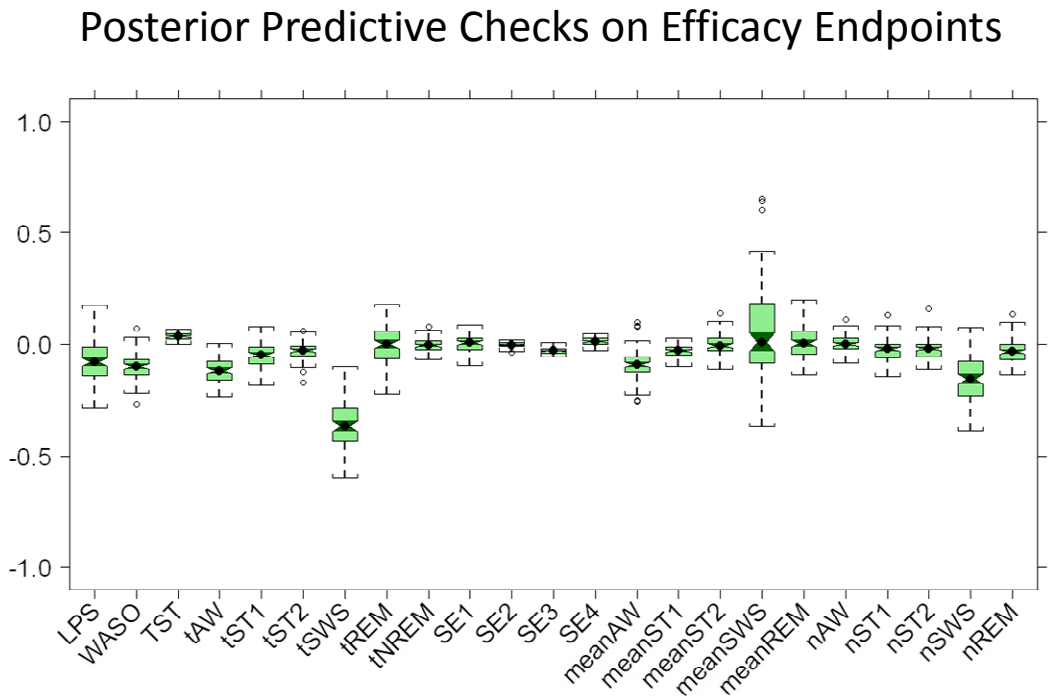


Figure 4.2: Results from posterior predictive check: distributions of the relative deviations of median efficacy endpoints in 100 simulated clinical studies from the medians in the real study.

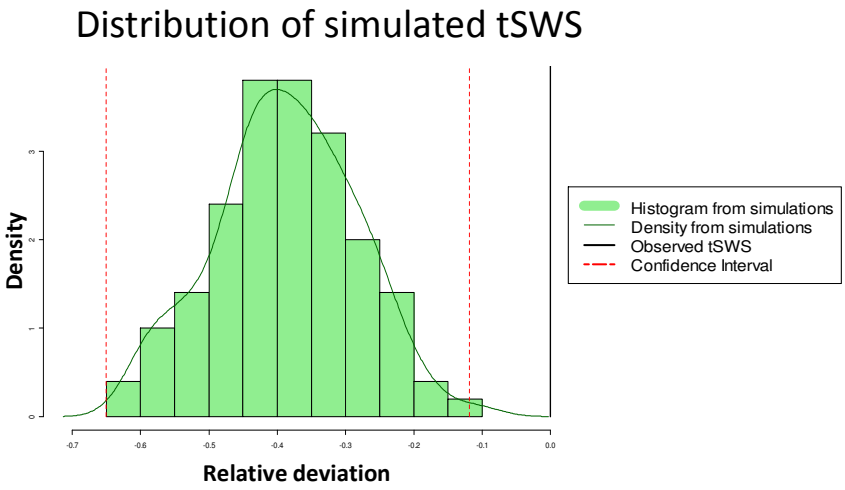


Figure 4.3: Distribution of relative deviations of median tSWS in simulated populations from the observed values.

4.5. Visual predictive check (VPC)

Since the model was built with the aim to properly describe the sleep internal architecture, it is important in the validation process to verify whether the model adequately describes sleep time-course along the night .

Visual predictive check (VPC) allows to test model capability to properly describe sleep stage and transition frequencies and therefore should be considered complementary to PPC, which mainly focus on parameters of clinical interest.

The concept governing this type of diagnostic method is the one which has been introduced in the theory section and applied for PPC: a statistic of interest is compared to a reference distribution obtained through simulation. However, in the VPC procedure the statistics under examination obtained from both raw and simulated data are evaluated over the independent variable, in this case the night time (the same procedure can be repeated over stage time, but night time is considered more relevant for a physiological perspective). The observed statistic and the corresponding confidence interval derived from simulations are therefore plotted against the independent variable in order to have a 'visual check' on the predictive model performance. This diagnostic is very helpful to evaluate potential model misspecifications over the independent variable.

The VPC implemented in this thesis regards two different statistics evolving with night time and obtainable from data:

a) stage frequencies, $\hat{\rho}_k, \ \forall\, k \in \mathcal{S}$,

b) transition frequencies, $\hat{f}_{km}$, $\forall\, k, m \in \mathcal{S}$,

according to Equations 3.1 and 3.2 in paragraph 3.1, respectively.

The 8-hour night time is divided into 10 intervals of equal width (48 minutes) and both transition and stage frequencies are computed on each of them. The choice of 10 intervals is based on two key features: to adequately follow sleep physiological variations across the night and to include a significant amount of data in each interval for computing confident frequencies.

VPC implementation was performed using the R package and some relevant code is presented in Appendix B.A brief explanation of this code is presented below.

Let's consider a specific dataset as presented in paragraph 4.2.1. The number of subjects in each stage is calculated for each epoch of the night: a summary dataset is created merging a 'Time' column (from 1 to 960) to the 'AW', 'ST1', 'ST2', 'SWS' and 'REM' columns, indicating the number of subjects in each of the sleep stages in each epoch. The number of transitions relative to each couple of sleep stages is also calculated over all the subjects at each epoch. As a result, 25 more columns are added to the summary dataset ('TRaw1', for example, indicates the number of subjects transitioning from AW to ST1 at every single epoch).

Time	AW	ST1	ST2	...	TRawaw	TRaw1	TRaw2	...	TRR3	TRRR
1	116	0	0	...	114	1	1	...	0	0
2	114	1	1	...	112	0	0	...	0	0
⋮										
959	44	17	36	...	2	0	...	...	0	0
960	49	14	34	...	2	0	...	...	0	0

The 'Time' column is thereafter divided into 10 parts of equal length and all variables are summed together in correspondence to each interval. In this way, the number of transitions from *k* to *m* and the number of occurrence of *k* ($\forall\, k, m \in \mathcal{S}$) are computed on the 10 night time intervals together with the final statistics (median and 95% confidence interval).

4.5.1. VPC: results

Figure 4.4 shows the results for the visual predictive check implemented on stage frequencies, $\hat{\rho}_k$. The plots show a general good agreement between the observed and the simulated statistics: the observed profiles computed on the 48-minute intervals of the night (red solid lines) fall within the confidence intervals (C.I.) obtained from the simulations (light blue areas) with very few exceptions (outliers). Both in the observed and simulated data, ST2 is the predominant stage all night long except for the first ~50 minutes; ST1, being a transitional state, covers a minimal part of the night; SWS mostly occurs between the second hour and the fourth hour, while it almost

disappears afterwards; REM sleep mainly occurs in the second half of the night; AW is strongly predominant in the first ~50 minutes and slightly regain importance in the last hours.

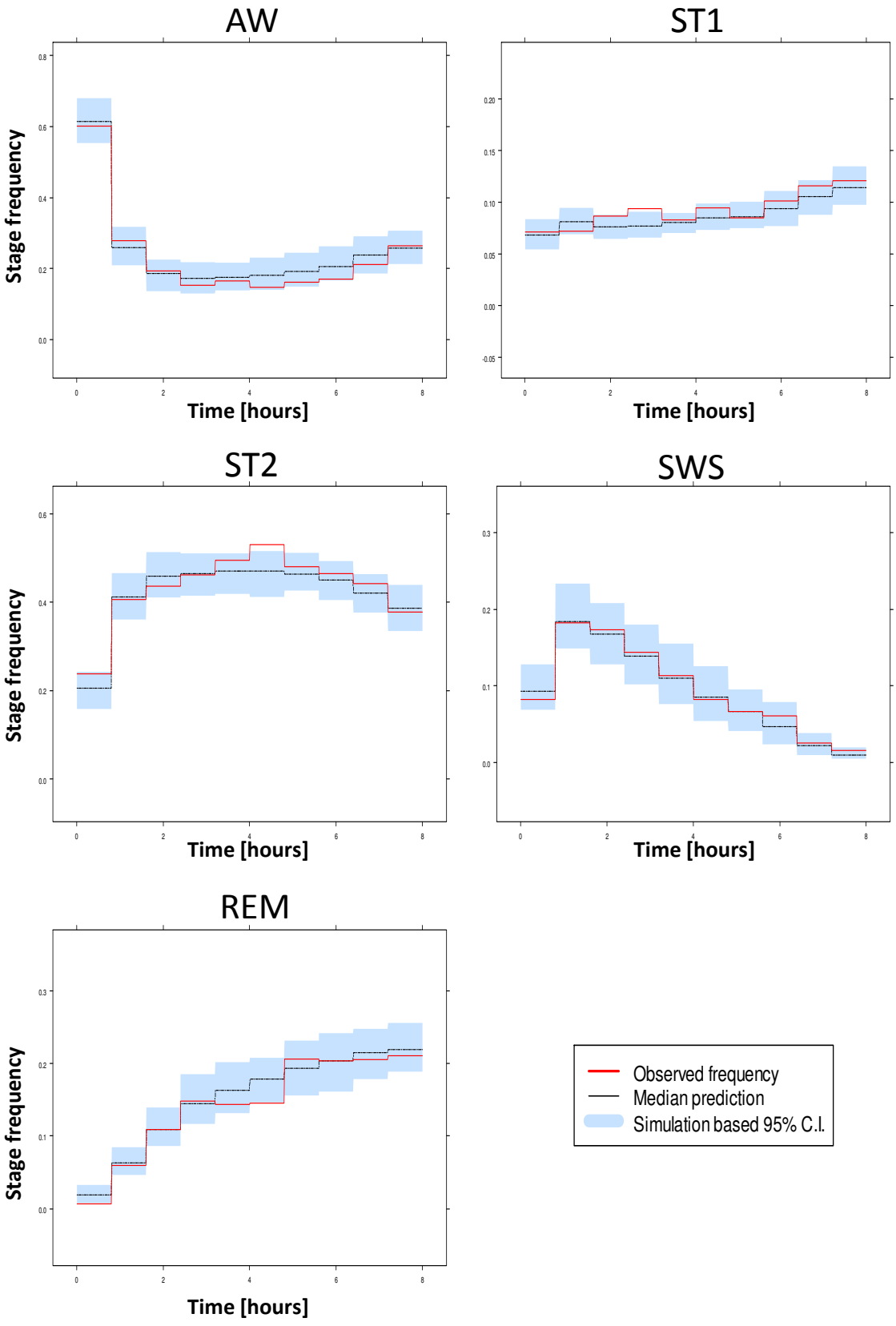


Figure 4.4: Results from visual predictive check on stage frequencies.

Figures 4.5, 4.6, 4.7, 4.8, 4.9 illustrate the predictive performance of the Markov-chain model with respect to the transition frequencies, $\hat{f}_{km}$, for the transitions from AW, ST1, ST2, SWS and REM respectively. A very good match is obtained between observed and simulated frequencies for all the sub-models: the observed frequencies (red solid lines) fall outside their simulated confidence interval (light blue areas) in very few occasions.

It is interesting to note the width of the confidence intervals. The transition frequencies from ST2 are those with the narrowest confidence intervals: in fact, a great amount of data is available for estimating transition probabilities from ST2. The *awake*, *stage1* and *stage2* sub-models are characterized by confidence intervals with approximately constant width, while the *sws* and *rem* sub-models present time-dependent confidence intervals. This is in agreement with the observed predominance of SWS in the first part of the night and of REM sleep in the second one: where data are more sparse a higher degree of uncertainty is shown, resulting in larger confidence intervals.

The plots also confirm that fixing some of the transition probabilities to zero in the model was an appropriate choice: in fact, their observed values (red solid lines) are actually very close to zero during the whole night time period.

In conclusion, the model seems appropriate to describe the physiological evolution of the considered sleep transitions along the night and VPC confirms to be a valuable tool for checking model adequacy to the observed process in its temporal domain.

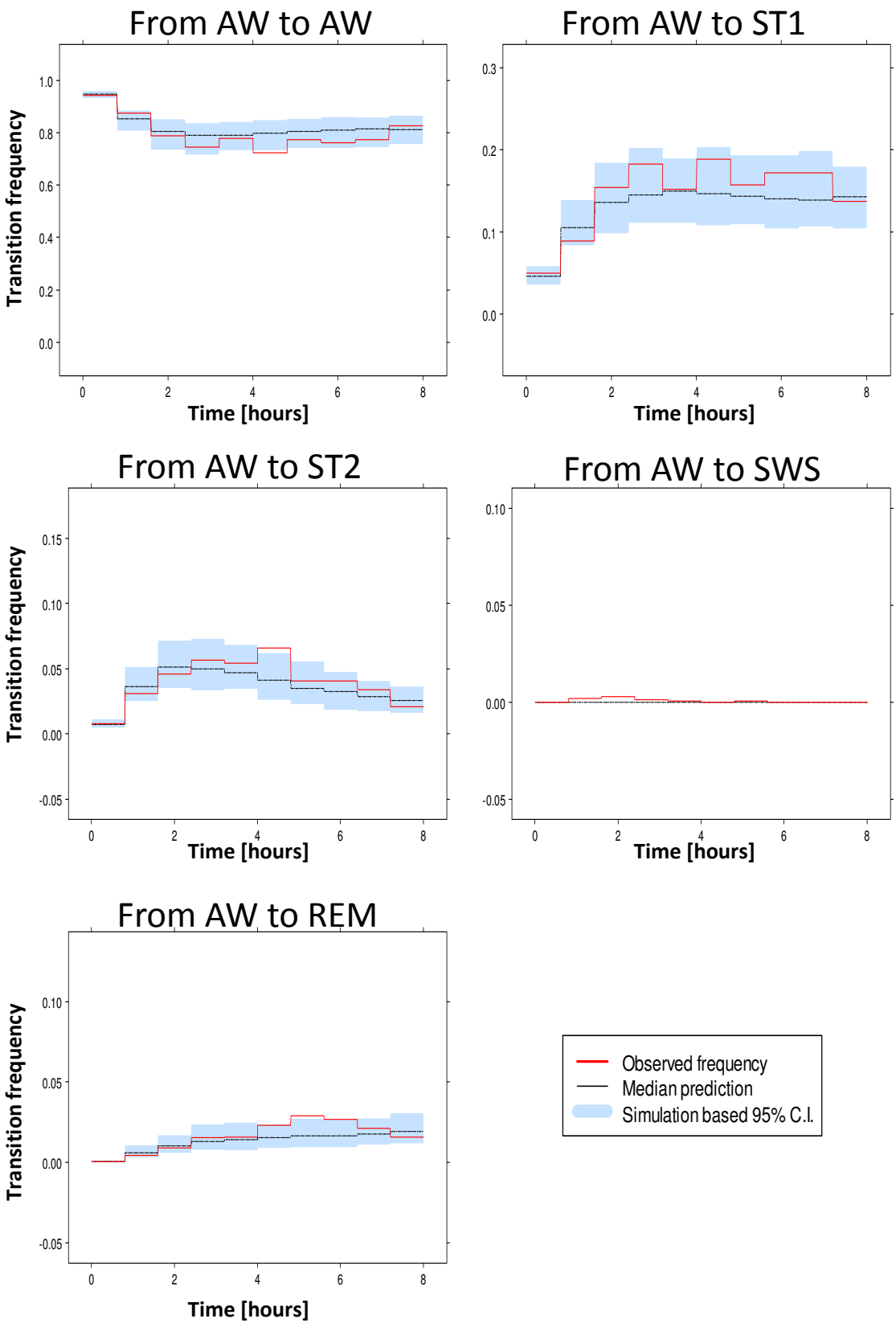


Figure 4.5: Results from visual predictive check on frequency of transitions from the AW state.

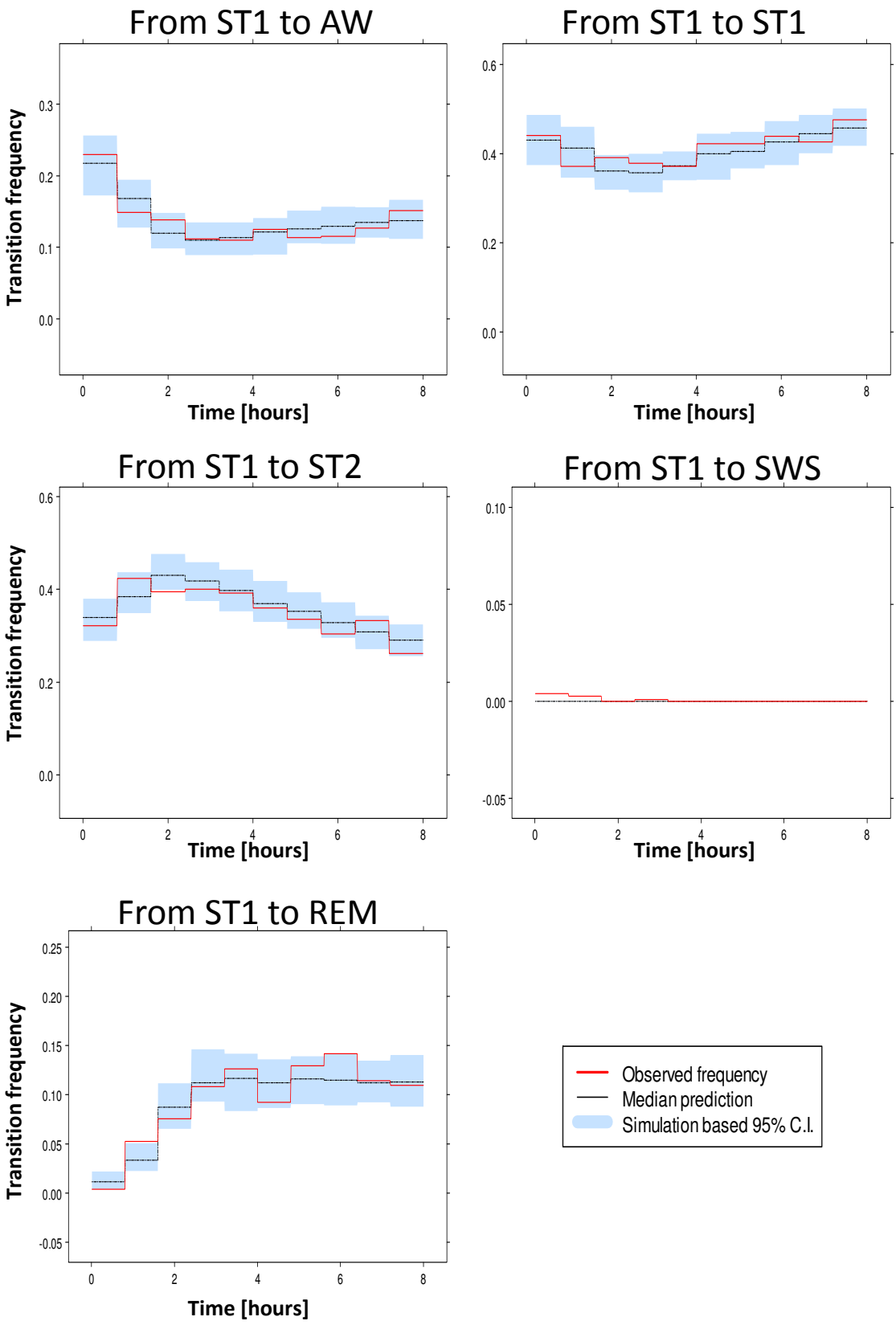


Figure 4.6: Results from visual predictive check on frequency of transitions from the ST1 state.

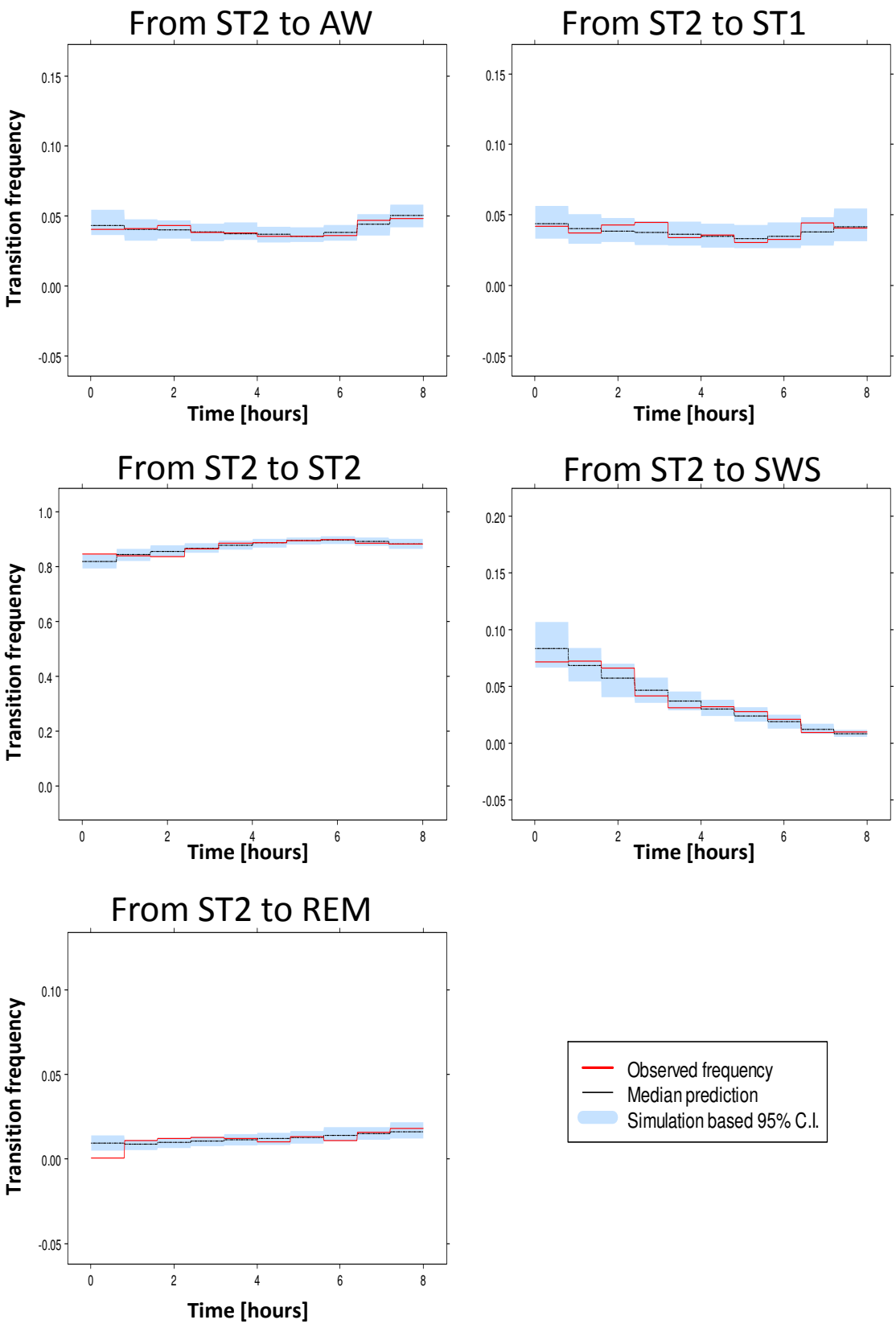


Figure 4.7: Results from visual predictive check on frequency of transitions from the ST2 state.

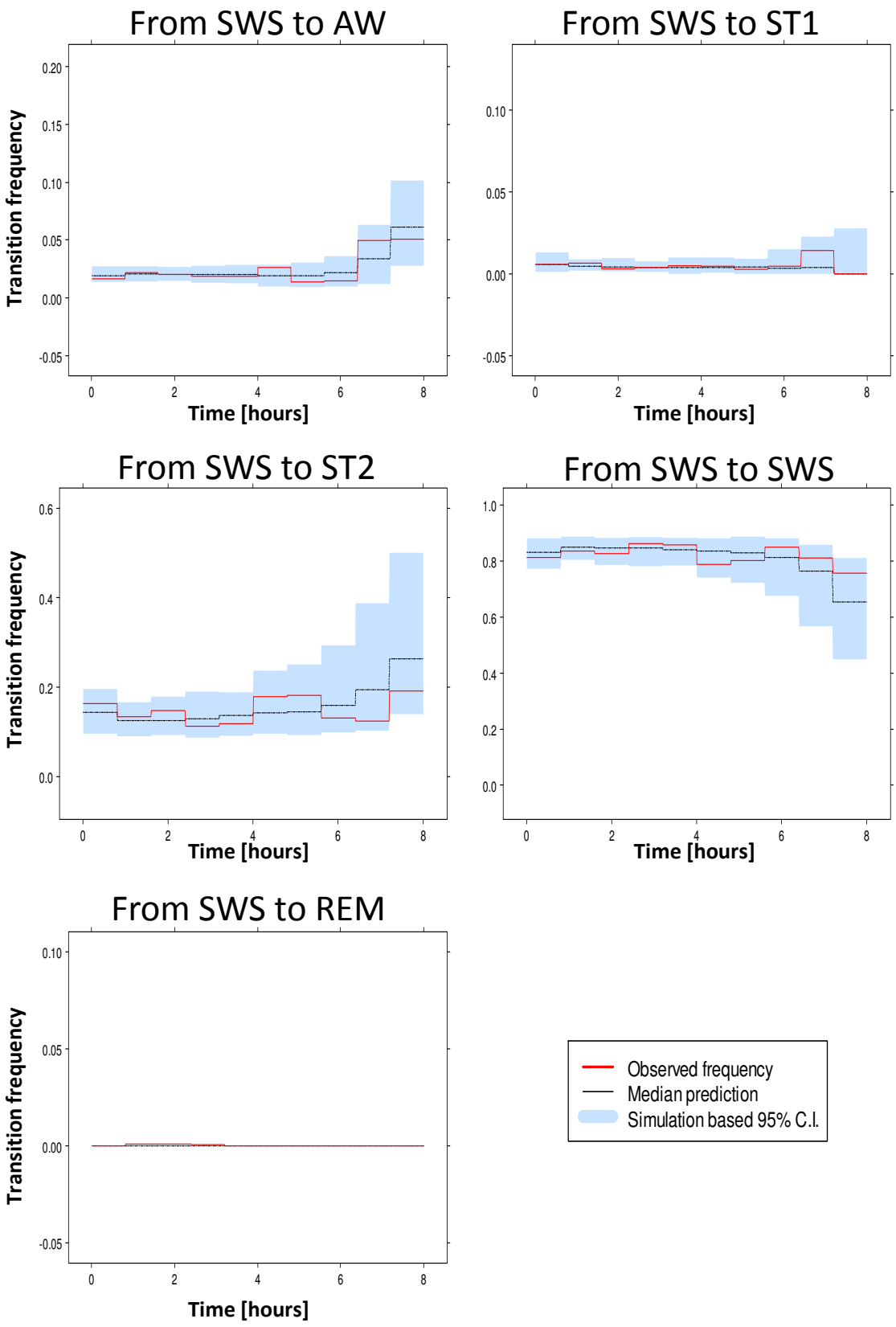


Figure 4.8: Results from visual predictive check on frequency of transitions from the SWS state.

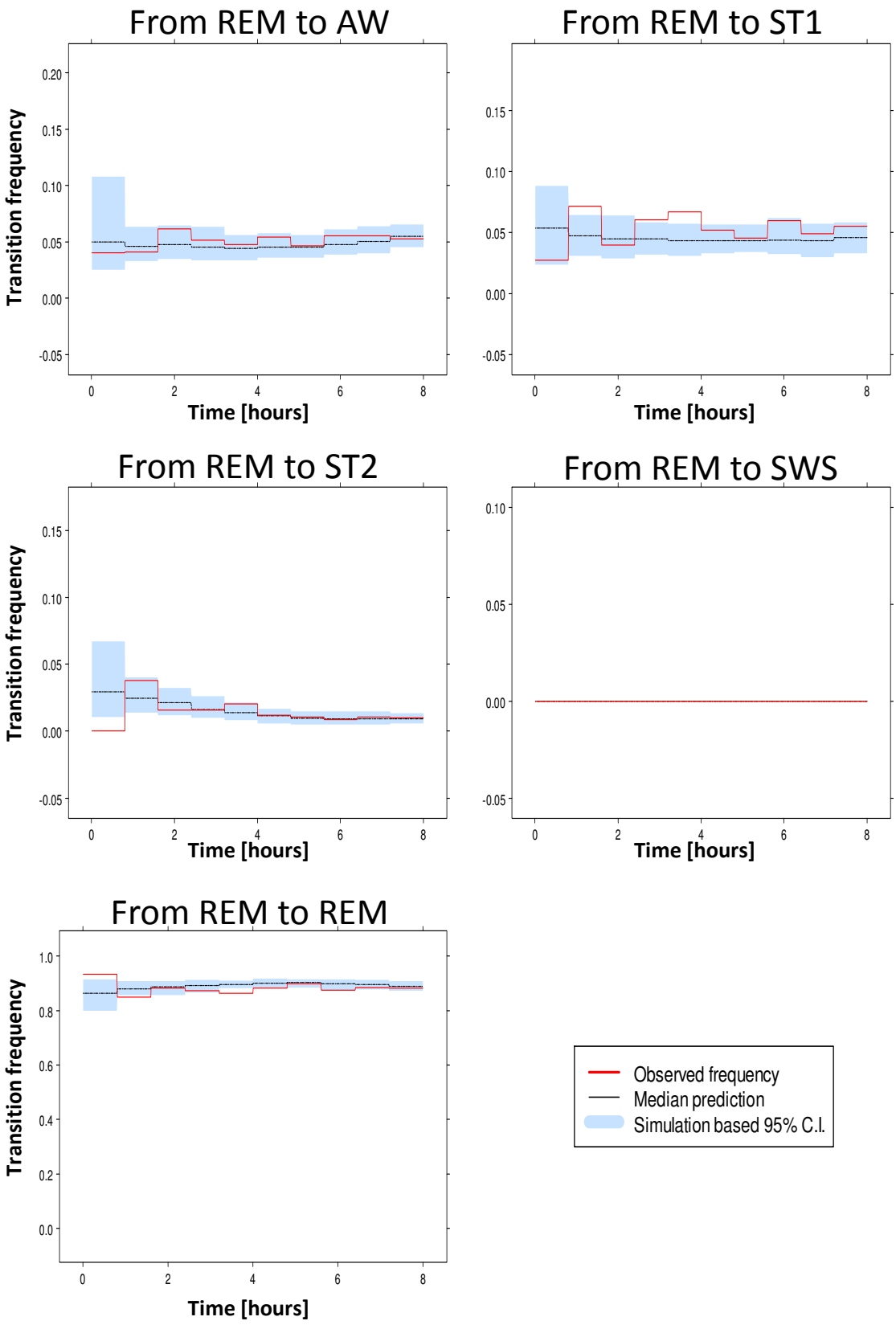
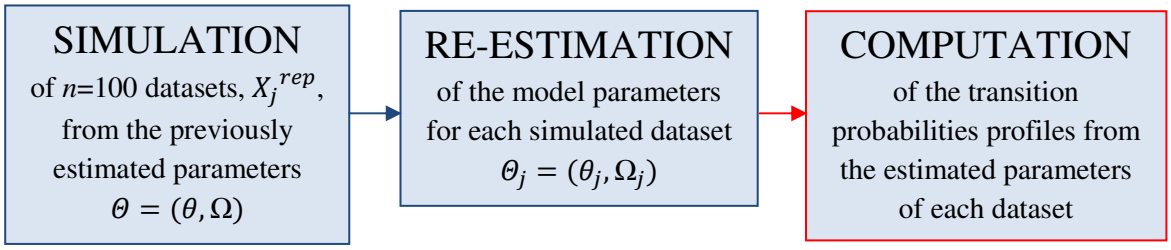


Figure 4.9: Results from visual predictive check on frequency of transitions from the REM state.

4.6. Visual estimation check (VEC)

Visual estimation check (VEC) is introduced in this thesis for the purpose to analyze the estimation of temporal profiles of typical and individual transition probabilities: profiles estimated from raw data are compared with the corresponding profiles estimated from the simulated datasets.

Differently from VPC where some statistics on the simulated datasets X_j^{rep} are produced, the simulated data are used for re-estimating the same Markov model, thus obtaining new estimates of the parameters, called Θ_j. This method has been already proposed as a validation approach where the distribution of Θ_j is compared with the original set of Θ [6,7]. The further implementation of this method proposed in this thesis is to transform the estimated parameters in the form of transition probabilities. In this way, we can still use a visual approach for the model validation (as in PPC and VPC) based on a meaningful parameters representation, as the transition probabilities are.



An automatic procedure is implemented for identifying each of the 5 sub-models on all the 100 simulated datasets. Five different summary datasheets are created by this procedure, in which population and inter-subject variability parameters are obtained for a specific sub-model using NONMEM software. In our case population parameters include the values of the typical logits at the night time break-points, the values of the stage time effects and the values for the internal break-points of night time and stage time. Between subject variability is estimated from covariance matrix and represents the individual deviations from the typical logits.

The transition probability profiles are then computed from the estimated parameters. In particular, the typical logit profiles for each sub-model are re-built from the population estimates. The individual deviations from the typical logits are drawn from Gaussian distributions with zero mean and covariance matrix estimated from the specific sub-model. Both typical and individual transition probability profiles are then computed according to Equation 3.8 from typical and individual logit profiles. The 5th and 95th percentiles of the individual transition probability profiles are therefore computed. Such procedure is followed for the original and all the simulated

datasets, producing an observed and n = 100 simulated profiles for the typical transition probabilities and the 5th and 95th percentiles on the BSV. The 95% confidence intervals for each of them are therefore computed from the simulated profiles.
Since this procedure implies computational burden it was applied only to the final model.

4.6.1. VEC: results

Figures 4.10, 4.11, 4.12, 4.13 and 4.14 illustrate the results from VEC performed on transition probabilities, for the transitions from AW, ST1, ST2, SWS and REM respectively. The black solid line represents the typical transition probability estimated from raw data and the pink area corresponds to the 95% confidence interval (C.I.) on the typical transition probability as derived from simulations. The dashed black lines represent the 5th and 95th percentiles of the inter-individual distribution and the blue area their 95% confidence intervals. Mixed color (violet) areas come from the superimposition of pink and blue confidence intervals (note that there is superimposition on the typical and 5th and/or 95th confidence intervals for the *awake*, *sws* and *rem* sub-models).
In general, a good agreement between profiles estimated from raw and simulated data is shown in these plots, with exception for the transitions from REM to ST2 and from REM to REM at the beginning of the night. Even though the observations of REM sleep in the first part of the night are almost absent (see Fig. 4.4), it is difficult to understand for which reason the transition probabilities estimated from observed data are different from the transition probabilities re-estimated from simulated data. One possible explanation is related to the estimation method in presence of shrinkage effect. In fact, the Laplacian method is less robust when high '*η-shrinkage*' is revealed [3]. A full description of the "shrinkage phenomena" was not an objective of this thesis and detailed information can be found in literature [46].
In nonlinear mixed-effect modeling, the estimated individual deviations from a typical parameter value are called '*Empirical Bayes Estimates*' (EBEs) as described in Chapter 2. EBEs of a specific inter-individual random effect η are used for computing η-shrinkage as follows:

$$1 - (\mathrm{SD}(\mathrm{EBEs}))/\omega,$$

where ω is the population model estimate of the SD in η.
In case of shrinkage the EBEs are biased and therefore the Laplace estimation method can suffer from this effect. According to literature [46] η-shrinkage higher than ~25%, is considered relevant.

The stage transitions mostly affected by shrinkage in the proposed Markov-chain model are those leaving SWS and REM sleep (see Table 4.1): η-shrinkage results ~50% for the logit value defined on the ratio between the probability of moving from SWS to ST1 sleep and the probability of staying in SWS, and > 25% for the logit value defined on the ratio between the probability of staying in REM sleep and the probability of moving to SWS.

Sub-model\logit	logit1	logit2	logit3	logit4
awake	0.17	0.14	0.23	.
stage1	0.17	0.10	0.18	.
stage2	0.17	0.10	0.08	0.28
sws	.	0.49	0.12	.
rem	0.15	0.14	0.27	.

Table 4.1: η-shrinkage values for all the sub-models.

Consequently, the only case in which high η-shrinkage could be considered a potential issue in the Laplacian method is the transition from REM to ST2 sleep. However, other high η-shrinkage effect as in the SWS stage did not result in any bias on VEC .

It is interesting to note that also the amplitude of confidence intervals on the transition probabilities is related to the amount of available information: where data are sparser their amplitude is larger. This can be seen for transitions of *sws* and *rem* sub-models: they show very similar time-dependency between amplitude of confidence intervals and amount of SWS and REM observations; moreover, their confidence interval width is higher, for example, than that computed for *stage2* sub-model, for which observations are much more abundant.

Furthermore, it is interesting to notice that confidence intervals in proximity of probability equal to 0 or 1 are generally narrower than those in proximity of probability = 0.5 (see, for example, the transitions from ST2 to ST1 or from SWS to ST2). This is due to the non-linear behavior of the logit transformation implemented in the model: an interval in the logit scale $(-\infty, +\infty)$ will be squashed in the probability scale [0, 1]. Therefore larger C.I. will appear in case of logit values near zero, while narrower C.I. will be present with logit values belonging to the higher and lower parts of the $(-\infty, +\infty)$ interval.

Concluding, the performed VEC allows to state once more that the model is suitable for describing the available sleep data, but also that the adopted estimation method is robust enough for estimating such model. VEC shows to be an appealing and meaningful visual diagnostics to apply with models on categorical data (but not only) and to put side by side to other indicators of potential estimator inadequacy (like the η-shrinkage computation).

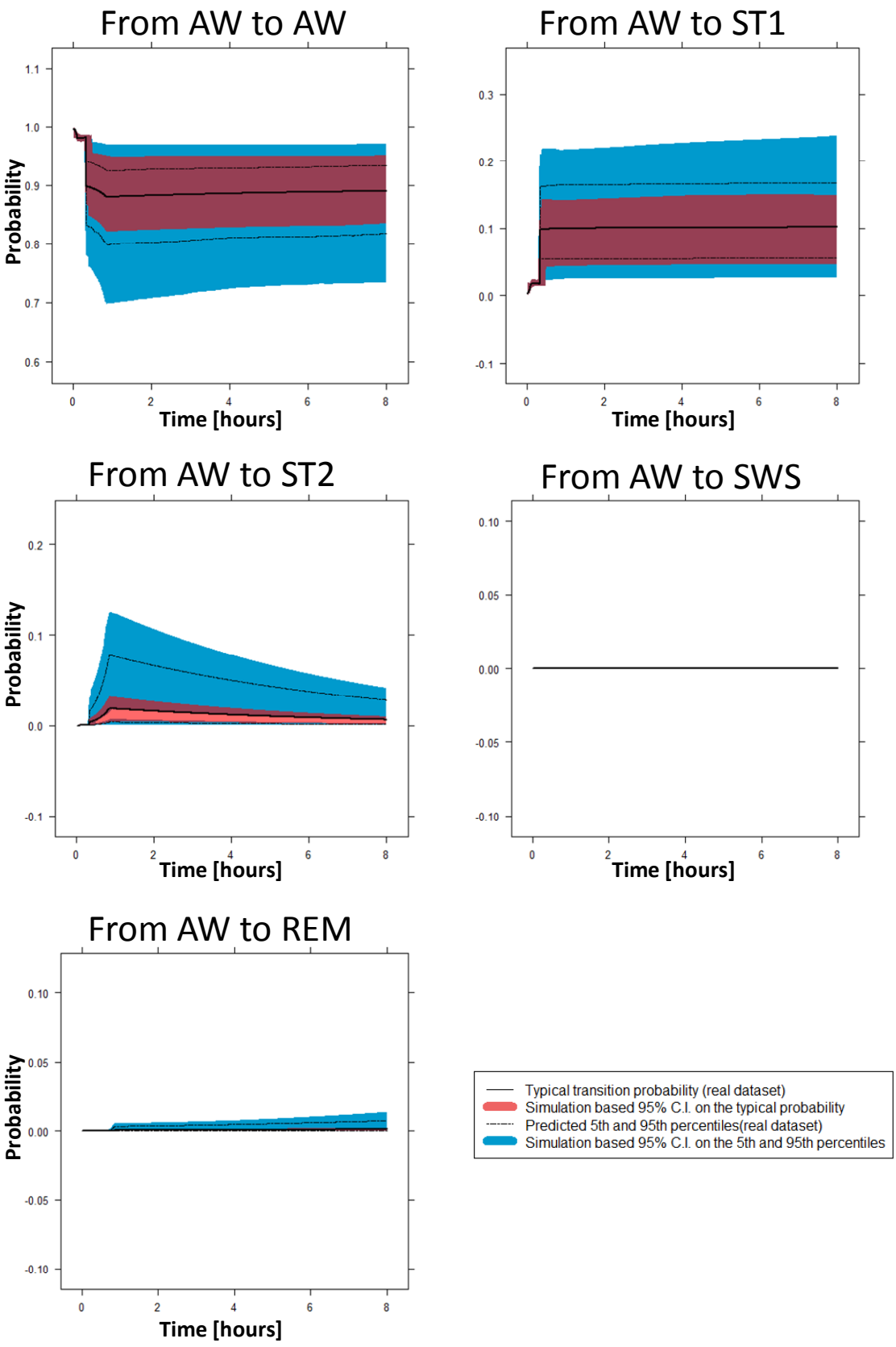


Figure 4.10: Results from visual estimation check on transitions from the AW state.

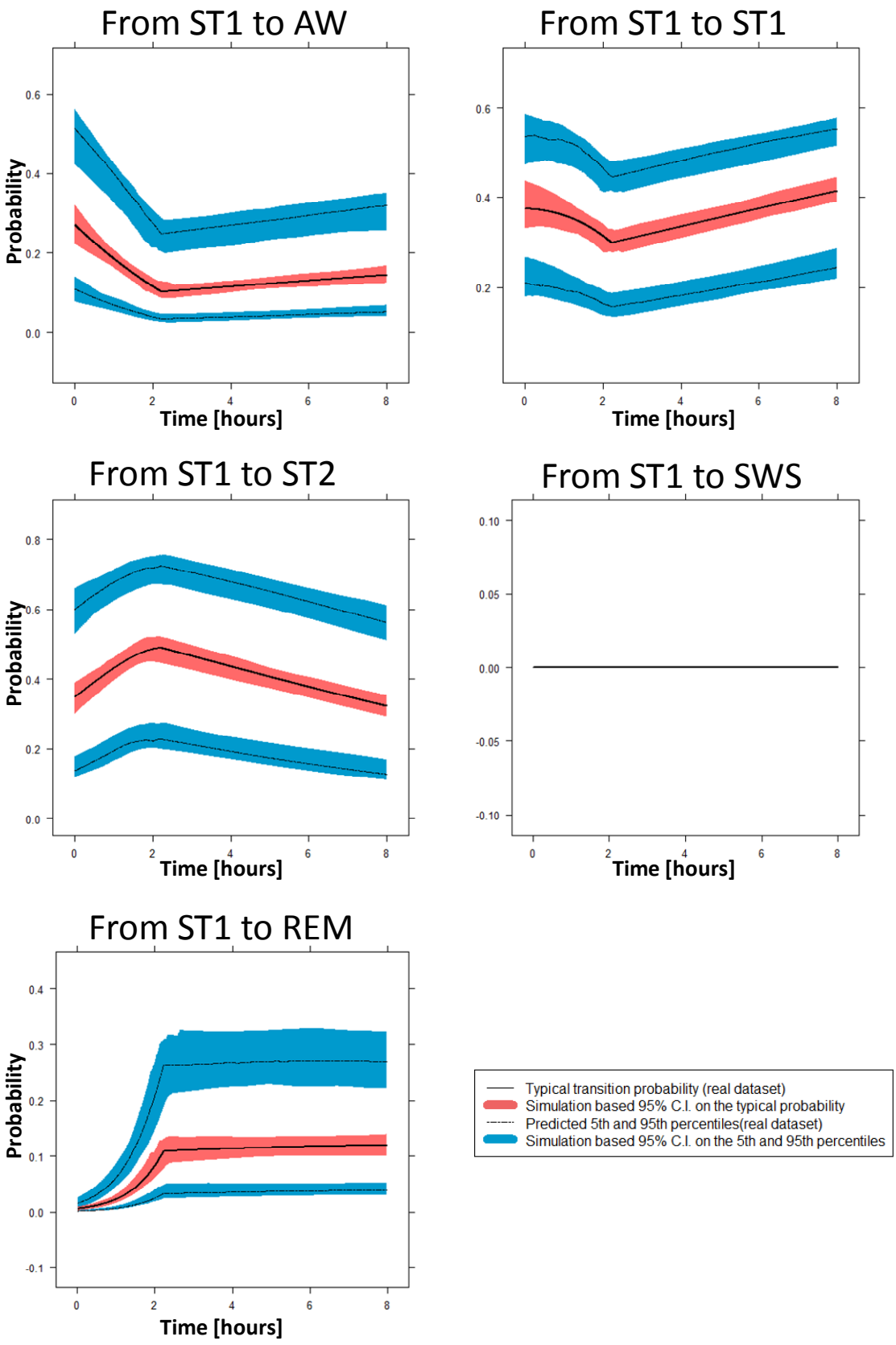


Figure 4.11: Results from visual estimation check on transitions from the ST1 state.

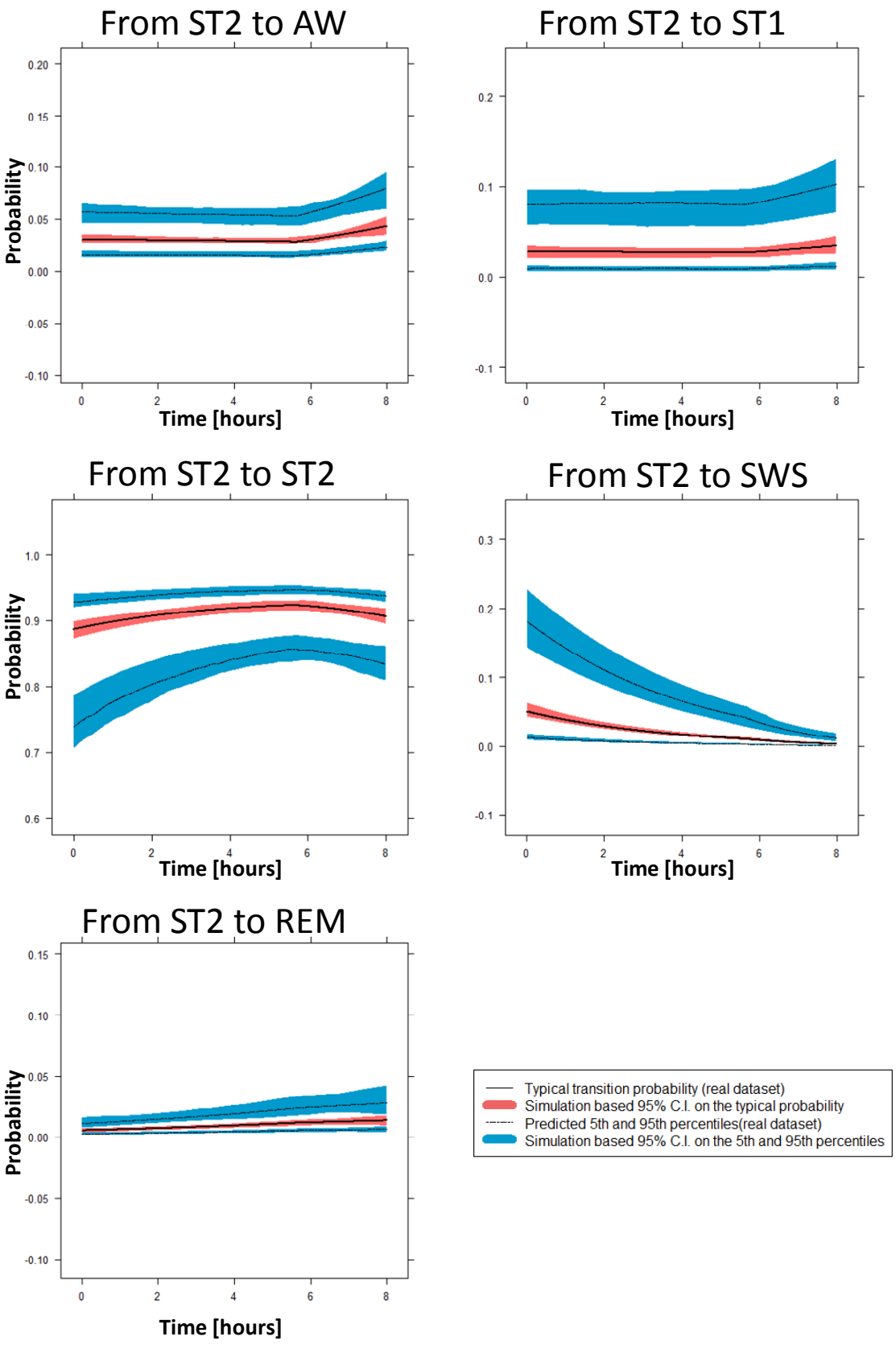


Figure 4.12: Results from visual estimation checks on transitions from the ST2 state.

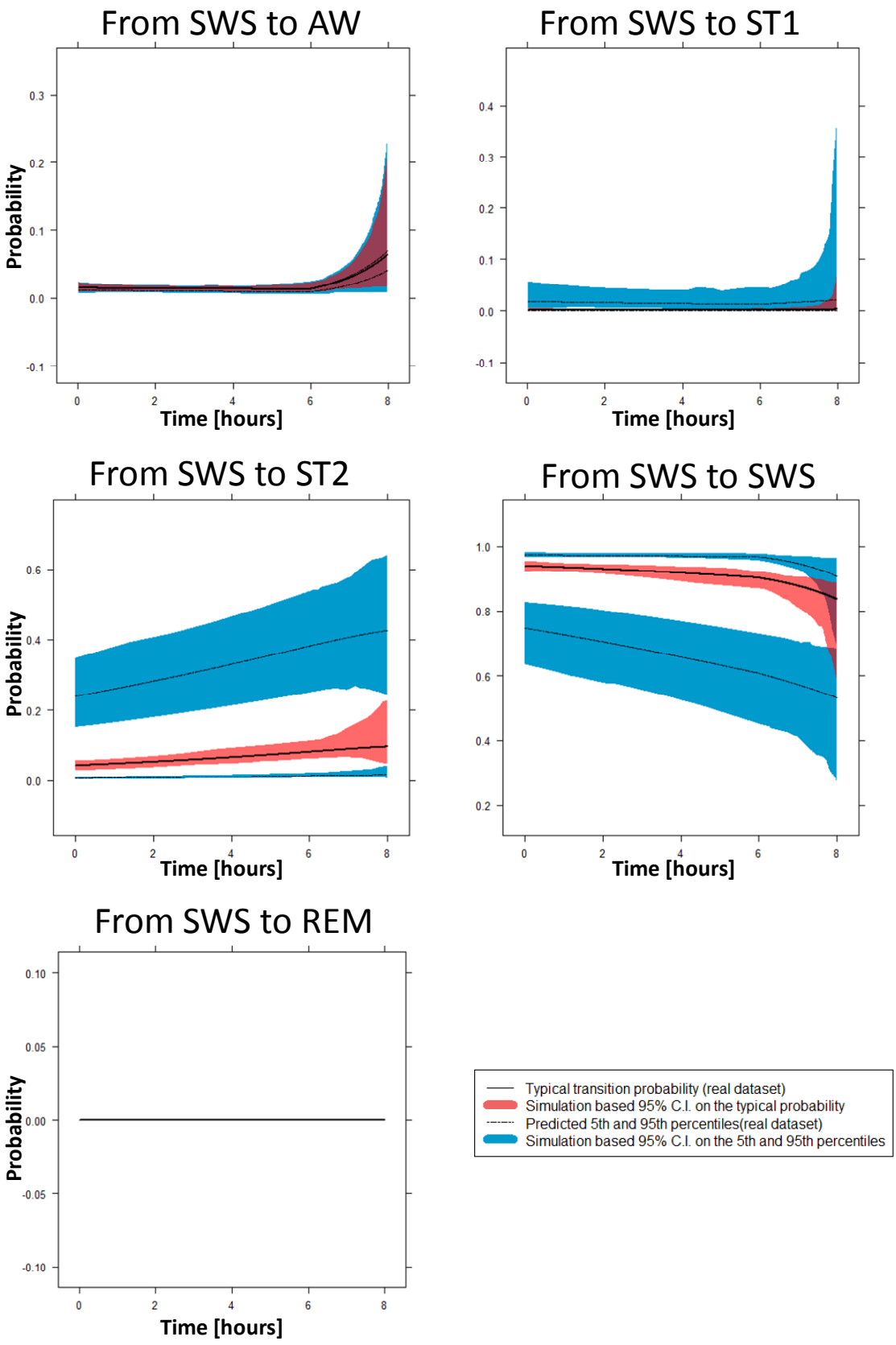


Figure 4.13: Results from visual estimation checks on transitions from the SWS state.

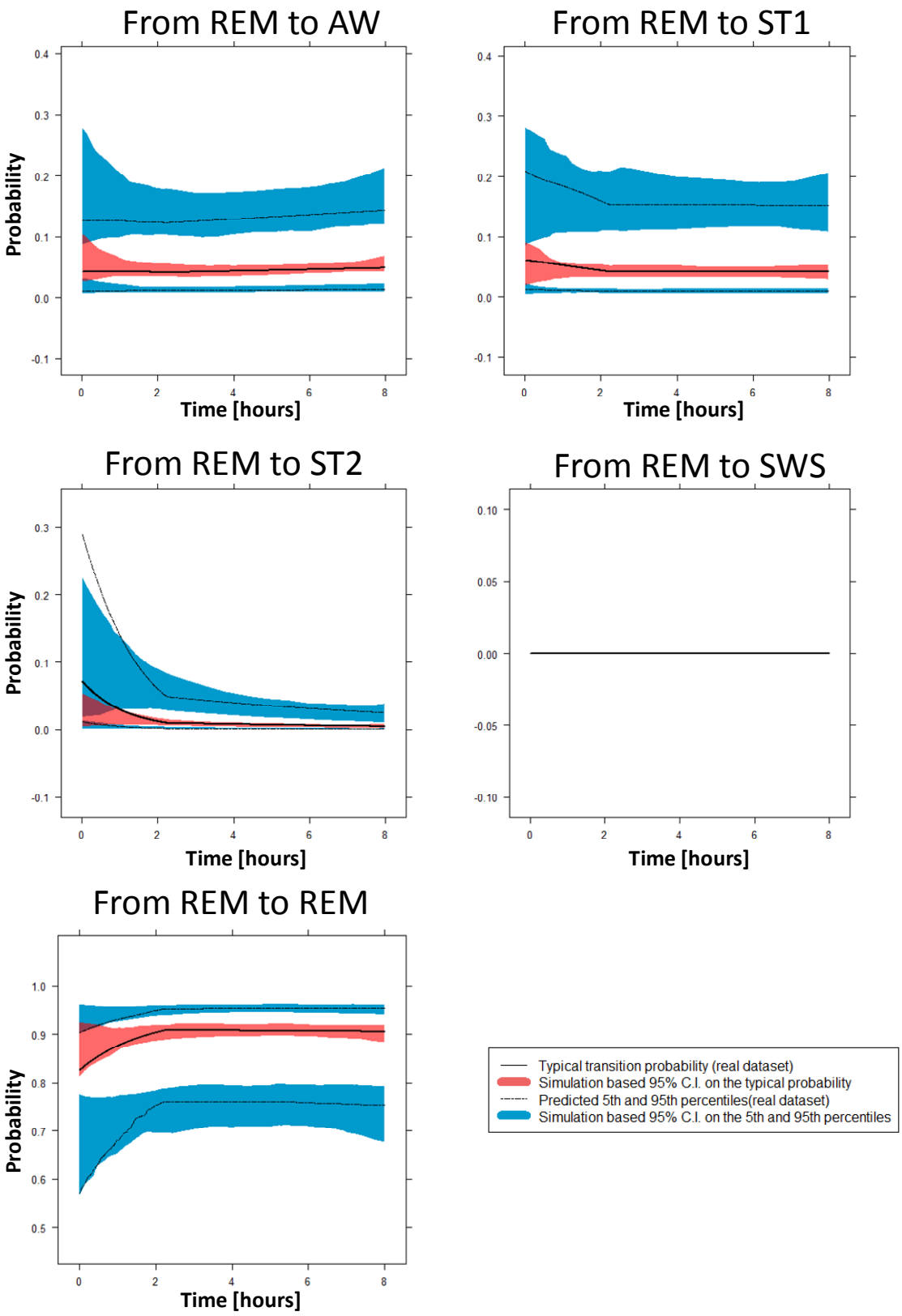


Figure 4.14: Results from visual estimation checks on transitions from the REM state.

4.7. Conclusions

In this Chapter different diagnostic methods were proposed to investigate the adequacy and the performance of the proposed Multinomial Markov-chain model [2] for describing sleep data. Firstly, diagnostic methods based on stochastic simulation were implemented to assess the capacity of the model to describe some important characteristics of the polysomnograpghic (PSG) data. In particular, 'Posterior Predictive Check' on the aggregated parameters (WASO, etc) and 'Visual Predictive Check' on the transition frequencies and the frequencies of occurrence of each stage were implemented. The results from both these methodologies showed that the proposed model is sufficiently robust for describing these data characteristics and the dynamic behaviour of the sleep process.

However, bias and imprecision in the desired predictions may also depend on weaknesses in the estimation methods. Therefore, A new diagnostic method derived by a recently proposed [6,7] approach for checking both model and estimation method performance has been introduced: the 'Stochastic Simulations followed by Estimation' (SSE) based diagnostics. The peculiarity of the method based on SSE as implemented in this thesis is that it allowed to perform a visual check on the characteristics of interest: in particular, the transition probabilities between sleep stages. This kind of visual comparison – applied for the first time in categorical models - has been called 'Visual Estimation Check' (VEC). The results from the implementation of such diagnostic corroborated the adequacy of the multinomial Markov-Chain model to describe sleep architecture and provided a valuable tool to evaluate the goodness of the proposed estimation method.

Discussion and Final Remarks

Polysomnography is a key measurement to objectively assess and properly diagnose different types of sleep disorders. Currently, the clinical use of polysomnography is mainly related to assess sleep aggregated characteristics (WASO, LPS, etc.). However, there is growing evidence suggesting that the maintenance of the sleep architecture is one of the critical aspects for sleep quality.

Clearly, mathematical models can be of great help when trying to quantify complex physiological processes and a Multinomial Mixed-Effect Markov-chain model approach has been proposed in this thesis to properly quantify the sleep architecture considered as a time-course pattern among specific sleep stages.

The aim of this thesis was therefore to evaluate the appropriateness of such Markov-chain model. The application of this model, in fact, is considered a valuable methodology for the analysis and interpretation of clinical studies in the insomnia indication and a proper evaluation of model performances was therefore considered of particular importance.

The analysis conducted in this work was focused, in particular, on the evaluation of model performances in predicting both the sleep aggregated parameters (LPS, WASO, etc.) used in the clinical practice to quantify the severity of sleep disorders and the sleep physiological pattern, whose maintenance has been demonstrated to be critical. It is worth mentioning that the novel mechanisms in the treatment of sleep disorders are aimed to reduce sleep onset and increase sleep maintenance preserving the physiological sleep architecture.

In addition, the ability of this model to properly describe the time-course of sleep can be a very helpful tool for clinical trial simulations, where the effect of different doses and treatment periods can be predicted.

Three different diagnostic methodologies have been presented and implemented in this thesis for checking the adequacy of the proposed model: the '*Posterior Predictive Check*'(PPC) and the '*Visual Predictive Check*'(VPC) based on stochastic simulation and the '*Visual Estimation Check*'(VEC), introduced with this thesis, derived by the 'stochastic simulation followed by estimation' (SSE) method. Such methodologies provided evidence of the good properties of the proposed Markov-chain model and their implementation in this thesis illustrates their usefulness for checking model adequacy. In particular, the PPC was demonstrated to be an important simulation-based diagnostic showing that the proposed model is suitable for predicting the aggregated characteristics of PSG data in a population of subjects. Furthermore, through the implementation of the VPCs on the transition frequencies and the stages frequencies of

occurrence, the model was demonstrated to be appropriate for describing the physiological evolution of the considered statistics along the night. The VPC method therefore confirms to be a good tool for checking model adherence to the observed data and to the process temporal dynamics. Finally, the results from the performed VEC allow to state once more that the model is suitable for describing the available sleep data, but also that the adopted estimation method is robust enough. Such diagnostics allow the evaluation of both model and estimation method performances in an easily understandable visual form. To my knowledge, this work is the first to implement the SSE in this way. The implemented VEC seems to be an appealing and meaningful visual diagnostics to apply with models on categorical data (but not only) and to put side by side to other indicators of potential estimator inadequacy (like the η-shrinkage computation).

In summary, the validation proposed in this thesis indicates that the proposed model can be considered a robust modeling framework for describing and predicting sleep architecture. This analysis has been conducted on data from a population of insomniac patients treated with placebo, but the natural future development will be to extend this model (and the related validation approach) to include the terms related to the drug effect to allow simulating complete clinical trials.

Appendix A

NONMEM code

A.1. 'Awake sub-model'

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
; AWAKE sub-model
; NONMEM model file for the transition from awake
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

$PROB trAW

$INPUT ID VIS TIME DV=STAG MDV0 MDV1 MDV2 MDV3 MDV5 MDV
       DROP DROP DROP DROP DROP STT EVID BPS0 DROP DROP DROP DROP MED
       DROP DROP DROP DROP SL DROP IS MIS

  ; TIME = epoch number (between 1 and 960)
  ; STT = stage_time (number of epochs)
  ; BPS0 = max_stage time observed in the data
  ; MED = median_stage time observed in the data
  ; SL = 0 if first sleep stage has not occurred yet, 1 otherwise
  ; IS = first epoch with SL==1
  ; MIS = maximum Initial Sleeplessness observed in the data


$DATA tr0Visit33MAD103894forNONMEMallTimes34tSTevidSLIS.csv IGNORE=@

  IGNORE=(MDV0.EQ.1)
  IGNORE=(STAG.GT.5)
  IGNORE=(STAG.EQ.3)

  ;dataset contains only observations immediately preceded by awake
   stage

$PRED

  TVG1A=THETA(1)
  TVG2A=THETA(2)
  TVG3A=THETA(3)
  TVG1B=THETA(4)
  TVG2B=THETA(5)
  TVG3B=THETA(6)
  TVG1C=THETA(7)
  TVG2C=THETA(8)
  TVG3C=THETA(9)

  STE1b=THETA(10)
  STE2b=THETA(11)
  STE3b=THETA(12)
```

```
  STE1c=THETA(13)
  STE2c=THETA(14)
  STE3c=THETA(15)

  BPA=IS
  BPC=960
  BPB=(BPC-BPA)*THETA(16)+BPA

  BPsa=1
  BPsc=BPS0
  BPsb=(BPsc-BPsa)*THETA(17)+BPsa

  BP1=2
  BP3=MIS    ;MAX initial sleeplessness (max(IS)-1)
  BP2=(BP3-BP1)*THETA(18)+BP1


  TVG11=THETA(19)
  TVG21=THETA(20)
  TVG31=THETA(21)
  TVG12=THETA(22)
  TVG22=THETA(23)
  TVG32=THETA(24)
  TVG13=THETA(25)
  TVG23=THETA(26)
  TVG33=THETA(27)

;--------------------------------
; stage_time effect interpolation
;--------------------------------

  IF (STT.LE.BPsb.AND.SL.EQ.1) THEN
    G1A=(TVG1A+ETA(1))*(BPsb-STT)/(BPsb-BPsa)+(TVG1A+ETA(1)+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2A=(TVG2A+ETA(2))*(BPsb-STT)/(BPsb-BPsa)+(TVG2A+ETA(2)+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3A=(TVG3A+ETA(3))*(BPsb-STT)/(BPsb-BPsa)+(TVG3A+ETA(3)+STE3b)*(STT-
BPsa)/(BPsb-BPsa)

    G1B=(TVG1B+ETA(1))*(BPsb-STT)/(BPsb-BPsa)+(TVG1B+ETA(1)+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2B=(TVG2B+ETA(2))*(BPsb-STT)/(BPsb-BPsa)+(TVG2B+ETA(2)+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3B=(TVG3B+ETA(3))*(BPsb-STT)/(BPsb-BPsa)+(TVG3B+ETA(3)+STE3b)*(STT-
BPsa)/(BPsb-BPsa)

    G1C=(TVG1C+ETA(1))*(BPsb-STT)/(BPsb-BPsa)+(TVG1C+ETA(1)+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2C=(TVG2C+ETA(2))*(BPsb-STT)/(BPsb-BPsa)+(TVG2C+ETA(2)+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3C=(TVG3C+ETA(3))*(BPsb-STT)/(BPsb-BPsa)+(TVG3C+ETA(3)+STE3b)*(STT-
BPsa)/(BPsb-BPsa)
  ENDIF

  IF (STT.GT.BPsb.AND.SL.EQ.1) THEN
    G1A=(TVG1A+ETA(1)+STE1b)*(BPsc-STT)/(BPsc-BPsb)+(TVG1A+ETA(1)+STE1c)
*(STT-BPsb)/(BPsc-BPsb)
    G2A=(TVG2A+ETA(2)+STE2b)*(BPsc-STT)/(BPsc-BPsb)+(TVG2A+ETA(2)+STE2c)
*(STT-BPsb)/(BPsc-BPsb)
    G3A=(TVG3A+ETA(3)+STE3b)*(BPsc-STT)/(BPsc-BPsb)+(TVG3A+ETA(3)+STE3c)
*(STT-BPsb)/(BPsc-BPsb)
```

```
    G1B=(TVG1B+ETA(1)+STE1b)*(BPsc-STT)/(BPsc-BPsb)+(TVG1B+ETA(1)+STE1c)
*(STT-BPsb)/(BPsc-BPsb)
    G2B=(TVG2B+ETA(2)+STE2b)*(BPsc-STT)/(BPsc-BPsb)+(TVG2B+ETA(2)+STE2c)
*(STT-BPsb)/(BPsc-BPsb)
    G3B=(TVG3B+ETA(3)+STE3b)*(BPsc-STT)/(BPsc-BPsb)+(TVG3B+ETA(3)+STE3c)
*(STT-BPsb)/(BPsc-BPsb)

    G1C=(TVG1C+ETA(1)+STE1b)*(BPsc-STT)/(BPsc-BPsb)+(TVG1C+ETA(1)+STE1c)
*(STT-BPsb)/(BPsc-BPsb)
    G2C=(TVG2C+ETA(2)+STE2b)*(BPsc-STT)/(BPsc-BPsb)+(TVG2C+ETA(2)+STE2c)
*(STT-BPsb)/(BPsc-BPsb)
    G3C=(TVG3C+ETA(3)+STE3b)*(BPsc-STT)/(BPsc-BPsb)+(TVG3C+ETA(3)+STE3c)
*(STT-BPsb)/(BPsc-BPsb)
  ENDIF

;--------------------------------
; logits interpolation before IS
;--------------------------------

  IF (TIME.GE.BP1.AND.TIME.LE.BP2.AND.SL.EQ.0) THEN

    ; individual logits
    G1=(TVG11+ETA(4))*(BP2-TIME)/(BP2-BP1)+(TVG12+ETA(4))*(TIME-
BP1)/(BP2-BP1)
    G2=(TVG21+ETA(5))*(BP2-TIME)/(BP2-BP1)+(TVG22+ETA(5))*(TIME-
BP1)/(BP2-BP1)
    G3=(TVG31+ETA(6))*(BP2-TIME)/(BP2-BP1)+(TVG32+ETA(6))*(TIME-
BP1)/(BP2-BP1)

    ; population logits
    G1p=TVG11*(BP2-TIME)/(BP2-BP1)+TVG12*(TIME-BP1)/(BP2-BP1)
    G2p=TVG21*(BP2-TIME)/(BP2-BP1)+TVG22*(TIME-BP1)/(BP2-BP1)
    G3p=TVG31*(BP2-TIME)/(BP2-BP1)+TVG32*(TIME-BP1)/(BP2-BP1)
  ENDIF

  IF (TIME.GT.BP2.AND.TIME.LE.BP3.AND.SL.EQ.0) THEN

   ; individual logits
   G1=(TVG12+ETA(4))*(BP3-TIME)/(BP3-BP2)+(TVG13+ETA(4))*(TIME-
BP2)/(BP3-BP2)
    G2=(TVG22+ETA(5))*(BP3-TIME)/(BP3-BP2)+(TVG23+ETA(5))*(TIME-
BP2)/(BP3-BP2)
    G3=(TVG32+ETA(6))*(BP3-TIME)/(BP3-BP2)+(TVG33+ETA(6))*(TIME-
BP2)/(BP3-BP2)

    ; population logits
    G1p=TVG12*(BP3-TIME)/(BP3-BP2)+TVG13*(TIME-BP2)/(BP3-BP2)
    G2p=TVG22*(BP3-TIME)/(BP3-BP2)+TVG23*(TIME-BP2)/(BP3-BP2)
    G3p=TVG32*(BP3-TIME)/(BP3-BP2)+TVG33*(TIME-BP2)/(BP3-BP2)
  ENDIF

;--------------------------------
; logits interpolation after IS
;--------------------------------

  IF (TIME.GE.BPA.AND.TIME.LE.BPB.AND.SL.EQ.1) THEN

    ; individual logits
    G1=G1A*(BPB-TIME)/(BPB-BPA)+G1B*(TIME-BPA)/(BPB-BPA)
    G2=G2A*(BPB-TIME)/(BPB-BPA)+G2B*(TIME-BPA)/(BPB-BPA)
```

```
    G3=G3A*(BPB-TIME)/(BPB-BPA)+G3B*(TIME-BPA)/(BPB-BPA)

    ; population logits
    G1p=TVG1A*(BPB-TIME)/(BPB-BPA)+TVG1B*(TIME-BPA)/(BPB-BPA)
    G2p=TVG2A*(BPB-TIME)/(BPB-BPA)+TVG2B*(TIME-BPA)/(BPB-BPA)
    G3p=TVG3A*(BPB-TIME)/(BPB-BPA)+TVG3B*(TIME-BPA)/(BPB-BPA)
  ENDIF

  IF (TIME.GT.BPB.AND.TIME.LE.BPC.AND.SL.EQ.1) THEN

    ; individual logits
    G1=G1B*(BPC-TIME)/(BPC-BPB)+G1C*(TIME-BPB)/(BPC-BPB)
    G2=G2B*(BPC-TIME)/(BPC-BPB)+G2C*(TIME-BPB)/(BPC-BPB)
    G3=G3B*(BPC-TIME)/(BPC-BPB)+G3C*(TIME-BPB)/(BPC-BPB)

    ; population logits
    G1p=TVG1B*(BPC-TIME)/(BPC-BPB)+TVG1C*(TIME-BPB)/(BPC-BPB)
    G2p=TVG2B*(BPC-TIME)/(BPC-BPB)+TVG2C*(TIME-BPB)/(BPC-BPB)
    G3p=TVG3B*(BPC-TIME)/(BPC-BPB)+TVG3C*(TIME-BPB)/(BPC-BPB)
  ENDIF

  ;-------------------------------------------------
  ; typical transition probabilities as anti-logit
  ;-------------------------------------------------

  PAWp=1/(1+EXP(G1p)+EXP(G2p)+EXP(G3p))
  P1p=EXP(G1p)/(1+EXP(G1p)+EXP(G2p)+EXP(G3p))
  P2p=EXP(G2p)/(1+EXP(G1p)+EXP(G2p)+EXP(G3p))
  P3p=0
  PRp=EXP(G3p)/(1+EXP(G1p)+EXP(G2p)+EXP(G3p))

  ;-------------------------------------------------
  ; individual transition probabilities as anti-logit
  ;-------------------------------------------------

  PAW=1/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P1=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P2=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P3=0
  PR=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3))

  ;---------------------------------
  ; likelihood
  ;---------------------------------

  Y=0
  IF (STAG.EQ.0) Y=PAW
  IF (STAG.EQ.1) Y=P1
  IF (STAG.EQ.2) Y=P2
  IF (STAG.EQ.5) Y=PR

  IWRES=1

$THETA                ;LOGIT:1,2,3  BP:A,B,C  BP_TIME:a,b,c

  -1; TVG1A
  -5; TVG2A
  -10 FIX ; TVG3A

  -1; TVG1B
  -1; TVG2B
```

```
  -4; TVG3B

  -.1; TVG1C
  -2; TVG2C
  -2; TVG3C

  -2; STE1b
  -3; STE2b
  -5; STE3b

  -5; STE1c
  -10 FIX; STE2c
  -10 FIX; STE3c

  (0,.1,1);  REL BPB

  (0,.1,1);  BPsb
  (0,.4,1);  REL BP2

  -5; TVG11
  -10 FIX; TVG21
  -10 FIX; TVG31

  -5; TVG12
  -5; TVG22
  -10 FIX; TVG32

  -1; TVG13
  -10 FIX; TVG23
  -10 FIX; TVG33

$OMEGA BLOCK(2)

  .1   ;G1i
  -.02 .5   ;G2i

$OMEGA

  .5  ;G3i
  0 FIX  ;G1aw
  0 FIX  ;G2aw
  0 FIX  ;G3aw

$ESTIMATION METHOD=COND LAPLACE LIKE MSFO=msf1

$COVARIANCE MATRIX=R PRINT=E

$TAB ID VIS TIME STAG
     ETA(1) ETA(2) ETA(3)
     PAWp P1p P2p P3p PRp PAW P1 P2 P3 PR PAWpm P1pm P2pm P3pm PRpm
     ONEHEADER NOPRINT NOAPPEND FILE=sdtab1

$TAB ID ETA(1) ETA(2) ETA(3)
     ONEHEADER NOPRINT NOAPPEND FILE=patab1
```

A.2.'Simulator'

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
; NONMEM control file for the simulation of new datasets
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

$PROB SHmf_296_292_288_297_290p1_34

$INPUT ID DV=VIS TIME DROP DROP DROP DROP DROP DROP MDV
       DROP DROP DROP DROP DROP DROP EVID

  ; TIME = epoch number (between 1 and 960)

$DATA treatment0Visit33MAD103894forNONMEMallTimes34tSTevid.csv IGNORE=@

$ABBREVIATED DERIV2=NO

$PRED                   ;LOGIT:1,2,3  BP:A,B,C  BP_STT:sa,sb,sc
                        ;STT=stage_time  STE=stage_effect
  REP=IREP

  AA=THETA(1)+ETA(1)+ETA(2)+ETA(3)+ETA(4)+ETA(5)+ETA(6)+ETA(7)+ETA(8)
  AA=AA+ETA(9)+ETA(10)+ETA(11)+ETA(12)+ETA(13)+ETA(14)+ETA(15)+ETA(16)

  MDV0=1
  MDV1=1
  MDV2=1
  MDV3=1
  MDV5=1

  IF (TIME.EQ.1) THEN
    PST=-1                      ;previous stage
    PSTT=-10                    ;previous stage time
    PSL=0                       ;previous sleep
    SL=0                        ;sleep
    IS=0                        ;prima epoca con SL==1
  ENDIF

  SL=0
  IF (PSL.EQ.0.AND.PST.GT.0) SL=1
  IF (PSL.EQ.1) SL=1
  IF (PSL.EQ.0.AND.SL.EQ.1) IS=TIME

  STT=PSTT+1       ; stage_time

  IF (PST.EQ.0) MDV0=0
  IF (PST.EQ.1) MDV1=0
  IF (PST.EQ.2) MDV2=0
  IF (PST.EQ.3) MDV3=0
  IF (PST.EQ.5) MDV5=0

  ; break-points
  BPA=2
  BPB=960
  BPC=960
  BPsa=1
  BP1=1
  BP2=2
```

```
BP3=1

; stage_time effect
STE4b=1
STE4c=1

; for stage2 sub-model
TVG4A=1
TVG4B=1
TVG4C=1
DEV4=1

; for awake sub-model
TVG11=1
TVG21=1
TVG31=1
TVG12=1
TVG22=1
TVG32=1
TVG13=1
TVG23=1
TVG33=1

;-------------------------------------------------------------
;  if the current state is AW the awake sub-model is taken
;  into account.
;  Initialization using the previously estimated THETAs
;-------------------------------------------------------------

IF (MDV0.EQ.0) THEN
  BPA=IS       ; first epoch with SL==1


  BPB=(BPC-BPA)*0.0679+BPA
  BPsb==(BPsc-BPsa)*0.024+BPsa
  BPsc=265   ; max stage_time

  BP1=2
  BP3=371              ;MAX initial sleeplessness
  BP2=(BP3-BP1)*0.0298+BP1

  STE1b=-2.49
  STE2b=-3.59
  STE3b=-5.67

  STE1c=-6.63
  STE2c=-10
  STE3c=-10

  TVG1A=-0.251
  TVG2A=-2.66
  TVG3A=-10
  TVG1B=-0.209
  TVG2B=-1.01
  TVG3B=-2.86
  TVG1C=-0.203
  TVG2C=-2.08
  TVG3C=-2.1

  TVG11=-5.76
  TVG21=-10
```

```
  TVG31=-10
  TVG12=-3.93
  TVG22=-7.63
  TVG32=-10
  TVG13=-4.86
  TVG23=-10
  TVG33=-10

  DEV1=ETA(1)
  DEV2=ETA(2)
  DEV3=ETA(3)
ENDIF

; stage1 sub-model
IF (MDV1.EQ.0) THEN
  BPB=268

  BPsb=3.24
  BPsc=20

  STE1b=-0.447
  STE2b=-0.52
  STE3b=-0.463

  STE1c=-0.634
  STE2c=-0.246
  STE3c=-3.31

  TVG1A=-0.321
  TVG2A=-0.0716
  TVG3A=-4.15
  TVG1B=-1.07
  TVG2B=0.492
  TVG3B=-1.01
  TVG1C=-1.05
  TVG2C=-0.251
  TVG3C=-1.24

  DEV1=ETA(4)
  DEV2=ETA(5)
  DEV3=ETA(6)
ENDIF

; stage2 sub-model
IF (MDV2.EQ.0) THEN
  BPB=676

  BPsb=5.62
  BPsc=143

  STE1b=-0.905
  STE2b=-0.894
  STE3b=-1.19
  STE4b=-0.993

  STE1c=-1.79
  STE2c=-6.44
  STE3c=0.927
  STE4c=-6.75

  TVG1A=-2.46
```

```
  TVG2A=-2.52
  TVG3A=-1.71
  TVG4A=-4.07
  TVG1B=-2.57
  TVG2B=-2.6
  TVG3B=-3.33
  TVG4B=-3.34
  TVG1C=-2.13
  TVG2C=-2.33
  TVG3C=-4.59
  TVG4C=-3.16

  DEV1=ETA(7)
  DEV2=ETA(8)
  DEV3=ETA(9)
  DEV4=ETA(10)
ENDIF

; sws sub-model
IF (MDV3.EQ.0) THEN
  BPB=715

  BPsb=5.9
  BPsc=103

  STE1b=-0.999
  STE2b=-0.939
  STE3b=-2.5

  STE1c=0.761
  STE2c=-3.53
  STE3c=-3.73

  TVG1A=-3.22
  TVG2A=-4.83
  TVG3A=-0.526
  TVG1B=-3.31
  TVG2B=-5.13
  TVG3B=0.142
  TVG1C=-1.65
  TVG2C=-4.55
  TVG3C=0.389

  DEV1=ETA(11)
  DEV2=ETA(12)
  DEV3=ETA(13)
ENDIF

; rem sub-model
IF (MDV5.EQ.0) THEN
  BPB=640

  BPsb=13.6
  BPsc=100

  STE1b=0.351
  STE2b=-0.238
  STE3b=-1.22

  STE1c=0.566
  STE2c=-1.22
```

```
    STE3c=1.8

    TVG1A=-3.12
    TVG2A=-2.87
    TVG3A=-3.11
    TVG1B=-3.25
    TVG2B=-3.02
    TVG3B=-4.6
    TVG1C=-2.96
    TVG2C=-2.98
    TVG3C=-4.56

    DEV1=ETA(14)
    DEV2=ETA(15)
    DEV3=ETA(16)
  ENDIF

  G1A=0
  G2A=0
  G3A=0
  G4A=0

  ;---------------------------------
  ; stage time effect interpolation
  ;---------------------------------

  IF (STT.GE.0.AND.STT.LE.BPsb.AND.SL.EQ.1) THEN
    G1A=(TVG1A+DEV1)*(BPsb-STT)/(BPsb-BPsa)+(TVG1A+DEV1+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2A=(TVG2A+DEV2)*(BPsb-STT)/(BPsb-BPsa)+(TVG2A+DEV2+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3A=(TVG3A+DEV3)*(BPsb-STT)/(BPsb-BPsa)+(TVG3A+DEV3+STE3b)*(STT-
BPsa)/(BPsb-BPsa)
    G4A=(TVG4A+DEV4)*(BPsb-STT)/(BPsb-BPsa)+(TVG4A+DEV4+STE4b)*(STT-
BPsa)/(BPsb-BPsa)

    G1B=(TVG1B+DEV1)*(BPsb-STT)/(BPsb-BPsa)+(TVG1B+DEV1+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2B=(TVG2B+DEV2)*(BPsb-STT)/(BPsb-BPsa)+(TVG2B+DEV2+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3B=(TVG3B+DEV3)*(BPsb-STT)/(BPsb-BPsa)+(TVG3B+DEV3+STE3b)*(STT-
BPsa)/(BPsb-BPsa)
    G4B=(TVG4B+DEV4)*(BPsb-STT)/(BPsb-BPsa)+(TVG4B+DEV4+STE4b)*(STT-
BPsa)/(BPsb-BPsa)

    G1C=(TVG1C+DEV1)*(BPsb-STT)/(BPsb-BPsa)+(TVG1C+DEV1+STE1b)*(STT-
BPsa)/(BPsb-BPsa)
    G2C=(TVG2C+DEV2)*(BPsb-STT)/(BPsb-BPsa)+(TVG2C+DEV2+STE2b)*(STT-
BPsa)/(BPsb-BPsa)
    G3C=(TVG3C+DEV3)*(BPsb-STT)/(BPsb-BPsa)+(TVG3C+DEV3+STE3b)*(STT-
BPsa)/(BPsb-BPsa)
    G4C=(TVG4C+DEV4)*(BPsb-STT)/(BPsb-BPsa)+(TVG4C+DEV4+STE4b)*(STT-
BPsa)/(BPsb-BPsa)
  ENDIF

  IF (STT.GT.BPsb.AND.SL.EQ.1) THEN
    G1A=(TVG1A+DEV1+STE1b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG1A+DEV1+STE1c)*(STT-BPsb)/(BPsc-BPsb)
    G2A=(TVG2A+DEV2+STE2b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG2A+DEV2+STE2c)*(STT-BPsb)/(BPsc-BPsb)
```

```
    G3A=(TVG3A+DEV3+STE3b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG3A+DEV3+STE3c)*(STT-BPsb)/(BPsc-BPsb)
    G4A=(TVG4A+DEV4+STE4b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG4A+DEV4+STE4c)*(STT-BPsb)/(BPsc-BPsb)

    G1B=(TVG1B+DEV1+STE1b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG1B+DEV1+STE1c)*(STT-BPsb)/(BPsc-BPsb)
    G2B=(TVG2B+DEV2+STE2b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG2B+DEV2+STE2c)*(STT-BPsb)/(BPsc-BPsb)
    G3B=(TVG3B+DEV3+STE3b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG3B+DEV3+STE3c)*(STT-BPsb)/(BPsc-BPsb)
    G4B=(TVG4B+DEV4+STE4b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG4B+DEV4+STE4c)*(STT-BPsb)/(BPsc-BPsb)

    G1C=(TVG1C+DEV1+STE1b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG1C+DEV1+STE1c)*(STT-BPsb)/(BPsc-BPsb)
    G2C=(TVG2C+DEV2+STE2b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG2C+DEV2+STE2c)*(STT-BPsb)/(BPsc-BPsb)
    G3C=(TVG3C+DEV3+STE3b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG3C+DEV3+STE3c)*(STT-BPsb)/(BPsc-BPsb)
    G4C=(TVG4C+DEV4+STE4b)*(BPsc-STT)/(BPsc-
BPsb)+(TVG4C+DEV4+STE4c)*(STT-BPsb)/(BPsc-BPsb)
  ENDIF

;---------------------------------
; logits interpolation
;---------------------------------

  G1=G1A
  G2=G2A
  G3=G3A
  G4=G4A

  IF (TIME.GE.BPA.AND.TIME.LE.BPB.AND.SL.EQ.1) THEN
    G1=G1A*(BPB-TIME)/(BPB-BPA)+G1B*(TIME-BPA)/(BPB-BPA)
    G2=G2A*(BPB-TIME)/(BPB-BPA)+G2B*(TIME-BPA)/(BPB-BPA)
    G3=G3A*(BPB-TIME)/(BPB-BPA)+G3B*(TIME-BPA)/(BPB-BPA)
    G4=G4A*(BPB-TIME)/(BPB-BPA)+G4B*(TIME-BPA)/(BPB-BPA)
  ENDIF
  IF (TIME.GT.BPB.AND.TIME.LE.BPC.AND.SL.EQ.1) THEN
    G1=G1B*(BPC-TIME)/(BPC-BPB)+G1C*(TIME-BPB)/(BPC-BPB)
    G2=G2B*(BPC-TIME)/(BPC-BPB)+G2C*(TIME-BPB)/(BPC-BPB)
    G3=G3B*(BPC-TIME)/(BPC-BPB)+G3C*(TIME-BPB)/(BPC-BPB)
    G4=G4B*(BPC-TIME)/(BPC-BPB)+G4C*(TIME-BPB)/(BPC-BPB)
  ENDIF

  IF (TIME.GE.BP1.AND.TIME.LE.BP2.AND.SL.EQ.0) THEN
    G1=(TVG11+DEV1)*(BP2-TIME)/(BP2-BP1)+(TVG12+DEV1)*(TIME-BP1)/(BP2-
BP1)
    G2=(TVG21+DEV2)*(BP2-TIME)/(BP2-BP1)+(TVG22+DEV2)*(TIME-BP1)/(BP2-
BP1)
    G3=(TVG31+DEV3)*(BP2-TIME)/(BP2-BP1)+(TVG32+DEV3)*(TIME-BP1)/(BP2-
BP1)
  ENDIF
  IF (TIME.GT.BP2.AND.SL.EQ.0) THEN
    G1=(TVG12+DEV1)*(BP3-TIME)/(BP3-BP2)+(TVG13+DEV1)*(TIME-BP2)/(BP3-
BP2)
    G2=(TVG22+DEV2)*(BP3-TIME)/(BP3-BP2)+(TVG23+DEV2)*(TIME-BP2)/(BP3-
BP2)
    G3=(TVG32+DEV3)*(BP3-TIME)/(BP3-BP2)+(TVG33+DEV3)*(TIME-BP2)/(BP3-
BP2)
```

```
ENDIF

;----------------------------------------
; transition probabilities as anti-logits
;----------------------------------------

PAW=2
P1=0
P2=0
P3=0

; awake
IF (MDV0.EQ.0) THEN
  PAW=1/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P1=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P2=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P3=0
  PR=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3))
ENDIF

; stage1
IF (MDV1.EQ.0) THEN
  PAW=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P1=1/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P2=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P3=0
  PR=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3))
ENDIF

; stage2
IF (MDV2.EQ.0) THEN
  PAW=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3)+EXP(G4))
  P1=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3)+EXP(G4))
  P2=1/(1+EXP(G1)+EXP(G2)+EXP(G3)+EXP(G4))
  P3=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3)+EXP(G4))
  PR=EXP(G4)/(1+EXP(G1)+EXP(G2)+EXP(G3)+EXP(G4))
ENDIF

; sws
IF (MDV3.EQ.0) THEN
  PAW=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P1=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P2=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P3=1/(1+EXP(G1)+EXP(G2)+EXP(G3))
  PR=0
ENDIF

; rem
IF (MDV5.EQ.0) THEN
  PAW=EXP(G1)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P1=EXP(G2)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P2=EXP(G3)/(1+EXP(G1)+EXP(G2)+EXP(G3))
  P3=0
  PR=1/(1+EXP(G1)+EXP(G2)+EXP(G3))
ENDIF

; probability scale
DEV1=PAW+P1
DEV2=DEV1+P2
DEV3=DEV2+P3
```

```
  ; Initialization of ST; first epoch is awake
  ST=0    ;stage

  AW=0
  ST1=0
  ST2=0
  ST3=0
  REM=0

  ; random call for the simulation step
  IF (ICALL.EQ.4) THEN
    CALL RANDOM (2,R)

    ; if R<PAW transition to AW is taken into account
    ; and the following stage is AW(0)
    IF (R.LT.PAW) THEN
      ST=0
      AW=1

    ; if PAW<=R<PAW+P1 transition to ST1 is taken into account
    ; and the following stage is ST1(1)
    ELSEIF (R.LT.DEV1.AND.R.GE.PAW) THEN
      ST=1
      ST1=1

    ; if PAW+P1<=R<PAW+P1+P2 transition to ST2 is taken into account
    ; and the following stage is ST2(2)
    ELSEIF (R.LT.DEV2.AND.R.GE.DEV1) THEN
      ST=2
      ST2=1

    ; if PAW+P1+P2<=R<PAW+P1+P2+P3 transition to SWS is
    ; taken into account and the following stage is SWS(3)
    ELSEIF (R.LT.DEV3.AND.R.GE.DEV2) THEN
      ST=3
      ST3=1

    ; if R>=PAW+P1+P2+P3 transition to REM is
    ; taken into account and the following stage is REM(5)
    ELSE
      ST=5
      REM=1
    ENDIF
  ENDIF

  CH=1                          ;change
  IF (ST.EQ.PST) CH=0                ; no change
  PSTT=STT*(CH-1)**2

  IF (TIME.EQ.1) STT=0
  PSL=SL
  PST=ST

$THETA 1

;AW
$OMEGA BLOCK(2) .134 -0.142 .831
$OMEGA 1.07
;ST1
$OMEGA BLOCK(3) .318 .18 .389 -0.0637 .138 .398
;ST2
```

```
$OMEGA .152 .456 .786 .239
;ST3
$OMEGA (0 FIX) 1.4 1.35
;REM
$OMEGA .584 .849 1.1

$SIM (123456) (123 UNIFORM) ONLY SUBPROBS=100 NOPREDICTION

$TAB ID VIS TIME ST MDV0 MDV1 MDV2 MDV3 MDV5 MDV
     AW ST1 ST2 ST3 REM STT EVID SL IS REP
     ONEHEADER NOPRINT NOAPPEND FILE=simulD1

$TAB ID TIME REP ETA(1) ETA(2) ETA(3) ETA(4) ETA(5) ETA(6) ETA(7) ETA(8)
    ETA(9) ETA(10) ETA(11) ETA(12) ETA(13) ETA(14) ETA(15) ETA(16)
     ONEHEADER NOPRINT NOAPPEND FILE=simulE1
```

Appendix B

R codes

B.1. Posterior Predictive Check (PPC)

```
DirData <- "F:\\eem52253\\tesi\\Dati\\"
DirSim <- 'F:\\eem52253\\tesi\\Dati\\simulation\\'

# raw dataset
d<-read.csv(paste(DirData,
            "treatment0Visit33MAD103894forNONMEMallTimes34.csv",sep=''))

ID<-unique(d$ID)
N<-length(ID)

kmax<-100
kmax<-kmax+1

# Initialization
LPS<-matrix(0,nrow=kmax,ncol=N)
WASO<-matrix(0,nrow=kmax,ncol=N)
TST<-matrix(0,nrow=kmax,ncol=N)
tAW<-matrix(0,nrow=kmax,ncol=N)
tREM<-matrix(0,nrow=kmax,ncol=N)
tNREM<-matrix(0,nrow=kmax,ncol=N)
tST1<-matrix(0,nrow=kmax,ncol=N)
tST2<-matrix(0,nrow=kmax,ncol=N)
tSWS<-matrix(0,nrow=kmax,ncol=N)
SE1<-matrix(0,nrow=kmax,ncol=N)
SE2<-matrix(0,nrow=kmax,ncol=N)
SE3<-matrix(0,nrow=kmax,ncol=N)
SE4<-matrix(0,nrow=kmax,ncol=N)
meanAW<-matrix(0,nrow=kmax,ncol=N)
meanST1<-matrix(0,nrow=kmax,ncol=N)
meanST2<-matrix(0,nrow=kmax,ncol=N)
meanSWS<-matrix(0,nrow=kmax,ncol=N)
meanREM<-matrix(0,nrow=kmax,ncol=N)
nAW<-matrix(0,nrow=kmax,ncol=N)
nST1<-matrix(0,nrow=kmax,ncol=N)
nST2<-matrix(0,nrow=kmax,ncol=N)
nSWS<-matrix(0,nrow=kmax,ncol=N)
nREM<-matrix(0,nrow=kmax,ncol=N)


for (k in 1:kmax) {
  if (k==kmax)
 {
 d<-read.csv(paste(DirData,
```

```
            "treatment0Visit33MAD103894forNONMEMallTimes34.csv",sep=''))
 }
else
 {
    d<-read.csv(paste(DirSim,
            "trt0visit3simulatoForNONMEM34data",k,".csv",sep=""))
  }

  n<-nrow(d)
  ID<-unique(d$ID)
  N<-length(ID)

  for (i in 1:N)
    {
    d1<-d[d$ID==ID[i],c(1,3,4)]
    ll <- nrow(d1)

    d1$SLEEP1<-NA                #=1 from LPS
    d1$SLEEP<-rep(0,ll)

    d1$STsucc1 <- rep(0,ll)
    d1$STsucc2 <- rep(0,ll)
    d1$STsucc3 <- rep(0,ll)
    d1$STsucc4 <- rep(0,ll)
    d1$STsucc5 <- rep(0,ll)
    d1$STsucc6 <- rep(0,ll)
    d1$STsucc7 <- rep(0,ll)
    d1$STsucc8 <- rep(0,ll)
    d1$STsucc9 <- rep(0,ll)
    d1$STsucc10 <- rep(0,ll)
    d1$STsucc11 <- rep(0,ll)
    d1$STsucc12 <- rep(0,ll)
    d1$STsucc13 <- rep(0,ll)
    d1$STsucc14 <- rep(0,ll)
    d1$STsucc15 <- rep(0,ll)
    d1$STsucc16 <- rep(0,ll)
    d1$STsucc17 <- rep(0,ll)
    d1$STsucc18 <- rep(0,ll)
    d1$STsucc19 <- rep(0,ll)

    d1$STsucc1[1:(ll-1)] <- d1$STAGE[2:ll]
    d1$STsucc2[1:(ll-2)] <- d1$STAGE[3:ll]
    d1$STsucc3[1:(ll-3)] <- d1$STAGE[4:ll]
    d1$STsucc4[1:(ll-4)] <- d1$STAGE[5:ll]
    d1$STsucc5[1:(ll-5)] <- d1$STAGE[6:ll]
    d1$STsucc6[1:(ll-6)] <- d1$STAGE[7:ll]
    d1$STsucc7[1:(ll-7)] <- d1$STAGE[8:ll]
    d1$STsucc8[1:(ll-8)] <- d1$STAGE[9:ll]
    d1$STsucc9[1:(ll-9)] <- d1$STAGE[10:ll]
    d1$STsucc10[1:(ll-10)] <- d1$STAGE[11:ll]
    d1$STsucc11[1:(ll-11)] <- d1$STAGE[12:ll]
    d1$STsucc12[1:(ll-12)] <- d1$STAGE[13:ll]
    d1$STsucc13[1:(ll-13)] <- d1$STAGE[14:ll]
    d1$STsucc14[1:(ll-14)] <- d1$STAGE[15:ll]
    d1$STsucc15[1:(ll-15)] <- d1$STAGE[16:ll]
    d1$STsucc16[1:(ll-16)] <- d1$STAGE[17:ll]
    d1$STsucc17[1:(ll-17)] <- d1$STAGE[18:ll]
    d1$STsucc18[1:(ll-18)] <- d1$STAGE[19:ll]
    d1$STsucc19[1:(ll-19)] <- d1$STAGE[20:ll]

    # at least 20 consecutive 'non-awake' stages
```

```
d1$SLEEP1 <- (d1$STAGE>0)*(d1$STsucc1>0)*(d1$STsucc2>0)*
        (d1$STsucc3>0)*(d1$STsucc4>0)*(d1$STsucc5>0)*(d1$STsucc6>0)*
       (d1$STsucc7>0)*(d1$STsucc8>0)*(d1$STsucc9>0)*(d1$STsucc10>0)*
    (d1$STsucc11>0)*(d1$STsucc12>0)*(d1$STsucc13>0)*(d1$STsucc14>0)*
    (d1$STsucc15>0)*(d1$STsucc16>0)*(d1$STsucc17>0)*(d1$STsucc18>0)*
    (d1$STsucc19>0)

#-----------------------------
# Latency to Persistent Sleep
#-----------------------------

LPS[k,i]<-min(min(which(d1$SLEEP1>0))-1,960)

if (LPS[k,i]<960)
    {
    d1$SLEEP1[(LPS[k,i]+2):ll]<-1
    }
pos<-min((which(d1$STAGE>0)),961)

if (pos<=ll)
    {
    # SLEEP=1 from the first epoch of non-awake
    d1$SLEEP[(pos):ll]<-1
    }

#-----------------------------
# Wake After Sleep Onset
#-----------------------------

WASO[k,i]<-sum(d1$STAGE==0 & d1$SLEEP1==1)

#-----------------------------
# Total Sleep Time
#-----------------------------

TST[k,i]<-sum(d1$STAGE!=0)


#----------------------------------
# Total time in each stage
#----------------------------------

tAW[k,i]<-sum(d1$STAGE==0)
tNREM[k,i]<-sum(d1$STAGE<5 & d1$STAGE>0)
tREM[k,i]<-sum(d1$STAGE==5)
tST1[k,i]<-sum(d1$STAGE==1)
tST2[k,i]<-sum(d1$STAGE==2)
tSWS[k,i]<-sum(d1$STAGE==3)

#-----------------------------------------------------------
# Sleep efficiency on 2-hour time intervals of the night
#-----------------------------------------------------------

SE1[k,i]<-sum(d1$STAGE!=0 & d1$Time<241)/240
SE2[k,i]<-sum(d1$STAGE!=0 & d1$Time%in%(241:480))/240
SE3[k,i]<-sum(d1$STAGE!=0 & d1$Time%in%(481:720))/240
SE4[k,i]<-sum(d1$STAGE!=0 & d1$Time%in%(721:960))/240

# mean extension of awake (after sleep onset)
p<-which(d1$STAGE==0 & d1$SLEEP1==1)
if (length(p)>0){
```

```
  p1<-c(0,1:nrow(d1),(nrow(d1)+1))
  p1<-p1[-(p+1)]
  pdelta<-p1[2:length(p1)]-p1[1:(length(p1)-1)]-1
  meanAW[k,i]<-mean(pdelta[pdelta!=0])
} else {
  meanAW[k,i]<-0
}

# mean extension of stage 1
p<-which(d1$STAGE==0 & d1$SLEEP1==1)
if (length(p)>0){
  p1<-c(0,1:nrow(d1),(nrow(d1)+1))
  p1<-p1[-(p+1)]
  pdelta<-p1[2:length(p1)]-p1[1:(length(p1)-1)]-1
  meanST1[k,i]<-mean(pdelta[pdelta!=0])
} else {
  meanST1[k,i]<-0
}

# mean extension of stage 2
    p<-which(d1$STAGE==0 & d1$SLEEP1==1)
if (length(p)>0){
  p1<-c(0,1:nrow(d1),(nrow(d1)+1))
  p1<-p1[-(p+1)]
  pdelta<-p1[2:length(p1)]-p1[1:(length(p1)-1)]-1
  meanST2[k,i]<-mean(pdelta[pdelta!=0])
} else {
  meanST2[k,i]<-0
}

# mean extension of sws
p<-which(d1$STAGE==0 & d1$SLEEP1==1)
if (length(p)>0){
  p1<-c(0,1:nrow(d1),(nrow(d1)+1))
  p1<-p1[-(p+1)]
  pdelta<-p1[2:length(p1)]-p1[1:(length(p1)-1)]-1
  meanSWS[k,i]<-mean(pdelta[pdelta!=0])
} else {
  meanSWS[k,i]<-0
}

# mean extension of rem
    p<-which(d1$STAGE==0 & d1$SLEEP1==1)
if (length(p)>0){
  p1<-c(0,1:nrow(d1),(nrow(d1)+1))
  p1<-p1[-(p+1)]
  pdelta<-p1[2:length(p1)]-p1[1:(length(p1)-1)]-1
  meanREM[k,i]<-mean(pdelta[pdelta!=0])
} else {
  meanREM[k,i]<-0
}


#-------------------------------------
# Number of transitions to each stage
#-------------------------------------

nAW[k,i]<-length(which(d1$STAGE!=0 & d1$STsucc1==0))
nST1[k,i]<-length(which(d1$STAGE!=1 & d1$STsucc1==1))
nST2[k,i]<-length(which(d1$STAGE!=2 & d1$STsucc1==2))
nSWS[k,i]<-length(which(d1$STAGE!=3 & d1$STsucc1==3))
```

```
    nREM[k,i]<-length(which(d1$STAGE!=5 & d1$STsucc1==5))

  }
}
```

B.2. Visual Predictive Check (VPC)

```
#-------------------------------------------------------------------------
# VPC: transition from AWAKE.
#-------------------------------------------------------------------------

library(lattice)
library(vdi)

DirData <- "F:\\eem52253\\tesi\\Dati\\"
DirSim <- 'F:\\eem52253\\tesi\\Dati\\simulation\\'


# samples per subject
N <- 960

# Time [hours]
Time <- (1:N)/120

# n=101 raw dataset
for (n in (1:101))
      {
      if (n==101)
      {
      s<-read.csv(paste(DirData,
            "treatment0Visit33MAD103894forNONMEMallTimes34.csv",sep=''))
      }
      else
      {
      s<-read.csv(paste(DirSim,
            "trt0visit3simulatoForNONMEM34data",n,".csv",sep=""))
      }

      #-----------------------------------
      # Counting transitions at each epoch
      #-----------------------------------

      # shift(post)
      p <- rbind(s[2:nrow(s),],s[1,])

      d1<-matrix(nrow=960,ncol=51)
      d1 <- data.frame(d1)
      names(d1)<- c("times","AW","ST1","ST2","ST3","REM","TRaw1",
                    "TRaw2","TRaw3","TRawR","TR1aw","TR12","TR13",
                    "TR1R","TR2aw","TR21","TR23","TR2R","TR3aw",
                    "TR31","TR32","TR3R","TRRaw","TRR1","TRR2",
                    "TRR3")

      d1[,1]<-1:N

      # columns d1
      # 1 times
      # 2:6 AW,ST1,ST2,SWS,REM: number of occurrences of each stage
      # 7:26 number of transitions from each stage to each other


      for (i in(1:5))
        {
        d1[,1+i] <- tapply(s[,10+i],s$Time,sum)
```

```
  }

 # temporary data.frame
 d1t <- s[,c('Time','ID')]

 d1t$TRaw1 <- 0
 d1t$TRaw2 <- 0
 d1t$TRaw3 <- 0
 d1t$TRawR <- 0
 d1t$TR1aw <- 0
 d1t$TR12 <- 0
 d1t$TR13 <- 0
 d1t$TR1R <- 0
 d1t$TR2aw <- 0
 d1t$TR21 <- 0
 d1t$TR23 <- 0
 d1t$TR2R <- 0
 d1t$TR3aw <- 0
 d1t$TR31 <- 0
 d1t$TR32 <- 0
 d1t$TR3R <- 0
 d1t$TRRaw <- 0
 d1t$TRR1 <- 0
 d1t$TRR2 <- 0
 d1t$TRR3 <- 0

 # transitions for each subject and each Time

 d1t$TRaw1[s$AW==1 & p$ST1==1 & p$MDV==0] <- 1
 d1t$TRaw2[s$AW==1 & p$ST2==1 & p$MDV==0] <- 1
 d1t$TRaw3[s$AW==1 & p$ST3==1 & p$MDV==0] <- 1
 d1t$TRawR[s$AW==1 & p$REM==1 & p$MDV==0] <- 1
 d1t$TR1aw[s$ST1==1 & p$AW==1 & p$MDV==0] <- 1
 d1t$TR12[s$ST1==1 & p$ST2==1 & p$MDV==0] <- 1
 d1t$TR13[s$ST1==1 & p$ST3==1 & p$MDV==0] <- 1
 d1t$TR1R[s$ST1==1 & p$REM==1 & p$MDV==0] <- 1
 d1t$TR2aw[s$ST2==1 & p$AW==1 & p$MDV==0] <- 1
 d1t$TR21[s$ST2==1 & p$ST1==1 & p$MDV==0] <- 1
 d1t$TR23[s$ST2==1 & p$ST3==1 & p$MDV==0] <- 1
 d1t$TR2R[s$ST2==1 & p$REM==1 & p$MDV==0] <- 1
 d1t$TR3aw[s$ST3==1 & p$AW==1 & p$MDV==0] <- 1
 d1t$TR31[s$ST3==1 & p$ST1==1 & p$MDV==0] <- 1
 d1t$TR32[s$ST3==1 & p$ST2==1 & p$MDV==0] <- 1
 d1t$TR3R[s$ST3==1 & p$REM==1 & p$MDV==0] <- 1
 d1t$TRRaw[s$REM==1 & p$AW==1 & p$MDV==0] <- 1
 d1t$TRR1[s$REM==1 & p$ST1==1 & p$MDV==0] <- 1
 d1t$TRR2[s$REM==1 & p$ST2==1 & p$MDV==0] <- 1
 d1t$TRR3[s$REM==1 & p$ST3==1 & p$MDV==0] <- 1

 # number of transitions for each Time
 for (i in (1:20))
   {
   d1[,6+i] <- tapply(d1t[,2+i],s$Time,sum)
   }

#--------------------------------
# data.frame on 10 Time intervals
#--------------------------------

# number of intervals
n <- 10
```

```
    # samples per interval
    k <- N/n

    # final data.frame
    m <- matrix(nrow=10,ncol=51)
    m <- data.frame(m)

    # temporary data.frame
    d1t <- d1[1:959,]

    # factor indicating the interval number
    f <- rep(1:10,each=k)
    d1t$f <- f[-960]

    # summing each column of transitions on the time intervals
     for (i in (2:26))
       {
       m[,i] <- tapply(d1t[,i],d1t$f,sum)
       }

     # interval number
     m[,1] <- 1:10
     names(m) <-  c("times","AW","ST1","ST2","ST3","REM","TRaw1",
                    "TRaw2","TRaw3","TRawR","TR1aw","TR12","TR13",
                    "TR1R","TR2aw","TR21","TR23","TR2R","TR3aw",
                    "TR31","TR32","TR3R","TRRaw","TRR1","TRR2","TRR3",
                    "Paw1","Paw2","Paw3","PawR","P1aw","P12","P13",
                    "P1R","P2aw","P21","P23","P2R","P3aw","P31","P32",
                    "P3R","PRaw","PR1","PR2","PR3","Pawaw","P11",
                    "P22","P33","PRR")

     # columns m
     # 1 times
     # 2:6 AW,ST1,ST2,SWS,REM: number of occurrences on each interval
     # 7:26 number of transitions from each stage to each other in each
     #      interval
     # 27:51 transition frequencies on the intervals

     # computing transition frequencies on the 10 intervals
     m$tot<-1
     for (i in (1:5))
       {
       m$somma<-0
       for (j in (1:4))
           {
           k <- 2+4*i+j

           m[,20+k]<-round((m[,k]/m[,1+i]),4)
           m[is.nan(m[,20+k]),20+k]<-NA
           m$somma<-m$somma+m[,20+k]
           }
       m[,46+i] <-round(m$tot-m$somma,4)

       }
     # …
     # m is therefore memorized into a summary file
     # and iterations continue
}# end for n
```

B.3. Visual estimation check(VEC)

```
#-----------------------------------------------------------------------
# VEC on the transition probabilities. Model AWAKE
#-----------------------------------------------------------------------

library(MASS)

# Linear interpolation
lininterp <- function(v1,v2,graph=F) {

  lin <- function(x) {
    if (x[3]!=x[1]) {
      m <- (x[4]-x[2])/(x[3]-x[1])
      k <- (x[3]*x[2]-x[1]*x[4])/(x[3]-x[1])
      return(m*(ceiling(x[1]):floor(x[3]))+k)
    } else return(x[2])
  }

  vv <- c(matrix(c(v1,v2),ncol=length(v1),byrow=TRUE))
  ww <- seq(from=1,by=2,length=length(v1)-1)
  zz <- as.data.frame(mapply(':',ww,ww+3))
  tt <- lapply(zz,function(x) lin(vv[x]))

  # only if piece-wise linear on two intervals
  if (vv[zz[1,2]]==round( vv[zz[1,2]])) {
   tt$V2<-tt$V2[-1]
  }

  names(tt) <- NULL
  tt <- c(unlist(tt))

  return(tt)

}


## quantiles
q25 <- function(x)
    {
    return(quantile(x,0.025))
    }

q975 <- function(x)
    {
    return(quantile(x,0.975))
    }

q5 <- function(x)
    {
    return(quantile(x,0.05))
    }

q95 <- function(x)
    {
    return(quantile(x,0.95))
    }
```

```
## m = model number:
  ## 1: AW
  m <- 1

  DirSim <- 'D:\\eem52253\\tesi\\Dati\\simulation\\'
  DirDati <- "D:\\eem52253\\tesi\\Dati\\"

  DF <- read.csv(paste(DirSim,'summary',m,'.csv',sep=''))

  # successfully estimated datasets
  Yes <- DF$n[DF$su=='Yes']
  n <- length(Yes)

  # number of failed minimizations
  No <- DF$n[DF$su=='No']
  NumNo <- length(No)

  # initialization: thetas and omegas
  nth <- length(grep('TH',names(DF)))/2
  nom <- length(grep('OM',names(DF)))/2

  TH <- matrix(0,nrow=n,ncol=nth)
  OM <- matrix(0,nrow=n,ncol=nom)

  # initialization: typical transition probabilities
  PAp <- matrix(nrow=n,ncol=959)
  P1p <- matrix(nrow=n,ncol=959)
  P2p <- matrix(nrow=n,ncol=959)
  P3p <- matrix(nrow=n,ncol=959)
  PRp <- matrix(nrow=n,ncol=959)

  # initialization: 95 th percentile on the BSV
  PApi95 <- matrix(nrow=n,ncol=959)
  P1pi95 <- matrix(nrow=n,ncol=959)
  P2pi95 <- matrix(nrow=n,ncol=959)
  P3pi95 <- matrix(nrow=n,ncol=959)
  PRpi95 <- matrix(nrow=n,ncol=959)

  # initialization: 5 th percentile on the BSV
  PApi5 <- matrix(nrow=n,ncol=959)
  P1pi5 <- matrix(nrow=n,ncol=959)
  P2pi5 <- matrix(nrow=n,ncol=959)
  P3pi5 <- matrix(nrow=n,ncol=959)
  PRpi5 <- matrix(nrow=n,ncol=959)


for (i in 1:n)
{
 #i=n raw dataset
    # index of the simulated dataset
    ii <- Yes[i]

    # theta
    for (j in (1:nth))
        {
        TH[i,j]<-DF[DF$n==ii,paste('TH',j,sep='')]
        }# end for j

    # omega
    for (j in 1:nom)
```

```
      {
      OM[i,j]<-DF[DF$n==ii,paste('OM',j,sep='')]
      } # end for j

    # night time BREAK-POINTS after the initial sleeplessness
    BPA <- medIS    # where medIS is the median IS observed in the data
    BPC <- 960
    BPB <- (BPC-BPA)*TH[i,16]+BPA

    BP <- c(BPA,BPB,BPC)

    # Stage time BREAK-POINTS
    BPsa <- 1
    BPsc <- maxSTT # maxSTT is the maximum STT observed in the data
    BPsb <- (BPsc-BPsa)*TH[i,17]+BPsa

    BPs <- c(BPsa,BPsb,BPsc)

    # Initial Sleeplessness BREAK-POINTS
    BP1 <- 2
    BP3 <- maxIS # where maxIS is the maximum IS observed in the data
    BP2 <- (BP3-BP1)*TH[i,18]+BP1

    BPi <- c(BP1,BP2,BP3)

    # Logit
    G1p <- matrix(0,nrow=1,ncol=959)
    G2p <- matrix(0,nrow=1,ncol=959)
    G3p <- matrix(0,nrow=1,ncol=959)


    # Linear interpolation of the Stage time effect, MED is the median
    # STT observed in the data
    G1Ap <-
lininterp(BPs,c(TH[i,1],(TH[i,1]+TH[i,10]),(TH[i,1]+TH[i,13])))[MED]
    G1Bp <-
lininterp(BPs,c(TH[i,4],(TH[i,4]+TH[i,10]),(TH[i,4]+TH[i,13])))[MED]
    G1Cp <-
lininterp(BPs,c(TH[i,7],(TH[i,7]+TH[i,10]),(TH[i,7]+TH[i,13])))[MED]

    G2Ap <-
lininterp(BPs,c(TH[i,2],(TH[i,2]+TH[i,11]),(TH[i,2]+TH[i,14])))[MED]
    G2Bp <-
lininterp(BPs,c(TH[i,5],(TH[i,5]+TH[i,11]),(TH[i,5]+TH[i,14])))[MED]
    G2Cp <-
lininterp(BPs,c(TH[i,8],(TH[i,8]+TH[i,11]),(TH[i,8]+TH[i,14])))[MED]

    G3Ap <-
lininterp(BPs,c(TH[i,3],(TH[i,3]+TH[i,12]),(TH[i,3]+TH[i,15])))[MED]
    G3Bp <-
lininterp(BPs,c(TH[i,6],(TH[i,6]+TH[i,12]),(TH[i,6]+TH[i,15])))[MED]
    G3Cp <-
lininterp(BPs,c(TH[i,9],(TH[i,9]+TH[i,12]),(TH[i,9]+TH[i,15])))[MED]

    # Linear interpolation of the logits in the second part of the night
    # (after IS)
    G1p[(ceiling(BPA)-1):959] <- lininterp(BP,c(G1Ap,G1Bp,G1Cp))
    G2p[(ceiling(BPA)-1):959] <- lininterp(BP,c(G2Ap,G2Bp,G2Cp))
    G3p[(ceiling(BPA)-1):959] <- lininterp(BP,c(G3Ap,G3Bp,G3Cp))

    # linear interpolation if the logits in the first part of the night
```

```
    G1p[1:(ceiling(BPA)-2)] <-
lininterp(BPi,c(TH[i,19],TH[i,22],TH[i,25]))[1:(ceiling(BPA)-2)]
    G2p[1:(ceiling(BPA)-2)] <-
lininterp(BPi,c(TH[i,20],TH[i,23],TH[i,26]))[1:(ceiling(BPA)-2)]
    G3p[1:(ceiling(BPA)-2)] <-
lininterp(BPi,c(TH[i,21],TH[i,24],TH[i,27]))[1:(ceiling(BPA)-2)]


    # antilogit
    PAp[i,] <- 1/(1+exp(G1p)+exp(G2p)+exp(G3p))
    P1p[i,] <- exp(G1p)/(1+exp(G1p)+exp(G2p)+exp(G3p))
    P2p[i,] <- exp(G2p)/(1+exp(G1p)+exp(G2p)+exp(G3p))
    P3p[i,] <- 0
    PRp[i,] <- exp(G3p)/(1+exp(G1p)+exp(G2p)+exp(G3p))


    #-----------------------------------------------------------
    ## Between subject variability: 95% prediction intervals
    #-----------------------------------------------------------

    # simulation of Ns individuals
    Ns <- 1000

    # covariance matrix
    COV <- matrix(c(OM[i,1],OM[i,2],OM[i,2],OM[i,3]),nrow=2)

    BSV1 <- mvrnorm(Ns,rep(0,2),COV)[,1]
    BSV2 <- mvrnorm(Ns,rep(0,2),COV)[,2]
    BSV3 <- rnorm(Ns,0,sqrt(OM[i,4]))
    # BSV4 <- fixed to zero
    # BSV5 <- fixed to zero
    # BSV6 <- fixed to zero

    BSV1m <- matrix(rep(BSV1,each=(960-(ceiling(BPA)-1)),times=1),
                    nrow=Ns,byrow=T)
    BSV2m <- matrix(rep(BSV2,each=(960-(ceiling(BPA)-1)),times=1),
                    nrow=Ns,byrow=T)
    BSV3m <- matrix(rep(BSV3,each=(960-(ceiling(BPA)-1)),times=1),
                    nrow=Ns,byrow=T)


    # individual logits: first part of the night
    G11 <- matrix(rep(G1p[1:(ceiling(BPA)-2)],Ns),nrow=Ns,byrow=T)
    G21 <- matrix(rep(G2p[1:(ceiling(BPA)-2)],Ns),nrow=Ns,byrow=T)
    G31 <- matrix(rep(G3p[1:(ceiling(BPA)-2)],Ns),nrow=Ns,byrow=T)

    # individual logits: second part of the night
    G12 <- matrix(rep(G1p[(ceiling(BPA)-1):959],Ns),
                  nrow=Ns,byrow=T)+BSV1m
    G22 <- matrix(rep(G2p[(ceiling(BPA)-1):959],Ns),
                  nrow=Ns,byrow=T)+BSV2m
    G32 <- matrix(rep(G3p[(ceiling(BPA)-1):959],Ns),
                  nrow=Ns,byrow=T)+BSV3m

    # individual logits
    G1 <- cbind(G11,G12)
    G2 <- cbind(G21,G22)
    G3 <- cbind(G31,G32)

    # individual probabilities
    PA <- 1/(1+exp(G1)+exp(G2)+exp(G3))
```

```
    P1 <- exp(G1)/(1+exp(G1)+exp(G2)+exp(G3))
    P2 <- exp(G2)/(1+exp(G1)+exp(G2)+exp(G3))
    P3 <- matrix(0,nrow=Ns,ncol=959)
    PR <- exp(G3)/(1+exp(G1)+exp(G2)+exp(G3))

    # Between subject variability distribution: 90% prediction interval
    # 95th percentile
    PApi95[i,] <-  apply(PA,2,q95)
    P1pi95[i,] <-  apply(P1,2,q95)
    P2pi95[i,] <-  apply(P2,2,q95)
    P3pi95[i,] <-  apply(P3,2,q95)
    PRpi95[i,] <-  apply(PR,2,q95)

    # 5th percentile
    PApi5[i,] <-  apply(PA,2,q5)
    P1pi5[i,] <-  apply(P1,2,q5)
    P2pi5[i,] <-  apply(P2,2,q5)
    P3pi5[i,] <-  apply(P3,2,q5)
    PRpi5[i,] <-  apply(PR,2,q5)

} #end for i

  #------------------------
  ## confidence intervals
  #------------------------

  # Typical transition probabilities
  PAp_down <- apply(PAp[-n,],2,q25)
  P1p_down <- apply(P1p[-n,],2,q25)
  P2p_down <- apply(P2p[-n,],2,q25)
  P3p_down <- apply(P3p[-n,],2,q25)
  PRp_down <- apply(PRp[-n,],2,q25)

  PAp_up <- apply(PAp[-n,],2,q975)
  P1p_up <- apply(P1p[-n,],2,q975)
  P2p_up <- apply(P2p[-n,],2,q975)
  P3p_up <- apply(P3p[-n,],2,q975)
  PRp_up <- apply(PRp[-n,],2,q975)

  # median prediction
  PAp_med <- apply(PAp[-n,],2,median)
  P1p_med <- apply(P1p[-n,],2,median)
  P2p_med <- apply(P2p[-n,],2,median)
  P3p_med <- apply(P3p[-n,],2,median)
  PRp_med <- apply(PRp[-n,],2,median)

  # 95th percentile BSV
  PAci95_down <- apply(PAci95[-n,],2,q25)
  P1ci95_down <- apply(P1ci95[-n,],2,q25)
  P2ci95_down <- apply(P2ci95[-n,],2,q25)
  P3ci95_down <- apply(P3ci95[-n,],2,q25)
  PRci95_down <- apply(PRci95[-n,],2,q25)

  PAci95_up <- apply(PAci95[-n,],2,q975)
  P1ci95_up <- apply(P1ci95[-n,],2,q975)
  P2ci95_up <- apply(P2ci95[-n,],2,q975)
  P3ci95_up <- apply(P3ci95[-n,],2,q975)
  PRci95_up <- apply(PRci95[-n,],2,q975)

  # median prediction
  PAci95_med <- apply(PAci95[-n,],2,median)
```

```
P1ci95_med <- apply(P1ci95[-n,],2,median)
P2ci95_med <- apply(P2ci95[-n,],2,median)
P3ci95_med <- apply(P3ci95[-n,],2,median)
PRci95_med <- apply(PRci95[-n,],2,median)

# 5th percentile
PAci5_down <- apply(PAci5[-n,],2,q25)
P1ci5_down <- apply(P1ci5[-n,],2,q25)
P2ci5_down <- apply(P2ci5[-n,],2,q25)
P3ci5_down <- apply(P3ci5[-n,],2,q25)
PRci5_down <- apply(PRci5[-n,],2,q25)

PAci5_up <- apply(PAci5[-n,],2,q975)
P1ci5_up <- apply(P1ci5[-n,],2,q975)
P2ci5_up <- apply(P2ci5[-n,],2,q975)
P3ci5_up <- apply(P3ci5[-n,],2,q975)
PRci5_up <- apply(PRci5[-n,],2,q975)

# median prediction
PAci5_med <- apply(PAci5[-n,],2,median)
P1ci5_med <- apply(P1ci5[-n,],2,median)
P2ci5_med <- apply(P2ci5[-n,],2,median)
P3ci5_med <- apply(P3ci5[-n,],2,median)
PRci5_med <- apply(PRci5[-n,],2,median)
```

Bibliography

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

2. R. Bizzotto, S. Zamuner, G. De Nicolao, M. O. Karlsson, R. Gomeni. Multinomial logistic estimation of Markov-chain models for modeling sleep architecture in primary insomnia patients, *J Pharmacokinet Pharmacodyn,* **DOI**: 10.1007/s10928-009-9148-2 (2010).

3. M. C. Kjellsson. Methodological Studies on Models and Methods for Mixed-Effects Categorical Data Analysis. *Acta Universitatis Upsaliensis. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 83* (2008).

4. M. O. Karlsson and R. M. Savic. Diagnosing Model Diagnostics, *Clinical Pharmacology & Therapeutics* **82**, 17–20 (2007).

5. M. Bergstrand, A. C. Hooker, M. O. Karlsson. Visual Predictive Checks for Censored and Categorical data, *PAGE **18 Abstr 1604 [www.page-meeting.org/?abstract=1604](2009)***.

6. S. Jönsson, M. C. Kjellsson, M. O. Karlsson. Estimating Bias in Population Parameters for Some Models for Repeated Measures Ordinal Data Using NONMEM and NLMIXED, *J. Pharmacokinetics and Pharmacodynamics* **31(4)**:299-320 (2004).

7. V. Duval, M. O. Karlsson. Impact of Omission or Replacement of Data below the Limit of Quantification on Parameter Estimates in a Two-Compartment Model, *Pharmaceutical Research* **19(12)**:1835-1840 (2002).

8. J. K. Walsh. Clinical and socioeconomic correlates of insomnia, *J Clin. Psychiatry* **65(8)**: 13-19 (2004).

9. A. Y. Avidan. Insomnia in the geriatric patient, *Clin. Cornerstone* **5**: 51-60 (2003).

10. T. Roth and T. Reehrs. Insomnia: epidemiology, characteristics, and consequences, *Clin. Cornerstone* **5**: 5-15 (2003).

11. H. Berger. Ueber das elektroenkephalogramm des menschen, *J Psychol Neurol* **40**: 160–179 (1930).

12. A.L. Loomis, N. Harvey, G.A. Hobart. Cerebral states during sleep, as studied by human brain potentials, *J Exp Psychol* **21**: 127–144 (1937).

13. E. Aserinsky, N. Kleitman. Regularly occurring periods of eye motility, and concomitant Phenomena, *Science* **118**: 273–274 (1953).

14. Dement, N. Kleitman. Cyclic variations in EEG during sleep and their relation to eye movement, body motility and dreaming, *Electroencephalogr Clin Neurophysiol* **9**: 673-690 (1957) .

15. M. Jouvet, F. Michel, J. Courjon. Sur un stade d'activite electrique cerebrale rapide au cours du sommeil physiologique, *C R Soc Biol (Paris)* **153**: 1024–1028 (1959).

16. A. Rechtschaffen and A. Kales. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*, Public Health Service, US Government Printing Office, Washington DC (1968).

17. C. Iber, S. Ancoli-Israel, A. Chesson, S.F. Quan . *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification*. Westchester, Ill: The American Academy of Sleep Medicine (2007).

18. American Electroenchephalographic Society. Guideline Fifteen: Guidelines for Polygraphic Assessment of Sleep-Related Disorders (Polysomnography), *J Clin. Neurophysiol* **11(1)**: 116-124 (1994).

19. T. Akerstedt, M. Billiard, M. Bonnet, G. Ficca, L. Garma, M. Mariotti, P. Salzarulo, H. Shulz, Awakening from sleep. *Sleep Med Rev* **6**:267-286 (2002).

20. M. Hirshkowitz. Normal human sleep: an overview, *Med Clin North Am* **88**: 551-565 (2004).

21. S.H. Feinsilver. Sleep in the elderly. What is normal?, *Clin Geriatr Med* **19**: 177-188 (2003).

22. T. Roehrs. Sleep physiology and pathophysiology, *Clin Cornerstone* **2**: 1-15 (2000).

23. M.A. Carskadon, W.C. Dement, M.M. Mitler, T. Roth, P.R. Westbrook, S. Keenan. Guidelines for the Multiple Sleep Latency Test (MSLT): a standard measure of sleepiness, *Sleep* **9**:519–524 (1986).

24. J. Horne. *Why we sleep*. Oxford University Press, Oxford, UK (1988).

25. Dinges D. *Proving the limits of functional capability: the effects of sleep loss on short duration tasks in sleep, arousal, and performance*. In: Broughton RJ, Ogilvie RD, editors.

26. J. Aschoff. Circadian rhythms in man, *Science* **148**: 1427–1432 (1965).

27. A.A. Borbely, P. Achermann. Concepts and models of sleep regulation: an overview, *J Sleep Res* **1**: 61 (1992).

28. K. Gumustekin, B. Seven, N. Karabulut, O. Aktas, N. Gursan, S. Aslan, M. Keles, E. Varoglu, S. Dane. Effects of sleep deprivation, nicotine, and selenium on wound healing in rats. *J Neurosci* **114(11)**: 1433-1442 (2004).

29. T.H. Turner, S.P.A. Drummond, J.S. Salamat, G.G. Brown. Effects of 42 hr sleep deprivation on component processes of verbal working memory [Electronic version], *Neuropsychology,* **21**: 787-795 (2007).

30. J. Born, J. Rasch, S. Gais. Sleep to remember [Electronic version], *Neuroscientist,* **12**: 410 (2006).

31. S. Ancoli-Israel, T. Roth. Characteristics of insomnia in the United States: Results of the 1991 National Sleep Foundation Survey. I. *Sleep* **22(2)**: S347–S353 (1999).

32. G.E. Simon, M. VonKorff. Prevalence, burden, and treatment of insomnia in primary care, *Am J Psychiatry* **154**: 1417-1423 (1997).

33. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*, American Psychiatric Association, Washington DC (2000).

34. D. E. Ford, D. B. Kamerow. Epidemiologic study of sleep disturbances and psychiatric disorders: an opportunity fo prevention?, *JAMA* **262**: 1479-1484 (1989).

35. P.P. Chang, D.E. Ford, L.A. Mead et al. Insomnia in young men and subsequent depression: the Johns Hopkins Precursors Study, *AM J Epidemiol* **146**: 105-114 (1997).

36. J. K. Walsh, C.L. Engelhardt. The direct economic costs of insomnia in the United States for 1995, *Sleep* **22(2)**: S386-S393 (1999).

37. M.K. Stoller. Economic effects of insomnia,. *Clin Ther* **16**: 873-897 (1994).

38. R. Elie, E. Ruther, I. Farr et al. Sleep latency is shortened during 4 weeks of treatment with zaleplon, a novel nonbenzodiazepine hypnotic, *J Clin Psychiatry* **60**: 536-544 (1999).

39. G. Hajak, B. Bandelow. Safety and tolerance of zolpidem in the treatment of disturbed sleep: A post-marketing surveillance of 16,944 cases, *Int. Clin Psychopharmacol* **13**: 157-167 (1998).

40. B. M. Stone, C. Turner, S. L. Mills, I. Paty, A. Patat, M. Darwish, P. Danjou. Noise-induced sleep maintenance insomnia: hypnotic and residual effects of zaleplon, *Br J Clin Pharmacol*. **53(2)**: 196–202 (2002).

41. L.B. Sheiner, S.L. Beal. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical data, *J Pharmacokinet Biopharm* **8**: 553-571 (1980).

42. S. L. Beal and L. B. Sheiner. *NONMEM User's Guides*, NONMEM Project Group, San Francisco (1998).

43. G. G. Gregory and R. Cabeza. A two-state stochastic model of REM sleep architecture in the rat, *J. Neurophysol.* **88**:2589-2597 (2002).

44. B. Kemp and H. A. C. Kamphuisen. Simulation of human hypnograms using a Markov chain model, *Sleep* **9**:405-414 (1986).

45. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*, Chapman and Hall, London (1995).

46. R. M. Savic, M. O. Karlsson. Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions**,** *The AAPS journal* **11(3)**:558-69 (2009).